



HAL
open science

Acquisition de la prosodie du français par des apprenants sinophones : Études acoustique, perceptive et électroencéphalographique

Lei Xi

► To cite this version:

Lei Xi. Acquisition de la prosodie du français par des apprenants sinophones : Études acoustique, perceptive et électroencéphalographique. Linguistique. Université Sorbonne Nouvelle (Paris 3), 2025. Français. ⟨NNT : >. ⟨tel-05143136⟩

HAL Id: tel-05143136

<https://theses.hal.science/tel-05143136v1>

Submitted on 3 Jul 2025

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.



Copyright - All rights reserved



Université Sorbonne Nouvelle – CNRS
École Doctorale 622 – Sciences du Langage

Sorbonne
Nouvelle
université des cultures

UMR 7018 / Laboratoire de Phonétique et Phonologie



Thèse pour l'obtention du grade de
DOCTEUR EN PHONÉTIQUE

**Acquisition de la prosodie du français par des
apprenants sinophones : Études acoustique,
perceptive et électroencéphalographique**

Présentée par

Lei Xi

Sous la direction de

Rachid RIDOUANE

Soutenance à Paris le 26 juin 2025 devant le jury composé de :

Elisabeth DELAIS-ROUSSARIE	DR, CNRS (rapporteuse)
Frederic ISEL	PR, Université Paris Nanterre (examinateur)
Sharon PEPPERKAMP	DR, CNRS (rapporteuse)
Claire PILLOT-LOISEAU	PR, Université Sorbonne Nouvelle (examinatrice)
Rachid RIDOUANE	DR, CNRS (directeur de thèse)

L'acquisition de la prosodie du français par des apprenants sinophones : Études acoustique, perceptive et électroencéphalographique

Bien que centrale dans la communication parlée, la prosodie reste un domaine moins exploré dans les travaux de recherche sur l'acquisition des langues secondes. Cette thèse traite de la segmentation et de l'organisation prosodiques en français chez des locuteurs natifs francophones et des apprenants sinophones de trois niveaux (débutants, intermédiaires, avancés).

La prosodie assume notamment deux rôles clés : délimiter les unités syntaxiques et structurer le flux temporel-rythmique. Le mandarin, utilisant des tons lexicaux, mobilise différemment ces éléments, comparé au français. Notre recherche cherche à montrer comment les apprenants sinophones acquièrent la prosodie française pour traiter des ambiguïtés syntaxiques locales, illustrées par des phrases à clôture précoce (CP : « *Chaque fois que le serpent mangeait, le rat se cachait.* ») ou tardive (CT : « *Chaque fois que le serpent mangeait le rat, le lapin se cachait.* »). Pour cela, des données acoustiques (31 locuteurs : 10 natifs, 21 apprenants), perceptives (60 auditeurs : 20 natifs, 40 apprenants) et électroencéphalographiques (60 participants : 20 natifs, 40 apprenants) ont été analysées, totalisant 151 individus. Les résultats acoustiques révèlent que dès le niveau intermédiaire, les apprenants utilisent la montée du F0 et l'allongement final pour signaler les frontières CP et CT. Concernant l'organisation prosodique, les valeurs de neuf paramètres quantitatifs (%V, ΔC et ΔV , VarcoV et VarcoC, nPVI-V et rPVI-C, CCI-C et CCI-V) ont révélé une même tendance dans l'évolution du patron rythmique : il devient de plus en plus stable à mesure que le niveau de français progresse. Sur le plan perceptif, les apprenants (et, dans une moindre mesure, les natifs) rencontrent des difficultés accrues avec les structures complexes comme la CP. Les données EEG montrent que les frontières CP et CT activent une *closure positive shift* (CPS) chez tous les groupes, confirmant leur traitement neurocognitif. Une neuroplasticité est suggérée par les variations topographiques des CPS entre natifs et apprenants durant le traitement de la CP. La *ramp-like negativity*, observée pour la CT, reflète un processus *top-down* chez les natifs contre un mécanisme *bottom-up* chez les apprenants lors de l'analyse des indices acoustiques.

En conclusion, cette recherche contribue à la compréhension de l'acquisition prosodique en L2 et met en lumière les différences comportementale et neurocognitive dans le traitement des frontières prosodiques localement ambiguës entre natifs et apprenants.

Mots clés : français, chinois mandarin, prosodie, frontière prosodique, acquisition, acoustique, perception, EEG, potentiel évoqué, CPS

Acquisition of French prosody by Chinese learners: acoustic, perceptive and electroencephalographic studies

Despite its central role in spoken communication, prosody remains a less explored area in second language acquisition research. This thesis investigates performance of French prosodic segmentation and organization in native speakers and Chinese L2 learners across three proficiency levels: beginner, intermediate, and advanced. Prosody serves two primary functions in language, namely, delimiting syntactic units and regulating the temporal flow of speech. Mandarin, using lexical tones, mobilizes prosodic features differently from French. This thesis aims to investigate how Chinese L2 learners acquire French prosody to deal with local syntactic ambiguities, as exemplified by sentences exhibiting early closure (EC: « *Chaque fois que le serpent mangeait, le rat se cachait.* », "Whenever the snake was eating, the rat would hide.") and late closure (LC: « *Chaque fois que le serpent mangeait le rat, le lapin se cachait.* », "Whenever the snake was eating the rat, the rabbit would hide."). The analysis involved acoustic (31 subjects: 10 natives, 21 learners), perceptual (60 participants: 20 natives, 40 learners), and electroencephalographic (60 participants: 20 natives, 40 learners) data, encompassing a total of 151 individuals.

The acoustic results reveal that learners employ F0 rising and final lengthening to signal EC and LC boundaries from the intermediate level onward. In terms of prosodic organization, nine rhythm metrics (%V, ΔC and ΔV , VarcoV and VarcoC, nPVI-V and rPVI-C, CCI-C and CCI-V) show a consistent trend in rhythmic pattern development: it becomes increasingly more stable as proficiency in French advances. Perceptually, learners (and to a lesser extent, natives) encounter challenges with complex structures like EC. The EEG data indicate that both EC and LC boundaries elicited a closure positive shift (CPS) across all participant groups, underscoring their neurocognitive processing. Neuroplasticity is implied by the distinct topography of CPS between natives and learners during EC processing. The ramp-like negativity observed in LC condition reflects a top-down process in natives versus a bottom-up mechanism in learners when deciphering prosodic cues.

In summary, this thesis enhances our comprehension of L2 prosodic acquisition and provides novel perspectives on the behavioral and neurocognitive differences in the interpretation of local ambiguous prosodic boundaries between native French speakers and Chinese L2 learners.

Keywords: French, Mandarin Chinese, prosody, prosodic boundary, acquisition, acoustics, perception, EEG, ERP, CPS

À mes parents

Remerciement

Au terme de cette thèse, je tiens à exprimer ma profonde gratitude à toutes les personnes qui, de près ou de loin, ont contribué à l'aboutissement de ce projet. Depuis longtemps, je me rendais compte qu'une telle thèse, qui peut être présentée finalement sous vos yeux, n'est jamais le travail d'une seule personne.

Tout d'abord, je tiens à remercier profondément Rachid Ridouane, mon directeur de thèse, qui m'a donné l'opportunité de faire ce doctorat à ses côtés en 2018. Il m'a permis de travailler sur un projet qui me passionnait et de suivre la voie que je souhaitais emprunter. Cette thèse est le fruit d'une collaboration de plus de six années avec lui. C'est à ses côtés que j'ai appris ce que c'est la recherche scientifique et que j'ai saisi l'importance de la rigueur et de la précision dans ce domaine. Tout au long de cette thèse, à travers sa direction, il a su partager avec moi ses brillantes expertises dans ses domaines de recherche, avec la plus grande patience et la plus grande responsabilité.

Je le remercie pour son encadrement, sa gentillesse, sa disponibilité permanente et pour les nombreux conseils et relectures qu'il m'a prodigués. Je le remercie en particulier pour sa prévoyance académique. La partie EEG de cette thèse est née d'une simple phrase qu'il m'avait dite alors encore au 19 rue des Bernardins : « Lei, tu pourrais peut-être te pencher sur la technique EEG. ». Je n'avais alors qu'une idée vague sur cette technique. Mais depuis, nous avons réfléchi ensemble à chaque détail de ce projet neurophysiologique de ma thèse : la formation, la conception des stimuli, le choix de la plateforme, l'élaboration du protocole, et bien plus encore. Rachid Ridouane a déployé tous les moyens pour soutenir pleinement la concrétisation de ce projet : personnes, ressources, possibilités de financement, etc. Bien que colossal et chronophage, ce projet EEG a été mené à son terme et a abouti à des résultats significatifs. Grâce à l'expérience acquise dans le cadre de ce projet, je réalise aujourd'hui l'importance d'un tel investissement. Cela revient en grande à la prévoyance de mon directeur, et à son conseil d'intégrer l'EEG comme une composante importante de ma thèse.

Je tiens à exprimer ma gratitude à Frédéric Isel. Je garderai pour toujours dans ma mémoire les nombreux rendez-vous que nous avons eus ensemble au 2^{ème} étage de la Maison de la Recherche pour discuter de mes données EEG. Durant ces années, j'ai appris énormément de ses expertises en EEG/ERP, en neurolinguistique, et en sciences cognitives. Je me souviendrai toujours d'un moment marquant : malgré une fracture au pied survenue lors d'un jogging, Frédéric Isel est tout de même venu à notre rendez-vous, spécialement pour discuter de ma thèse. Je l'ai croisé en bas de la Maison de la Recherche, avançant lentement avec une canne en direction de l'entrée. Cette scène, émouvante, aura toujours sa place quelque part dans ma mémoire. En plus de son encadrement, je tiens à remercier Frédéric Isel, pour son accueil au sein du laboratoire MoDyCo (UMR 7114, Université Paris Nanterre), ce qui m'a permis de réaliser mon expérience EEG. Je remercie également Karin Heidlmayr, Caroline Bogliotti, Pierre Vigneron et Aikaterini Premeti, pour leur aide précieuse lors de l'implémentation et de la passation de cette expérience.

Je tiens à remercier Claire Pillot-Loiseau, qui m'a non seulement initié à la discipline de la phonétique et qui a également encadré mon premier projet de recherche dans

ce domaine. Je comptais parmi ses élèves lorsque j'étais en Master Sciences du Langage à la Sorbonne Nouvelle. Les connaissances que j'ai acquises lors de ses séminaires sur la didactique de la prononciation et sur l'acquisition phonétique en L2 ont constitué l'embryon du projet qui a donné naissance à cette thèse.

Je tiens aussi à remercier Prof. Dr. Petra Schumacher, Directrice du Laboratoire de Linguistique Expérimentale (XLinC Lab) de l'Université de Cologne, pour m'avoir accueilli pendant un mois au sein de son laboratoire. Je remercie également Dr. Ingmar Brilmayer, Mme Claudia Kilter et Mme Brita Rietdorf du XLinC Lab, qui m'ont formé, via zoom ou sur place à Cologne, à la manipulation de l'appareil d'électroencéphalographie, au traitement des données, ainsi qu'à l'élaboration du protocole de ma propre expérience. C'est grâce à ce séjour que j'ai pu devenir aujourd'hui un jeune chercheur autonome capable de mener une étude EEG et son analyse de A à Z. Cela sera sans aucun doute très bénéfique pour la suite de mon parcours académique.

Je tiens à remercier Mme Elisabeth Delais-Roussarie, Mme Sharon Peperkamp, Mme Claire Pillot-Loiseau, et M. Frederic Isel pour l'honneur qu'ils me font d'être dans mon jury de thèse.

Mes remerciements vont aussi aux 172 participants (dont 21 pilotes) qui ont accepté de participer volontairement aux différentes expériences menées dans le cadre de cette thèse. Sans leur contribution, il aurait été impossible de constituer l'ensemble des données à la fois riche et exhaustif qui constitue le fondement de ma thèse.

Je tiens à remercier le Labex EFL, qui m'a soutenu à travers une bourse de mobilité sortante. Sans cette bourse, je n'aurais pas pu avoir l'occasion de séjourner et de me former au XLinC Lab à Cologne. L'axe 1 du Labex EFL m'a également soutenu financièrement, en permettant de prendre en charge le dédommagement des participants EEG, et à travers un contrat comme ingénieur d'étude à la fin de ma thèse.

Pour conclure, je tiens à exprimer ma profonde gratitude envers mes parents pour leur soutien financier infaillible tout au long de mon parcours. J'ai commencé mes études à Paris en 2015 par une Licence en Sciences du Langage à Sorbonne Université Lettres. Depuis cette date, et pendant 10 années consécutives, ils m'ont soutenu sans jamais faillir. Grâce à leur soutien, j'ai pu mener une vie très confortable tout au long de mes études à Paris, sans jamais avoir à me préoccuper de gagner ma vie, pour subvenir à mes besoins.

Table des matières

Introduction générale	1
1.1 Objectifs de la thèse	1
1.2 Structure de la thèse	2
Chapitre 1	5
Partie théorique	5
1.1 Prosodie et frontière prosodique : définitions et fonctions	6
1.1.1 Définitions de la prosodie.....	8
1.1.2 Éléments matériels de la prosodie	9
1.1.3 Pluralité des fonctions de la prosodie	11
1.1.4 Synthèse	19
1.2 Aspects mélodiques de la prosodie en chinois mandarin et en français ..	20
1.2.1 Aspects mélodiques de la prosodie en chinois mandarin	20
1.2.2 Aspects mélodiques de la prosodie en français	30
1.2.3 Synthèse : comparaison des aspects mélodiques entre le mandarin et le français	41
1.3 Aspects temporels de la prosodie	42
1.3.1 Structures syllabiques	43
1.3.2 Systèmes métrique, rythmique et macrorythmique	48
1.3.3 Synthèse.....	79
1.4 Sur l'acquisition de la prosodie du français par des apprenants sinophones	81
1.4.1 Pourquoi la prosodie est importante dans l'acquisition de la phonétique d'une L2	81
1.4.2 Travaux antérieurs sur l'acquisition de la prosodie du français par des apprenants sinophones.....	83
1.4.3 Synthèse	91
1.5 L'électroencéphalographie (EEG) et son application aux études de la prosodie et de l'acquisition langagière	92
1.5.1 L'électroencéphalographie (EEG) et le potentiel évoqué (ERP).....	92
1.5.2 Composantes ERP liées au traitement langagier et prosodique	95
1.5.3 Études sur l'acquisition d'une L2 avec l'EEG.....	99
1.5.4 Synthèse	107
Partie expérimentale.....	109
Chapitre 2 Étude acoustique	110
2.1 Introduction et questions de recherche	110
2.2 Méthodes	111

2.2.1 Locuteurs et corpus.....	111
2.2.2 Traitement et analyses des données	113
2.3 Résultats.....	116
2.3.1 Segmentation mélodique	116
2.3.2 Aspects temporels	123
2.4 Discussion	142
2.4.1 Segmentation prosodique.....	142
2.4.2 Organisation prosodique.....	146
2.5 Annexes	150
2.5.1 Questionnaires LEAP-Q à remplir par les francophones natifs	150
2.5.2 Questionnaires LEAP-Q à remplir par les apprenants sinophones.....	151
Chapitre 3 Étude perceptive	153
3.1 Introduction	153
3.2 Méthodologie	155
3.2.1 Les stimuli.....	155
3.2.2 Participants et procédure	160
3.3 Résultats.....	161
3.3.1 Résultats généraux chez les trois groupes de participants	161
3.3.2 Analyse des résultats en fonction des transitivités des verbes.....	166
3.4 Discussion	171
3.5 Annexes	176
3.5.1 Les stimuli utilisés dans le premier bloc de l'expérience perceptive, ainsi que les réponses données respectivement par les trois groupes de participants pour les deux conditions prosodiques CP et CT dans la tâche de complétion	176
3.5.2 Les stimuli utilisés dans le deuxième bloc de l'expérience perceptive, ainsi que les réponses données respectivement par les trois groupes de participants pour les deux conditions prosodiques CP et CT dans la tâche de complétion	178
Chapitre 4 Étude électroencéphalographique	180
4.1 Introduction et questions de recherche	180
4.2 Méthode	181
4.2.1 Participants	181
4.2.2 Stimuli	182
4.2.3 Procédure.....	183
4.3 Prétraitement et analyses.....	184
4.3.1 Prétraitement des données	184
4.3.2 Analyses des données	185

4.4 Résultats des effets CPS.....	190
4.4.1 CPS chez les francophones natifs.....	192
4.4.2 CPS chez les apprenants sinophones avancés.....	202
4.4.3 CPS chez les apprenants sinophones intermédiaires	212
4.4.4 Comparaison des effets CPS en fonction de niveaux de français	222
4.5 Discussion	225
4.6 Annexes	235
4.6.1 Consentement de participation à l'expérience électroencéphalographique	235
4.6.2 Questionnaires à remplir par les participants francophones natifs et les sinophones natifs apprenants du français.....	238
4.6.3 Les 150 stimuli utilisés dans l'expérience électroencéphalographique	243
4.6.4 Avis du Comité d'Éthique de la Recherche, Université Sorbonne Nouvelle (CER-USN-01-2023)	249
4.6.5 Certificat du Service de Protection des Données du CNRS (2-22087) ..	250
4.6.6 Signaux ERP sur les 42 électrodes dans les trois ROIs pour la comparaison entre FrontièreCP vs NonFrontièreCP et la comparaison entre FrontièreCT vs NonFrontièreCT dans chacun des trois groupes de participants.....	251
Chapitre 5 Synthèse et perspectives.....	257
5.1 Synthèse des résultats et contributions principaux.....	258
5.2 Limites des méthodologies et perspectives.....	263
Bibliographie.....	267

Liste des Figures

Figure 1.1 Place de la prosodie dans le système de production et de compréhension de la parole.....	7
Figure 1.2 : Schéma des correspondances entre les composantes formelles et matérielles de la prosodie (Di Cristo, 2004 : 82).....	11
Figure 1.3 : Tons lexicaux du chinois mandarin standard : variations de contours mélodiques représentées par les valeurs moyennes de F0 (haut) et par l'échelle à cinq niveaux (bas) (Zhu & Wang, 2015 : 507).....	21
Figure 1.4 : Trois patrons mélodiques de base du chinois mandarin (avec I pour la modalité déclarative, II pour la question polaire avec la structure syntaxique énonciative, III pour l'interrogative) (Lin, 2007 : 229).....	23
Figure 1.5 : Schéma d'organisation et de segmentation mélodiques pour un exemple d'énoncé en chinois mandarin (modifié de Cao (2000a)).....	29
Figure 1.6 : Modèle de Jun et Fougeron (2000), avec deux niveaux prosodiques (IP, AP) et patron prototypique de l'AP.....	34
Figure 1.7 : Cinq patrons mélodiques de surface possibles proposés dans Jun et Fougeron (2000) pour le phrasé prototypique /LHiLH*/.....	34
Figure 1.8 : Prototype du syntagme intonatif du français proposé par Post (2000 : 227).....	35
Figure 1.9 : Architecture des différents modèles minimalistes d'organisation des patrons mélodiques du français (a : Hirst & Di Cristo (1984) ; b : Jun & Fougeron (1995) ; c : Hirst & Di Cristo (1984) ; d : Post (2000)) (tirée de Di Cristo, 2016 : 118).	36
Figure 1.10 : Architecture des modèles d'organisation des patrons mélodiques du français à trois niveaux (a : Di Cristo & Hirst (1993) ; b : Jun & Fougeron (2000)) (tirée de Di Cristo, 2016 : 123).....	37
Figure 1.11 : Équivalences entre les dénominations différentes pour les niveaux des phrasés.....	38
Figure 1.12 : Trois phrases ambiguës en français avec les mêmes informations segmentales mais différenciées uniquement par les indices prosodiques (« :: » pour indiquer l'allongement à la fin du syntagme, « : » pour l'allongement à la fin du lexique, « Δ » pour syllabe initiale, « • » pour syllabe finale, « * » pour mot grammatical) (tirée de Vaissière & Michaud (2006 : 54)).....	40
Figure 1.13 : Cycles respiratoires (inspiration-expiration) sans et avec production de la parole (Martin, 2018 : 45).	50
Figure 1.14 : Continuum rythmique proposé par Dauer (1983 : 60).	55
Figure 1.15 : Utilisation des paramètres quantitatifs rythmiques pour la catégorisation rythmique des langues du monde (gauche : %V et ΔC de Ramus et al. (1999) ; droite : nPVI-C et nPVI-V de Grabe & Low (2002)).....	65
Figure 1.16 : Schéma représentant les deux catégories rythmiques idéales basées sur le modèle CCI.....	69

Figure 1.17 : Trois niveaux hiérarchiques prosodiques du chinois mandarin et leurs propriétés acoustico-rythmiques (adaptés de Cao (2001)).	74
Figure 1.18 : Comparaisons des valeurs %V vs ΔC (haut) et ΔV vs ΔC (bas) parmi 15 langues du monde (avec le chinois mandarin entouré en orange et le français entouré en violet)	80
Figure 1.19 : Communication entre neurones conduisant au potentiel électrique.	93
Figure 1.20 : Ordre d'émergence des composantes ERP au cours d'acquisition de la langue maternelle (Friederici, 2005 : 483).	100
Figure 2.1 : Exemple de segmentation pour la phrase	114
Figure 2.2 : Comparaisons des différences de F0 (en Hz) pour V-FrontièreCP et V-NonFrontièreCP pour les quatre groupes de locuteurs.	118
Figure 2.3 : Comparaisons des différences de F0 (en Hz) pour V-FrontièreCT vs V-NonFrontièreCT pour les quatre groupes de locuteurs.	121
Figure 2.4 : Comparaisons des durées des voyelles (en ms) pour V-FrontièreCP vs V-NonFrontièreCP pour les quatre groupes de locuteurs.	124
Figure 2.5 : Comparaisons des durées des voyelles (en ms) pour V-FrontièreCT vs V-NonFrontièreCT pour les quatre groupes de locuteurs.	127
Figure 2.6 : Comparaison des valeurs de %V pour les quatre groupes de locuteurs.	130
Figure 2.7 : Distribution des valeurs de ΔV (gauche) et ΔC (droite) pour les quatre groupes de locuteurs.	131
Figure 2.8 : Comparaison des valeurs de ΔV et ΔC pour les quatre groupes de locuteurs.	132
Figure 2.9 : Distribution des valeurs de VarcoV (gauche) et VarcoC (droite) pour les quatre groupes de locuteurs.	134
Figure 2.10 : Comparaison des valeurs de VarcoV et VarcoC pour les quatre groupes de locuteurs.	135
Figure 2.11 : Distribution des valeurs de nPVI-V et rPVI-C chez les quatre groupes de locuteurs.	137
Figure 2.12 : Comparaison des valeurs de nPVI-V et rPVI-C pour les quatre groupes de locuteurs.	138
Figure 2.13 : Distribution des valeurs de CCI-C (gauche) et CCI-V (droite) pour les quatre groupes de locuteurs.	139
Figure 2.14 : Comparaison des valeurs de CCI-V et CCI-C pour les quatre groupes de locuteurs.	141
Figure 2.15 : La production de la phrase « <i>Chaque fois que le comédien fredonnait le morceau, la mélodie était jolie à l'oreille.</i> » dans la condition CT, par un apprenant débutant.	145
Figure 2.16 : La production de la phrase « <i>Chaque fois que le professeur enseignait la théorie littéraire, son raisonnement devenait moins obscur.</i> » dans la condition CT, par un apprenant débutant.	145

Figure 3.1 : Structure syntaxique arborescente pour une phrase avec la frontière CP (haut) CT (bas).	156
Figure 3.2 : Comparaisons des différences de F0 (Hz) des voyelles en fonction des conditions prosodiques dans les syntagmes verbaux (SV, gauche) et nominaux (SN, droite).	157
Figure 3.3 : Comparaisons des différences de durée (ms) des voyelles en fonction des conditions prosodiques dans les syntagmes nominaux (SN, gauche) et verbaux (SV, droite).	158
Figure 3.4 : Signaux acoustiques et contours mélodiques d'un exemple représentant	159
Figure 3.5 : Deux blocs de stimuli utilisés dans l'étude de perception.	160
Figure 3.6 : Comparaison de nombre de réponses correctes données par les trois groupes (natifs vs avancés vs intermédiaires) pour la condition prosodique CP (haut) et la condition prosodique CT (bas).	162
Figure 4.1 : Exemple de positions des quatre déclencheurs intégrés dans une paire de stimuli ambigus.	182
Figure 4.2 : Exemple du discours temporel d'un stimulus (partie gauche) et positions des 64 électrodes du système Biosemi ActiveTwo (partie droite).	184
Figure 4.3 : Électrodes dans les trois régions d'intérêt (ROIs) prédéfinies dans notre étude EEG.	185
Figure 4.4 : Illustration de l'approche de latence d'une zone fractionnée.	186
Figure 4.5 : Différences d'amplitudes moyennes entre les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> dans les trois FTs et les trois ROIs chez les participants francophones natifs	193
Figure 4.6 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les participants francophones natifs	194
Figure 4.7 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les participants francophones natifs	195
Figure 4.8 : Différences d'amplitudes moyennes entre les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> dans les trois FTs et les trois ROIs chez les participants francophones natifs	196
Figure 4.9 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les participants francophones natifs	197
Figure 4.10 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les participants francophones natifs	197
Figure 4.11 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> dans la FT [562-924] ms chez les participants francophones natifs	199

Figure 4.12 : Les composantes CPS pour <i>FrontièreCP</i> (haut, bleu) et <i>FrontièreCT</i> (bas, vert) sur l'électrode représentative Oz chez les participants francophones natifs .	200
Figure 4.13 : Différences d'amplitudes comparées entre « <i>FrontièreCP-NonFrontièreCP</i> » (FT [455-720] ms) vs « <i>FrontièreCT-NonFrontièreCT</i> » (FT [562-924] ms) chez les participants francophones natifs .	201
Figure 4.14 : Différences d'amplitudes moyennes entre les conditions prosodiques	203
Figure 4.15 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les apprenants avancés .	204
Figure 4.16 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les apprenants avancés .	204
Figure 4.17 : Différences d'amplitudes moyennes entre les conditions prosodiques	206
Figure 4.18 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les apprenants avancés .	207
Figure 4.19 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les apprenants avancés .	208
Figure 4.20 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> dans la FT [634-1500] ms chez les apprenants avancés .	209
Figure 4.21 : Les composantes CPS pour <i>FrontièreCP</i> (haut, bleu) et <i>FrontièreCT</i> (bas, vert) sur l'électrode représentative Fz chez les apprenants avancés .	210
Figure 4.22 : Différences d'amplitudes comparées entre « <i>FrontièreCP-NonFrontièreCP</i> » (FT [610-797] ms) vs « <i>FrontièreCT-NonFrontièreCT</i> » (FT [634-1500] ms) chez les apprenants avancés .	211
Figure 4.23 : Différences d'amplitudes moyennes entre les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> dans les trois FTs et les trois ROIs chez les apprenants intermédiaires .	213
Figure 4.24 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les apprenants intermédiaires .	214
Figure 4.25 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les apprenants intermédiaires .	214
Figure 4.26 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> dans la FT [572-843] ms chez les apprenants intermédiaires .	215
Figure 4.27 : Différences d'amplitudes moyennes entre les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> dans les trois FTs et les trois ROIs chez les apprenants intermédiaires .	216

Figure 4.28 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les apprenants intermédiaires	217
Figure 4.29 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les apprenants intermédiaires	217
Figure 4.30 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> dans la FT [596-1500] ms chez les apprenants intermédiaires	218
Figure 4.31 : Les composantes CPS pour <i>FrontièreCP</i> (haut, bleu) et <i>FrontièreCT</i> (bas, vert) sur l'électrode représentative CPz chez les apprenants intermédiaires ..	219
Figure 4.32 : Différences d'amplitudes comparées entre « <i>FrontièreCP-NonFrontièreCP</i> » (FT [572-843] ms) vs « <i>FrontièreCT-NonFrontièreCT</i> » (FT [596-1500] ms) chez les apprenants intermédiaires	220
Figure 4.33 : Résultat du test de clusters basés sur les permutations pour la comparaison entre « <i>FrontièreCP – NonFrontièreCP</i> » vs « <i>FrontièreCT–NonFrontièreCT</i> » dans la FT [570-1500] ms chez les apprenants intermédiaires	221
Figure 4.34 : Résultat du test de clusters pour les différences « <i>FrontièreCP–NonFrontièreCP</i> »	223
Figure 4.35 : Effets de négativité (entourés en pointillés noirs) dans la condition « <i>NonFrontièreCP</i> »	226
Figure 4.36 : Effets de négativité (entourés en pointillés noirs) dans la condition « <i>NonFrontièreCP</i> »	227
Figure 4.37 : Comparaison des cartes topographiques de la condition <i>FrontièreCP</i> dans la FT [450-840] ms entre les francophones natifs (gauche) et les apprenants intermédiaires (droite).	230
Figure 4.38 : Effets de négativité (entourés en pointillés noirs) dans la condition « <i>NonFrontièreCP</i> »	231
Figure 4.39 : Peu d'effets de négativité (entourés en pointillés noirs) dans la condition « <i>NonFrontièreCP</i> » sur les 42 électrodes des 3 ROIs chez les apprenants intermédiaires	232
Figure 4.40 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les participants francophones natifs . (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).	251
Figure 4.41 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les participants francophones natifs (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).	252
Figure 4.42 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les apprenants avancés	253

Figure 4.43 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les apprenants avancés	254
Figure 4.44 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques <i>FrontièreCP</i> vs <i>NonFrontièreCP</i> chez les apprenants intermédiaires . (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).....	255
Figure 4.45 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques <i>FrontièreCT</i> vs <i>NonFrontièreCT</i> chez les apprenants intermédiaires . (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).....	256

Liste des Tableaux

Tableau 1.1 : Fonctions de la prosodie selon Di Cristo (2013).	14
Tableau 1.2 : Exemples des morphèmes [ma] et [pa] portant les quatre tons lexicaux et leurs significations différentes (tirés de Lin, 2007 : 89).	22
Tableau 1.3 : Conclusions générales dans des études sur les patrons d'organisation et de segmentation prosodiques du chinois mandarin.	25
Tableau 1.4 : Quelques modèles principaux d'organisation hiérarchique de l'intonation du français.	32
Tableau 1.5 : Combinaisons syllabiques possibles en chinois mandarin (Chao, 2011 : 59).	46
Tableau 1.6 : Nombre de possibilités de combinaisons syllabiques en chinois mandarin (Duanmu, 2008 : 97).	47
Tableau 1.7 : Proportions des structures syllabiques en français selon différentes études.	47
Tableau 1.8 : Catégorisation rythmique des langues basée sur l'isochronie (Ramus, 1999 : 41).	53
Tableau 1.9 : Catégorisation rythmique des langues par Grabe et Low (2002 : 518).	54
Tableau 1.10 : Typologie prosodique basée sur la proéminence et sur l'unité rythmique (Jun, 2005 : 444).	57
Tableau 1.11 : Typologie prosodique développée à partir de Jun (2005) (Jun, 2014 : 535).	59
Tableau 1.12 : Synthèse sur les paramètres quantitatifs rythmiques (adaptée de Loukina et al. (2011) entre autres).	63
Tableau 1.13 : Catégorisation des langues basée sur les niveaux I et II du modèle CCI (Tableau établi en fonction des études de Bertinetto & Bertini, 2010) ; Romano & Mairano (2010) ; Bertinetto et al. (2012)).	71
Tableau 1.14 : Valeurs des trois paramètres quantitatifs rythmiques pour le chinois mandarin (adaptées de Romano & Mairano (2010 : 96)).	75
Tableau 1.15 : Patrons microrhythmique et macrorhythmique du chinois mandarin (extrait de Jun (2005)).	76
Tableau 1.16 : Valeurs des trois paramètres quantitatifs rythmiques pour le français européen (adaptées de Romano & Mairano (2010 : 96)).	77
Tableau 1.17 : Patrons microrhythmique et macrorhythmique du français (extrait de Jun (2005 : 444)).	78
Tableau 1.18 : Comparaison des particularités macrorhythmiques entre le chinois mandarin et le français.	81
Tableau 1.19 : Écarts de la prononciation sur le plan intonatif en français chez des apprenants sinophones (adaptés de Lee (1987)).	84
Tableau 1.20 : Écarts de la prononciation prosodique en français chez des apprenants sinophones (adaptés de Shen (1990)).	85

Tableau 1.21 : Écarts de la prononciation prosodique en français chez des apprenants sinophones (adaptés de Landron (2007)).	86
Tableau 1.22 : Écarts de la prononciation prosodique en français chez des apprenants sinophones (adaptés de Kockaert & Li (2008)).	88
Tableau 1.23 : Écarts de la prononciation prosodique en français chez des apprenants sinophones selon Xi et Pillot-Loiseau (2019) et Xi et al. (2020).	91
Tableau 1.24 : Différentes phases d'acquisition morphosyntaxique révélées par composantes ERP chez les apprenants tardifs (Steinhauer et al., 2009).	102
Tableau 1.25 : Les quatre études existantes dans la littérature sur l'utilisation de la CPS pour étudier l'acquisition de la prosodie d'une L2.	105
Tableau 2.1 : Informations (sexe et âge moyen) sur les 31 sujets ayant participé à l'expérience acoustique.	112
Tableau 2.2 : Les cinq paires de phrases utilisées dans l'expérience acoustique.	113
Tableau 2.3 : Exemple d'une paire de phrases localement ambiguës dans l'expérience de production de la parole (« # » représente la frontière prosodique).	115
Tableau 2.4 : Significations des abréviations utilisées dans l'étude de production de la parole.	115
Tableau 2.5 : Significations des paramètres rythmiques mesurés dans cette étude.	116
Tableau 2.6 : Statistiques descriptives des différences de F0 (en Hz) pour V-FrontièreCP et V-NonFrontièreCP pour les quatre groupes de locuteurs.	118
Tableau 2.7 : Résultats du modèle linéaire mixte pour les différences de F0 entre V-FrontièreCP vs V-NonFrontièreCP en fonction des facteurs « positions des voyelles » et « niveaux de français ».	119
Tableau 2.8 : Résultats des tests post-hoc de Tukey pour les différences de F0 entre V-FrontièreCP vs V-NonFrontièreCP dans chaque groupe de locuteurs.	119
Tableau 2.9 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de F0 des V-FrontièreCP pour les quatre groupes de locuteurs.	120
Tableau 2.10 : Statistiques descriptives des différences de F0 (en Hz) pour V-FrontièreCT vs V-NonFrontièreCT pour les quatre groupes de locuteurs.	121
Tableau 2.11 : Résultats du modèle linéaire mixte pour les différences de F0 entre V-FrontièreCT vs V-NonFrontièreCT en fonction des facteurs « positions des voyelles » et « niveaux de français ».	122
Tableau 2.12 : Résultats des tests post-hoc de Tukey pour les différences de F0 entre V-FrontièreCT vs V-NonFrontièreCT dans chaque groupe de locuteurs.	122
Tableau 2.13 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de F0 des V-FrontièreCT pour les quatre groupes de locuteurs.	123
Tableau 2.14 : Statistiques descriptives des durées des voyelles (en ms) pour V-FrontièreCP et V-NonFrontièreCP pour les quatre groupes de locuteurs.	124
Tableau 2.15 : Résultats du modèle linéaire mixte pour les durées entre V-FrontièreCP vs V-NonFrontièreCP en fonction des facteurs « positions des voyelles » et « niveaux de français ».	125

Tableau 2.16 : Résultats des tests post-hoc de Tukey pour les durées entre V-FrontièreCP vs V-NonFrontièreCP dans chaque groupe de locuteurs.....	125
Tableau 2.17 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de durées des V-FrontièreCP pour les quatre groupes de locuteurs.....	126
Tableau 2.18 : Statistiques descriptives des durées des voyelles (en ms) pour V-FrontièreCT et V-NonFrontièreCT pour les quatre groupes de locuteurs.	127
Tableau 2.19 : Résultats du modèle linéaire mixte pour les durées entre V-FrontièreCT vs V-NonFrontièreCT en fonction des facteurs « positions des voyelles » et « niveaux de français ».	128
Tableau 2.20 : Résultats des tests post-hoc de Tukey pour les durées entre V-FrontièreCT vs V-NonFrontièreCT dans chaque groupe de locuteurs.....	128
Tableau 2.21 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de durées des V-FrontièreCT pour les quatre groupes de locuteurs.....	129
Tableau 2.22 : Statistiques descriptives des valeurs %V pour les quatre groupes de locuteurs.....	129
Tableau 2.23 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de %V pour les quatre groupes de locuteurs.....	130
Tableau 2.24 : Statistiques descriptives des valeurs ΔV et ΔC pour les quatre groupes de locuteurs.	131
Tableau 2.25 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de ΔV et ΔC pour les quatre groupes de locuteurs.	133
Tableau 2.26 : Statistiques descriptives des valeurs VarcoV et VarcoC pour les quatre groupes de locuteurs.....	133
Tableau 2.27 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de VarcoV et VarcoC pour les quatre groupes de locuteurs.....	136
Tableau 2.28 : Statistiques descriptives des valeurs nPVI-V et rPVI-C pour les quatre groupes de locuteurs.....	136
Tableau 2.29 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de nPVI-V et rPVI-C pour les quatre groupes de locuteurs.....	137
Tableau 2.30 : Statistiques descriptives des valeurs CCI-V et CCI-C chez les quatre groupes de locuteurs.....	139
Tableau 2.31 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de CCI-V et CCI-C pour les quatre groupes de locuteurs.	140
Tableau 3.1 : Trois types de phrases utilisées dans l'expérience perceptive.....	156
Tableau 3.2 : Différences de F0 et de durée de la même voyelle dans les conditions prosodiques CP et CT.	157
Tableau 3.3 : Informations (sexe et âge) des participants dans l'expérience perceptive.	160
Tableau 3.4 : Statistiques descriptives des résultats obtenus dans la tâche de complétion en fonction des conditions prosodiques (CP vs CT) pour les trois niveaux de participants (natifs vs avancés vs intermédiaires).....	163

Tableau 3.5 : Résultats des tests post-hoc de Tukey pour les interactions entre chaque condition prosodique et chaque niveau de groupe de participants. (Les interactions significatives sont mises en gras.).....	163
Tableau 3.6 : Catégorisation des verbes en deux groupes en fonction de leur transitivité.....	167
Tableau 3.7 : Verbes donnés lieu à des réponses incorrectes chez les trois groupes de participants.....	169
Tableau 3.8 : Réponses correctes (cases vides) et incorrectes (cases en rouge) pour la tâche de complétion des 20 phrases (P1-P20) dans le premier bloc pour les 10 sujets (S1 à S10) des trois groupes.....	177
Tableau 3.9 : Réponses correctes (cases vides) et incorrectes (cases en rouge) pour la tâche de complétion des 20 phrases (P1 à P20) dans le deuxième bloc pour les 10 sujets (S1 à S10) des trois groupes.	179
Tableau 4.1 : Informations sur les trois groupes de participants de l'expérience EEG.	181
Tableau 4.2 : Les dénominations et les significations des quatre conditions prosodiques de l'expérience EEG.	183
Tableau 4.3 : Onsets, pics et offsets (identifiés par l'approche de latence de la zone fractionnée [0-1500] ms) du résidu positif de la différence « Frontière – NonFrontière » pour chaque condition prosodique de chaque groupe de participants.	187
Tableau 4.4 : Trois FTs de la composante CPS pour chaque condition prosodique et chaque groupe de participants.....	187
Tableau 4.5 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques FrontièreCP vs NonFrontièreCP chez les participants francophones natifs	192
Tableau 4.6 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques FrontièreCT vs NonFrontièreCT chez les participants francophones natifs	196
Tableau 4.7 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques FrontièreCP vs NonFrontièreCP chez les apprenants avancés	202
Tableau 4.8 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques FrontièreCT vs NonFrontièreCT chez les apprenants avancés	205
Tableau 4.9 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques FrontièreCP vs NonFrontièreCP chez les apprenants intermédiaires	212
Tableau 4.10 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques FrontièreCT vs NonFrontièreCT chez les apprenants intermédiaires	216
Tableau 4.11 : FT pour les tests de clusters « FrontièreCP–NonFrontièreCP » dans les comparaisons par paire.	222

Tableau 4.12 : FTs pour les tests de clusters « FrontièreCT–NonFrontièreCT » dans les comparaisons par paire.	224
Tableau 5.1 : Comparaison des distributions topographiques des composantes CPS en fonction de conditions prosodiques et de niveaux de français.	259
Tableau 5.2 : Position du déclencheur (ligne rouge) pour extraire la composante P600 dans Nickels et Steinhauer (2018).	264
Tableau 5.3 : Positions des déclencheurs à insérer sur l’onset du dernier syntagme verbal (positions indiquées par la ligne pointillée rouge) dans nos stimuli pour étudier le traitement entre la prosodie (CPS) et la syntaxe (P600) (déclencheurs actuels : lignes pointillées noires).	265

Introduction générale

La prosodie de la parole constitue un champ de recherche vaste, diversifié et complexe, englobant des phénomènes variés tels que le ton, l'accent, la quantité, la métrique, le rythme ou encore l'intonation. Ces éléments prosodiques, également appelés prosodèmes, jouent des rôles essentiels et multifonctionnels dans la communication verbale. Parmi ces nombreuses fonctions, la prosodie occupe une place centrale dans la segmentation des phrases en unités de traitement syntaxiques, sémantiques ou pragmatiques. Cependant, les mécanismes de segmentation prosodique diffèrent d'une langue à l'autre. Cette fonction ne peut être assurée « *que si le locuteur dispose de "schémas" préstockés des évolutions prosodiques habituellement associées, dans la langue, à la structure linguistique de l'énoncé* » (Dubois *et al.*, 2001 : 387). Ainsi, lors de l'apprentissage d'une langue seconde ou étrangère, et en particulier après l'âge adulte, utiliser efficacement la prosodie, notamment pour encoder et décoder les messages linguistiques, représente un défi considérable. Cette thèse aborde ce défi en étudiant la segmentation et l'organisation prosodiques en français, sous les angles acoustique, perceptif et neurocognitif, tant chez les francophones natifs que chez les apprenants sinophones. Dans la recherche sur l'acquisition sonore d'une langue seconde (L2), l'attention s'est majoritairement concentrée sur le niveau segmental. En revanche, l'étude de l'acquisition de la prosodie en L2 est longtemps restée marginale et peu explorée, probablement en raison de la complexité intrinsèque des phénomènes prosodiques.

1.1 Objectifs de la thèse

La présente thèse s'intéresse à l'acquisition de la prosodie du français par des apprenants sinophones, en s'intéressant plus spécifiquement à montrer comment les apprenants sinophones encodent et décodent des phrases syntaxiquement ambiguës. Pour ce faire, nous avons utilisé un corpus de phrases caractérisées par des frontières prosodiques à deux positions différentes : une clôture précoce (CP), comme dans « *Chaque fois que le serpent mangeait, le rat se cachait* », et une clôture tardive (CT), comme dans « *Chaque fois que le serpent mangeait le rat, le lapin se cachait* ». À travers trois approches expérimentales complémentaires, nous avons étudié et comparé les stratégies utilisées par des locuteurs francophones natifs et

des apprenants sinophones de différents niveaux pour désambiguïser ces phrases sur les plans acoustique, perceptif et neurocognitif. La première expérience, consacrée à la production de la parole, porte sur la segmentation et l'organisation prosodiques. Les données acoustiques, recueillies auprès de 31 locuteurs (10 francophones natifs et 21 apprenants sinophones), permettent d'analyser les compétences des participants à signaler les frontières prosodiques et à désambiguïser les énoncés syntaxiques à l'aide d'indices mélodiques et temporels. La deuxième expérience, centrée sur la perception de la parole, explore le traitement des frontières prosodiques ambiguës à l'aide d'une tâche comportementale. Nous avons recueilli des données auprès de 60 auditeurs (20 francophones natifs et 40 apprenants sinophones) afin de déterminer dans quelle mesure ils utilisent les indices acoustiques pour distinguer les différentes frontières prosodiques. Ces données offrent des indications précieuses sur les préférences et les stratégies de traitement des participants des frontières CP et CT. La troisième expérience, basée sur une approche électroencéphalographique (EEG), examine le traitement neurocognitif des phrases localement ambiguës. Nous avons analysé, en particulier, la composante CPS (*closure positive shift*), un potentiel évoqué associé spécifiquement au traitement des frontières prosodiques. Les données EEG, recueillies auprès de 60 participants (20 francophones natifs et 40 apprenants sinophones), ont permis d'évaluer si cette composante est déclenchée chez les natifs et les apprenants et d'en comparer les différences. Ces résultats apportent une compréhension du traitement neurophysiologique des frontières prosodiques ambiguës en français, tout en éclairant le processus d'acquisition prosodique chez les apprenants sinophones. Ces trois approches expérimentales ont permis de constituer un large corpus de données multimodales, qui offre une vue d'ensemble sur le parcours d'acquisition prosodique en français par des apprenants sinophones. Cette thèse constitue également une contribution à la littérature existante en présentant le corpus le plus vaste à ce jour, avec un total de 151 participants, pour étudier l'acquisition de la prosodie en français.

1.2 Structure de la thèse

Cette thèse est structurée en cinq chapitres. Le Chapitre 1 introduit le cadre théorique de cette thèse. Nous y présentons les concepts fondamentaux liés à la

prosodie : ses définitions, ses caractéristiques acoustiques et rythmiques, ainsi que ses fonctions dans la communication. L'accent est ensuite mis sur les différences entre le français et le mandarin en termes de patrons mélodiques et temporels, notamment en ce qui concerne l'organisation et la segmentation prosodiques. Une revue de la littérature existante sur l'acquisition de la prosodie française par des apprenants sinophones est également proposée, bien que les études sur ce sujet soient limitées. Enfin, la méthode électroencéphalographique et la composante CPS sont introduites, ainsi que leur pertinence pour l'étude de la prosodie en L2. Le Chapitre 2 présente les résultats de l'expérience de production. Nous y analysons comment les locuteurs natifs et les apprenants produisent des phrases localement ambiguës, en mettant l'accent sur les indices mélodiques et temporels utilisés pour signaler les frontières prosodiques CP et CT. Nous nous sommes concentrés principalement sur l'allongement final et la montée de F0 à la position des dernières syllabes précédant ces frontières prosodiques. En ce qui concerne l'organisation prosodique, nous analysons et comparons les valeurs de neuf paramètres quantitatifs rythmiques, afin d'évaluer les patrons temporels des locuteurs francophones natifs et des apprenants.

Le Chapitre 3 est consacré à l'expérience perceptive, qui explore le traitement des frontières prosodiques ambiguës CP et CT. À l'aide d'une tâche de complétion, et sur la base du même type de phrases analysées dans le chapitre 2, notre objectif est de déterminer si les auditeurs francophones et sinophones sont capables, en se basant uniquement sur les indices prosodiques, de déterminer les fonctions syntaxiques d'un ensemble de syntagmes nominaux et désambiguïser ainsi les relations syntaxiques entre constituants prosodiques.

Le Chapitre 4 se concentre sur l'expérience électroencéphalographique. Nous y analysons le traitement neurocognitif des frontières prosodiques CP et CT à travers la composante CPS. Les propriétés neurophysiologiques du CPS (latence, amplitude, distribution topographique) sont examinées, et les différences entre les groupes sont discutées.

Enfin, le Chapitre 5 synthétise les résultats des trois expériences et propose une discussion générale. Nous résumons les difficultés rencontrées par les apprenants dans le traitement comportemental et neurocognitif des phrases ambiguës, tout en montrant comment les données des trois approches expérimentales se complètent

pour offrir une vision approfondie de l'acquisition prosodique. Nous concluons en évoquant les limites de cette recherche et en proposant des perspectives pour des études futures.

Chapitre 1

Partie théorique

Ce chapitre est composé de cinq sections : nous commencerons par situer la prosodie dans le champ de la communication langagière. Ensuite nous aborderons les définitions de la prosodie, ses matérialités et ses fonctions, notamment les aspects cognitifs de la prosodie et de la frontière prosodique dans la communication langagière.

Les deux sections suivantes sont consacrées aux deux aspects matériels les plus importants de la prosodie : l'aspect mélodique et l'aspect temporel. Nous débuterons par définir les concepts de mélodie et de temporalité. Ensuite, nous examinerons comment ces notions se manifestent dans l'organisation et la segmentation prosodiques en chinois mandarin et en français. Les principales différences entre les systèmes tonals et temporels de ces deux langues seront ensuite mises en lumière.

La quatrième section portera sur les travaux antérieurs concernant l'acquisition de la prosodie du français par des apprenants sinophones. Bien que ce sujet ait été peu étudié et que nous ne disposons que d'un nombre limité d'études, nous pouvons néanmoins synthétiser les conclusions et les découvertes des travaux existants. Cela nous permettra d'ouvrir de nouvelles perspectives pour notre recherche dans la partie expérimentale.

Dans la dernière section, l'accent sera mis sur la méthode électroencéphalographique (EEG). Après une présentation succincte de cette technique, nous passerons en revue certaines composantes ERP actuellement utilisées dans la recherche en sciences du langage. Nous nous focaliserons ensuite sur un potentiel évoqué spécifique : la CPS (*closure positive shift*). Cet effet ERP reflète le traitement cognitif de la frontière prosodique. Nous aborderons sa découverte, ses propriétés électrophysiologiques et son application dans la recherche sur l'acquisition de la prosodie d'une L2.

1.1 Prosodie et frontière prosodique : définitions et fonctions

La communication verbale, en apparence simple et spontanée, est en réalité un processus d'une complexité remarquable. Chaque échange verbal implique une multitude de niveaux d'analyse, allant de la production des sons par l'appareil phonatoire et l'articulation phonétique, à l'organisation syntaxique, jusqu'à l'interprétation sémantique et pragmatique des énoncés. Ces niveaux interagissent de manière simultanée et dynamique, influencés par des facteurs linguistiques et extralinguistiques, ainsi que par les intentions communicatives des interlocuteurs. Il est nécessaire de décomposer ces mécanismes pour mieux comprendre comment, par des signes sonores partagés, les locuteurs humains parviennent à transmettre des messages, des idées, des émotions et d'autres informations complexes. Cette exploration révèle non seulement la diversité des langues, mais aussi les capacités cognitives sophistiquées qui sous-tendent le langage humain.

Au cours de ces mécanismes complexes, le locuteur dispose d'une série de ressources verbales et non-verbales pour construire l'énoncé, établir un canal d'échange, transmettre son message et véhiculer son intention communicative. Quant à l'auditeur, il lui incombe de décoder toutes les ressources mises en place. Dans ses travaux sur la prosodie, Di Cristo (2013) a établi un schéma qui montre ce double processus d'« encodage-décodage » dans l'échange verbal. En particulier, il situe la place de la prosodie au cœur de ce processus en la qualifiant de « compilateur prosodique » (Figure 1.1) :

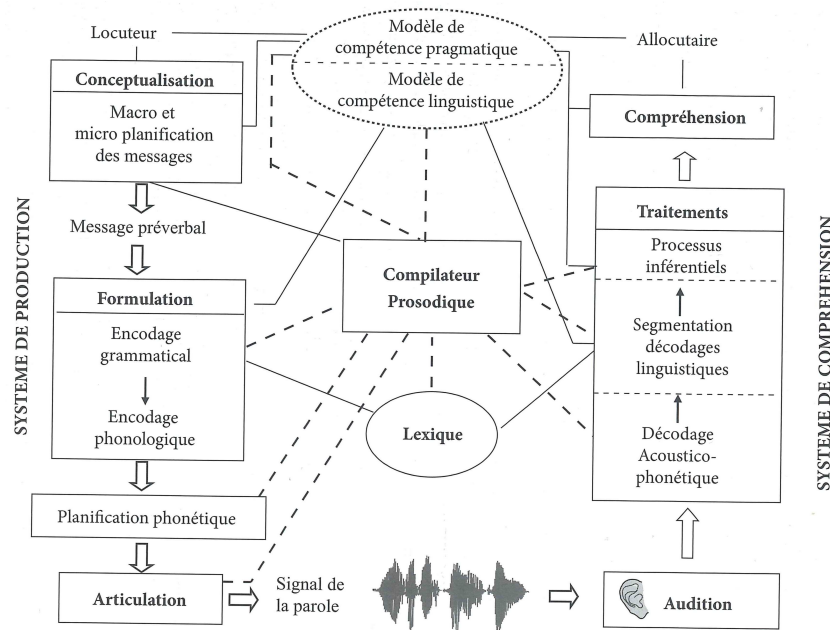


Figure 1.1 Place de la prosodie dans le système de production et de compréhension de la parole. (Di Cristo, 2013 : 44).

La Figure 1.1 illustre le fait que la prosodie intervient, avec d'autres éléments linguistiques, dans tous les niveaux du processus de « production – perception » de la parole : de la planification cognitive du message mental au décodage acoustico-phonétique du message verbal en passant par l'articulation, l'émission et le traitement du signal sonore.

La prosodie est un indice indispensable dans l'ensemble du processus de la communication, mais elle n'est qu'un élément parmi toutes les ressources verbales et non-verbales disponibles. Cependant, la prosodie elle-même peut constituer un phénomène d'une grande complexité. Dans le système langagier, la prosodie varie en allant de pair avec d'autres éléments linguistiques (par exemple, la syllabe, le lexique, ou le syntagme) en fonction des langues différentes. Lorsque l'on passe à un contexte énonciatif précis, la prosodie subit davantage de variations (de l'ordre émotionnel, attitudinal, etc.) afin de satisfaire la transmission de l'intention véhiculée dans le message du locuteur.

Cette complexité est due non seulement à la nature intrinsèquement sophistiquée de la communication verbale, mais également aux phénomènes que recouvre le terme de « prosodie ». Dans la section suivante, nous commencerons par donner une définition exhaustive et claire de « la prosodie de la parole ».

1.1.1 Définitions de la prosodie

La prosodie est souvent décrite comme « la musique de la parole ». Elle nous fournit une multitude d'informations, qu'elles soient purement linguistiques, émotionnelles, attitudinales, ou liées à l'identité personnelle du locuteur. Dans la communication langagière, un énoncé construit par des lexiques en suivant certaines règles grammaticales ou syntactiques nous transmet ce qui est dit, tandis que la prosodie nous permet de savoir comment cela est dit (Fónagy, 1971¹, 1983 ; Hirst, 2024²). Pour citer Di Cristo (2004 : 89) : « *Les éléments prosodiques exercent au niveau du mot (prosodie lexicale) et au-delà de ce dernier (prosodie supra-lexicale ou post-lexicale) un faisceau de fonctions grammaticales, para-grammaticales et extra-grammaticales, qui se rapportent à « ce qui est dit », à « la façon dont cela est dit », ainsi qu'à l'identité du sujet parlant, et qui s'avèrent déterminantes pour signaler la structure des énoncés et du discours et pour procéder à leur interprétation sémantique et pragmatique* ». La place de la prosodie dans la parole est illustrée de manière pertinente par une formule proposée par Hirst (2024) : étant donné que la parole humaine est constituée de « ce qui est dit » et de « comment cela est dit » (« *Speech = Text + Prosody* »), il en découle que « *Prosody = Speech - Text* » (concept également discuté chez Martin (2018)).

Comme l'a rappelé Nootboom (1997), le mot « prosodie » vient du grec ancien, où il désignait une « chanson chantée avec l'instrument musical ». Plus tard, le mot a été utilisé pour désigner la « science de la versification » et la « loi du mètre » (Nootboom, 1997). Dans le domaine des sciences du langage d'aujourd'hui, le mot « prosodie », comme un terme collectif, se réfère, si on le définit de manière simple, aux variations de F0, d'intensité, de tempo et de rythme (Crystal, 2008). Dans son ouvrage, Di Cristo (2013) a rassemblé plusieurs définitions proposées par différents chercheurs et, en se basant sur les définitions existant dans la littérature, il a formulé

¹ Double codage de la vive voix chez Fónagy : “*The speech sound is a result of a DOUBLE CODING. The choice of the phoneme is followed by the selection of a speech sound representing this phoneme, and CONVEYING at the same time an INDEPENDENT attitudinal MESSAGE. The encoding of phoneme-sequences is based upon an essentially arbitrary linguistic code; secondary phonetic messages, however, seem to be 'natural', deeply rooted in human habits, founded on a PARALINGUISTIC code.*” (Fónagy, 1971: 191).

² “*... there are still many other things that we get from a spoken message that we don't get from a written one. It is a truism that the way we say something is just as important as, and sometimes even more important than, what we say. And, to a considerable extent, we can say that the way we say something is due to the prosody of what we say.*” (Hirst, 2024: 5).

sa propre définition synthétique de la prosodie. Cette définition englobe les éléments prosodiques, les niveaux d'analyse et leurs fonctions. Il nous semble essentiel de reprendre ici cette définition, car nous l'adoptons dans le cadre de ce travail de thèse : « *La prosodie est une branche de la linguistique consacrée à l'analyse des propriétés **formelles** (phonologie prosodique), de la **matérialité** (phonétique prosodique) et de la fonctionnalité des éléments non verbaux de l'expression orale, non coextensifs aux phonèmes, tels que l'accent, les tons, l'intonation, la quantité, le tempo et les pauses, [éléments constituant la prosodie] que l'on qualifie collectivement de **prosodèmes**. La matérialité physique des prosodèmes se manifeste par les variations de la fréquence fondamentale (F0), de la durée et de l'intensité (**paramètres prosodiques acoustiques**), ces variations étant perçues par l'auditeur comme des changements respectifs de hauteur, de longueur et de volume sonore (**paramètres prosodiques auditifs**). Les informations prosodiques ont la particularité d'être polysémiques et de véhiculer à la fois des informations d'ordre linguistique, paralinguistique et extralinguistique, qui se révèlent déterminantes pour l'interprétation des énoncés du discours et de la conversation » (ibid. : 21).*

En nous basant sur la définition de Di Cristo (2013), dans les sections suivantes, nous présenterons, respectivement et de manière plus précise, les « prosodèmes », leurs paramètres matériels et leurs fonctions dans la parole.

1.1.2 Éléments matériels de la prosodie

Dans cette section, nous ferons le point sur les éléments matériels de la prosodie et mettrons en évidence ceux qui seront examinés dans la partie expérimentale. Les trois paramètres principaux de la prosodie sont la fréquence fondamentale (F0), la durée et l'intensité, « *qui représentent les paramètres objectifs de la prosodie. Ces variations sont perçues par l'auditeur comme des changements de hauteur (ou de mélodie), de longueur et de sonie, qui constituent les paramètres subjectifs de la prosodie* » (Di Cristo, 2004 : 88). Ces trois paramètres³ seront brièvement présentés ci-dessous à partir des points de vue acoustique et auditif :

³ Outre ces trois éléments principaux, d'autres paramètres matériels sont également pris en compte dans l'étude de la prosodie. Parmi ceux-ci, on trouve le spectre (timbre), qui est associé aux voyelles accentuées vs inaccentuées, ainsi que la microprosodie, qui désigne l'interaction entre la fréquence fondamentale (F0) et la réalisation du matériau segmental (Di Cristo, 2004, 2013).

- **Fréquence fondamentale (F0) :**

La F0 est la fréquence à laquelle les plis vocaux vibrent, qui est une « *conjonction d'activités musculaires et de phénomènes aérodynamiques* » (Di Cristo, 2013 : 84). La valeur de F0, notée en Hertz (Hz), varie en fonction de la périodicité et elle se traduit par la formule $F0=1/T$, où T représente la durée de la période. Son équivalence perceptive est couramment désignée par le terme « *pitch* » en anglais, ou sous le terme « hauteur » en français, et peut être notée en octaves, fractions d'octaves ou demi-tons. Les phénomènes prosodiques – ton lexical et intonation – renvoient ainsi à la variation de F0, mais ils se distinguent par leur domaine d'application (pour une présentation détaillée, cf. section 1.2).

- **Durée :**

La durée, généralement exprimée en milliseconde (ms), constitue un paramètre essentiel dans l'analyse prosodique. Dans l'étude de la prosodie, elle peut être mesurée à différents niveaux, tel que le phone, la syllabe, le débit ou encore la pause, etc. Ces mesures fournissent des informations « *sur la durée intrinsèque des phonèmes, ainsi que sur l'organisation prosodique des niveaux syllabique, lexical, syntaxique, etc.* » (Lacheret-Dujour & Beaugendre, 1999 : 243).

- **Intensité :**

Dans les travaux en prosodie, l'intensité n'a pas attiré autant d'attention que les deux paramètres « fréquence fondamentale » et « durée »⁴. L'intensité est « *proportionnelle à l'amplitude de la vibration des cordes vocales. Cette amplitude dépend essentiellement du niveau de pression sous-glottique qui est lié, à la fois, à la force exercée par la pression de l'air pulmonaire et au degré de résistance qu'opposent les cordes vocales à cette pression* » (Di Cristo, 2013 : 85). L'équivalent perceptif de l'intensité est appelé sonie (*loudness*). Elle est généralement exprimée en décibels (dB), une échelle logarithmique qui mesure d'intensité en fonction d'une valeur de référence.

⁴ « *L'intensité correspond au paramètre le plus négligé dans les recherches en prosodie. Cette lacune n'est bien sûr pas liée à la difficulté d'extraction du paramètre sur le signal de parole, mais résulte probablement de trois raisons principales. D'une part, l'intensité est souvent considérée comme une simple co-variable de la fréquence fondamentale... D'autre part, on a pendant longtemps attribué à ce paramètre un rôle tout à fait secondaire dans l'organisation prosodique du français. Enfin, l'étude de l'intensité demande des précautions méthodologiques particulières.* » (Lacheret-Dujour & Beaugendre, 1999 : 243).

Pour les trois paramètres matériels principaux de la prosodie, Di Cristo (2004) a établi un schéma de correspondances entre les éléments substantiels et formels. Ce schéma, reproduit ci-dessous (Figure 1.2), illustre les correspondances entre les composantes formelles et matérielles de la prosodie :

Caractéristiques abstraites		Caractéristiques concrètes: mesurables (aspects physiques et sensoriels)		
<i>Lexicales</i>	<i>supra-lexicales</i>	<i>Production</i> (physiologie)	<i>Transmission</i> (Acoustique)	<i>Perception</i> (Audition)
Quantité	Tempo, Pauses	chronométrage des gestes articulatoires	durée des événements et des silences	longueur
Accent lexical	Accents supra-lexicaux. Rythme	effort articulatoire	F0, amplitude spectre, durée	mélodie, intensité timbre, longueur
Ton	Intonation	activité musculaire activité respiratoire	fréquence fondamentale (F0)	mélodie

Figure 1.2 : Schéma des correspondances entre les composantes formelles et matérielles de la prosodie (Di Cristo, 2004 : 82).

Ce schéma (Figure 1.2) montre la complexité des phénomènes prosodiques : bien qu'il n'y ait que trois éléments matériels, la prosodie de la parole peut les exploiter pour des usages variés. La F0 peut être utilisée pour le ton, l'intonation et l'accentuation ; la durée peut servir comme indice pour le rythme, le débit, et la pause. La F0 et la durée sont également les deux indices matériels les plus importants que nous examinerons dans la partie expérimentale de ce travail.

1.1.3 Pluralité des fonctions de la prosodie

Parler des fonctions de la prosodie n'est pas une tâche aisée en raison de sa grande pluralité et de sa richesse. « *Les significations que la prosodie contribue à véhiculer sont hautement diversifiées, car elles informent à la fois sur le contenu explicite et implicite des messages, ainsi que sur les intentions communicatives, les attitudes, l'état émotif et l'identité de leurs auteurs.* » (Di Cristo, 2013 : 170). Dans la littérature, de nombreux travaux ont proposé une synthèse des fonctions de la prosodie (par exemple, Cutler *et al.*, 1997 ; Speer & Blodgett, 2006 ; Lacheret, 2011 ; Dahan, 2015 ; Pratt, 2018, etc.). Bien que ces études aient apporté des contributions significatives, elles n'ont pas toujours couvert tous les aspects de manière exhaustive. Di Cristo

(2013) se distingue par une approche plus détaillée, en abordant les fonctions prosodiques dans presque tous les aspects.

Comme nous l'avons mentionné dans l'introduction de l'étude, la prosodie recouvre divers phénomènes. Dans cette section, qui constitue encore une présentation générale de la prosodie, nous nous concentrons sur les fonctions générales de la prosodie (nous aborderons les rôles spécifiques de chaque élément prosodique dans les sections 1.2 et 1.3). Dans un premier temps, nous revisiterons brièvement l'étude sur les fonctions de la prosodie de Di Cristo (2013). Ensuite, l'accent sera mis sur les fonctions cognitives de la prosodie, un aspect insuffisamment exploré par les chercheurs, mais crucial pour notre étude. Les fonctions cognitives de la prosodie seront mises en évidence, non seulement pour leur importance dans la langue orale, mais également dans la lecture, dans la mémoire, etc.

Avant d'entamer cette discussion, il nous semble pertinent d'explicitier que, comme les éléments prosodiques ne sont pas totalement indépendants les uns des autres, il en va de même pour les fonctions de la prosodie : certaines fonctions prosodiques pourraient être conjointement assurées par le rythme et la proéminence, ou par la proéminence et l'intonation. Comme l'a souligné Di Cristo (2013 : 183), il est parfois difficile d'isoler les fonctions de la prosodie, « *qui sont le plus souvent cumulatives et entremêlées* ».

1.1.3.1 Typologie fonctionnelle de la prosodie proposée par Di Cristo (2013)

Étudier les fonctions de la prosodie peut être appréhendé à partir d'angles différents. Dans la littérature, nous disposons de nombreux travaux complets et détaillés sur ce sujet, entre autres, Cutler *et al.* (1997), Wennerstrom (2001), Di Cristo (2013), Gussenhoven et Chen (2020). Ceux de Di Cristo (2013) se démarquent par leur exhaustivité et leur niveau de détail. Il a analysé les fonctions de la prosodie en suivant une échelle qui répartit les différentes fonctions entre deux extrémités de « linguisticité ». En fonction de cette échelle, les types de fonctions sont aménagés en trois grands groupes : linguistique, paralinguistique et extralinguistique. Dans son analyse, il a proposé d'abord deux fonctions « générales » de la prosodie : une fonction dite « *bootstrapping* » (Pratt, 2018), qui permet aux jeunes enfants d'acquérir la langue maternelle et une autre fonction de guidage, qui sert de cadre

précoce de planification avant d'émettre un énoncé à l'oral chez le locuteur. En plus de ses rôles fondamentaux dans l'acquisition de la langue maternelle et dans la planification énonciative, la prosodie fonctionne comme un repère qui permet aux locuteurs d'identifier des langues inconnues sans même connaître un mot de cette langue (Ohala & Gilbert, 1979 ; Ramus & Mehler, 1999 ; Mary & Yegnanarayana, 2008). À la suite des fonctions générales, Di Cristo (2013) a traité les fonctions spécifiques en suivant l'échelle de « linguisticité » et en allant du plus linguistique au moins linguistique. Nous récapitulons les fonctions mentionnées par Di Cristo (2013) dans le Tableau 1.1 et nous pouvons voir comment elles sont catégorisées :

Tableau 1.1 : Fonctions de la prosodie selon Di Cristo (2013).

Fonctions	Exemples	
Fonction accentuelle	Distinction ; saillance ; cadre métrique ; segmentation ; focalisation	
Fonction tonale au niveau lexical ou supralexical	Tons lexicaux ou intonation	
Fonctions structurales	Au niveau de la phrase ou énoncé	Actualisation ou indication de la hiérarchie syntaxique par la frontière prosodique ; désambiguïsation
	Au niveau du message	Statuts informationnels (thème vs. propos ou topique vs. commentaire)
	Au niveau du discours	Organisation, parenthésage, liage ⁵
Fonctions énonciative et illocutoire	Signaler la force énonciative : actes d’assertion, de question et de commandement, l’ordre, la requête, l’apostrophe, etc.	
Fonctions expressive et impulsive	Expressions des aspects attitudinal, affectif, sentimental et émotionnel	
Fonction interactionnelle	Participation au jeu des échanges interpersonnels dans le dialogue ou la conversation ; indication de relations sociales	
Fonctions indicielle ou identificatrice	Donner des renseignements sur les profils personnels des locuteurs (identité individuelle ; appartenance à une communauté géographique, socioculturelle ou professionnelle ; accent étranger, etc.)	
Fonctions de contextualisation et écologique	Les indices prosodiques perceptibles permettent aux locuteurs de construire le contexte communicationnel et de l’actualiser en variant des marques prosodiques.	

Dans la section suivante, nous nous concentrerons sur les fonctions cognitives de la frontière prosodique : un sujet qui n’est pas suffisamment traité et dont les conclusions sont dispersées dans la littérature. Nous essayerons de résumer les

⁵ « La prosodie contribue à segmenter le discours en unité et à indiquer les relations que ces unités contractent au sein de la macrostructure... D’une part, la prosodie, à l’instar des connecteurs verbaux, a la faculté de contribuer à signaler, par elle-même, certaines relations de discours. D’autre part, il appartient à la prosodie d’indiquer occasionnellement si un item lexical peut être interprété comme un connecteur ou pas. » (Di Cristo, 2013 : 210, 219). À partir de ce sens-là, la prosodie permet de maintenir la cohérence et la cohésion discursives.

aspects importants concernant le rôle cognitif que joue la prosodie, en nous intéressant plus particulièrement à la frontière prosodique.

1.1.3.2 Fonctions cognitives de la prosodie

Les fonctions cognitives jouées par la prosodie sont assez peu traitées dans la littérature, bien que, comme l'a souligné Di Cristo (2013 : 171), « *la prosodie participe en réalité d'un dispositif cognitivement enraciné* ». Les recherches actuelles n'ont pas encore permis de dresser un panorama complet des aspects cognitifs de la prosodie. Comme l'a indiqué Cutler (1998)⁶, la plupart des études ont traité des fonctions prosodiques telles que la segmentation syntaxique, l'accentuation informationnelle, la modalité intonative en passant sous silence d'autres fonctions. Dans cette section, en nous basant principalement sur les études de Cutler *et al.* (1997), de Speer et Blodgett (2006), de Pratt (2018) et de certaines d'autres, nous essayerons d'établir un plan détaillé des fonctions cognitives de la prosodie dans la communication orale ou écrite. Afin de mieux synthétiser des points de vue différents dans les études disponibles, les fonctions cognitives seront regroupées en trois catégories fonctionnelles : dans un premier temps, ce sont l'encodage et le décodage ; et puis la segmentation et le liage et, enfin, la prosodie dans la mémoire et la lecture silencieuse.

- Encodage, décodage et guidage

D'un point de vue général, le processus de la communication langagière chez l'être humain peut posséder une double façade : d'une part, un processus d'encodage (la production de la parole) et d'autre part, un processus de décodage (la compréhension de la parole). Ce processus existe dans tous les niveaux de la communication, que ce soit au niveau segmental, suprasegmental, sociolectal ou même individuel. Le décodage est désigné, dans certains cas par des psycholinguistes, sous le terme anglais « *parsing* », qui a été au début utilisé dans l'analyse syntaxique. Mais depuis une vingtaine d'années, le rôle important de la prosodie dans le traitement de la parole émergeait et certains chercheurs ont aussi proposé le « *parsing* » prosodique (Beckman, 1996 ; Schafer, 1997).

⁶ "Moreover, as Cutler *et al.* conclude, the literature is very unbalanced: some topics have been repeatedly examined, in studies differing only in minor details, while other topics have been ignored completely." (Cutler, 1998: 41).

Si nous regardons la prosodie du point de vue cognitif, toutes les fonctions abordées dans la section précédente peuvent être considérées comme un processus cognitif d'encodage ou de décodage acoustico-phonétiques. Par exemple, encoder ou décoder un ton lexical pour exprimer un sens lexical ou accéder au sens, encoder ou décoder une modalité intonative pour émettre ou reconnaître une question ou une exclamation, encoder ou décoder la prosodie émotionnelle ou attitudinale selon un contexte communicatif spécifique pour exprimer ou connaître l'état mental du locuteur, etc.

En plus d'encoder ou de décoder une unité linguistique dont le domaine est plus ou moins grand, la prosodie dans son ensemble a une fonction de planification et de guidage (Ferreira, 1993), qui est liée étroitement avec le processus « encodage – décodage ». Cette fonction intervient toujours au premier plan dans les activités langagières : « *Parmi les fonctions de la prosodie que l'on peut qualifier de générales, les fonctions de guidage de la production et de la perception de la parole ont assurément un rôle de premier plan à jouer... La constitution d'une structure prosodique mentale précoce sert de cadre général à la planification phonologique et phonétique des syntagmes et des phrases* » (Di Cristo, 2013 : 173-174). Ce processus intervient notamment en amont du traitement syntaxique⁷ (Cutler *et al.*, 1997 ; Speer & Blodgett, 2006 ; Martin, 2013). Au niveau de la production (« encodage »), après la planification mentale préalable, le locuteur peut gérer l'inspiration et la prise de souffle en fonction de la frontière syntaxique ou prosodique, de la longueur de la phrase, du débit qu'il veut adopter, etc. (Lacheret-Dujour & Beaugendre, 1999 ; Martin, 2015, 2018). Au niveau de la perception (« décodage »), le guidage prosodique permet aux auditeurs de suivre la chaîne de parole continue, se manifestant par la reconstitution d'une hiérarchie prosodique au fur et à mesure que les sons de la parole se produisent linéairement dans le temps⁸ (Langus *et al.*, 2012). Dans une situation communicative bruyante, l'auditeur peut également suivre la

⁷ "Prosody has an immediate rather than delayed effect on a range of syntactic parsing decisions, suggesting that prosody can determine the incremental construction of a syntactic representation on the fly during sentence comprehension. These effects have implications for the architecture of the language processing mechanism, suggesting that prosodic structure is primary to the process of recovering sentence structure." (Speer & Blodgett, 2006: 513).

⁸ "... that listeners can keep track of prosodic cues from different levels of the prosodic hierarchy, that they perceive prosody as organized hierarchically... the information contained in the prosody of speech allows listeners to discover hierarchical constituent structure, one of the core properties of human language." (Langus *et al.*, 2012: 302).

chaîne de parole du locuteur en tirant profit de la prosodie, ce qui est désigné comme l'effet de « *cocktail party* » (Pollack & Pickett, 1957).

- **Segmentation, liage et désambiguïsation**

La segmentation prosodique est une procédure consistant à diviser l'énoncé en unités constituantes prosodiques, tandis que le liage renvoie à la manière dont ces unités sont connectées entre elles. Tout cela contribue à la désambiguïsation en clarifiant le sens du discours.

La fonction de segmentation possède un lien direct et inséparable avec les fonctions cognitives prosodiques. La segmentation est primordiale pour la compréhension de la parole, non seulement dans le langage oral, mais également pour la mémoire et la lecture (cf. *infra* et Frazier *et al.* (2006)). La segmentation et le liage sont en effet les deux faces d'une même pièce car une segmentation (que ce soit dans la production ou la perception) de la parole enchaînée effectuée par la prosodie implique nécessairement que certains éléments ou morceaux de la parole sont en même temps soudés par unités, ce qui est une sorte de liage. Pareillement, la désambiguïsation syntaxique par la prosodie se fait à travers la segmentation et le liage.

- **La prosodie et la mémoire**

L'importance de la prosodie pour la mémoire a été déjà démontrée dans la littérature (par exemple, O'Connell *et al.*, 1968 ; Leonard, 1973 ; Harriman & Buxton, 1979 ; Frankish, 1995 ; Paris *et al.*, 2000 ; Rosner *et al.*, 2004 ; Boucher, 2006, etc.). La prosodie a l'effet de renforcer la mémoire de travail et de maintenir l'information émise au cours de la chaîne de parole pour le traitement ultérieur. C'est principalement la segmentation prosodique qui décompose la chaîne de parole en morceaux conformes aux règles grammaticales et syntactiques, lesquels facilitent le stockage dans la mémoire à court terme de l'auditeur et puis la compréhension. Lorsque la prosodie est absente dans la parole synthétisée ou monotone, l'auditeur doit fournir davantage d'efforts pour traiter les informations sonores⁹ (Pratt, 2018).

⁹ "... the grouping of spoken utterances into phrases may take advantage of this feature in the auditory buffer. The temporal phrasing of speech may increase the efficiency of auditory memory, and thus play a significant role in comprehension. Prosody is hypothesized to be a critical component of the subvocalization routine, and thus presents consequences for the storage and processing of material in the phonological loop... there is a long history of research supporting its effect on memory for speech. The effect of prosody on memory is clearly demonstrated in studies showing that synthetic speech produced without prosodic cues adversely affects memory and comprehension... Natural prosody in speech and oral reading is more effective in aiding memory than monotone reading... When prosody is

- La prosodie dans la lecture silencieuse (et à voix haute)

Lors de la lecture silencieuse, on ne lit pas mot par mot, mais par groupes de mots¹⁰ et « le lecteur ne peut s'empêcher d'entendre intérieurement certaines syllabes comme accentuées » (Martin, 2018 : 85), un processus qui s'appelle « subvocalisation », ou « parole intérieure », ou encore « endophasie ». L'importance de la prosodie dans la lecture (que ce soit à voix haute ou silencieuse) est principalement liée à la segmentation, et puis la segmentation est étroitement liée à la mémoire pour favoriser la compréhension : d'une part, la compétence prosodique donne aux lecteurs la capacité de segmenter correctement une chaîne de caractères écrits, qui sont simplement séparés par des espaces et des ponctuations et qui sont privés des propriétés sonores évidentes comme l'intonation, la pause, l'accentuation dans le langage oral et ; d'autre part, la bonne segmentation facilite ensuite la désambiguïsation, le stockage, la compréhension et l'intégration au niveau discursif supérieur¹¹ (Fodor, 2002¹² ; Martin, 2018 ; Pratt, 2018). Pendant ces dernières années, beaucoup d'autres études psycholinguistiques ont été effectuées pour examiner le rôle spécifique des éléments prosodiques dans la lecture silencieuse (cf. Breen (2014)) pour une revue de littérature complète sur ce sujet).

La lecture à voix haute, tout comme la lecture silencieuse, exige aussi que le lecteur soit capable de faire la bonne segmentation syntactique ou prosodique. Pour les apprenants au cours de l'acquisition d'une langue étrangère, cette capacité a un lien direct avec la fluidité dans les lectures à voix haute et silencieuse (Pratt, 2018). En se basant sur des théories et des conclusions mises en lumière par des études sur ce

disrupted or removed completely, greater strain is placed on memory and memory-based tasks, an effect seen even more robustly in complex and demanding parsing contexts. (Pratt, 2018: 375).

¹⁰ « ... un mot isolé ne peut être prononcé, même silencieusement, qu'en formant un groupe accentuel... la segmentation en groupe accentuels est inévitable. » (Martin, 2018 : 82, 85).

¹¹ Ce point revient à la fonction que joue la prosodie quant à la mémoire : *“Appropriately chunking text into syntactically and semantically related groups reflects cognitive restructuring of the input and leads to successful encoding into memory... Early studies in reading fluency noted the correlation of fluency with comprehension, also noting that greater reading fluency is associated with more “appropriate” prosodic phrasing and contour... Prosody may be seen as an intermediary between fluency and comprehension, such that individuals who demonstrate appropriate prosody are more likely to exhibit better comprehension as well... The ability to group words into appropriate units is a key aspect of fluent reading, and non-fluent readers tend to either read word-by-word or group words differently than typical oral speech patterns.”* (Pratt, 2018: 376).

¹² *Implicit Prosody Hypothesis (IPH): “In silent reading, a default prosodic contour is projected onto the stimulus, and it may influence syntactic ambiguity resolution. Other things being equal, the parser favors the syntactic analysis associated with the most natural (default) prosodic contour for the construction.”* (Fodor, 2002).

sujet, des implications didactiques ont été aussi proposées quant à l'entraînement de la lecture en langue étrangère (par exemple, Pegolo, 1985).

1.1.4 Synthèse

Dans cette première section, nous avons d'abord situé l'importance de la prosodie dans la communication verbale : la prosodie intervient dans l'ensemble du processus communicatif verbal, en allant ensemble avec d'autres ressources linguistiques, pour véhiculer le message du locuteur. La complexité de la prosodie ou des phénomènes prosodiques a été soulignée tout au long de cette section :

- 1) Tout d'abord en raison des multiples phénomènes qu'englobe le terme « prosodie » : nous avons retenu la définition fournie par Di Cristo (2013) car elle offre la description la plus claire et complète ;
- 2) Ensuite, en raison des paramètres matériels associés à la prosodie : en soulignant les principaux éléments matériels liés à la prosodie, nous avons aussi mentionné leurs manifestations à travers différents phénomènes (ton, accentuation, rythme, intonation, etc.) et à divers niveaux (substantiel, formel, articulatoire, acoustique et perceptive) ;
- 3) Enfin, en raison de la pluralité des fonctions que la prosodie assume dans la parole : après avoir synthétisé sous forme de tableau la discussion exhaustive de Di Cristo (2013) sur ce sujet, nous avons mis les fonctions cognitives en évidence : notamment la segmentation et le liage prosodiques, qui fonctionnent ensemble pour désambiguïser la syntaxe et organiser le discours. Étant un point important qui sera examiné dans la partie expérimentale, nous avons également insisté sur comment la segmentation et le liage peuvent aider le traitement des informations linguistiques dans la mémoire, dans la lecture, ainsi que dans la production et perception de la parole.

Dans les sections suivantes 1.2 et 1.3, nous examinerons deux aspects matériels de la prosodie : mélodie (section 1.2) et temporalité (section 1.3). Dans les deux sections, en plus de la présentation générale de ces deux paramètres matériels, les patrons mélodiques et temporels du français et du chinois mandarin seront aussi introduits et comparés.

1.2 Aspects mélodiques de la prosodie en chinois mandarin et en français

Dans cette section, nous regarderons principalement les aspects mélodiques de la prosodie, ce qui implique l'analyse acoustique de la F0, en s'intéressant plus spécifiquement au chinois mandarin et au français. Nous commencerons par le chinois mandarin (section 1.2.1), en explorant les tons lexicaux, l'organisation et la segmentation prosodiques, ainsi que ses patrons mélodiques de base. À travers cette section, nous donnerons une vue globale sur comment la F0 est utilisée dans cette langue. L'accent sera notamment mis sur ses particularités mélodiques en matière d'organisation et de segmentation prosodiques, un sujet à la fois complexe et peu décrit (et parfois de manière controversée) dans la littérature, probablement en raison de la relation complexe entre tons lexicaux et intonation. Pour les aspects mélodiques du français (section 1.2.2), nous disposons de beaucoup d'études concernant ce sujet et nous synthétiserons donc les approches principales et les modèles connus qui jalonnent les grandes évolutions théoriques au cours de ces dernières décennies, sans oublier, bien sûr, les patrons mélodiques dans l'organisation et la segmentation prosodiques en français. Nous essayerons, dans la section 1.2.3, de fournir une comparaison des aspects mélodiques en chinois mandarin et en français, en mettant en lumière les différences majeures quant à l'organisation et la segmentation prosodiques entre les deux langues.

1.2.1 Aspects mélodiques de la prosodie en chinois mandarin

1.2.1.1 Méloies (F0) en chinois mandarin

La manipulation de la F0 revêt une importance particulière en chinois mandarin, car elle est utilisée à la fois pour les tons lexicaux et pour l'intonation. Dans cette section consacrée aux aspects mélodiques du chinois mandarin, il est important de mentionner d'abord le système des tons lexicaux du chinois mandarin, même si cela ne fait pas l'objet de cette présente étude. Ensuite, nous passerons aux particularités des patrons mélodiques en chinois mandarin, avec l'accent mis sur l'organisation et la segmentation mélodiques dans cette langue.

1.2.1.2 Tons lexicaux en chinois mandarin

Le chinois mandarin est une langue à tons, signifiant que chaque morphème peut se modifier en fonction du changement de sa F0 (Yip, 2002), ou plus précisément pour

le cas du chinois mandarin, du changement du contour de la F0. Les tons lexicaux sont donc dotés d'une fonction lexicale distinctive.

Toutes les langues chinoises sont des langues à tons et les nombres de tonèmes sont différents en fonction de la variété géographique (Zhu & Wang, 2015). Dans la Figure 1.3, les variations des contours mélodiques pour les quatre (ou parfois cinq, si le ton neutre est compté) tons lexicaux du chinois mandarin standard sont représentées. Le Tableau 1.2 présente des exemples de morphèmes, illustrant comment les quatre tons lexicaux confèrent des significations différentes à ces items.

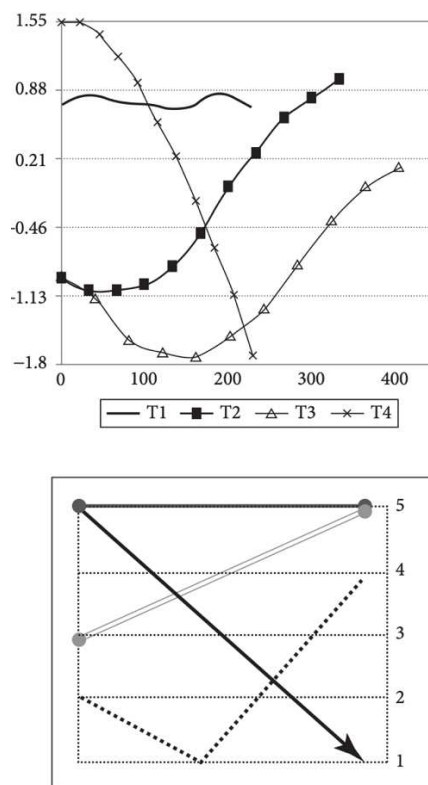


Figure 1.3 : Tons lexicaux du chinois mandarin standard : variations de contours mélodiques représentées par les valeurs moyennes de F0 (haut) et par l'échelle à cinq niveaux (bas) (Zhu & Wang, 2015 : 507).

Tableau 1.2 : Exemples des morphèmes [ma] et [pa] portant les quatre tons lexicaux et leurs significations différentes (tirés de Lin, 2007 : 89).

Tonal contrasts in SC

Tone number	Pitch pattern	Pitch value	Example
1	high level	55	[ma] ₅₅ mā 'mother' 妈
2	high rising	35	[ma] ₃₅ má 'hemp' 麻
3	low falling-rising	214	[ma] ₂₁₄ mǎ 'horse' 马
4	high falling	51	[ma] ₅₁ mà 'to scold' 骂

Additional examples for the tonal contrasts in SC

Tone number	Pitch pattern	Pitch value	Example
1	high level	55	[pa] ₅₅ bā 'eight' 八
2	high rising	35	[pa] ₃₅ bá 'pull out' 拔
3	low falling-rising	214	[pa] ₂₁₄ bà 'handle' 把
4	high falling	51	[pa] ₅₁ bà 'father' 爸

1.2.1.3 Relation et interaction entre tons et intonation en chinois mandarin

En chinois mandarin, en plus des tons lexicaux, la F0 est également utilisée, comme dans toute autre langue, pour l'intonation. Une question centrale et commune pour la recherche sur la phonétique du chinois mandarin (mais aussi pour d'autres langues à tons), est l'interaction entre tons et intonation : la F0 étant utilisée par les tons lexicaux, comment les différentes modalités intonatives se manifestent-elles ? Un constat similaire et général, révélé dans beaucoup d'études, semble établir un consensus concernant cette question : les différentes modalités intonatives semblent être signalées par des registres mélodiques différents (ou l'étendue de la F0 (*pitch range*) (Xu, 2015)) (cf. Shen (1989)¹³ et Lin (2007) pour une présentation détaillée¹⁴). Proposé pour la première fois par Shen (1989) et repris par Lin (2007), le schéma ci-dessous (Figure 1.4) illustre les variations mélodiques pour trois

¹³ "It is clear that the terminal shape does not constitute the cue for discriminating the various intonation patterns; rather, it is the pitch register which is the clue to their differentiation. My analysis proves the validity of Siertsema's claim (1962: 392): "In the tone language, ... the ... intonation rests on the range of the intervals rather than on the tonal contour, which is fixed by the lexical ... structures of the utterance." Thus we may assume that generally in tone language it is the register that has intonational function, whereas in the non-tone languages it is the contour." (Shen, 1989: 72).

¹⁴ "... the first strategy is to manipulate the pitch level: (i) the pitch level of the whole utterance may be raised higher or be depressed lower; and/or (ii) the high pitch of a tone gets higher and the low pitch of a tone gets even lower (cf. Chao 1933: 131-2; Shen 1989a, b). The second strategy is to add a H or L pitch after the tone features of the sentence final syllable (Chao 1933: 132-3)." (Lin, 2007: 230).

modalités intonatives de base, permettant de voir les changements du registre mélodique en fonction de ces modalités. Au début des phrases, les registres aux points de départ pour les trois modalités se situent à peu près au même niveau, avec les interrogatives (types II et III) se situant légèrement plus haut. Vers la fin des phrases, le registre mélodique de la question polaire (type II) s'élève progressivement vers une région fréquentielle plus haute, tandis que l'interrogative (type III) descend vers un registre bas, atteignant le même niveau que la déclarative (type I). L'interrogative (type III) se distingue de la déclarative (type I) uniquement par la partie du registre haut avant la fin de la phrase.

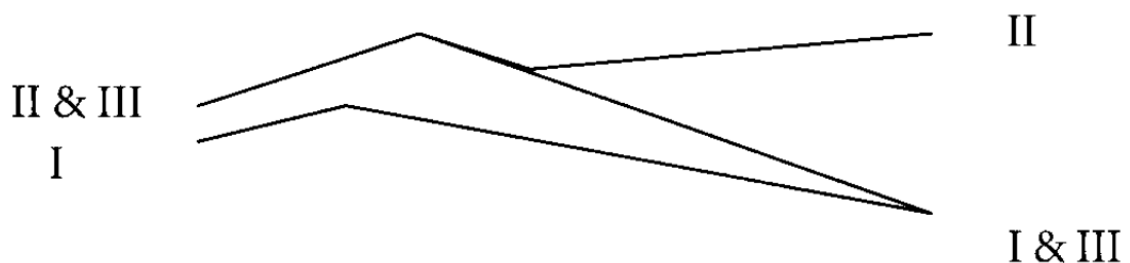


Figure 1.4 : Trois patrons mélodiques de base du chinois mandarin (avec I pour la modalité déclarative, II pour la question polaire avec la structure syntaxique énonciative, III pour l'interrogative) (Lin, 2007 : 229).

Dans tous les cas, le contour global de la F0 de chaque syllabe ou morphème est préservé et ne peut subir une grande variation mélodique (notamment quand il s'agit d'un ton lexical montant à la fin d'une déclarative ou d'un ton lexical descendant à la fin d'une interrogative, on admet uniquement une modification légère de montée ou de descente pour l'interrogative ou la déclarative) de sorte que les auditeurs parviennent toujours à reconnaître le ton précis quelle que soit la modalité intonative¹⁵ (Cao, 2004a, 2004b ; Lin, 2007 ; Zhang *et al.*, 2020).

¹⁵ "For example, the high level tone (T1) tends to be slightly falling in a statement and slightly rising in a question, but the tone is not altered beyond recognition..." (Lin, 2007: 229). "Standard Chinese opts to cue lexical tones at the cost of potential intonation misidentification (Ho 1977; Shih 1988; Liu and Xu 2005). A falling tone at the end of a question maintains its falling f0 contour (but would not fall as low as in declaratives), while a rising tone at the end of a statement is realized with its characteristic rising f0 contour (but at a relatively lowered f0 level compared to that in questions). Neither the global f0 raising nor the local f0 rise distort the lexical tone contours." (Zhang *et al.*, 2020: 342).

1.2.1.4 Patrons mélodiques en chinois mandarin : organisation et segmentation

La plupart des études sur les intonations en chinois mandarin traitent de l'interrogative et du focus informationnel (cf. Xu (2015) et Zhang *et al.* (2020)) pour une revue de littérature). Nous n'aborderons pas toutes les modalités intonatives (cf. Figure 1.4, et aussi Shen (1989) et Lin (2007) pour les patrons mélodiques de base), car cette étude se concentre sur la manière dont la F0 se manifeste dans l'organisation et la segmentation prosodiques. Comme nous l'avons évoqué dans la section précédente (section 1.2.1.2), la variation du contour mélodique est largement contrainte par la fonction primordiale et distinctive des tons lexicaux. Lorsque nous parlons de l'organisation et plus spécifiquement de la segmentation mélodique, il est important de mentionner l'aspect temporel, même si cela fera l'objet d'une présentation plus détaillée dans la section 1.3.

La segmentation et le regroupement prosodiques en chinois mandarin, notamment pour la désambiguïsation au niveau syntaxique, est un sujet qui a fait déjà l'objet de quelques études. Pour autant, nous ne disposons pas de suffisamment de travaux pour connaître la manière précise dont les locuteurs du chinois mandarin utilisent la prosodie à cette fin. En effet, les conclusions dans les études concernant ce sujet sont plus ou moins controversées, notamment en ce qui concerne l'ordre de préférence pour les indices acoustiques privilégiés pour faire la segmentation prosodique (cf. *infra*). Nous présentons dans ce qui suit les études disponibles qui traitent de ce sujet et récapitulons les points importants en regardant respectivement l'organisation et la segmentation mélodiques en chinois mandarin. Étant donné l'interaction complexe entre tons lexicaux et intonation (cf. section 1.2.1.3), tous deux liés à la F0, les chercheurs dans la plupart des études ont indiqué que la pause (ou/et la durée) est toujours l'indice le plus utilisé que ce soit pour l'organisation ou la segmentation prosodique en chinois mandarin tandis que la F0 (plus précisément le contour de la F0) joue un rôle plus limité¹⁶. Shin (2000), par exemple, a examiné la variation du contour mélodique dans plusieurs modalités

¹⁶ "In Mandarin Chinese, boundary strength is realized with gradient means rather than categorical ones, differentiated mainly in pre-boundary lengthening and optional silent pause, but not F0 (Xu and Wang, 2009; Wang *et al.*, 2018b). Although pitch reset has been found at a strong boundary (Dutch: de Pijper and Sandeman, 1994; Swerts, 1997; English: Ladd, 1988), F0 plays a limited role to distinguish boundary strength in Mandarin Chinese when tones and focus are carefully controlled (Xu and Wang, 2009; Wang *et al.*, 2018b)." (Wang *et al.*, 2022: 4).

intonatives différentes. De manière générale, le patron mélodique garde une tendance relativement fixe, qui est la déclinaison¹⁷. Si nous voulons examiner de plus près les configurations mélodiques précises du chinois mandarin, il existe dans la littérature de nombreux travaux pertinents, notamment les travaux de Cao (2000a, 2001, 2004a, 2004b), Yang et Wang (2002), Wang *et al.* (2005), Xu et Wang (2009), Chen et Tseng (2012), Wang (2016), Wang *et al.* (2022). Les études de Shen (1992, 1993), Chow (2005), Yang *et al.* (2014), ainsi qu’Ip et Cutler (2022) se sont intéressées à la façon dont les indices mélodiques (et prosodiques) peuvent être utilisés pour désambiguïser les sens des phrases ambiguës. Les conclusions principales concernant l’organisation et la segmentation prosodiques dans ces travaux ont été tirées et récapitulées dans le Tableau 1.3. Comme nous l’avons précédemment indiqué, il est important d’inclure certaines particularités des aspects temporels pour mieux illustrer le rôle limité de la F0 dans l’organisation et la segmentation de la prosodie du chinois mandarin.

Tableau 1.3 : Conclusions générales dans des études sur les patrons d’organisation et de segmentation prosodiques du chinois mandarin.

Étude	Organisation	Segmentation
Ip & Cutler (2022)	\ (non examinée)	- 1 ^{er} indice : pause ; - Allongement vocalique avant la frontière ; - F0 ne joue aucun rôle.
Wang <i>et al.</i> (2022)	Tendance de « downstep »	- La F0 n’est pas utilisée pour marquer la frontière prosodique.
Wang <i>et al.</i> (2018)	Non étudiée spécifiquement, mais au niveau du patron mélodique : F0 élevée au début d’un syntagme et F0 abaissée à la fin de ce syntagme ¹⁸ .	- L’allongement final et la pause sont utilisés pour signaler la frontière. La combinaison des deux indices est tout à fait suffisante pour percevoir les frontières de tailles différentes ¹⁹ .

¹⁷ “‘Declination’ refers to the overall downward tendency of F0 throughout the course of an utterance.” (Ladd, 1983: 45).

¹⁸ “The boundary effect shows mostly in the preboundary word, in the form of phrase-final F0 lowering at a strong boundary. Phrase-initial F0 raising occurs only when a boundary is strong, and mostly only in terms of minimum F0.” (Wang *et al.*, 2018: 48).

¹⁹ “In the present study, we found that by combining the two durational cues, i.e., preboundary lengthening and silent pause, as proposed in Xu (2009), not only can the 3 boundaries be well distinguished, but also the degree of similarity between different boundaries is clearly marked.” (Wang *et al.*, 2018: 49).

		<ul style="list-style-type: none"> - La variation de F0 est limitée aux frontières de tailles différentes.²⁰ - Le ton lexical avant la frontière est hyperarticulé, ce qui peut servir comme indice pour la perception de la frontière.
Wang (2016)	<ul style="list-style-type: none"> - Déclinaison à l'intérieur du syntagme intonatif ; - Plus on s'approche de la fin de la phrase, plus la variation du registre mélodique est grande (« <i>soft-landing model</i> »)²¹. 	Déclinaison du registre évident sur le mot prosodique avant la frontière prosodique ²¹ .
Xu (2015)	\	F0 ne joue pas de rôle. L'allongement à la fin de la frontière prosodique est un indice largement utilisé ²² .
Yang <i>et al.</i> (2014)	\	<ul style="list-style-type: none"> - Dans la production : Pause > Réhaussement de registre > Allongement final - Dans la perception : Pause > Allongement final = Réhaussement de registre
Xu & Wang (2009)	La durée joue le rôle le plus important dans le regroupement tandis que la F0 n'a pas de fonction particulière.	\
Wang <i>et al.</i> (2005)	Déclinaison à l'intérieur du syntagme intonatif.	<ul style="list-style-type: none"> - Pause ; - Réhaussement de registre.
Cao (2001, 2004a, b)	Tendance générale de déclinaison ²³ .	<ul style="list-style-type: none"> - Réhaussement de registre aux frontières prosodiques.
Yang & Wang (2002)	Déclinaison à l'intérieur du syntagme prosodique.	<ul style="list-style-type: none"> - Frontière du mot prosodique :

²⁰ "It is surprising that, when all the other factors are largely controlled, the effect of boundary on F0 is limited and sometimes even not sensitive enough to distinguish all the three boundaries." (Wang *et al.*, 2018: 48).

²¹ "The *soft-landing model* is proposed on the basis of the top line of the pitch contour. In this study, the pitch range, which is the difference between the top and bottom point, is also investigated, and it is found that **the pitch range of the earlier prosodic words are small, with that of the last the largest**, which is due to the 'final lowering' effect. Final lowering, the lowering of pitch at the end of an intonational phrase, has been observed in many languages, like Spanish and Yoruba... Words near the end of the intonation phrase may have 'final lowering' effect, which have demarcative function, i.e., marking the end of a sentence. As a result, the pitch range of the final prosodic word is the greatest." (Wang, 2016: 13).

²² "For Mandarin, domain-final lengthening is also consistently found (Tseng *et al.* 2005; Yuan *et al.* 2006; Xu and Wang 2009). A further question about boundary marking is whether it also involves other cues. There is some evidence that at least at the phrasal level, no direct F0 marking is involved in Mandarin (Xu and Wang 2009). At the sentential level, however, the issue is still unclear. Sentence-final lengthening is even more extensive than that at the phrasal level, thus using the same cue as sentence-medial boundaries (Tseng *et al.* 2005; Yuan *et al.* 2006)." (Xu, 2015: 497).

²³ "There does contain an underlying form of entire intonation, which presents as the declined trend of pitch register. This trend is slowly downward and upward alternately due to the resetting at different prosodic boundaries." (Cao, 2004a: 37). "Resetting of pitch register, i.e. a dislocation of pitch register taking place between rhythmic units is another important boundary marker... The uninterrupted declination tendency of pitch register in a prosodic phrase or intonation phrase usually has a connective function signaling coherence within the unit, while a break, i.e., resetting in the declination tendency, has a demarcative function signaling a unit boundary." (Cao, 2001: 67, 74).

		1) Un léger réhaussement de registre ; 2) Allongement final. - Frontière au syntagme prosodique ou intonatif : 1) Pause ; 2) Réhaussement de registre ; 3) Allongement final.
Shen (1992, 1993)	\	- Shen (1992) : Pause et allongement final ; - Shen (1993) : Descente du registre mélodique, abaissement d'intensité, pause et allongement final.

Parmi toutes ces études disponibles, trois ont particulièrement retenu notre attention car elles abordent directement le cœur de notre recherche, à savoir la désambiguïsation syntaxique par la prosodie : Ip et Cutler (2022), Yang *et al.* (2014), et Chow (2005). Il nous semble pertinent de résumer ces trois études ici pour donner une vision plus détaillée sur comment la prosodie du chinois mandarin contribue à la désambiguïsation syntaxique.

- *Ip et Cutler (2022)*

Deux types de phrases ambiguës (clôtures précoce vs tardive), contenant les mêmes informations segmentales et qui ne peuvent se distinguer que par leurs différentes positions de frontières prosodiques, ont été utilisés pour les expériences de production et de perception. Leurs résultats ont montré que dans la production des clôtures précoces par rapport aux clôtures tardives, la pause est l'indice le plus saillant à la position de clôture, suivie de l'allongement avant la frontière. La F0 ne joue en revanche aucun rôle dans cette distinction. Dans la perception des deux types de clôtures, Ip et Cutler (2022) ont mené des tests avec des stimuli où la pause était soit conservée soit supprimée. Les auditeurs sinophones ont éprouvé des difficultés particulières à distinguer les deux types de clôtures lorsque la pause était absente. Selon les auteurs, privilégier constamment la pause pour la segmentation prosodique permet de réduire l'utilisation de la F0 et de la durée²⁴.

²⁴ "Disambiguation is thus more a feature of Mandarin speech processing than of English, and actually inserting a pause between words or phrases is a way of disambiguating sequences without altering either F0 or segmental durations." (Ip & Cutler, 2022: 38).

- *Yang et al. (2014)*

Yang *et al.* (2014) ont mené deux expériences perceptives dans lesquelles ils avaient examiné les effets de trois indices (pause, allongement final, réhaussement de registre) dans la perception et la différenciation des phrases ambiguës. Des stimuli avec plusieurs conditions prosodiques ont été créés : avec un seul indice (pause ou allongement final ou réhaussement de registre), deux indices (pause et allongement final) ou trois indices. Par leurs expériences, un ordre d'importance des trois indices acoustiques dans la perception des frontières prosodiques et dans la désambiguïsation syntaxique par la prosodie a été établi : d'abord la pause, comme déjà révélé dans beaucoup d'autres études, occupe la place prépondérante quant à la segmentation prosodique ; ensuite, ce sont l'allongement final et le réhaussement de registre, les deux ayant une fonction relativement équivalente²⁵.

- *Chow (2005)*

Chow (2005) a mené deux expériences, l'une sur la production et l'autre sur la perception, en utilisant des paires de phrases partageant les mêmes informations segmentales. Ces expériences visaient à examiner comment les trois indices prosodiques contribuent à résoudre l'ambiguïté syntaxique, en accordant une attention particulière à la F0, un aspect souvent ignoré dans les études antérieures. En plus de la pause et de l'allongement final, le rôle important du réhaussement de registre pour la segmentation prosodique²⁶ a été empiriquement démontré. Les résultats de Chow (2005) ont ainsi montré que, d'une part, les trois indices acoustiques avaient une importance équivalente dans la production, et que, d'autre part, dans la perception des phrases ambiguës, en l'absence des indices temporels (allongement final et pause), le réhaussement de registre était suffisant pour que les auditeurs sinophones perçoivent les frontières prosodiques.

Après avoir passé en revue la littérature disponible concernant les patrons mélodiques liés à l'organisation et à la segmentation prosodiques en chinois

²⁵ "...perception of an IPB is heavily dependent on the presence of pauses, even to the extent that it may overrule the contribution of other parameters such as pre-boundary lengthening and pitch reset." (Yang *et al.*, 2014: 7).

²⁶ "In a tonal language such as Mandarin, the melodic contour of the utterance is largely defined by the combination of individual lexical tones associated with each syllable. Despite the crowded tone space dominated by lexical tones, experimental results indicate that pitch reset is an effective marker of prosodic boundaries in speech production and perception... for conveying and recognizing prosodic boundaries." (Chow, 2005: 211, 212).

mandarin, nous constatons qu'un consensus a émergé parmi les chercheurs ayant analysé divers corpus. Nous résumons ici les principales caractéristiques identifiées :

- **Concernant l'organisation** : au niveau mélodique, les contours de la F0 étant principalement associés aux tons lexicaux, son mouvement reste très restreint. L'organisation mélodique au niveau de la phrase se manifeste par une déclinaison à l'intérieur du syntagme prosodique (cf. cercle rouge dans la Figure 1.5).

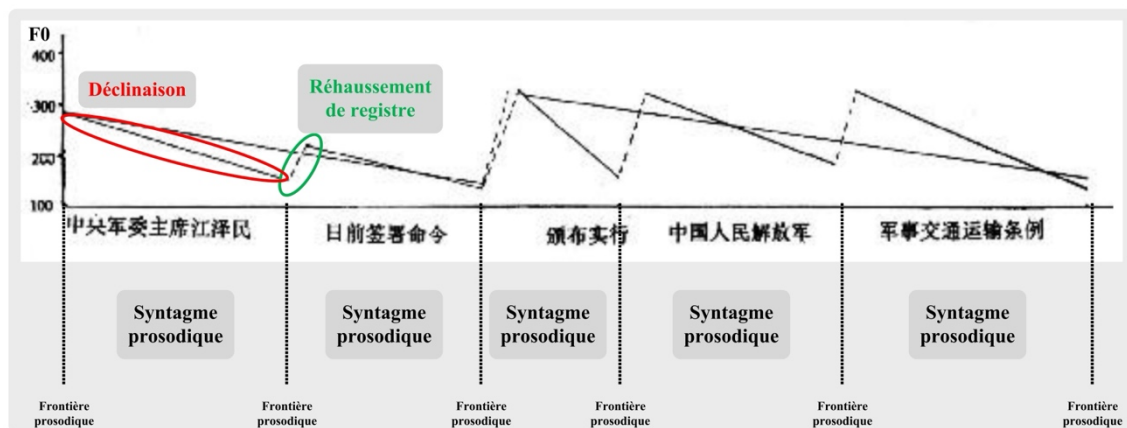


Figure 1.5 : Schéma d'organisation et de segmentation mélodiques pour un exemple d'énoncé en chinois mandarin (modifié de Cao (2000a)).

Bien que la déclinaison soit une particularité pour l'organisation mélodique, elle n'est pas propre au chinois mandarin. Cette caractéristique mélodique est universellement présente dans les langues du monde (dont le français), et résulte des contraintes physiologiques liées à la production de la parole (Vaissière, 1983).

- **Concernant la segmentation** :

- 1) La pause est le seul indice nécessaire et à la fois suffisant pour faire percevoir la frontière prosodique en chinois mandarin, la F0 et l'allongement final restant optionnels ;
- 2) Si l'aspect mélodique participe à la segmentation prosodique, il se manifeste par le réhaussement de registre (cercle vert dans la Figure 1.5). Plus précisément, à l'intérieur d'un syntagme prosodique, la F0 descend au niveau de son registre (i.e. déclinaison) et après la frontière prosodique, au début du syntagme prosodique suivant, elle remonte à un niveau du registre plus haut (*pitch resetting*²⁷), et puis redescend, et ainsi de suite. Le réhaussement de registre au début d'un syntagme

²⁷ "Pitch reset refers to the interruption of the declining sentence melodic contour at a prosodic boundary location. The pitch register of the syllable after the boundary is said to reset to a higher point where the contour continues to decline towards the end of the sentence (Pijper & Sanderman, 1994; Swerts 1996)." (Chow, 2005: 183).

prosodique est également considéré comme une des caractéristiques prosodiques universelles dans la production de la parole, contrainte, tout comme la déclinaison, par le processus physiologique²⁸. Comme proposé par Vaissière (1983), le réhaussement de registre, en conjonction avec la déclinaison, fonctionne pour lier les éléments segmentaux en syntagmes et délimiter leurs frontières²⁹. Le mécanisme décrit par Vaissière (1983) correspond précisément au fonctionnement de l'organisation et de segmentation prosodiques en chinois mandarin.

- 3) L'ordre de préférence des trois indices prosodiques pour la segmentation fait débat dans la littérature. Parmi ces indices, la pause est incontestablement considérée par la majorité des chercheurs comme le plus important. L'importance du réhaussement de registre et de l'allongement final est beaucoup plus discutée : selon les études et les corpus, le réhaussement de registre est parfois jugé plus significatif que l'allongement final, et parfois c'est l'inverse.

1.2.2 Aspects mélodiques de la prosodie en français

Depuis des décennies, le système intonatif du français était bien étudié par nombre de chercheurs (voir Chapitre 8 dans Di Cristo (2016) pour une revue complète sur diverses approches d'étude à partir du début du 20^{ème} siècle). La littérature abonde de modèles théoriques pour représenter et analyser les aspects phonologiques et phonétiques des intonations du français (Delattre, 1966b pour les dix modalités intonatives de base ; Jun & Fougeron, 1998, 2000, 2002 ; Di Cristo, 1976, 1978, 1998, 2016 ; Hirst & Di Cristo, 1984, entre autres). Nous aborderons ci-dessous l'organisation et la segmentation des patrons mélodiques du français en nous référant à quelques modèles importants.

²⁸ "A direct consequence of this tendency for F_0 to decline is that F_0 values tend to be lower near the end of breath-groups than at the beginning. After a pause, they are generally reset. We may suppose that resetting of F_0 values after a respiratory pause is a natural physiological tendency, although the relation between resetting and breathing is not entirely clear." (Vaissière, 1983: 57).

²⁹ "Together with the declination tendency and pauses, the resetting of F_0 may contribute to characterizing sentences and clauses as acoustic entities. It has been shown (by Maeda [1976] for English; by Carlson and Granstrom [1973] for Swedish; and by Fujisaki et al. [1979] for Japanese) that resetting of the baseline can be simulated between two phrases, even in the absence of a pause. The resetting is used as a boundary marker, and the degree of resetting indicates the importance of the boundary." (Vaissière, 1983: 57).

1.2.2.1 Organisation

Lorsque nous parlons de l'organisation mélodique du français, il est incontournable de mentionner la notion de « phrasé ». Le phrasé est une notion empruntée au domaine de la musique pour désigner dans le domaine de la prosodie (notamment du français) un constituant ou un morceau ou une unité prosodique : « *le concept de phrasé prosodique... présente des analogies avec celui de phrasé musical, qui se rapporte à la façon dont les notes d'une partition sont regroupées en mesures, conformément aux intentions des compositeurs et de leurs interprètes. Dès lors qu'il est question d'un découpage et du groupement du matériau linguistique, les chercheurs sont enclins à convoquer des notions comme celles de domaine et de constituance* » (Di Cristo, 2016 : 77). Le phrasé est une unité importante d'analyse prosodique pour le français, une langue étiquetée comme « *une langue à frontières par excellence* » (Vaissière, 2010) (nous reviendrons sur cet aspect plus en détail dans la section suivante (1.2.2.3) quand nous évoquons la segmentation prosodique en français). L'organisation des constituants prosodiques du français était bien étudiée depuis des décennies. Plusieurs modèles prosodiques, avec des terminologies diverses, explicitent les organisations hiérarchiques des éléments prosodiques du français par « phrasé ». Nous rappelons ci-dessous (Tableau 1.4) quelques modèles d'organisation, par ordre chronologique, et comparons les niveaux prosodiques et les terminologies différentes dans ces modèles.

Tableau 1.4 : Quelques modèles principaux d'organisation hiérarchique de l'intonation du français.

Études	Niveaux prosodiques	Exemple	Observations
Delattre (1966b)	Continuation majeure (CM) Continuation mineure (Cm)	(1) Mes voisins <i>CM</i> ont vendu <i>Cm</i> leur maison <i>Cm</i> en Bretagne <i>FIN</i> (2) Mes voisins <i>CM</i> ont vendu <i>Cm</i> leur maison <i>CM</i> en Bretagne <i>FIN</i>	
Di Cristo (1976, 1978)	Groupe Accentuel (GA) Groupe Intonatif (GI)	(1) Mes voisins <i>GI</i> ont vendu <i>GA</i> leur maison <i>GA</i> en Bretagne <i>GI</i> (2) Mes voisins <i>GI</i> ont vendu <i>GA</i> leur maison <i>GI</i> en Bretagne <i>GI</i>	- sur la syllabe finale, un contour mélodique montant ou descendant (si c'est la fin d'une phrase) ; - un écart mélodique nettement plus accusé entre la syllabe finale tonique et la syllabe prétonique ; - un modèle qui prend l'intonation (GI) et la proéminence (GA) en considération.
Hirst & Di Cristo (1984) Di Cristo & Hirst (1993) Di Cristo (1998)	Unité Tonale (UT) Unité Rythmique (UR) Unité Intonative (UI)	(1) La se) _{UT} crétaire) _{UT} du pré) _{UT} sident) _{UT}]UI a bien dé) _{UT} coré) _{UT} la salle) _{UT}]UI (2) La se crétaire) _{UT} du pré sident) _{UT}]UI a bien) _{UT} dé coré) _{UT} la salle) _{UT}]UI	- Gabarit tonal UT : LH - Gabarit tonal UI : LL ou LH
Jun & Fougeron (1995, 2000, 2002) Michelas & D'Imperio (2009, 2010, 2011) D'Imperio & Michelas (2014)	Syntagme Accentuel (AP : « Accentual Phrase ») Syntagme Intermédiaire (ip : « intermediate phrase ») Syntagme Intonatif (IP : « Intonational Phrase »)	(1) La mamie] _{AP} de Rémy] _{AP/ip} demandait l'institutrice. (2) La mamie] _{AP} des amis] _{AP} de Rémy] _{AP/ip} demandait l'institutrice.	- Forme prototypique du syntagme accentuel : AP : /LHiLH*/, avec « LHi » associé à la frontière de gauche et « H* » à la dernière syllabe de la frontière de droite ; - ip : L- ou H- ; - IP : L% ou H%.

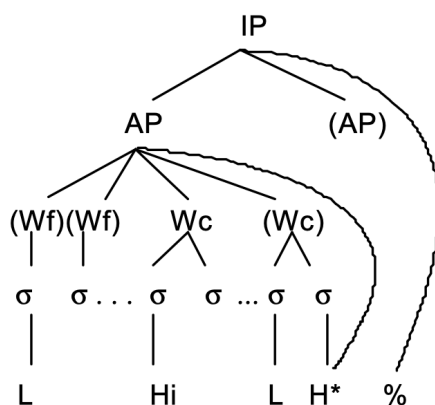
Parmi les modèles d'organisation prosodique récapitulés dans le Tableau 1.4, il est opportun d'aller plus en détail sur quelques-uns afin de donner une description de décomposition plus fine sur les phrasés mélodiques possibles du français.

D'abord, il s'agit de modèles qui s'inscrivent dans le cadre de l'approche autosegmentale métrique (l'approche AM) (Pierrehumbert, 1980 ; Beckman & Pierrehumbert, 1986) et qui ont évolué au fil des ans pour aboutir à une version adaptée spécifiquement à la langue française. Ces modèles adoptent le système ToBI (Delais-Roussarie *et al.*, 2015) comme protocole de transcription pour représenter la variation de la F0 de surface et le mouvement du contour mélodique dans une phrase donnée.

- *Modèle de Jun et Fougeron (1995, 2000, 2002)*

Les travaux de Jun et Fougeron (1995, 2000, 2002) sont sans doute le cadre théorique le plus connu parmi les études sur le système intonatif du français. Dérivé de l'approche AM (Pierrehumbert, 1980 ; Beckman & Pierrehumbert, 1986) et adapté au français, le modèle de Jun et Fougeron (2000) a proposé que l'intonation du français est composée, comme d'autres langues accentuelles (l'anglais, l'allemand, etc.), d'une séquence de tons hauts (H) et bas (B). Dans une phrase donnée, les tons H et B sont liés à des syllabes spécifiques, considérées comme cibles tonales (*tonal targets*).

Dans le cadre de l'approche AM, l'intonation du français est hiérarchiquement rangée en deux (ou trois) niveaux prosodiques : principalement syntagme accentuel (AP, *Accentual Phrase*) et syntagme intonatif (IP, *Intonational Phrase*) (Figure 1.6), avec éventuellement un troisième niveau : syntagme intermédiaire (ip, *Intermediate Phrase*) entre les deux (Michelas & D'Imperio, 2009, 2010, 2011 ; D'Imperio & Michelas, 2014).



IP: Intonation Phrase AP: Accentual Phrase
 Wf: function word Wc: content word
 σ: syllable %: Intonation phrase boundary

Figure 1.6 : Modèle de Jun et Fougeron (2000), avec deux niveaux prosodiques (IP, AP) et patron prototypique de l'AP.

Le patron prototypique du syntagme accentuel en français, qui est considéré comme un phrasé sous-jacent, est représenté par la suite tonale /LHiLH*/. La partie « LHi » marque la frontière gauche, considérée souvent comme « accent initial ou secondaire » et la partie « LH* » marque la frontière droite, considérée comme « accent final ou primaire », ou équivalente de « *pitch accent* » dans le cadre du système intonatif des langues anglo-saxonnes. Étant seulement une représentation sous-jacente, le phrasé mélodique /LHiLH*/ connaît plusieurs variantes dans la production de surface.

	/L Hi L H*/
	a. [L (Hi L) H*]
	b. [L (Hi) L H*]
	c. [L Hi (L) H*]
	d. [(L) Hi L H*]
	e. [L Hi (L) L*]

Figure 1.7 : Cinq patrons mélodiques de surface possibles proposés dans Jun et Fougeron (2000) pour le phrasé prototypique /LHiLH*/.

Le modèle et les patrons mélodiques proposés dans les travaux de Jun et Fougeron (1995, 2000, 2002) constituent une avancée notable par rapport aux modèles de Di Cristo (1976, 1978) et de Hirst et Di Cristo (1984). Le modèle de Jun et Fougeron (1995, 2000, 2002), en se basant sur son phrasé prototypique, a permis d'établir un inventaire des variantes possibles qui ont des cibles tonales précises pour la production de surface (Figure 1.7).

- *Post (2000)*

Contrairement aux approches précédemment présentées, qui décomposent l'intonation du français en divers phrasés, le schéma tonal proposé dans la thèse de Post (2000) est un modèle qui vise à représenter l'organisation mélodique d'une phrase comme un seul ensemble. « *L'étude de Jun & Fougeron est partielle et s'intéresse principalement à des questions de phrasé intonatif. Celle de Post est plus complète et sa finalité est de construire la grammaire à états-finis de l'intonation du français* » (Di Cristo, 2016 : 192). L'organisation mélodique d'un syntagme intonatif dans le modèle de Post (2000) est élargie par rapport à celui de Jun et Fougeron (1995, 2000) et est composée d'une série de primitives tonales : tons de frontière de gauche (%L, %H) et de droite (L%, H%, 0%) ; accents mélodiques (H* et H+H*) ; et le ton facultatif L qui se trouve éventuellement entre deux tons H (Figure 1.8).

$$\text{The Intonation Phrase:} \\ \left\{ \begin{array}{c} \%L \\ \%H \end{array} \right\} \quad (H^* (L))_0 \quad \left\{ \begin{array}{c} H^* \\ H + H^* \end{array} \right\} \quad \left\{ \begin{array}{c} L\% \\ H\% \\ 0\% \end{array} \right\}$$

Figure 1.8 : Prototype du syntagme intonatif du français proposé par Post (2000 : 227).

1.2.2.2 Synthèse sur les modèles d'organisation mélodique et phrasés en français

Dans cette section, nous avons passé en revue quelques modèles des patrons intonatifs du français, qui ont jalonné l'évolution de la recherche sur l'organisation hiérarchique mélodique du français. Les Figures 1.9 et 1.10, tirées de Di Cristo (2016), illustrent les patrons mélodiques avec dénominations diverses et leurs organisations dans ces modèles, d'abord avec deux niveaux (minimalistes) (Figure 1.9) et puis avec trois niveaux (Figure 1.10).

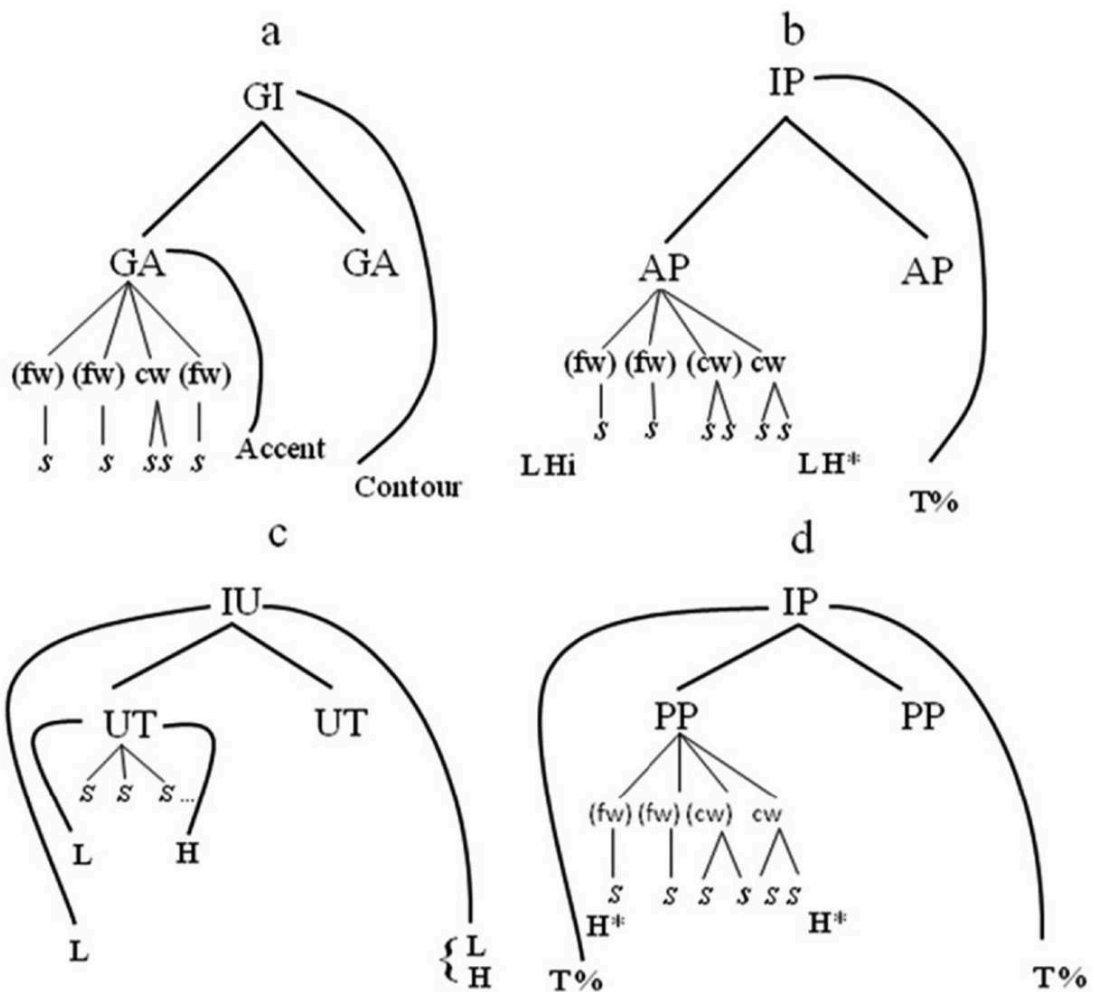


Figure 1.9 : Architecture des différents modèles **minimalistes** d'organisation des patrons mélodiques du français (a : Hirst & Di Cristo (1984) ; b : Jun & Fougeron (1995) ; c : Hirst & Di Cristo (1984) ; d : Post (2000)) (tirée de Di Cristo, 2016 : 118).

Les modèles illustrés dans la Figure 1.9 ont été appelés « minimalistes » car ils ne comprennent que deux niveaux de phrasés prosodiques. Comme nous l'avons déjà évoqué, un troisième niveau a été proposé pour caractériser les patrons mélodiques du français de manière plus fine : il s'agit notamment du niveau « ip » (*intermediate phrase*, « syntagme intermédiaire ») ou du niveau « unité rythmique » chez Di Cristo et Hirst (1993) (appelé « mot prosodique » chez Di Cristo (1998)) (Figure 1.10).

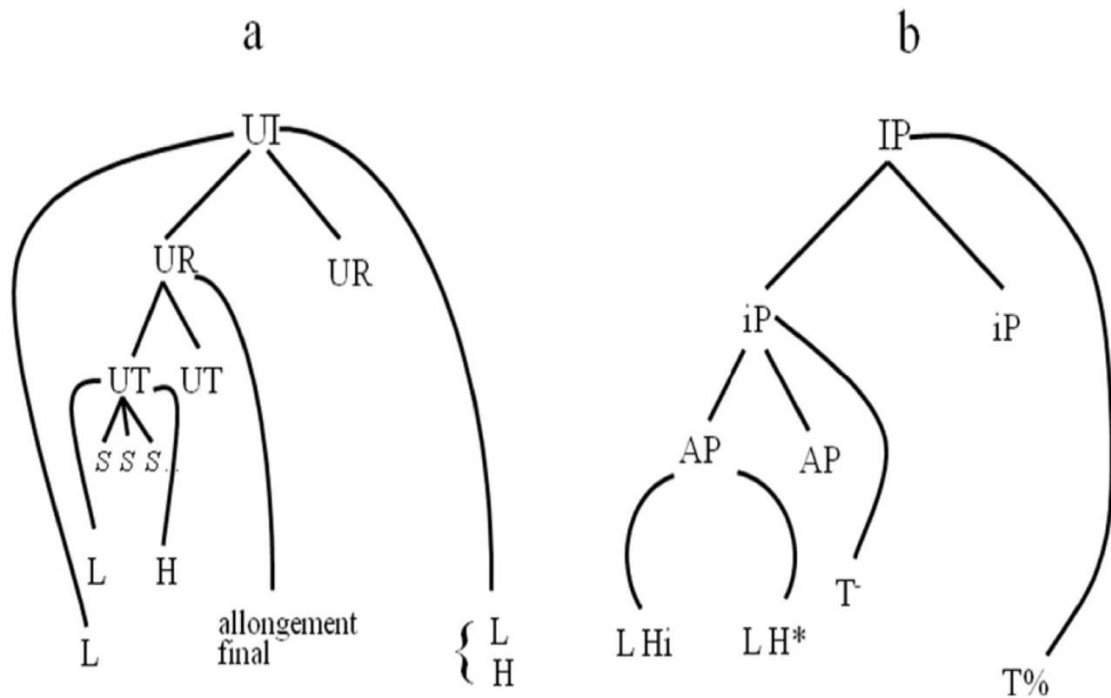


Figure 1.10 : Architecture des modèles d'organisation des patrons mélodiques du français à **trois niveaux** (a : Di Cristo & Hirst (1993) ; b : Jun & Fougeron (2000)) (tirée de Di Cristo, 2016 : 123).

Même si des prosodistes ont établi des modèles avec des terminologies différentes pour dénommer les phrasés et les niveaux mélodiques, l'organisation mélodique du français se manifeste par la même particularité : « *un groupe de syllabes délimité par un accent final (ou primaire)* » (Di Cristo, 2016 : 117). Ce groupe de syllabes, qu'il soit désigné comme « *syntagme accentuel* » par Jun et Fougeron (1995), « *unité tonale* » dans Hirst et Di Cristo (1984) ainsi que Di Cristo (1998), ou d'autres (cf. Figure 1.11 pour les équivalences terminologiques entre les différents modèles), est « *généralement jugé comme quasi-équivalent par la plupart des prosodistes français, constitue en fait des réminiscences de la notion plus ancienne de Groupe Accentuel (GA)* » (Di Cristo, 2016 : 140).

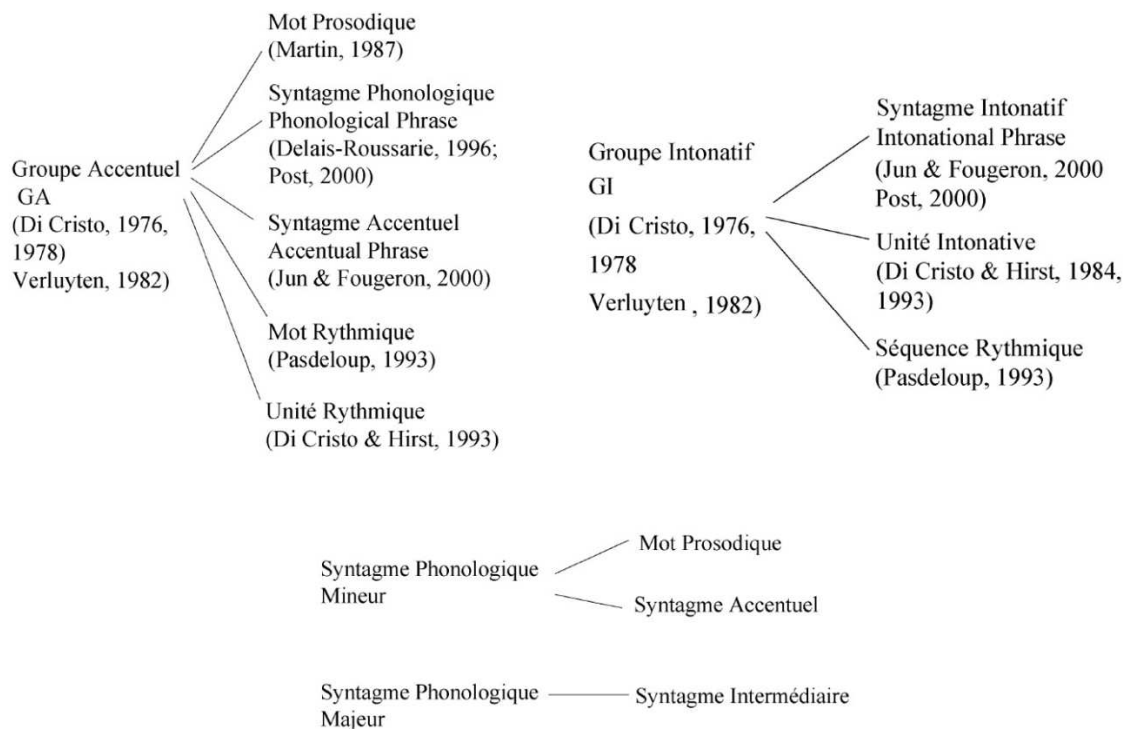


Figure 1.11 : Équivalences entre les dénominations différentes pour les niveaux des phrasés dans les modèles divers de la prosodie du français (tirée de Di Cristo, 2016 : 141).

1.2.2.3 Segmentation

Contrairement à l'anglais, qui est perçu comme une langue à accent lexical (Vaissière & Michaud, 2006), le français est défini par Vaissière (1991, 2010) comme « une langue à frontières »³⁰. Cela s'explique par les indices acoustiques des frontières, tels que l'augmentation de la F0 et l'allongement initial ou final, qui se manifestent aux niveaux des phonèmes, des syllabes, des lexiques et des syntagmes supra-lexicaux en français. Le français est aussi considéré comme une langue mélodiquement montante (*a rising language*), avec l'allongement final. Aux oreilles des francophones natifs, le patron mélodique du français est marqué par des montées de continuation à la fin des syntagmes prosodiques³¹.

³⁰ « Parmi les langues à intonation, nous avons proposé de considérer le français comme « une langue à frontières », où domine l'aspect démarcatif de la prosodie, par opposition aux langues dites à accent, comme l'anglais (Vaissière, 1991), où domine son aspect culminatif. » (Vaissière, 2010 : 16).

³¹ "French is generally considered as a 'rising' language with final lengthening. An insightful account of French prosodic phrasing was proposed by Delattre 1966 (usefully complemented by Fónagy 1980). The French ear is trained to perceive continuation at the end of a prosodic phrase. In this sense, French is a 'rising' language: each prosodic phrase inside a sentence tend to end with a sharp rise (Delattre's continuation majeure), or a smaller rise, or a high F₀ value (Delattre's continuation mineure)." (Vaissière & Michaud, 2006: 49).

En français, les locuteurs utilisent plusieurs marqueurs acoustiques pour signaler la frontière prosodique : allongement vocalique, montée mélodique (ou descente)³², pause, voire rehaussement de la ligne de base³³ (Delattre, 1966a ; Delais-Roussarie, 1993 ; Di Cristo, 1998 ; Vaissière, 1974, 1975, 1991, 2020³⁴ ; Vaissière, 2002 ; Vaissière & Michaud, 2006, Delais-Roussarie *et al.*, 2011, D’Imperio *et al.*, 2020). En général, c’est à la dernière syllabe, dont la position coïncide avec la frontière droite d’un syntagme accentuel ou/et d’un syntagme intonatif³⁵, que les locuteurs émettent des indices acoustiques et que les auditeurs s’attendent à les percevoir.

Parmi tous les indices acoustiques disponibles, Simon et Christodoulides (2016), ainsi que Christodoulides (2018) ont examiné les poids des indices (*cue-weighting*) dans la perception des frontières prosodiques, d’une part par le test de perception des indices prosodiques des frontières chez des auditeurs, et d’autre part par le classificateur automatique basé sur les valeurs des indices extraites des grands corpus. Leurs résultats ont montré une certaine homogénéité dans l’importance des indices acoustiques qui marquent les frontières prosodiques du français : 1) la pause est l’indice le plus significatif ; 2) elle est suivie par l’allongement de la durée de la dernière syllabe avant la frontière prosodique : plus le niveau de la frontière prosodique dans la hiérarchie est élevé, plus la durée syllabique est longue (Christodoulides, 2018) ; 3) enfin, le mouvement mélodique est l’indice le moins important³⁶. Il convient de noter que la préférence pour ces indices acoustiques n’est

³² “...a sharp fall in F0 before a pause indicates the end of a sense group in final position in a sentence; a rise followed by a pause indicates the end of a sense group in a nonfinal position in the sentence.” (Vaissière, 1974: 213).

³³ « ...la frontière gauche des syntagmes intonatifs est marquée par la présence d’un rehaussement tonal et la frontière droite par un allongement de la durée, voire une pause. » (Delais-Roussarie *et al.*, 2011 : 38). “The French ip is marked by a return to the register line (‘total reset’) determined by the first H* peak within the IP (D’Imperio and Micheals 2014). This internal ‘reset’ appears at an IP-internal right edge (hence in non-final ip’s), usually marking a major syntactic boundary, blocking the iterative downstep reported in Post (2000a; see also Delais-Roussarie *et al.* 2011).” (D’Imperio *et al.*, 2020: 247).

³⁴ « Le principe démarcatif de base en français est simple : à l’intérieur de l’énoncé, plus la dernière syllabe d’un mot est longue, plus elle est montante et plus la frontière est perçue comme forte, et la présence d’une pause renforce cette perception. » (Vaissière, 2020 : 112).

³⁵ « En français, la congruence entre accent lexical (qui se situe sur la dernière syllabe pleine des mots lexicaux, c’est-à-dire leur frontière droite) et les marques intonatives de continuation (sur les frontières droites des constituants en anglais) rend le marquage des frontières droites plus régulières et évidentes que dans les autres langues ». (Vaissière, 2010 : 16).

³⁶ « ...la pause silencieuse est l’indice le plus décisif dans la perception d’une frontière, suivi par l’allongement de la syllabe finale de mot. En parole naturelle, c’est ensuite la syntaxe (fin d’unité de rection puis fin de séquence) qui joue un rôle. La hauteur mélodique ne semble pas être un facteur décisif. » (Simon & Christodoulides, 2016 : 104). “We have shown that the main acoustic correlates of prosodic boundary strength are the presence of a subsequent silent pause and pre-boundary lengthening, in this order of importance. Pitch movements (relative pitch and intra-syllabic pitch

pas toujours classée de la même manière. Vaissière et Michaud (2006) ont donné un exemple avec 3 phrases ambiguës pour illustrer comment les locuteurs francophones natifs utilisent les indices acoustiques pour faire la segmentation prosodique et ensuite désambigüiser les sens des phrases (Figure 1.12). Dans l'illustration donnée dans leurs travaux (Vaissière & Michaud, 2006 ; Vaissière, 2020), ce sont la montée mélodique et l'allongement final de la syllabe avant la frontière qui servent d'indices principaux pour lever l'ambiguïté.

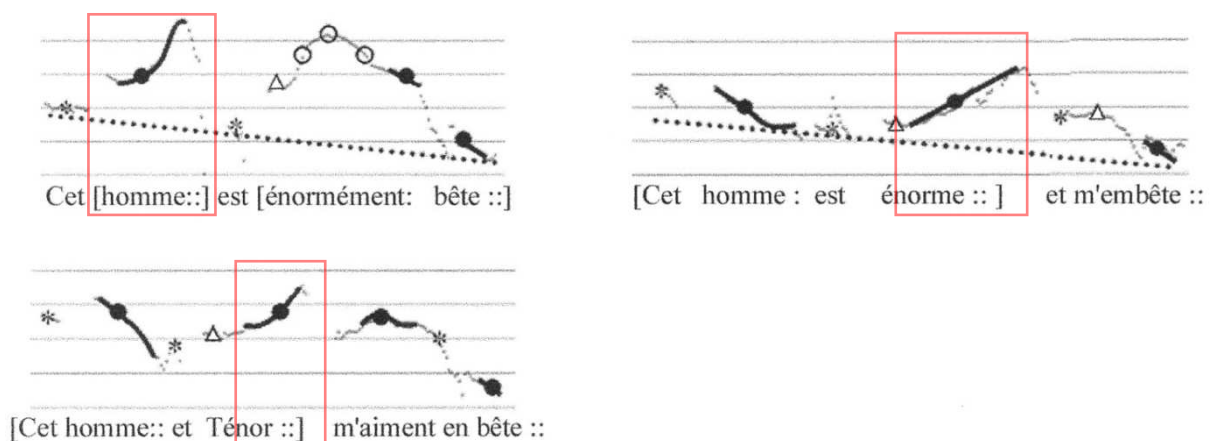


Figure 1.12 : Trois phrases ambiguës en français avec les mêmes informations segmentales mais différenciées uniquement par les indices prosodiques (« :: » pour indiquer l'allongement à la fin du syntagme, « : » pour l'allongement à la fin du lexique, « Δ » pour syllabe initiale, « • » pour syllabe finale, « * » pour mot grammatical) (tirée de Vaissière & Michaud (2006 : 54)).

Dans la Figure 1.12, nous pouvons voir que les segmentations prosodiques (dernières syllabes avant les frontières entourées en rouge), notamment quand il s'agit du regroupement des mots en unités de sens et de la désambigüisation syntaxique par la prosodie, les frontières (majeures) sont marquées conjointement par une montée mélodique et un allongement final³⁷. Selon Vaissière et Michaud (2006), une durée plus longue de la dernière syllabe, combinée à un mouvement mélodique (donc la F0) montant, rend la frontière prosodique plus marquée perceptivement.

Même si les locuteurs disposent d'une série d'indices acoustiques, tous ne sont pas exploités dans la production de la parole. En effet, l'absence d'un indice spécifique

movement) are indicative of prosodic boundary function, rather than strength; however, stronger prosodic boundaries (e.g. period boundaries in the Rhapsodie corpus) tend to correlate with larger pitch movements. (Christodoulides, 2018: 1545).

³⁷ "...the syllable (more precisely the syllable rhyme) with the strongest rising slope corresponds to a major boundary (Delattre), and is accompanied by lengthening." (Vaissière & Michaud, 2006: 54).

peut être compensée par l'hyperarticulation d'un autre. Le renforcement d'un ou deux indices prosodiques sélectionnés suffit pour que la frontière soit perçue par les auditeurs. Ainsi, selon Vaissière (2020 : 112) : « *la durée relative de la première syllabe suffit à distinguer entre « bordures » [bɔ̃ʁdy::ʁ] et « bords durs » [bɔ̃:ʁdy::ʁ], « Jean-Pierre et Jacques » et « Jean, Pierre et Jacques », sans qu'il soit besoin de modifier F0 ».*

1.2.3 Synthèse : comparaison des aspects mélodiques entre le mandarin et le français

Après avoir examiné les études sur les patrons mélodiques du mandarin et du français, nous pouvons identifier les points importants concernant les aspects mélodiques de ces deux langues. Cela nous permet de mettre en évidence le contraste existant entre l'organisation et la segmentation mélodiques dans ces langues.

- Pour le mandarin :

Nous pouvons constater un consensus sur les manifestations des aspects mélodiques dans cette langue et synthétiser ses caractéristiques concernant l'organisation et la segmentation prosodiques :

- 1) Le contour de la F0 est crucial pour le mandarin, car il est intrinsèquement lié aux tons lexicaux. La variation permise dans les contours au niveau intonatif est très limitée, afin que les auditeurs puissent reconnaître les différentes significations des morphèmes qui portent des tons lexicaux divers, indépendamment de la modalité intonative ;
- 2) Les modalités intonatives se distinguent par des registres différents : par exemple, les phrases interrogatives se manifestent par un registre mélodique plus haut sur l'ensemble de l'énoncé comparativement aux phrases déclaratives ;
- 3) Comparativement au français, l'organisation mélodique en mandarin est notamment caractérisée par la déclinaison à l'intérieur du syntagme prosodique ;
- 4) Concernant la segmentation, la pause est l'indice principal, tandis que le rôle de la F0 est très limité et elle n'est pas privilégiée dans la segmentation prosodique. La frontière prosodique est marquée mélodiquement par un ajustement du registre de la F0 : le registre descend à l'intérieur d'un syntagme ou d'une unité prosodique. Lorsqu'il s'agit d'une frontière et du début d'un nouveau syntagme,

la F0 remonte à un registre mélodique plus haut, puis redescend dans son registre, et ainsi de suite ;

- 5) La déclinaison à l'intérieur du syntagme prosodique et le réhaussement de registre à la position de la frontière prosodique constituent donc les deux caractéristiques les plus saillantes du patron mélodique du mandarin.

- Pour le français :

Les patrons mélodiques en français sont aménagés par phrasés en plusieurs couches : d'abord de manière minimaliste avec seulement deux niveaux (« CM » et « cm » chez Delattre (1966b) ; « GA » et « GI » chez Di Cristo (1976, 1978) ; « UT » et « UI » chez Hirst & Di Cristo (1984) ; « AP » et « IP » chez Jun & Fougeron (1995, 2000, 2002)). Des organisations avec trois couches ont été proposées plus tard (« UR » chez Di Cristo & Hirst (1993) ; « ip » chez Michelas & D'Imperio (2009, 2010, 2011) ; D'Imperio & Michelas (2014)).

Pour la segmentation prosodique, le français, tout comme le mandarin, privilégie la pause comme principal indice. Deux autres caractéristiques acoustiques saillantes sont la montée de la F0 et l'allongement final à la dernière syllabe avant la frontière prosodique. Ainsi, pour signaler la frontière prosodique, les marques acoustiques communes aux deux langues sont la pause et l'allongement final³⁸. Les marques mélodiques se manifestent différemment : en français, cela se traduit par une montée mélodique, impliquant une variation de contour, tandis qu'en mandarin, c'est par un réhaussement de registre, impliquant un changement de niveau du registre, avec un contour peu variable en raison de sa fonction distinctive pour les tons lexicaux.

1.3 Aspects temporels de la prosodie

Dans cette section, nous aborderons les aspects temporels de la prosodie : les systèmes métriques, temporels et la pause seront les éléments principaux. Nous illustrerons comment le patron prosodique est temporellement organisé par le rythme et d'autres phénomènes temporels dans les deux langues.

³⁸ Même si l'allongement final est un indice commun dans les deux langues, en français, les durées syllabiques des allongements finaux sont positivement corrélées avec les niveaux hiérarchiques prosodiques (voir par exemple Fougeron & Keating, 1997) alors qu'en chinois mandarin, ce n'est pas exactement le même cas. Nous expliciterons cette particularité temporelle du chinois mandarin dans la section suivante (1.3.2.6) quand nous évoquons les aspects temporels de la prosodie.

La syllabe, en tant qu'unité fondamentale de la structuration prosodique (notamment, rythmique), sera le premier élément que nous aborderons. Ensuite, nous examinerons le système rythmique (métrique), suivi du système macro-rythmique. Nous évoquerons également les fonctions du rythme dans le langage. Concernant l'étude du rythme, la notion d'isochronie est essentielle, car elle constitue la base de la typologie rythmique traditionnelle ainsi que des paramètres quantitatifs. Cependant, cette typologie traditionnelle est remise en cause par de nombreux chercheurs, et l'utilisation des paramètres quantitatifs dans l'étude du rythme du langage est également critiquée, tant pour l'étude de langues naturelles que pour l'analyse de l'interlangue chez les apprenants de langues étrangères. C'est pour cela qu'il nous semble important d'introduire le modèle *Control/compensation model* proposé par Bertinetto et Bertini (2008), qui permet de mieux décrire les patrons rythmiques langagiers et de rendre compte des différences temporelles que les paramètres quantitatifs traditionnels ne permettent pas de révéler.

Suite de la présentation générale sur le rythme et sa mesure, nous décrivons les systèmes rythmiques du chinois mandarin et du français. Après avoir exploré les patrons rythmiques des deux langues, nous nous attarderons sur le phénomène de la pause, un élément prosodique important mais souvent sous-estimé ou négligé dans l'étude prosodique. Enfin, la synthèse de cette section établira une comparaison entre le chinois mandarin et le français sur les plans métrique, rythmique et temporel en mettant en lumière les principales différences entre les deux langues.

1.3.1 Structures syllabiques

Dans cette section, l'accent sera mis sur la syllabe. Dans un premier temps, nous définirons la syllabe et son rôle dans la prosodie, puis nous présenterons les structures syllabiques du chinois mandarin et du français. Cette présentation des structures syllabiques dans les deux langues nous permettra de mieux décrire leurs patrons rythmiques.

1.3.1.1 La syllabe : unité de structuration prosodique

Définir la « syllabe » n'est pas une tâche aisée : plusieurs chercheurs ont proposé diverses définitions selon des perspectives phonétique, phonologique et

psychologique (pour une revue sur les définitions de la « syllabe », cf. Angoujard, 1997 ; Labrune, 2005 ; Meynadier, 2001 ; Ridouane *et al.*, 2011). La syllabe est considérée depuis longtemps comme une unité fondamentale de la structuration prosodique³⁹ (van Heuven, 1994) et comme une unité de base dans la segmentation de la parole enchaînée. Lors de l'acquisition de leur langue maternelle, les locuteurs natifs développent une façon de segmenter la chaîne de parole en fonction de la structure syllabique typique de leur langue (Mehler *et al.*, 1981 ; Cutler *et al.*, 1983, 1986, 1992).

La syllabe est indissociable de la prosodie. En tant qu'élément suprasegmental (Selkirk, 1984), elle joue un rôle essentiel dans l'analyse de divers phénomènes prosodiques, notamment le métrique et le rythme, mais aussi l'accent et la proéminence. Tous ces aspects peuvent interagir de manière significative avec la syllabe.

La raison pour laquelle la place de la syllabe est importante dans l'étude prosodique, c'est que la syllabe est une unité linguistique de taille intermédiaire (*ibid.*) entre le niveau segmental et le niveau suprasegmental (plus précisément, selon la hiérarchie prosodique générale, la couche syllabique est un niveau intermédiaire entre la more et le pied (Labrune, 2005)). Le lien étroit que la syllabe tisse entre le niveau segmental et le niveau suprasegmental se manifeste par : d'une part, au niveau segmental, des consonnes et/ou des voyelles se combinent en suivant certaines contraintes phonologiques d'une langue donnée pour former des syllabes et puis des mots ; d'autre part, en se combinant, des syllabes constituent d'autres syntagmes prosodiques du niveau supérieur : le cadre métrique du lexique et puis le mot

³⁹ Un point de vue qui a été critiqué par plusieurs chercheurs (voir par exemple Wheeldon & Lahiri, 1997). Ainsi pour Labrune (2005) : « Depuis longtemps, la syllabe est implicitement tenue comme l'élément clé de l'architecture prosodique sans que rien, finalement, ne justifie cette primauté (ni même que celle-ci ne soit reflétée sur le plan formel puisque la pratique courante est aujourd'hui de faire figurer la more comme premier degré de la hiérarchie). Il est pour le moins curieux que dans un modèle qui se veut hiérarchique, la more, d'une part, et le pied, d'autre part, soient tous deux définis par rapport à un niveau qui leur est intermédiaire, à savoir la syllabe... Mais que faire alors des cas comme le japonais, le gokana ou le bella coola... où les syllabes ne jouent aucun rôle ? Dans ces conditions, on peut penser que seul le mot prosodique – qui est aussi, rappelons-le, l'unique élément de la hiérarchie pourvu d'une réalité à la fois phonologique, morphologique, syntaxique et lexicale – est universellement présent. C'est donc à partir de ce domaine prosodique majeur que devraient être définis et construits, suivants des paramètres variables selon les langues, les constituants de rangs inférieurs, quand ils se révèlent pertinents. » (*ibid.* : 114).

prosodique, le syntagme phonologique, etc. (Selkirk, 1984⁴⁰). Non seulement la structure syllabique d'une langue donnée permet de catégoriser ce dernier selon l'isochronisme (cf. *infra*), mais le niveau supérieur formé (le groupe rythmique, notamment) est également une manifestation directe du patron rythmique ou phrasé prosodique. En un mot, la syllabe fonctionne comme un intermédiaire, qui permet d'organiser les niveaux segmental et suprasegmental et puis de les lier selon la hiérarchie prosodique (cf. *prosodic licensing* chez Itô, 1988 ; Fant *et al.*, 1991⁴¹). Ainsi, la syllabe possède un lien étroit avec les phénomènes rythmiques (Fry, 1964 ; Dauer, 1983⁴², Ramus *et al.*, 1999) : par exemple, dans Patel (2008), la diversité des structures syllabiques est considérée comme le premier facteur qui contribue au niveau typologique aux différents patrons rythmiques sur le plan phonologique ; Fox (2000) a mis en lumière la relation étroite qui existe entre la structure syllabique et la durée vocalique. Nous verrons plus loin dans cette section, que les différentes structures syllabiques conduisent à des patrons rythmiques variés dans la production, ce qui se manifeste par certaines valeurs des paramètres quantitatifs rythmiques et compliquerait la catégorisation rythmique des langues basée sur l'isochronie (cf. *infra*).

C'est pour tout cela qu'avant d'aborder le patron rythmique, il est nécessaire d'évoquer la syllabe dans un premier temps, afin d'explicitier et de différencier les structures syllabiques du chinois mandarin et du français. Cela nous permettra, dans la partie expérimentale, de mieux comprendre et expliquer le patron rythmique interlangagier produit par des apprenants sinophones.

1.3.1.2 Structures syllabiques en chinois mandarin

La syllabe du chinois mandarin est relativement simple et facile à identifier : en général, un caractère, ou un morphème correspond à une syllabe (Duanmu, 2008).

⁴⁰ "The syllable is the paradigm case of a unit of prosodic constituent structure, and so will provide a point of reference in discussing the status of other such hierarchical units in the theory." (Selkirk, 1984: 22).

⁴¹ "The duration of a syllable depends, on the one hand, on its complexity and specific phonological composition and, on the other hand, on stress and overall tempo and position within a phrase or a stress group." (Fant *et al.*, 1991: 356).

⁴² "... in prosodic analysis (Firth, 1948) syllable structure and vowel reduction are considered to be prosodies and therefore are naturally related to stress and rhythm... It is precisely the language structure with all its language specific segmental variation that is responsible for perceived differences in language rhythm." (Dauer, 1983: 59).

Étant donné que le chinois mandarin est une langue isolante et qu'il présente peu de flexion, la syllabe conserve ses propriétés phonétiques, notamment le contour de F0 pour maintenir le sens associé au ton lexical (Duanmu, 2014). La structure syllabique maximale est « CGVX » tandis que la structure syllabique minimale peut se composer d'une seule voyelle « V », ou même d'une seule « C »⁴³ (Duanmu, 1999, 2007, 2008, 2014). Duanmu (1999) a identifié 12 types de structures syllabiques possibles dans le chinois mandarin. De son côté, Chao (2011) a également recensé toutes les combinaisons susceptibles de former une syllabe en chinois mandarin (Tableau 1.5) :

Tableau 1.5 : Combinaisons syllabiques possibles en chinois mandarin (Chao, 2011 : 59).

Medial	Initial	Final												
Zero		y	a	e	ai	ei	au	ou	an	en	ang	eng	ong	el
	b p m f		巴	.么 ^m	掰	碑	包	谋 ^{m'}	班	奔	帮	崩		□!
	d t n l		搭	.的	呆	得 ^v	刀	兜	单	嫩 ^{n'}	当	登	东	
	tz ts s		兹	咂	则 [']	灾	贼 [']	糟	邹	簪	簪	脏	增	宗
	j ch sh r		知	扎	遮	斋	这 [\]	招	周	沾	真	张	争	中
g k h O		嘎	哥		该	给 ^v	高	钩	干	根	刚	庚	工	
-i		i	ia	ie	iai		iau	iou	ian	in	iang	ing	iong	
	b p m -		逼	□!	憋		标	谬 ^{m'}	边	宾		冰		
	d t n l		低		爹		刁	丢	颠	林 ^{l'}	娘 ^{n'}	丁		
j ch sh O		鸡	家	接	崖 ^{o'}		交	撤	尖	今	将	经	局	
-u		u	ua	uo	uai	uei			uan	uen	uang	ueng		
	b p m f		不		波									
	d t n l		都		多		堆		端	敦				
	tz ts s		租		撮		堆		钻	尊				
	j ch sh r		朱	抓	桌	摔 ^{sh}	迫		专	淳	庄			
g k h		姑	瓜	锅	乖	归		官	滚 ^v	光	翁 ^{o'}			
-iu		iu			iue				iuan	iun				
	- - n l		女 ^{n'}		略 ^{l'}				李 ^{l'}	淋 ^{l'}				
j ch sh		居		厥				捐	均					

⁴³ "The maximal Chinese syllable, regardless of the dialect, is often thought to contain four positions, or CGVX, where C is a consonant, G is a glide, and VX is a diphthong, or a short vowel plus a consonant, or a long vowel... The VX portion can be called the rime... In Chinese, most syllables have an onset. For example, in Standard Chinese, there are 404 syllables (ignoring tones, interjections, and dialectal loans), and only 14 of them lack an onset, or 3%. Most onset-less syllables are content words whose rimes are VX. The only onset-less syllable whose rime is V is an interjection [a]. In summary, Chinese has CV syllables. Occasionally, V syllables are also found, as well as some C syllables." (Duanmu, 2014: 423-425).

Duanmu (2008) a calculé toutes les combinaisons possibles en fonction des nombres d'éléments segmentaux en chinois mandarin. Cela se traduit par le schéma suivant (Tableau 1.6) :

Tableau 1.6 : Nombre de possibilités de combinaisons syllabiques en chinois mandarin (Duanmu, 2008 : 97).

Possible combinations of syllables in Standard Chinese		
	Choices	Notes
C	19	One of 18 Cs, or no C
G	4	One of [j, u, ɥ], or no G
V	5	One of five vowels
X	5	One of [i, u, n, ɲ] or no X
Total	1,900	without tonal contrasts
	7,600	with tonal contrasts (four tones per syllable)

1.3.1.3 Structures syllabiques en français

En français, la structure syllabique principale est constituée de CV (Léon, 2012), et les syllabes ouvertes représentent une part importante de toutes les combinaisons syllabiques possibles⁴⁴. Selon des statistiques réalisées par différents chercheurs en utilisant différents corpus (Wioland, 1991, cité dans Lauret, 2007 ; Goldman *et al.*, 1996 ; Delattre (1966c) cité dans Léon, 2012 ; Pagel *et al.*, 2012), les structures syllabiques en français sont composées de :

Tableau 1.7 : Proportions des structures syllabiques en français selon différentes études.

Étude Combinaisons	Pagel <i>et al.</i> (2012)	Delattre (1966c) cité dans Léon (2012)	Lauret (2007)	Goldman <i>et al.</i> (1996)
CV	55,6%	59,9%	55%	46,4%
CCV	13,9%	14,2%	14%	11,1%
V	9,8%	(non renseignée)	10%	20,1%
Syllabe ouverte (Totale)	79,3%	74,1%	79%	77,6%
CVC	13,5%	17,1%	13%	13,4%
CCVC	2,6%	(non renseignée)	2,5%	3%
CVCC	1,5%	(non renseignée)	1,5%	(non renseignée)
VC	1,3%	1,9%	1,3%	4,5%
Syllabe fermée (Totale)	18,9%	19%	18,3%	20,9%

⁴⁴ « Paul Passy avait déjà montré que le français privilégie dans la parole les syllabes ouvertes... Sur 10 syllabes, 8 sont terminées par une voyelle. Cette proportion de 80% est voisine de celle que Delattre a trouvée dans son corpus (59,9 + 14,2 = 74,1 %). » (Léon, 2012 : 135). D'ailleurs, « en y ajoutant, la structure CVC, plus des 3/4 des syllabes sont représentées. » (Sauzedde, 2014 : 92).

À partir du Tableau 1.7, qui récapitule les proportions des structures syllabiques du français, nous pouvons voir très clairement que la structure syllabique CV prédomine dans toutes les combinaisons possibles, et que les syllabes ouvertes (V, CV, CCV) sont les plus attestées en français.

1.3.2 Systèmes métrique, rythmique et macrorythmique

Dans cette section, nous aborderons les différents aspects liés au rythme. Nous commencerons par explorer l'omniprésence du rythme dans l'univers ainsi que ses manifestations dans le langage humain. Ensuite, les définitions et les fonctions des concepts « métrique », « rythme », « macrorythme », et « phrasé » seront présentées et différenciées. Nous ne manquerons pas d'évoquer l'isochronie, une notion incontournable dans l'étude rythmique, et puis la catégorisation rythmique traditionnelle basée sur l'isochronie. Par la suite, nous aborderons les paramètres quantitatifs rythmiques, qui demeurent la méthodologie la plus adoptée dans les études empiriques sur le rythme. Après avoir souligné les limites de l'utilisation de ces paramètres quantitatifs dans l'étude rythmique, nous présenterons un modèle d'analyse, le modèle CCI (*Control/Compensation Index*), proposé par Bertinetto et Bertini (2008), qui servira de cadre pour notre analyse des productions rythmiques des apprenants dans la partie expérimentale. Enfin, cette section conclura par une description des systèmes rythmiques du chinois mandarin et du français.

1.3.2.1 Phénomènes rythmiques : universalité et horloge interne

Le rythme, au sens général, est un phénomène omniprésent dans l'univers et chez l'être humain (Fraisse, 1974 ; Fox, 2000 ; Patel, 2008 ; Di Cristo, 2016). Avant d'examiner en détail le rythme dans la parole, cette introduction préliminaire mettra en évidence son omniprésence, d'une part dans le monde, et d'autre part, ses manifestations dans le langage humain.

En général, le phénomène de rythme, qui évoque la récurrence d'un événement, englobe les notions suivantes : d'un côté, l'égalité temporelle, qui implique une temporalité, une régularité, une répétitivité, une périodicité, ou une cadence

équilibrée, etc. ; de l'autre, le groupement, l'organisation et la hiérarchisation (Di Cristo, 2016 ; une présentation précise chez Lacheret-Dujour & Beaugendre, 1999⁴⁵). Le rythme est d'abord un phénomène universel : cela se manifeste par sa relation avec le fonctionnement corporel et le monde extérieur : « *le phénomène rythmique se rapporte à la fois à des fonctions vitales (les rythmes cardiaques et respiratoires), à la façon d'appréhender l'écoulement du temps, ainsi qu'à tout un ensemble d'activités hautement diversifiées, comme la danse, la musique, la mimogestualité et diverses formes d'expressions écrites, graphiques et orales* » (Di Cristo, 2016 : 53). En ce qui concerne spécifiquement les langues du monde, il est universellement observé que la structure temporelle des langues se situe généralement dans une fourchette de 2-8 Hz, ce qui est lié au circuit cérébral responsable du système moteur de la production langagière chez l'être humain (Poeppel & Assaneo, 2020).

Comme ce qui est indiqué plus haut, le rythme est considéré par certains comme une horloge interne ou biologique ancrée chez l'être humain, « *qui nous permettrait de reporter le rythme naturel biologique sur d'autres activités, telles que la musique par exemple* » (Tortel, 2009 : 68), et plus spécifiquement le langage. Dans les systèmes corporel et cognitif de l'être humain, cette horloge interne se manifeste par les aspects articulatoire, physiologique et cognitif (Di Cristo, 2016) :

- d'un point de vue articulatoire : le rythme de la parole peut se manifester par l'articulation lorsque l'on parle (Allen, 1975 ; Fletcher, 2010 ; Turk & Shattuck-Hufnagel, 2014 ; Poeppel & Assaneo, 2020). Les raisons de la rythmicité articulatoire dans la production pourraient être, d'un côté, les articulateurs de la parole ne sont pas indépendants et ils fonctionnent de manière synergique, et de l'autre, le

⁴⁵ « *La notion de rythme est, depuis l'Antiquité, irréductiblement liée à celles d'équilibre, de symétrie, de cadence, de disposition, de proportion, enfin de figure : « comme le mètre, le rythme suppose une structure répétitive et la proximité dans le temps de ces structures proportionnelles » (J. Molino et J. Tamine 1982). Le rythme est mouvement périodique et cadencé, il est configuration des mouvements ordonnés dans la durée (E. Benveniste 1966) et se définit par la régularité du retour, c'est-à-dire par une forte prévisibilité. Le premier concept qui puisse être rattaché à celui de rythme est donc celui de régularité : les événements rythmiques n'existent que parce qu'ils sont organisés de façon régulière dans le temps ; l'activité rythmique est une activité temporelle, structurante : la structure se trouve toujours liée à une périodicité et la périodicité est toujours organisation des structures (P. Fraisse 1974). Ainsi se dégagent deux caractéristiques fondamentales du rythme : la périodicité et la forme ; il s'agit de la mise en place de structures temporelles minimales et récurrentes, qui semble correspondre à une tendance naturelle chez l'homme à produire des formes rythmiques simples, composées d'un nombre limité d'éléments (ibid.).* » (Lacheret-Dujour & Beaugendre, 1999 : 34).

mouvement articulatoire est contraint par un certain degré de liberté, que ce soit au niveau spatial ou au niveau temporel⁴⁶.

- d'un point de vue physiologique : « *la production de la parole, par laquelle s'exprime la langue parlée, est d'abord contrainte par l'activité respiratoire qui, du fait de son caractère cyclique, est manifestement rythmique* » (Di Cristo, 2016 : 53). Ce point de vue est illustré par la Figure 1.13 (Martin, 2018) :

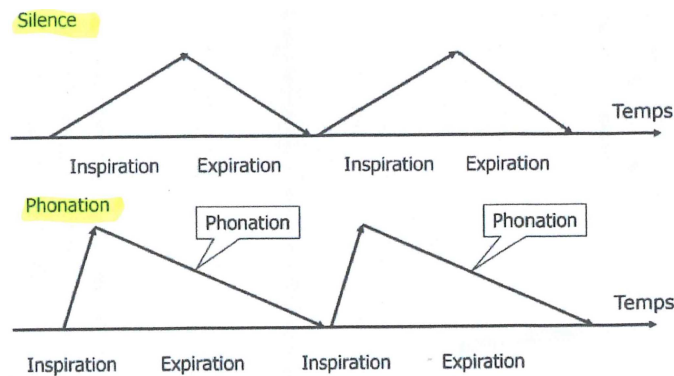


Figure 1.13 : Cycles respiratoires (inspiration-expiration) sans et avec production de la parole (Martin, 2018 : 45).

La périodicité physiologique dans le langage se manifeste aussi par comment la syllabe est physiologiquement produite : Abercrombie (1967) a proposé deux modalités physiologiques régulières concernant la production syllabique : la pulsation thoracique (*chest pulses*), qui suggère la récurrence de la syllabe (donc, isochronie syllabique, cf. *infra.*) et la pulsation accentuelle (*stress pulses*), qui suggère la périodicité de la syllabe accentuée (donc, isochronie accentuelle, cf. *infra.*).

- d'un point de vue cognitif : des chercheurs ont aussi trouvé des preuves qui démontrent la rythmicité langagière. Lors de la perception de la parole, le cortex auditif a la capacité de suivre la structure rythmique de la chaîne parlée. Cette capacité permet aux oscillations neuronales de synchroniser avec la chaîne rythmique et faciliter ainsi la segmentation de la parole pour décoder et traiter les informations sonores (Martin, 2018). Ce processus de synchronisation s'appelle l'entraînement auditif et le résultat de synchronisation correspond à la structure

⁴⁶ Au niveau spatial: “*speech does not rely on the kinematics of individual articulators: it emerges from the cooperative dynamics between articulators in order to achieve a common goal. We suggest that during natural speech production the rhythm of the interplay between the articulators is preserved instead of the rhythm of the isolated elements*” ; au niveau temporel: “*the rhythmicity across the domains of description derives from the same cause: the motor gestures of speech are sequentially executed at a relatively regular time interval.*” (Poeppel & Assaneo, 2020).

temporelle universelle que nous avons évoquée plus haut (2-8 Hz) (Poeppel & Assaneo, 2020). En effet, par exemple, à travers des expériences phonétiques et neurophysiologiques, Martin (2012, 2014, 2015, 2018, 2020) a trouvé des preuves similaires pour le français : en français, « *la durée des syllabes varie de 100 ms à 250 ms... des durées de groupes accentuels variant de 1 à 8 syllabes, l'intervalle entre syllabes accentuées successives variant de 250 ms environ (groupe accentuel d'une seule syllabe) à quelque 1250 ms (groupe de 8 syllabes)* » (ibid., 2018 : 53, 60) ; et au niveau cognitif, cette fourchette temporelle régulière coïncide avec les ondes cérébrales thêta et delta : « *on constate que les variations de durée entre deux syllabes inaccentuées successives s'inscrivent assez bien (très bien même...) dans celles des ondes cérébrales thêta, de 100 ms à 250 ms. De même, l'intervalle minimal de 250 ms entre deux syllabes accentuées successives correspond à la période minimale des ondes delta. Quant à la période maximale de l'onde delta, elle correspond à l'intervalle maximal 1250 ms des groupes accentuels* » (ibid. : 245). D'autres travaux similaires ont révélé le lien entre le rythme de la parole et l'oscillation neuronale (Turk & Shattuck-Hufnagel, 2013), indiquant que la périodicité de certaines oscillations neuronales se synchronise avec la périodicité des événements sonores dans la parole.

1.3.2.2 Différentes terminologies : « métrique », « rythme », « isochronie », « macrorythme » et « phrasé »

Dans cette section, nous explorerons une série de terminologies et faire une distinction entre « métrique », « rythme », « macrorythme », « phrasé » et « isochronie ». Nous introduirons aussi la typologie rythmique traditionnelle basée sur l'isochronie et sur les paramètres quantitatifs. L'accent sera ensuite mis sur les limites du critère « isochronie », ce qui nous permettra ensuite de mieux illustrer la nécessité de notions de « macrorythme » et de « phrasé » dans l'étude du rythme langagier. Dans le cadre de la différenciation terminologique que nous ferons, une synthèse sera présentée à la fin, dans la section 1.3.2.2.3, afin de fournir une ébauche plus complète et détaillée de la définition du rythme.

1.3.2.2.1 « Métrique », « rythme » et « isochronie »

Nous commencerons par définir le rythme dans son sens propre en lien avec le langage. En général, dans l'étude de la prosodie, on ne peut parler du « rythme » sans évoquer la notion de « métrique ». En effet, le rythme et la métrique se réfèrent à des concepts similaires à deux niveaux différents : « *la métrique tend à être interprétée comme un dispositif structurel sous-jacent (ou abstrait) et le rythme, comme les manifestations concrètes de ce dispositif, au niveau des structures de surface qui encodent la prononciation effective* » (Di Cristo, 1999 ; 2013 ; 2016 : 55)⁴⁷. Dans la majorité des langues, la structure métrique est composée du pied, « *unité phonologique... formé d'une syllabe accentuée et d'un nombre limité de syllabes inaccentuées* » (*ibid.*). La proposition de la notion de « métrique » permet de « *dépasser la querelle historique concernant l'opposition entre ordre temporel et ordre accentuel et d'intégrer ces derniers comme deux composantes fondamentales et indissociables du rythme* » (*ibid.*, 1999 : 146). Dans le champ de la métrique, les linguistes américains Mark Liberman, Alan Prince et Bruce Hayes, dans les années 1980s, ont proposé la phonologie métrique, un courant théorique de la mouvance de la phonologie non linéaire, qui consiste à décrire phonologiquement les règles et les principes du matériau rythmique, et les relations entre prééminence et regroupement. Les structures métriques se classifient généralement en deux catégories : une structure iambique, où la syllabe accentuée est précédée de syllabes inaccentuées (autrement dit, tête à droite) ; et une structure trochaïque, où la syllabe accentuée est suivie de syllabes inaccentuées (autrement dit, tête à gauche) (Hayes, 1985, *the Iambic-Trochaic Law* ; cf. Langus *et al.*, 2017 pour une revue).

Comme nous l'avons déjà évoqué plus haut, le rythme dans son sens général renvoie à la régularité ou la périodicité d'un événement. Dans la langue, cette régularité se manifeste par l'isochronie, un phénomène périodique linguistique « *caractérisé par le retour à des intervalles réguliers de temps marquées* » (Di Cristo, 2016 : 64). En se basant sur la périodicité langagière désignée comme l'isochronie, Pike (1945) et

⁴⁷ Pour une explication plus claire : « *... la métrique fait-elle référence... à un mode d'organisation sous-jacent, plus conventionnel et plus contraint que celui du rythme réalisé. En effet, alors que la métrique est supposée soumise à des principes stricts de régularité formelle, le rythme est présumé davantage motivé par des forces fonctionnelles que par des contraintes d'agencement rigoureux. Cette différence peut s'expliquer en grande partie par le fait que la construction du rythme est sensible à l'influence des événements qui sont contemporains de la production de la parole, dans les diverses situations de communication. Les influences qu'exercent ces facteurs externes tendent ainsi à outrepasser les contraintes que pourrait imposer une ordonnance métrique stricte.* » (Di Cristo, 2013 : 121).

Abercrombie (1965, 1967) ont établi une typologie rythmique qui regroupe les langues du monde d'abord en deux catégories : isochronie syllabique (*syllable-timed*), qui sont les langues où les syllabes sont produites à intervalles réguliers ; et isochronie accentuelle (*stress-timed*), les langues où les durées des intervalles interaccentuels sont quasi égales. La troisième catégorie, isochronie moraïque (*mora-timed*), a été proposée par Bloch (1950), Han (1962) et Ladefoged (1975) (cf. Otake (2015) pour une revue) et cette catégorie regroupe les langues, comme le japonais, dont les durées moraïques sont quasi égales. Selon la typologie traditionnelle, les langues du monde peuvent ainsi être catégorisées comme suit (Tableau 1.8) :

Tableau 1.8 : Catégorisation rythmique des langues basée sur l'isochronie (Ramus, 1999 : 41).

Langue	Référence
Langues accentuelles	
Anglais	Pike (1945)
Arabe	Abercrombie (1967)
Néerlandais et les langues germaniques en général	Ladefoged (1975)
Russe	Abercrombie (1967)
Polonais	Rubach et Booij (1985)
Thaï	Luangthongkum (1977), cité par Dauer (1983)
Langues syllabiques	
Espagnol	Pike (1945)
Français	Abercrombie (1967)
Italien	Bertinetto (1981)
Catalan	Mehler et coll. (1993)
Grec	Arvaniti (1994)
Telegu	Abercrombie (1967)
Yoruba	Abercrombie (1967)
Hindi	O'Connor (1973), cité par Dauer (1983)
Langues moraïques	
Japonais	Ladefoged (1975)
Tamoul	Steever (1987)
Alyawarra, Aranda et Paiute	Haraguchi et Tanaka (1992) cités par Nazzi (1997)

Dans une étude plus récente, Grabe et Low (2002) ont établi une classification légèrement différente (Tableau 1.9).

Tableau 1.9 : Catégorisation rythmique des langues par Grabe et Low (2002 : 518).

Language	Classification
British English	Stress-timed (Classe, 1939, Pike, 1946, Abercrombie, 1967)
German	Stress-timed (Kohler, 1982)
Dutch	Stress-timed (Ladefoged, 1975, Smith, 1976)
Thai	Stress-timed (Luangthongkum, 1977)
Tamil	Syllable-timed (Corder, 1973, Asher, 1985)
Spanish	Syllable-timed (Pike, 1946, Hockett, 1958)
French	Syllable-timed (Abercrombie, 1967, Catford, 1977)
Singapore English	Syllable-timed (Tongue, 1979, Platt and Weber, 1980)
Japanese	Mora-timed (Bloch, 1942, Han, 1962)
Polish	Mixed (Dauer, 1987, Nespors, 1990)
Catalan	Mixed (Dauer, 1983, Nespors, 1990)
Estonian	Unclassified
Greek	Unclassified
Luxembourg	Unclassified
Malay	Unclassified
Mandarin	Unclassified
Rumanian	Unclassified
Welsh	Unclassified

Dans le cadre de la catégorisation traditionnelle basée sur l'isochronie, on considérait qu'il y avait une distinction nette entre les trois catégories. Toutefois, cette notion a été fortement critiquée par des chercheurs qui soutiennent que la typologie traditionnelle s'appuie davantage sur des impressions auditives que sur des validations empiriques, et que les expérimentations réalisées n'ont pas permis d'établir une classification claire (Roach, 1982 ; Dauer, 1983 ; Bertinetto, 1989 ; Cutler, 1991 ; Ramus, 1999 ; Ramus *et al.*, 1999 ; Fox, 2000 ; Grabe & Low, 2002 ; Patel, 2008 ; Arvaniti, 2009 ; Turk & Shattuck-Hufnagel, 2013 ; Di Cristo, 2016 ; White & Malisz, 2020, etc.). Roach (1982), par exemple, a indiqué qu'il n'y a pas de langue à isochronie strictement syllabique ou accentuelle, mais que les langues ont un patron temporel syllabique ou accentuel plus ou moins prédominant. Dauer (1983) a été le premier à proposer que toutes les langues, situées sur un continuum, sont plus ou moins isochroniquement syllabiques ou accentuelles⁴⁸(Figure 1.14), plutôt que de se regrouper dans des catégories strictes et distinctes. Ce continuum rythmique a été empiriquement démontré dans l'étude de Grabe et Low (2002).

⁴⁸ "We can say that a language is more or less stress-based, depending on how large a role stress plays in that language, especially with respect to the three areas described above [la structure syllabique, la réduction vocalique et l'accentuation]. Languages such as English in which stress has important consequences in all these areas, and possibly others as well, can be called truly stress-based. Languages can be compared to each other along the dimension of having a more or less stress-based rhythm." (Dauer, 1983: 59).

Cependant, il a été critiqué par Arvaniti (2009), qui a souligné que l'étude de Grabe et Low (2002) n'a pas été testé le continuum avec un nombre suffisant de langues, notamment avec celles dites moins « prototypiques ».

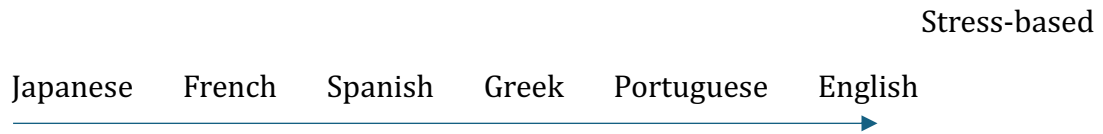


Figure 1.14 : Continuum rythmique proposé par Dauer (1983 : 60).

1.3.2.2.2 « Macrorhythme » et « phrasé »

L'approche d'isochronisme évoquée ci-dessus, qui est quelque peu datée et qui catégorise le patron rythmique à partir d'un niveau inférieur dans la hiérarchie prosodique, ne parvient pas à décrire de manière exhaustive la typologie des rythmes des langues du monde. Afin de fournir une description complète sur le rythme, il est primordial d'aborder deux autres concepts « macrorhythme » et « phrasé », qui sont souvent méconnus ou moins explorés dans la recherche sur le rythme du langage.

Di Cristo (2016) a indiqué que la notion de rythme englobe deux dispositifs : le simple et le complexe (des idées qui rejoignent celles de Dauer, 1983⁴⁹, Arvaniti, 2009⁵⁰) : « *les dispositifs simples s'activent à segmenter le signal d'entrée en groupements et à extraire des régularités temporelles sous la forme d'alternances de battements forts et faibles* » (Di Cristo, 2016 : 54). C'est l'aspect primaire qui constitue la définition traditionnelle du rythme : un phénomène de périodicité temporelle basée sur une unité prosodique de base (« structures temporelles minimales et récurrentes », i.e., syllabe, accent ou more, cf. *supra*)⁵¹. Il rajoute que « *les dispositifs complexes s'appliquent à hiérarchiser les groupements et à promouvoir une représentation abstraite des niveaux de proéminence relatifs aux battements* »

⁴⁹ "In all languages we would expect syllables to be grouped into larger units, even if the basis for grouping is something other than a stress beat. Repetition of particular sounds, syllables, grammatical markers, or pitch patterns, might also be used to group syllables into larger units. If rhythmic grouping takes place in all languages, then the differences summed up by the terms "stressed-timed" and "syllable-timed" refer to what goes on within rhythmic groups, the characteristics of successive syllables and their interrelationships, which are ultimately a product of the entire linguistic system." (Dauer, 1983: 60).

⁵⁰ "... we need to reconsider our view of speech rhythm so that it is less focused on timing, which should in principle be examined as a distinct phenomenon. Instead, it appears advantageous to adopt a conception of rhythm that goes beyond timing and rhythmic types but rests instead on grouping and patterns of prominence." (Arvaniti, 2009: 61).

⁵¹ « Les notions de temporalité et de régularité sont présentées comme cruciales, dans la mesure où le rythme est défini avant tout comme une organisation imposée à l'écoulement du temps. » (Di Cristo, 2016 : 54).

(*ibid.*). Le processus régulier de regroupement et ce qu'il constitue dans une langue donnée peuvent aussi être considérés comme un événement périodique ou rythmique, marqué par une récurrence régulière de proéminence.

C'est pourquoi nous souhaitons introduire les notions de « macrorythme » et de « phrasé », qui désignent le résultat du regroupement et qui donnent une description plus complète sur le rythme langagier, en tenant compte de l'alternance de proéminence⁵².

Le macrorythme est une notion proposée par Jun (2005, 2012, 2014). Elle désigne « le patron rythmique supérieur à l'unité lexicale »⁵³. Étant donné que le domaine du macrorythme est plus grand que celui du microrythme, défini par Jun (2005) comme l'unité syllabique, accentuelle ou moraïque, l'unité macrorhythmique contient plus de variations. Cela dit, certaines formes périodiques peuvent être observées. La notion de « macrorythme » est de plus en plus clarifiée dans ses travaux ultérieurs : elle permet de mieux rendre compte de la régularité langagière en tenant compte la variation tonale (Jun, 2012, 2014)⁵⁴ aux niveaux syntagmatique et énonciatif. Le concept de « macrorythme » donne la possibilité, d'une part, d'établir une nouvelle typologie rythmique ou prosodique (Tableau 1.10) en se basant sur la proéminence, et d'autre part, de mieux décrire les problèmes et les difficultés dans la production du patron rythmique chez des apprenants d'une L2, plutôt que de se limiter à leur étude via la méthodologie dominante portant sur les paramètres quantitatifs rythmiques (cf. *infra*).

⁵² "The temporal structure of segmental, "micro-rhythmic" units and prosodic, "macro-rhythmic" units both play an important role in rhythm production and perception. If we assume that typological templates of some kind do exist, measuring durations of segmental intervals alone will never suffice to capture systematic tendencies in the acoustic substance of those templates." (Rathcke & Smith, 2015: 2844).

⁵³ "The rhythmic patterns derived by the postlexical prosodic unit, which I call a 'macro' rhythmic unit..." (Jun, 2005: 442).

⁵⁴ "Macro-rhythm is a rhythm created from a pitch contour, i.e., tonal rhythm... Macro-rhythm is a tonal rhythm, a rhythm perceived by changes in f_0 ... Therefore, a stronger degree of macro-rhythm would be created by a sequence of alternating low and high tones (L-H-L-H-L-H)... Classifying languages based on the degree of macro-rhythm can allow us to explain how languages have similar or different phrase-medial intonation patterns regardless of their prominence type and word prosody... The parameter of macro-rhythm captures the regularity of phrase-medial intonation patterns across languages." (Jun, 2012: 535-536; 2014: 524, 534, 539).

Tableau 1.10 : Typologie prosodique basée sur la proéminence et sur l'unité rythmique (Jun, 2005 : 444).

Language	Prosody					Rhythmic/prosodic unit					
	Prominence					Lexical			Postlexical		
	tone	stress	LPA	head	edge	mora	syll	foot	AP	ip	IP
English		x		x				x		x	x
German		x		x				x		x ^a	x
Dutch		x		x				x			x+1
Greek		x		x				x		x	x
Italian		x		x			x	(x) ^b		x	x
Spanish		x		x			x			(x) ^c	x
Portuguese		x		x			x				x
Arabic		x		x				x		x	x
Farsi		x		x	x		x		x	x	x
BGW		x		x	x			x	(x) ^d		x+1
Swedish		x	x	x				x			x
Sb-Croat.		x	x	x	x			x			x
Chickasaw		x	x	x	x			x	x ^e		x
Japanese			x	x	x	x			x	(x) ^f	x
Basque			x	x	x		x		x	x	x
French				x	x		x		x	(x) ^g	x
Bengali				x ^h	x		x		x		x
Korean					x		x		x		x
Mandarin	x	x		(x) ⁱ			x			x	x
Cantonese	x						x				x
Kinande	x				x		x		x		x

Pour autant, cette première version de la typologie de Jun (2005) n'a pas pris en considération la configuration tonale au niveau prosodique supérieur⁵⁵. Selon Jun (2012), cette catégorisation (Tableau 1.10) n'a pas non plus permis de mettre en évidence les différences et les similitudes au niveau du patron mélodique entre les langues. C'est pour cette raison qu'elle a proposé un nouveau niveau prosodique : « macrorhythme » (2012, 2014), qui permet de mettre en évidence la prosodie

⁵⁵ "Specifically, it could not distinguish languages that have different global tonal patterns of utterances but belong to the same type of prominence marking, nor capture the similarity of global tonal pattern of language belonging to a different type of prominence marking... This suggests that a global tonal pattern of an utterance is another prosodic dimension that is orthogonal to the types of prominence marking, motivating the need to add the complexity or regularity of phrase-medial tonal patterns as a parameter of prosodic typology." (Jun, 2012: 535).

lexicale, le degré de rythmicité et en même temps, la configuration mélodique⁵⁶. En se basant sur les positions des proéminences, la nouvelle typologie avec trois types de patrons macrorhythmiques a donc été proposée par Jun (2012) :

- 1) Langues à proéminence de tête (*head-prominence languages*)⁵⁷ : des langues où la proéminence se trouve à la tête du syntagme, comme le suédois, l'anglais et les langues à tons : e.g., le chinois mandarin ;
- 2) Langues à proéminence de tête et queue (*head/edge-prominence languages*) : des langues où la proéminence est marquée tant à la tête du syntagme qu'à la queue du syntagme : e.g., le français ;
- 3) Langues à proéminence de queue (*edge-prominence languages*) : des langues où seule la proéminence à la fin du syntagme marque une frontière (e.g., le coréen).

Les trois catégories ci-dessus sont ensuite subdivisées en fonction du degré du macrorhythmisme (fort, moyen, faible, cf. Jun (2012, 2014) pour les critères de subdivision) (Tableau 1.11) :

⁵⁶ "In sum, macro-rhythm captures the tonal aspects of phrasal prosody of languages, and combining prominence types, word prosody, and the degree of macro-rhythm provides new directions in prosodic typology." (Jun, 2012: 538).

⁵⁷ "Head-prominence languages where the prominence at phrase level is marked by the head of a phrase, which is derived from the head of a word... languages that have a pitch accent, regardless of whether the pitch accent location or type is determined lexically or postlexically are head-prominence languages. Tone languages are also head-prominence languages because the prominence is derived from the tonal specification of a specific syllable in the lexicon. Though the degree of macro-rhythm is gradual, we can divide the head-prominence languages roughly in three macro-rhythm groups: Strong, Medium, and Weak." (Jun, 2012: 537).

Tableau 1.11 : Typologie prosodique développée à partir de Jun (2005) (Jun, 2014 : 535).

Prom. type	Word prosody	Macro-rhythm		
		Strong	Medium	Weak
Head	Stress	Brazilian Portuguese, Castilian Spanish, Catalan, Egyptian Arabic, Greek, Italian, Samoan	Dutch, English, German, Jamaican Creole English, Lebanese Arabic	European Portuguese, Wolof, Kuot
	Tone/lexical pitch accent	(<i>Kihehe</i> , <i>Safwa^a</i> , <i>Somali^b</i>)	(<i>Isthmus Zapotec^c</i> , <i>Kimatuumbi^a</i> , <i>Skou^d</i>)	Cantonese
	Both	Papiamentu, Swedish		Mandarin
	None			
Head/Edge	Stress	Bengali, Georgian, Kiche, Tamil, Tongan	Bininj Gun-wok, Dalabon	
	Tone/lexical pitch accent	Japanese, Leketio Basque		
	Both	Serbo-Croatian	Chickasaw ⁷	
	None	French, Standard Basque		
Edge	None	Accentless dialects of Japanese, Halh Mongolian, Oirat Mongolian, Seoul Korean, West Greenlandic		

Nous avons déjà introduit la notion de « phrasé » dans la section 1.2.2.1. En la situant dans le domaine du rythme du langage, nous pensons qu'elle implique exactement la même idée que la notion de macrorythme : englober un domaine prosodique supérieur à l'unité de base (que ce soit la syllabe, l'accent ou la more) et tenir compte du regroupement par proéminence dans l'analyse rythmique. Comme l'a souligné Martin (2018 : 20), le phrasé est défini comme « *le découpage des phrases en groupes « accentuels »* » et « *un énoncé est donc constitué d'une suite de groupes accentuels* ». Encore une fois, la proéminence est intégrée comme un élément récurrent et périodique dans le rythme de la parole.

1.3.2.2.3 Sur la définition du rythme : une synthèse

Après avoir revu les différentes définitions du rythme proposées par des chercheurs ainsi que les particularités associées à ce concept, nous souhaitons proposer ici une synthèse. En nous basant sur les observations théoriques dans la littérature, nous essayons d'établir une ébauche de définition plus exhaustive du rythme, qui inclut également le phénomène d'accent.

Le rythme est d'abord un phénomène langagier hiérarchique. Dans le cadre de cette étude, nous partageons pleinement l'idée exprimée par divers chercheurs : la

structure rythmique de la parole est composée de plusieurs niveaux hiérarchiques (Wagner, 2010). Il ne s'agit pas simplement d'une question temporelle (isochronisme) (Fox, 2000) : le traitement du rythme de la parole est composé en effet de deux grands dispositifs : simple (structures temporelles et récurrentes au niveau d'unité prosodique minimale) et complexe (structures temporelles et récurrentes au niveau de proéminence) (Di Cristo, 2016). L'un des aspects souvent négligés dans la façon traditionnelle de définir le rythme, est la notion d'organisation ou de regroupement par le biais de la proéminence⁵⁸. Cependant, comme l'a indiqué Kohler (2009), de plus en plus de chercheurs prennent conscience de l'importance de l'accent dans l'étude du rythme. Les propriétés acoustico-phonétiques des proéminences, qui s'ancrent dans les positions prédéfinies du cadre métrique, permettent de regrouper les unités rythmiques de base, de les hiérarchiser selon les frontières ou les contraintes syntactiques et ensuite, de former la tête et/ou la queue du niveau rythmique supérieur (Selkirk, 1984⁵⁹). Par ailleurs, la forme sous laquelle les éléments ou les morceaux s'organisent constitue un événement régulier et rythmique, autrement dit, « *une certaine forme de répétitivité ou de périodicité (par laquelle un schéma se reproduit à l'identique de lui-même à des intervalles de temps que l'on présuppose sensiblement égaux)* » (Di Cristo, 2016 : 54). C'est pour cette raison que les notions de « macrorythme » et de « phrasé » ont été introduites, permettant ainsi de décrire cet événement régulier et d'enrichir la définition du rythme.

Le rythme langagier ne peut pas être défini uniquement en fonction de l'isochronisme temporel : 1) d'abord, le rythme langagier ne se limite pas à l'élément temporel ; 2) ensuite, il ne s'agit pas d'un fait purement temporel en raison des variations induites par plusieurs facteurs (linguistiques, paralinguistiques et extralinguistiques). Le rythme langagier implique en effet la récurrence

⁵⁸ Nous pouvons trouver l'idée similaire chez Hayes (1985 : 429) : "*there appear to be parallels between rhythmic structure and linguistic stress patterns. In particular, a defining characteristic of rhythm is the recurrence of events at regular intervals. Stress systems are likewise "designed" to place stresses at equal intervals.*"; chez Patel (2008) : "*Thus I will define rhythm as the systematic patterning of sound in terms of timing, accent, and grouping... that may have nothing to do with isochrony.*" (Patel, 2008: 96, 122).

⁵⁹ "... we assert that the rhythmic structure of the utterance does not reduce to the concatenation of the rhythmic structure of its component words, but rather that it reflects as well the arrangement of words into syntactic constituents, which is to say that there is a rhythmic organization proper to the phrasal level." (Selkirk, 1984: 142).

d'événements, mais cet événement périodique repose à la fois sur la durée ET la fréquence fondamentale (F0), sous l'influence de l'interaction entre divers facteurs (Dasher & Bolinger, 1982, Dauer, 1983, Patel, 2008). Ces facteurs incluent la proéminence, la structure syllabique, la réduction vocalique, le débit de parole, le style discursif, etc., sans oublier la diversité ou la variation due aux niveaux lexical, syntactique, contextuel ou idiosyncratique. D'après la discussion précédente, la syllabe ne peut être le seul élément examiné et la durée ne peut non plus être le seul facteur à considérer dans l'étude du rythme (bien que la durée soit le facteur le plus pertinent) (cf. Pamies, 2010⁶⁰).

Enfin, le principe d'isochronisme ne serait, selon certains chercheurs (par exemple, Allen, 1975 ; Turk & Shattuck-Hufnagel, 2013), qu'une manifestation métrique (ou phonologique), qui se situe à un niveau abstrait. Par ailleurs, le système perceptif de l'être humain a la capacité de transformer un signal non-périodique en un signal quasi-périodique (Turk & Shattuck-Hufnagel, 2013). Dans la situation communicative, le rythme de la parole se manifeste plutôt par une tendance vers l'isochronie. Comme l'a indiqué l'étude de Poeppel et Assaneo (2020), il existe une certaine périodicité translinguistique, qui est non seulement une impression perceptive, mais qui a également une manifestation acoustique⁶¹. Cette périodicité langagière n'est pas une stricte isochronie (Cutler, 1991), mais se situe dans une fourchette de 2-8 Hz, en tenant compte de tous les facteurs variables externes (i.e. langues, dialectes, âges, sexes, conditions de production, etc.).

1.3.2.3 Fonctions du rythme

À l'instar de tous les éléments prosodiques, le rythme joue plusieurs fonctions dans la parole. Dans cette section, nous examinerons les rôles spécifiques que le rythme

⁶⁰ « ... un éventuel rythme linguistique (régulier ou pas) organise des unités qui sont linguistiques elles aussi, ce qui, dans de nombreuses langues, exclut la durée en tant que telle, et donne un rôle essentiel au nombre et à la disposition des éléments rythmiquement pertinents. Ceux-ci dépendent aussi d'autres contraintes morpho-phonologiques, voire syntaxiques, ce qui favorise donc encore plus l'hypothèse d'une organisation très asymétrique du rythme discursif, à supposer qu'il existe, comme nos intuitions nous font penser, mais dont aucune description précise et bien fondée ne témoigne pour l'instant. » (Pamies, 2010 : 255).

⁶¹ "Regardless of the type of analysis employed, all studies converge on the same conclusion: the speech envelope possesses an overall 1/f noise spectrum, which, when removed, reveals an increase in power for frequencies between 2 and 8 Hz, with a notable peak between 4 and 5 Hz. Critically, these features are preserved across speakers, languages and speaking conditions. To summarize, speech is not only perceived as quasi-rhythmic but the produced acoustic signal itself shows striking temporal regularities." (Poeppel & Assaneo, 2020: 2).

assume. Tout d'abord, le rythme, tout comme la prosodie, fonctionne d'une part comme un « *bootstrapping* » pour faciliter l'acquisition de la prosodie de la langue maternelle et la distinction entre différents patrons rythmiques chez des très jeunes enfants (Cutler, 1994a ; Nazzi *et al.*, 1998 ; Ramus *et al.*, 1999 ; Patel, 2008 ; Abboub *et al.*, 2016). D'autre part, il fonctionne comme un guidage (Kohler, 2009 ; Cushing & Dellwo, 2010 ; Turk & Shattuck-Hufnagel, 2014), qui permet de planifier un cadrage temporel et mélodique préalable de l'énoncé, tout en aidant les auditeurs à prédire ce qui va suivre dans l'énoncé.

Ensuite, le rythme constitue, par le biais de la chaîne syllabique, le squelette temporel de la langue (Allen, 1975), à travers duquel le rythme de la parole remplit les fonctions de segmentation et de regroupement (Lehiste, 1972 ; Cutler & Butterfield, 1992 ; Cutler, 1994b ; Gilbert *et al.*, 2011). Comme nous l'avons mentionné plus haut, la segmentation et le regroupement sont deux faces d'une même pièce, car c'est par la segmentation que l'on peut déterminer quels éléments sont regroupés (la même idée chez Wagner, 2010⁶² ; Di Cristo, 2013⁶³). En variant la durée et/ou la F0 (en fonction de la langue), la prééminence et la pause étant réalisées, la chaîne de parole peut être organisée et regroupée selon l'intention communicative du locuteur (Astésano, 2016).

1.3.2.4 Mesure du rythme

Dans cette section, nous présenterons les différentes façons de mesurer le rythme. Dans la première sous-section, nous introduirons les paramètres quantitatifs les plus adoptés dans l'étude du rythme sur une langue naturelle donnée ou sur des interlangues chez des apprenants d'une L2. L'isochronie rythmique est partiellement démontrée par les paramètres quantitatifs dans Ramus (1999). Cependant, ceux-ci ont fait l'objet de nombreuses critiques concernant leur validité, en raison, d'une part, de la variabilité flexible du rythme de la parole, influencée par

⁶² Dans Wagner (2010 : 120) : « *Rhythmic grouping and rhythmic structure are closely related, since the relation between stronger and weaker beats causes grouping processes – as they are depicted in metrical trees.* »

⁶³ Dans Di Cristo (2013), il indique que dans le cadre de la phonologie métrique, les prééminences et la formation des groupes participent « *d'une représentation conjuguée* ». L'usage de cette expression signifie que l'attribution d'une prééminence à une syllabe donnée entraîne ipso facto la création d'un groupement ou d'un domaine et que la formation d'un groupement, entraîne pareillement la projection d'une prééminence, qui constitue alors la tête métrique (ou l'élément gouverneur) de ce groupement (ou de ce domaine). » (Di Cristo, 2013 : 123).

le style discursif, le débit, l'effet idiosyncratique, etc., et, d'autre part, du fait que leur principe considère le rythme langagier comme un simple phénomène temporel. C'est pourquoi, dans la deuxième sous-section, nous présenterons le modèle CCI, un cadre d'analyse qui permet de prendre en considération la prééminence dans l'analyse rythmique.

1.3.2.4.1 Les paramètres quantitatifs rythmiques classiques

Nous avons évoqué la catégorisation rythmique basée sur l'isochronie dans la section 1.3.2.2.1, ainsi que le continuum rythmique proposé par Dauer (1983). Le manque d'une base solide pour la typologie d'isochronie a poussé certains chercheurs à chercher des indices quantitatifs pour démontrer empiriquement la validité de la catégorisation traditionnelle (Figure 1.15). C'est pour cela qu'une série de paramètres quantitatifs (*rhythm metrics*) a été proposée et largement adoptée dans l'étude du rythme de la parole et de l'acquisition du rythme des langues étrangères.

En général, une dizaine de paramètres ont été développés pour la mesure rythmique. Certains d'entre eux sont plus souvent utilisés, alors que d'autres le sont moins. Nous récapitulons dans le Tableau 1.12 les paramètres quantitatifs rythmiques, ainsi que les références et leurs significations, classés selon leur fréquence d'utilisation :

Tableau 1.12 : Synthèse sur les paramètres quantitatifs rythmiques (adaptée de Loukina *et al.* (2011) entre autres).

Paramètre	Références	Signification
%V	Ramus (1999) ; Ramus <i>et al.</i> (1999)	« La proportion des intervalles vocaliques au sein de la phrase, calculée comme la somme des durées des intervalles vocaliques divisée par la durée totale de la phrase. » (Ramus, 1999 : 47). Il est « l'indice acoustique de la complexité des structures syllabiques et de la présence/absence de réduction vocalique à la fois. » (Pillot-Loiseau & Xie, 2018 : 5).
ΔV	Ramus (1999) ; Ramus <i>et al.</i> (1999)	« L'écart-type des durées d'intervalles vocaliques au sein de la phrase » (Ramus, 1999 : 47) et il reflète « la présence/absence de réduction vocalique dans une langue donnée : les langues autorisant la réduction vocalique (anglais) ont ainsi une valeur de ΔV plus grande que les autres qui l'autorisent moins (français). » (Pillot-Loiseau & Xie, 2018 : 5).
ΔC	Ramus (1999) ; Ramus <i>et al.</i> (1999)	« L'écart-type des durées d'intervalles consonantiques au sein de la phrase » (Ramus, 1999 : 47) et il reflète « la complexité des structures syllabiques d'une langue. Sa valeur sera inférieure pour les langues comme le français dont les structures syllabiques sont assez simples. » (Pillot-Loiseau & Xie, 2018 : 5).

VarcoV VarcoC VarcoVC	Dellwo (2006) ; White & Mattys (2007a, b) ; Liss <i>et al.</i> (2009)	ΔV et ΔC normalisés
rPVI-V rPVI-C rPVI-VC (ou CV) ⁶⁴	Low <i>et al.</i> (2000) ; Grabe & Low (2002) ; Barry <i>et al.</i> (2003) ; Liss <i>et al.</i> (2009)	« <i>La mesure des variabilités brutes des intervalles vocaliques/consonantiques successifs</i> » (Pillot-Loiseau & Xie, 2018 : 5) ou intervalle du groupe « voyelle + consonne » (ou « consonne + voyelle »). $rPVI = \left[\sum_{k=1}^{m-1} d_k - d_{k+1} / (m - 1) \right]$
nPVI-V nPVI-C nPVI-VC (ou CV)	Low <i>et al.</i> (2000) ; Grabe & Low (2002) ; Asu & Nolan (2005) ; Liss <i>et al.</i> (2009)	rPVI normalisé : « <i>les valeurs de nPVI sont plus élevées dans les langues à rythmicité accentuelle, et inférieures pour les langues de rythme syllabique et moraique, car les intervalles vocaliques sont plus variés dans des langues à isochronie accentuelle.</i> » (Pillot-Loiseau & Xie, 2018 : 5). $nPVI = 100 \left[\sum_{k=1}^{m-1} \left \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right / (m - 1) \right]$
ΔS	Chen & Zechner (2011)	L'écart-type des durées d'intervalles syllabique
VarcoS	Mok & Dellwo (2008)	ΔS normalisé
rPVI-S	Mok & Dellwo (2008)	La mesure des variabilités brutes des intervalles syllabiques successifs
nPVI-S	Mok & Dellwo (2008)	rPVI-S normalisé
VI	Deterding (2001)	<i>Variability Index</i> : durée syllabique consécutive normalisée (sauf la syllabe finale)
Vdur/Cdur	Barry & Russo (2003)	Ratio entre durée vocalique et durée consonantique
YARD	Wagner & Dellwo (2004)	<i>Yet Another Rhythm Determination</i> : variabilité de la durée syllabique d'une séquence
Rhythm Ratio (RR)	Gibbon & Gut (2001)	Un quasi-substitut du paramètre PVI ⁶⁵

⁶⁴ "We chose not to use pure syllable durations, as in Gut (2003), in order to preserve the acoustic nature of the metrics and to avoid imposing phonological constraints such as syllabification rules on the speech string. We combined each vocalic interval with the subsequent consonantal interval (rather than vice versa, as in Barry *et al.*, 2003), on the basis that phonological theory and psychological research indicate that it is the nature of the syllable rhythm (the vocalic nucleus plus consonantal coda) that gives rise to the perception of syllable weight." (Liss *et al.*, 2009: 1337).

⁶⁵ "... the Rhythm Ratio (RR), which calculates the average ratio of all adjacent syllable or vowel pairs as a percentage (multiplied by 100) ..." (Gut, 2012: 84) et "There is no substantive difference between the PVI and the RR." (Gibbon & Gut, 2001).

VtoV	Pettorino <i>et al.</i> (2013)	Intervalle entre deux VOPs (<i>Vowel Onset Points</i>)
------	-----------------------------------	--

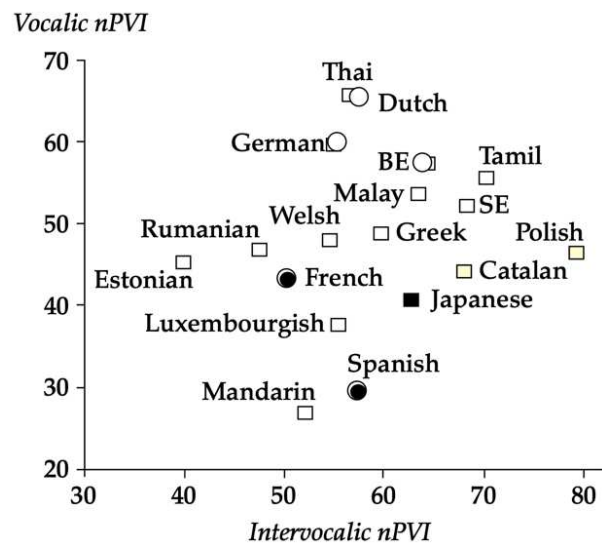
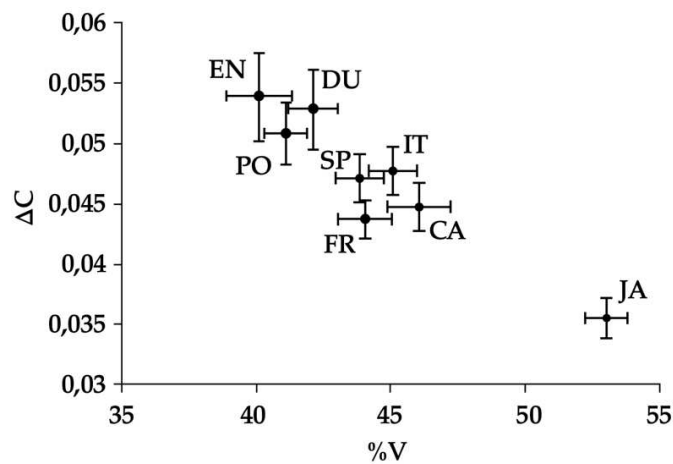


Figure 1.15 : Utilisation des paramètres quantitatifs rythmiques pour la catégorisation rythmique des langues du monde (gauche : %V et ΔC de Ramus *et al.* (1999) ; droite : nPVI-C et nPVI-V de Grabe & Low (2002)).

D'autres paramètres sont aussi proposés par des chercheurs selon leur but de l'étude : « *mean_rpviv* », « *mean_npviv* » (« *r* » pour brute ; « *n* » pour normalisée), « *med_rpviv* », « *med_npviv* » (« *med* » pour médian), etc.⁶⁶ (Ferragne & Pellegrino, 2004) ; « %VO », « *deltaUV* », « *nPVI-VO* » et « *rPVI-UV* » qui permettent de mesurer les durées des intervalles voisés et non-voisés et qui proposent une nouvelle catégorisation rythmique en fonction de ces critères (cf. Fourcin & Dellwo, 2009).

⁶⁶ Significations des paramètres : par exemple, « *mean_rpviv* » : "mean difference in duration between all pairs of consecutive vocalic intervals (the <r> stands for 'raw')." (Ferragne & Pellegrino, 2004). À l'instar de cela, une série de paramètres consonantiques ont été aussi proposés.

Afin de donner une description complète sur le patron rythmique d'une langue donnée en recourant aux paramètres quantitatifs rythmiques, il est préférable de calculer les valeurs d'un nombre suffisant de paramètres (Loukina *et al.*, 2009) (cf. Knight (2011) pour un avis contraire⁶⁷). Cependant, l'utilisation des paramètres quantitatifs comme moyen de catégorisation rythmique a été critiquée :

- a) Même si elle avait été examinée avec des paramètres quantitatifs (voir Ramus, 1999 ; Ramus *et al.*, 1999), l'hypothèse d'isochronie rythmique n'a pas été totalement empiriquement validée. Cela implique que le rythme de la parole n'est pas simplement un phénomène isochronique temporel et que les paramètres quantitatifs ne permettent pas de révéler toutes les caractéristiques rythmiques ;
- b) La mesure rythmique basée sur les paramètres quantitatifs ne prend en compte que le facteur temporel tandis que le rythme langagier suggère d'autres facteurs que la durée (Barry *et al.*, 2009), notamment la F0 (Cumming, 2008, 2010) (cf. aussi notre synthèse dans la section 1.3.2.2.3) ;
- c) Dans la littérature, les valeurs calculées pour une langue donnée peuvent considérablement varier d'une étude à l'autre (Wiget *et al.*, 2010) ;
- d) Des langues peuvent être catégorisées différemment en fonction des paramètres choisis (Loukina *et al.*, 2011). Par ailleurs, certains paramètres sont plus sensibles aux différences entre patrons rythmiques, donc plus performants, alors que certains autres paramètres le sont moins (Loukina *et al.*, 2009) ;
- e) Dans l'étude sur la production rythmique d'une langue étrangère, un paramètre donné se manifeste de manière diverse : en fonction des catégories rythmiques des L1 et L2 (ou Ln) (source et cibles), du protocole expérimental, du style de la parole (lue vs spontanée), des niveaux variés des apprenants au cours d'acquisition (qui est étroitement lié au débit de parole), etc. (par exemple, Gut, 2012 ; Lai *et al.*, 2013 ; Ordin & Polyanskaya, 2015a). Tout cela conduit à une comparaison translinguistique plus difficile et floue (Arvaniti *et al.*, 2008 ; Ferjan *et al.*, 2008 ; Arvaniti, 2012), voire invalide (Gut, 2012, cf. Knight, 2011 pour une revue sur ce point). Toutefois, les valeurs des

⁶⁷ Certains chercheurs proposent le contraire : "... it makes sense to use only those metrics that are the most stable, reliable and valid, such as %V." (Knight, 2011: 278).

paramètres quantitatifs pourraient, dans certains cas, refléter le processus d'acquisition rythmique : plusieurs études ont montré que des valeurs chez des apprenants avancés sont plus proches de celles des locuteurs natifs comparés aux apprenants débutants (par exemple, Stockmal *et al.*, 2005 ; Gutiérrez Díez *et al.*, 2008). Dans ce cas-là, la recherche sur l'acquisition du rythme d'une langue étrangère peut bénéficier des paramètres quantitatifs, mais uniquement à titre indicatif (Grenon & White, 2008). En aucun cas, ces paramètres ne constituent un moyen complet et précis.

Parmi tous les paramètres listés dans le Tableau 1.12, il y a quelques paramètres qui sont considérés comme robustes face à la variabilité de la parole :

- %V, PVI-CV (Barry *et al.*, 2003) qui permettent de capturer la relation entre consonnes et voyelles ;
- %V est robuste face à la variation du débit de parole (Russo & Barry, 2008a, 2008b). La combinaison des paramètres %V, VarcoV et nPVI-V permet de mieux interpréter la catégorisation rythmique des langues et aussi le transfert rythmique de la langue maternelle à la langue seconde (White & Mattys, 2007a) ;
- Les paramètres basés sur les durées vocaliques (%V et ΔV) semblent plus pertinents que ceux basés sur les durées consonantiques, notamment %V qui reste plus ou moins stable comparé aux autres paramètres (Knight, 2011) ;
- Les paramètres coefficients sont plus robustes face à des corpus différents (Arvaniti *et al.*, 2008).

L'analyse à travers les valeurs quantitatives considère le rythme de la parole comme un événement purement temporel, ce qui est à l'évidence inapproprié. Comme nous l'avons rappelé à plusieurs reprises, la surface de production rythmique est une combinaison de plusieurs facteurs intrinsèques (e.g., structure syllabique, proéminence, allongement final)⁶⁸ et extrinsèques (e.g., débit de parole, longueur de la phrase, style de parole, etc.) (Arvaniti & Rodriguez, 2013). C'est pour cette raison

⁶⁸ "It seems likely that metrics simply capture the timing cues that listeners use to differentiate languages, whilst various other factors influence how humans perceive speech rhythm. For example, prosodic timing processes such as phrase-final lengthening (e.g. White & Mattys 2007a, Prieto *et al.* 2010), levels above and below the syllable (e.g. Allen & Hawkins 1980: 232), and utterance-specific coupling of nested levels of prosody (e.g. Cummings & Port 1998; Cummings 2002, 2003) must also be taken into account for a full understanding of how we perceive speech rhythm." (Knight, 2011: 278).

que dans la section suivante nous voulons introduire le modèle CCI et sa quantification pour l'analyse rythmique.

1.3.2.4.2 Le modèle CCI (*Control/Compensation Index*)

Dans la section 1.3.2.2.3, nous avons proposé une synthèse de la définition du rythme de la parole selon laquelle le rythme, au lieu d'être un fait purement temporel, est en effet un phénomène qui englobe d'autres facteurs (comme la F0, la proéminence, etc.) et se manifeste à plusieurs niveaux (par exemple, les deux grands dispositifs chez Di Cristo, 2016 ; le microrythme et le macrorythme chez Jun, 2005, 2012, 2014 ; Selkirk, 1984). Dans la section 1.3.2.4.1, nous avons revisité les paramètres quantitatifs rythmiques, la méthodologie la plus adoptée dans l'étude du rythme, ainsi que les critiques formulées à leur rencontre. Afin de donner une meilleure mesure du patron rythmique, tant chez les locuteurs natifs que chez les apprenants d'une langue étrangère, nous souhaitons introduire le modèle CCI (*Control/Compensation*) (Bertinetto & Bertini, 2008, 2010).

Ce modèle se base sur la souplesse ou la flexibilité des mouvements des articulateurs dans la production de la parole (Bertinetto *et al.*, 2012). Les langues se distinguent dans la manière dont les voyelles et les consonnes se combinent. Dans les langues dites « *controlling* » (ci-après, « contrôlantes », où la chaîne d'articulation est relativement figée, les voyelles et les consonnes nécessitent un effort articulaire quasi identique, ce qui, dans un cas idéal, se traduit par des durées similaires (représenté par la zone blanche dans la Figure 1.16). À l'inverse, les langues dites « *compensating* » (ci-après, « compensatrices ») se caractérisent par une articulation plus flexible, permettant plus de variabilités, comme la coarticulation ou la réduction vocalique (représenté par la zone grise claire dans la Figure 1.16).

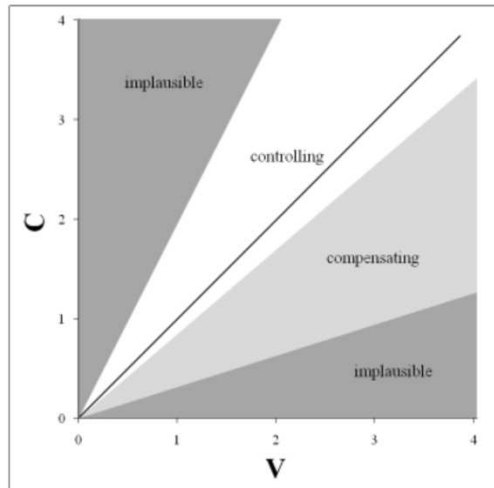


Figure 1.16 : Schéma représentant les deux catégories rythmiques idéales basées sur le modèle CCI (Bertinetto *et al.*, 2012 : 520).

En effet, les catégories « contrôlante » et « compensatrice » ne représentent que deux cas idéaux comme indiqué dans les études de Bertinetto et Bertini (2008, 2010). Le modèle CCI s'apparente à l'idée du continuum rythmique chez Dauer (1983), où les termes « contrôlante » et « compensatrice » se réfèrent respectivement aux deux extrémités du continuum rythmique « isochronie syllabique » (contrôlante) et « isochronie accentuelle » (compensatrice) (Bertinetto & Bertini, 2008). Cette correspondance a été démontrée quantitativement dans l'étude de Mairano et Romano (2011).

Deux niveaux d'analyse peuvent être établis dans le cadre du CCI : le niveau I, niveau phonotactique, se basant sur l'alternance vocalique vs consonantique (le niveau microrythmique) ; et le niveau II, niveau énonciatif, se basant sur l'alternance de proéminence vs non-proéminence (le niveau macrorythmique)⁶⁹. Comme pour le niveau I, il existe dans le niveau II une distinction entre « contrôlante » et « compensatrice » en fonction de la rigidité ou la flexibilité du cadrage métrique.

Le calcul de l'index du niveau I du modèle CCI est modifié sur la base du paramètre PVI. La différence entre l'index du modèle CCI et le paramètre PVI, est que l'index CCI ajoute le nombre de segments dans un même intervalle vocalique ou

⁶⁹ "... the nature of the level-I oscillators is universally specified, for all languages consist of recurring Vs and Cs. This, however, does not hold for level-II. Here again we have to do with physical objects – namely the more or less regularly recurring sentence accents – but the actual implementation of the level-II oscillating system varies from language to language, depending on the specific set of acoustic cues making up the accents' physical substance... The parameter of word-stress flexibility should thus be regarded as the level-II equivalent of the level-I C/C-contiuum: the more mobile the stress is, the more flexible (COMPS) the accentual structure, for the speaker may then have a larger degree of freedom in regulating the inter-accentual distances." (Bertinetto & Bertini, 2010: 61, 63).

consonantique (Bertinetto & Bertini, 2010), ce qui permet de prendre en considération la complexité de la structure phonotactique (par exemple, l'intervalle vocalique peut être composé d'une seule voyelle ou d'une voyelle longue ou d'une séquence vocalique ; l'intervalle consonantique peut être composé d'une seule consonne, ou d'une gémérée, ou d'une séquence consonantique) :

$$CCI = \frac{100}{m-1} \sum_{k=1}^{m-1} \left| \frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}} \right|$$

(m : nombres d'intervalles ; d : durée ; n : nombres de segments dans un intervalle donné)

Concernant la quantification du niveau II, elle se base sur l'étude de O'Dell et Nieminen (1999) selon laquelle les patrons rythmiques des langues du monde peuvent aussi se différencier par les durées entre les intervalles accentuels (*average phrase distance*). La quantification du niveau II est calculée à l'aide de la formule suivante (Bertinetto *et al.*, 2012 : 520) :

$$I(n) = \frac{r}{r\omega_1 + r\omega_2} + \frac{1}{r\omega_1 + r\omega_2} n^{70}$$

Cette formule permet de décrire la relation entre les durées des intervalles accentuels et le nombre de syllabes contenant dans un intervalle donné⁷¹.

À la suite des travaux de Bertinetto et de ses collaborateurs (Bertinetto & Bertini, 2010 ; Bertinetto *et al.*, 2012), des études ultérieures ont comparé la différence des niveaux I et II entre quelques langues en utilisant le modèle CCI (Tableau 1.13) :

⁷⁰ "‘Averaged Phase Distance’ theory: *I* stands for ‘duration of the inter-accentual intervals’, *n* for ‘number of syllable-peaks’, ω_1 and ω_2 for the angular frequency – or velocity – of the two oscillators (assumed to be constant and expressing the ‘natural’ rhythm of each oscillator), and *r* indicates their relative-strength parameter, i.e. the degree of dominance of the accentual phrase over the syllable." (Bertinetto *et al.*, 2012: 521).

⁷¹ "If the resulting *r* is greater than 1, then the overarching (accentual) first oscillator predominates; if *r* is less than 1, the subordinate (syllable-peak) second oscillator prevails." (Bertinetto & Bertini, 2010: 67).

Tableau 1.13 : Catégorisation des langues basée sur les niveaux I et II du modèle CCI (Tableau établi en fonction des études de Bertinetto & Bertini, 2010) ; Romano & Mairano (2010) ; Bertinetto *et al.* (2012)).

Niveau I	Niveau II	Langues
COMPS (compensatrice)	CONTR (contrôlante)	Polonais (Structure phonotactique complexe ; cadre métrique rigide)
CONTR	CONTR	Italien, Français (Bertinetto & Bertini, 2010) (Structure phonotactique simple ; cadre métrique rigide)
COMPS	COMPS	Anglais (Structure phonotactique complexe ; cadre métrique flexible)
CONTR	COMPS	Chinois (Bertinetto <i>et al.</i> , 2012) (Structure phonotactique simple ; cadre métrique flexible)

Par ailleurs, comme l'ont rappelé Bertinetto *et al.* (2012), le modèle CCI est un cadre d'analyse qui est aussi sensible à la variation comme le débit de parole ou le style de parole. Ainsi, ce modèle d'analyse peut également révéler la différence des patrons rythmiques qui ont des débits de parole variés (cf. l'étude de Bertinetto & Bertini (2010)).

La raison pour laquelle nous souhaitons adopter l'approche CCI dans l'analyse des données rythmiques dans la partie expérimentale, est que cette mesure peut rendre compte de la différence rythmique en tenant compte des syllabes non-accentuées et accentuées dans la parole. Cette approche permet donc de compléter l'analyse rythmique faite par les paramètres quantitatifs classiques. Comme nous l'avons souligné précédemment, la proéminence fait partie des phénomènes rythmiques et nous devons l'inclure dans notre analyse rythmique dans la partie expérimentale. Le modèle CCI répond donc à ce besoin⁷².

1.3.2.4.3 Sur la mesure du rythme : une synthèse

Dans les deux sections précédentes consacrées à la mesure du rythme, nous avons revu deux méthodologies principales pour mesurer le rythme de la parole : les paramètres quantitatifs acoustiques traditionnels et le modèle CCI. Nous avons

⁷² L'idée de départ pour proposer le modèle CCI : "*Specifically, although the phonological role of word stress differs from language to language, one may assume that sentence accents are universally present as rhythm regulators, whatever their language-specific phonetic implementation...* [donc la relation étroite entre le rythme et la proéminence]... *The suggestion underlying the model proposed here is that no single measure can capture the actual behavior of any language: both level-I and level-II should be taken into account.* [donc seuls les paramètres quantitatifs rythmiques ne suffisent pas]". (Bertinetto & Bertini, 2010: 62, 64).

indiqué que les paramètres traditionnels capturent les caractéristiques temporelles au niveau microrhythmique mais l'analyse menée dans le cadre de cette méthodologie est influencée par une série de facteurs, notamment le débit de parole. En plus, cette mesure ne parvient pas à donner une description complète des phénomènes rythmiques, surtout au niveau macrorhythmique comme la proéminence et le phrasé. Comme nous l'avons évoqué dans la synthèse de la définition du rythme, au moins deux niveaux (ou deux dispositifs) d'analyse sont nécessaires. L'ajout du modèle CCI, comme un nouveau moyen, permettra de compléter l'analyse au niveau macrorhythmique et d'approfondir la compréhension de l'organisation rythmique dans la parole, notamment pour étudier la réalisation rythmique de l'interlangue.

1.3.2.5 Pauses

La pause est un élément temporel important parmi les aspects temporels de la prosodie et elle peut être définie comme une période ou phase d'interruption d'articulation et de phonation au cours de la production de la parole (Trouvain & Braun, 2020). Les pauses dans la parole peuvent être classées typologiquement en deux catégories : pauses silencieuses (pendant laquelle aucune activité de phonation ne se passe) et pauses remplies (souvent marquées par des interjections « euh », « um ») ; ou en pauses grammaticales, dont les productions correspondent aux jonctures de la structure syntaxico-prosodique hiérarchisée d'un énoncé ou discours donné⁷³ (plus le niveau hiérarchique est haut, plus la durée de la pause est longue) et pauses non grammaticales, qui sont le résultat du processus de planification de la parole (par exemple, les pauses liées au temps nécessaire pour planifier un mot ou un énoncé à venir) (Krivokapić *et al.*, 2020).

L'occurrence des pauses dans la parole est liée au débit : plus le débit de parole est lent, plus les pauses sont fréquentes⁷⁴ (Krivokapić, 2007 ; Krivokapić *et al.*, 2020). Par ailleurs, dans la parole produite par les apprenants d'une L2, l'occurrence des pauses est généralement plus nombreuse et la durée moyenne est relativement plus

⁷³ "Research on the effects of syntactic structure on pause occurrence has shown that pauses often occur at major syntactic boundaries and that the more complex the syntactic constituent, the more likely a pause is to occur (see for example Cooper & Paccia-Cooper, 1980; Grosjean *et al.*, 1979; Strangert, 1991 for Swedish)." (Krivokapić, 2007: 164).

⁷⁴ "L2 speech is often characterized by a slower tempo than L1 speech, due to a slower articulation rate and more and longer pauses (Pürschel 1975 and Wiese 1983 for German learners of English; Trofimovich and Baker 2006 for Korean learners of English)." (Trouvain & Braun, 2020: 612).

longue par rapport aux pauses dans la parole des locuteurs natifs (Trouvain & Braun, 2020 ; Riazantseva, 2001). Ainsi, l'occurrence des pauses et ses durées dépendent d'une série de facteurs : structures prosodique et syntaxique, longueur du syntagme, débit de parole, locuteurs, styles de parole, etc.

La pause est un élément important dans l'étude de la prosodie, car, comme nous l'avons explicité plus haut, elle constitue un indice universel pour marquer les frontières prosodiques (voir par exemple Stack & Watson, 2023). D'autre part, elle est un indicateur pour la compétence prosodique dans la production chez les apprenants d'une L2 : la position, la distribution, l'occurrence et la durée des pauses dans la parole peuvent refléter non seulement la performance en fluidité de la parole en L2 ⁷⁵ (Kahng, 2018 ; De Jong, 2016 ; Riazantseva, 2001), mais également la capacité à segmenter la parole continue par la pause.

1.3.2.6 Le rythme du chinois mandarin

Du point de vue de la catégorisation traditionnelle des rythmes des langues du monde, le chinois mandarin est considéré comme une langue d'isochronie syllabique (Jun, 2005 ; Lin & Wang, 2007 ; Mok & Dellwo, 2008 ; Mok, 2009 ; Romano *et al.*, 2011). Cependant, certaines études remettent cela en question : par exemple, plusieurs études (Cao, 2000b ; Peng, 2006) ont indiqué que le chinois mandarin n'est pas une langue à isochronie au niveau syllabique à proprement parler, car les durées syllabiques varient en fonction des types syllabiques et des tons lexicaux. Néanmoins, une tendance isochronique se manifeste au niveau macrorhythmique : le syntagme phonologique contient un nombre de syllabes plus ou moins fixe : soit 7 ± 2 , ce qui est corroboré par de nombreuses études (Miller, 1956 ; Dittmann & Llewellyn, 1967 ; Pegolo, 1985 ; Kohno & Tsushima, 1989 ; Kohno & Tanioka, 1990). Cela suggère une tendance universelle dans l'encodage du signal de parole, partagée par les langues du monde.

Cao (2001) a exploré le patron macrorhythmique, en soutenant d'abord qu'il n'existe pas de vraie isochronie syllabique et que la régularité rythmique ne revient que dans

⁷⁵ "For instance, in Bosker *et al.* (2013), both pause frequency and pause length were negatively correlated with fluency ratings. However, in Kormos and Deñes (2004) fluency ratings did not correlate with pause frequency but did correlate with pause length." (Kahng, 2018: 570). "...fluency refers to "the ability to talk at length with few pauses, the ability to fill time with talk" (Fillmore, 1979, p. 93). This characteristic clearly refers to speaking with a minimum amount of hesitation and pausing, and thus reflects the temporal aspect of fluency." (Riazantseva, 2001: 498).

certaines positions prosodiques. Elle a donné une description acoustique plus détaillée sur le patron rythmique du chinois mandarin : selon elle, la structure rythmique du chinois mandarin est divisée en trois niveaux : mot prosodique, syntagme prosodique⁷⁶ (ou « unité rythmique » dans son étude) et syntagme intonatif. À chacun de ces niveaux, des particularités acoustiques se manifestent comme montrées par la Figure 1.17:

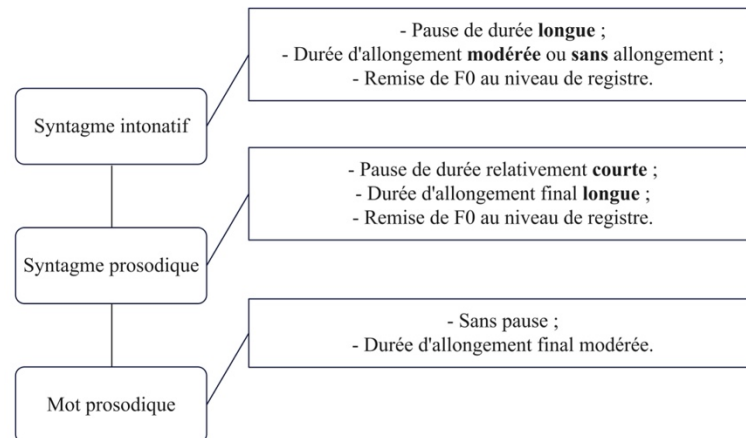


Figure 1.17 : Trois niveaux hiérarchiques prosodiques du chinois mandarin et leurs propriétés acoustico-rythmiques (adaptés de Cao (2001)).

Cao (2001) a conclu que dans le chinois mandarin, comme dans beaucoup d'autres langues, la pause et l'allongement final (et le mouvement mélodique) jouent conjointement le rôle de démarcation. Mais, l'allongement final dans un syntagme prosodique intermédiaire (syntagme « sémantiquement non terminal » selon l'expression de l'auteure) est plus long que dans un syntagme intonatif (syntagme « sémantiquement terminal ») (même idée chez Wang *et al.* (2019) qui a examiné les durées des pauses et des allongements finaux à différents niveaux de la hiérarchie prosodique⁷⁷). Au contraire, les durées des pauses sont positivement corrélées aux niveaux hiérarchiques prosodiques : plus le niveau prosodique de la frontière est haut, plus la durée de la pause est longue. Par ces caractéristiques révélées dans son étude, Cao (2001) a soutenu que des marqueurs comme l'allongement final ou la

⁷⁶ "Rhythmic structure of Mandarin Chinese contains three basic layers, namely prosodic word, prosodic phrase and intonation phrase." (Cao, 2001: 76).

⁷⁷ "...pre-boundary syllable duration stops increasing beyond break level 2 in Mandarin, but continues to increase in English. Meanwhile, the duration of silent pause significantly increases beyond break level 2 in Mandarin, as if to compensate for the lack of continuous syllable lengthening." (Wang *et al.*, 2019: 2906).

pause sont non seulement dotées d'une fonction démarcative, mais leurs manifestations aux niveaux prosodiques différents servent également à garder une cohésion de la hiérarchie discursive⁷⁸.

Du point de vue de paramètres quantitatifs métriques, Romano et Mairano (2010) ont mesuré les valeurs de %V, ΔV et ΔC de 15 langues en utilisant le texte *La bise et le soleil* et ils ont obtenu les valeurs suivantes pour le chinois mandarin (Tableau 1.14) :

Tableau 1.14 : Valeurs des trois paramètres quantitatifs rythmiques pour le chinois mandarin (adaptées de Romano & Mairano (2010 : 96)).

Langue	%V	ΔV	ΔC
Chinois mandarin	51,2	43,67	43,88

Cependant, nous avons indiqué à plusieurs reprises que les paramètres quantitatifs traditionnels ne suffisent pas à définir ou à analyser le patron rythmique d'une langue donnée, et que d'autres facteurs devront être pris en compte, comme l'ont suggéré Romano et Mairano (2010)⁷⁹. Il nous semble donc nécessaire de synthétiser les études dans lesquelles le patron rythmique du chinois mandarin a été décrit en prenant compte d'autres facteurs que la durée segmentale.

Parmi les études disponibles, nous avons d'abord les travaux de Jun (2014, 2005) qui donnent une description relativement complète. Dans son cadre théorique (Jun, 2014, 2005), le patron rythmique du chinois mandarin a été décrit aux niveaux microrythmique, macrorythmique et accentuel (Tableau 1.15). À noter que dans son travail de 2014, le chinois mandarin a été présenté comme une langue à

⁷⁸ "...the quantitative difference which exists between these pauses also plays the role of connection between units in combination with certain distribution of pre-boundary lengthening. Therefore, a longer pre-boundary lengthening occurring at the end of a prosodic phrase is not surprising, because there is a relative pause there at the same time. Such a combination typically reflects **the special status of this intermediate unit in speech rhythm**. That is, on one hand, prosodic phrase is a most important and common division related to perceived rhythm, hence it is usually marked by a short silent pause; on the other hand, however, **prosodic phrase is a connective bridge functioning between prosodic word and intonation phrase**, so it must be marked by a significant pre-boundary lengthening, so as to show more connection, instead of demarcation, in this position." (Cao, 2001: 76).

⁷⁹ "Preliminary testing showed that duration could be a good correlate, giving a physical estimate of rhythmic (maybe derived) properties, but we believe that most direct correlates should take into account (distance, extension and shape of) peaks and general melodic profiles." (Romano & Mairano, 2010: 106).

proéminence de tête (qui se trouve à la position postlexicale) avec une macrorhythmicité faible⁸⁰.

Tableau 1.15 : Patrons microrhythmique et macrorhythmique du chinois mandarin (extrait de Jun (2005)).

Patron rythmique	Proéminence					Unité rythmique					
	lexicale			post-lexicale		lexicale			post-lexicale		
Chinois mandarin	ton	accent	« LPA » ⁸¹	tête	Frontière	more	syllabe	pied	AP	ip	IP
	X	X		X			X			X	X

Adoptant le modèle CCI, l'étude de Zhi *et al.* (2011) a montré que le chinois mandarin est une langue contrôlante et que l'accent énonciatif est moins évident. Bertinetto et Bertini (2010) ont indiqué que des langues comme le chinois ou le japonais, dont l'accent lexical est relativement fixe, se classent dans le quatrième type (niveau I : contrôlante ; niveau II : compensatrice) du Tableau 1.13. Bertinetto *et al.* (2012) ont ensuite précisé que pour le chinois mandarin, le rôle relativement faible de l'accent énonciatif au niveau II est compensé par l'oscillation syllabique parfaitement stable et rigide au niveau I⁸². Les auteurs ont conclu que les sinophones natifs sont plutôt sourds à l'accent énonciatif de leur propre langue maternelle. L'étude de Wang (2004) a également donné une description, qui rejoignait un peu les remarques de Bertinetto *et al.* (2012) et de Jun (2014) et qui correspondait à notre impression empirique : concernant le patron rythmique du chinois mandarin, au lieu d'une alternance de battements forts vs faibles, le patron se manifeste plutôt par deux particularités : 1) les segments phonétiques subissent la même force de battement, sans distinction d'accents forts vs faibles ; 2) avec les segments égaux au niveau de battement, on forme des groupes rythmiques, qui se séparent entre eux par pauses. Ainsi, à l'intérieur d'un groupe, les segments entre eux sont plutôt collés et serrés, alors qu'entre les groupes rythmiques, ou bien entre le dernier segment du groupe

⁸⁰ Selon le raisonnement de Jun (2014 : 530), "free tone or limited tone languages such as Cantonese or Yoruba would have Weak macro-rhythm since each syllable/word can carry any of the tone types, thus not likely to have any regular alternation of High and Low within a phrase", dont le chinois mandarin fait partie.

⁸¹ "lexical pitch-accent"

⁸² "Presumably, in languages like Chinese level-II plays a reduced role, with the phonotactic component doing most of the rhythmic job. Separate inspection of spontaneous and read materials adds further details to the picture... the situation is fairly stable in BC [Beijing Chinese]." (Bertinetto *et al.*, 2012: 522).

précédent et le premier segment du groupe suivant, c'est plutôt relâché, ou moins collé. Wang (2004) a pris le terme anglais « *looseness* » (*ibid.* : 28) pour caractériser ce patron rythmique en chinois mandarin. Dans ce contexte, la pause est un marqueur principal dans la prosodie pour signaler les syllabes groupées ainsi que les « relâchements », et donc les frontières prosodiques.

En combinant les descriptions ci-dessus sur le rythme du chinois mandarin (l'isochronie traditionnelle, le cadre macrorythmique, le modèle CCI et l'alternance de relâchements), nous pouvons dire qu'au niveau microrythmique (ou niveau I du modèle CCI), cette langue possède une alternance syllabique régulière relativement forte (isochronie syllabique) malgré la variation dans la parole, et au niveau macrorythmique (ou niveau II du modèle CCI), cette langue a une régularité moins évidente et les frontières entre syllabes groupées sont principalement signalées par les pauses.

1.3.2.7 Le rythme du français

Du point de vue de la typologie rythmique traditionnelle, le français est également considéré comme une langue à isochronie typiquement syllabique (Tableau 1.16) (Pike, 1945 ; Fant *et al.*, 1991⁸³ ; Ramus, 1999 ; Romano & Mairano, 2010).

Tableau 1.16 : Valeurs des trois paramètres quantitatifs rythmiques pour le français européen (adaptées de Romano & Mairano (2010 : 96)).

Langue	%V	ΔV	ΔC
Français européen	49,4	41,04	39,86

Cependant, la nature d'isochronie syllabique du français ne faisait pas consensus parmi les chercheurs, certaines études affirmant le contraire (Delattre, 1966a ; Crompton, 1980 ; Wenk & Wioland, 1982 ; Nishinuma & Duez, 1989). En effet, la catégorisation traditionnelle isochronique n'a pas pris en considération le patron macrorythmique, qui est supérieur au niveau syllabique (Fletcher, 1991).

⁸³ "In French the resistance to articulatory reductions reduces the stressed/unstressed contrast. Furthermore, the relatively low impact of non-terminal stresses in French is associated with smaller F0 modulations than in the major, usually pre-pause stresses. These are in our view the main arguments for referring to French as syllable timed." (Fant *et al.*, 1991: 363).

En français, l'unité rythmique devrait être le groupe rythmique iambique (Fletcher, 2010)⁸⁴. En effet, l'allongement final de la dernière syllabe du groupe rythmique est une particularité très saillante du français, noté par beaucoup de chercheurs dans la littérature (par exemple, Delattre, 1939 ; Vaissière, 1991, 2010 ; Astésano, 2016). Et donc, comme pour le chinois mandarin, la régularité rythmique du français se manifeste plutôt au niveau macrorythmique : c'est l'alternance des syllabes non accentuées vs accentuées, qui constitue une unité rythmique iambique (*trailer-timed*, Wenk & Wioland, 1982), et c'est cette récurrence iambique qui donne l'impression que le français est isochronique syllabique (Fletcher, 1991). Cette forte périodicité a été démontrée dans plusieurs études : en français, la durée moyenne des intervalles entre accents est à peu près de 555 ms (Fant *et al.*, 1991) ; l'étude de Jun et Fougeron (2002) a aussi indiqué que l'AP (syntagme accentuel, marqué par l'accent primaire) du français contient en moyenne 2,3-2,6 mots ou 3,5-3,9 syllabes. Ainsi, dans ce contexte, le français serait une langue dans laquelle les isochronies syllabique et accentuelle coexistent (Astésano, 2016).

En effet, le français n'est pas uniquement marqué à la queue. Selon la typologie établie par Jun (2005, 2014), le français est une langue à proéminences de tête et de queue, présentant une forte macrorhythmicité (Tableau 1.17). La proéminence dans ce type de langues a généralement une position relativement fixe : par exemple, pour le français, ce sont deux accents mélodiques LH* qui marquent simultanément la tête et la queue de l'AP (*ibid.*, 2014).

Tableau 1.17 : Patrons microrythmique et macrorythmique du français (extrait de Jun (2005 : 444)).

Patron rythmique	Proéminence					Unité rythmique					
	lexicale			post-lexicale		lexicale			post-lexicale		
Français	ton	accent	« LPA » ⁸⁵	tête	frontière	more	syllabe	pied	AP	ip	IP
					X	X		X		X	X

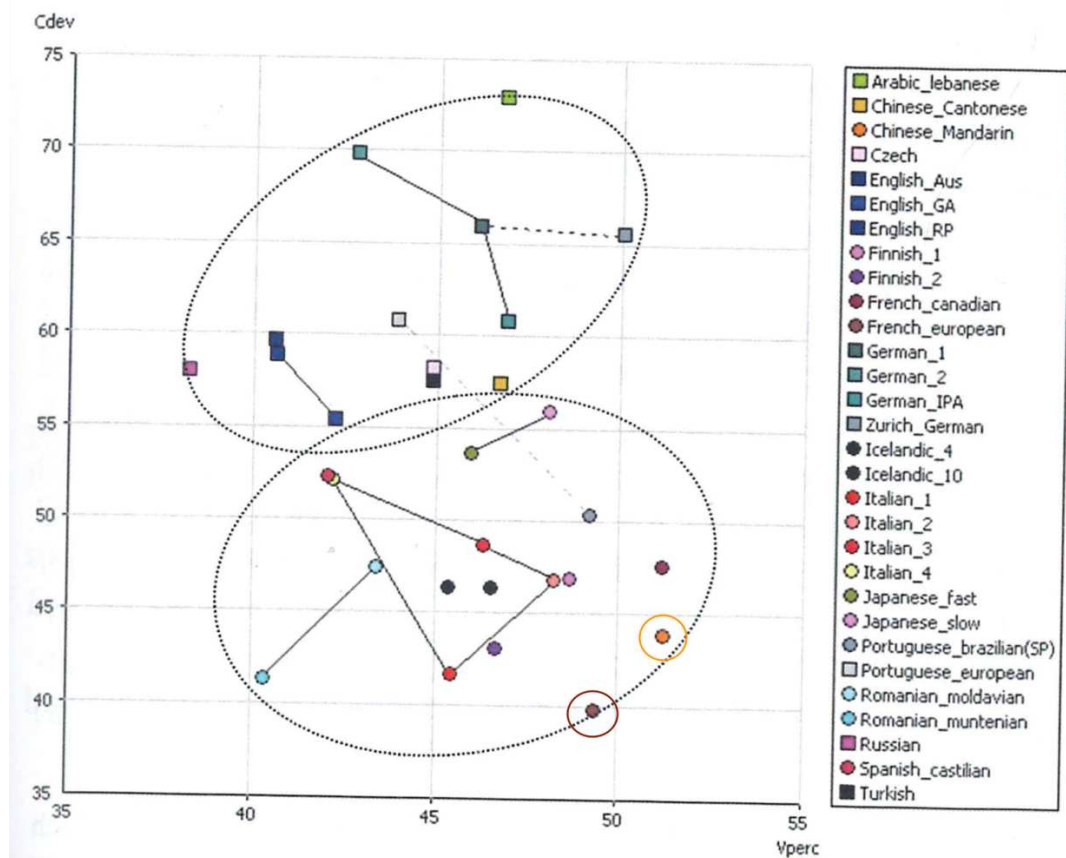
⁸⁴ "According to Grammont, the rhythmic units of French are not individual syllables, accented or unaccented, but groups of syllables that form right-headed iambic rhythmic feet or "groupes rythmiques." accordingly, Wenk and Wioland (1982) describe French rhythm as "trailer-timed..." (Fletcher, 2010: 558).

⁸⁵ lexical pitch-accent

1.3.3 Synthèse

Dans cette section, nous nous sommes concentrés sur les aspects temporels de la prosodie. Nous avons d'abord souligné l'omniprésence des phénomènes temporels, puis fourni les définitions du rythme de la langue humaine et de ses mesures, pour conclure avec la présentation des rythmes (micro et macro) du chinois mandarin et du français.

Du point de vue d'isochronisme au niveau microrhythmique, le chinois mandarin et le français sont majoritairement considérés comme des langues à isochronie syllabique, bien qu'il existe des études qui remettent cette idée en question. En comparant les trois valeurs du chinois mandarin et du français obtenues dans la même étude (cf. Tableaux 1.14 et 1.16) et la Figure 1.18 qui regroupe les 15 langues dans Romano et Mairano (2010 : 97, 99), nous pouvons constater qu'au moins en termes de paramètres quantitatifs rythmiques, le chinois mandarin et le français présentent globalement un patron rythmique similaire :



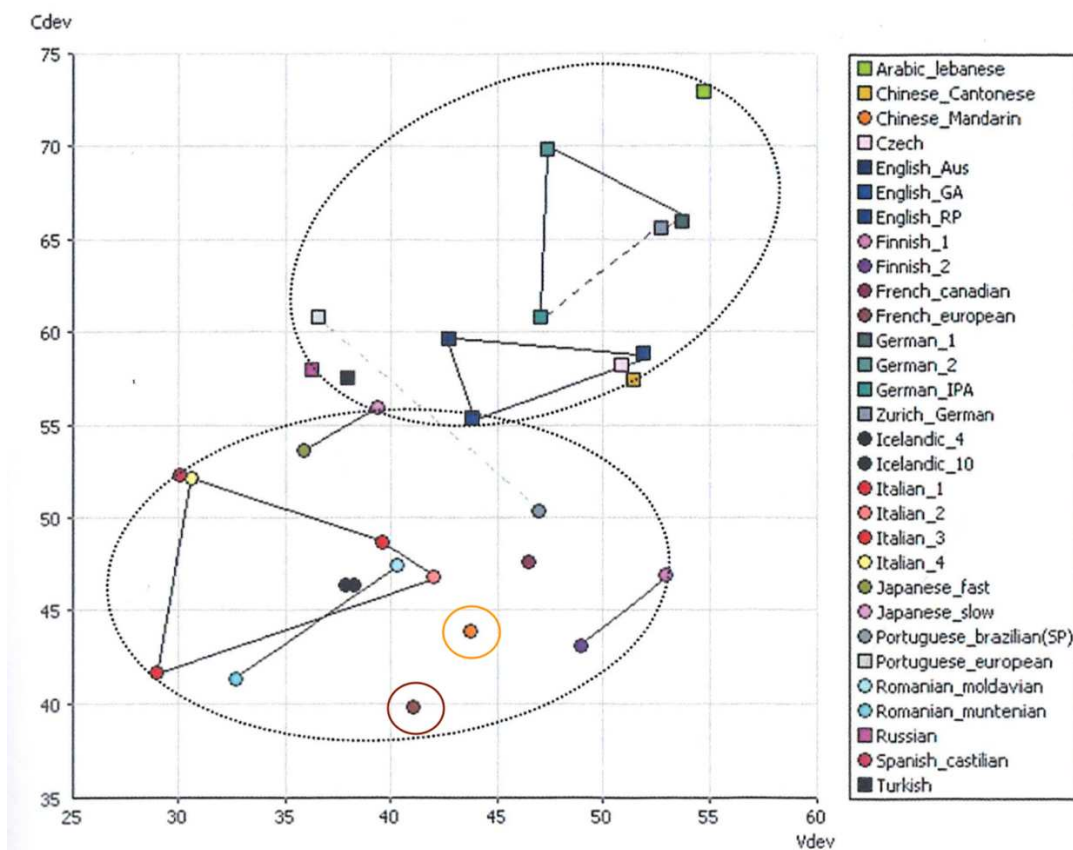


Figure 1.18 : Comparaisons des valeurs %V vs ΔC (haut) et ΔV vs ΔC (bas) parmi 15 langues du monde (avec le chinois mandarin entouré en orange et le français entouré en violet) (tirée de Romano & Mairano (2010 : 97, 99)).

Pour nous, nous pensons que les deux langues ne sont pas strictement isochroniques syllabiques, en raison de leurs manifestations différentes au niveau macrorhythmique⁸⁶ (Dauer, 1983) et de la large variation observée au niveau de la parole (bien qu'il existe dans certains cas une tendance vers l'isochronie). Comme le révèlent plusieurs études, c'est la récurrence plus marquée au niveau macrorhythmique qui engendre l'isochronie auditive dans les deux langues. Ainsi, ce sont les propriétés acoustiques au niveau macrorhythmique⁸⁷ qui permettent de mieux différencier les patrons rythmiques des deux langues :

⁸⁶ "It seems likely that rhythmic grouping takes place even in languages which have been called syllable-timed." (Dauer, 1983: 51).

⁸⁷ L'aspect mélodique a été aussi inclus dans le Tableau 1.18 comme particularité macrorhythmique car comme nous l'avons insisté à plusieurs reprises, la variation régulière de F0 fait partie aussi de la rythmicité d'une langue.

Tableau 1.18 : Comparaison des particularités macrorhythmiques entre le chinois mandarin et le français.

Particularité macrorhythmique	Le chinois mandarin	Le français
Niveau temporel	Syntagme inférieur : Allongement final fort + pause courte ; Syntagme supérieur : Allongement final léger + pause longue	Allongement final saillant : plus le niveau hiérarchique est haut, plus la durée d'allongement est longue (Shattuck-Hufnagel & Turk, 1996).
Niveau mélodique	Macrorhythmicité faible ; Proéminence marquée à tête ; Déclinaison et réhaussement de registre.	Macrorhythmicité forte ; Proéminence marquée à tête et à queue ; Récurrence de LHiLH*.

Le Tableau 1.18 et les contrastes rythmiques entre le français et le chinois mandarin que nous avons mis en évidence serviront de pistes pour nos analyses dans la partie expérimentale, et permettront de mieux cibler des sources éventuelles des difficultés rencontrées par les apprenants sinophones dans l'acquisition du rythme du français.

1.4 Sur l'acquisition de la prosodie du français par des apprenants sinophones

1.4.1 Pourquoi la prosodie est importante dans l'acquisition de la phonétique d'une L2

La prosodie d'une L2 et son acquisition revêtent une grande importance pour tout apprenant. De manière générale, l'acquisition de la prosodie d'une L2 se compose de deux aspects (Borrell & Salsignac, 2002 : 164) : au niveau de la perception, « *l'apprenant-auditeur a tout intérêt à prendre en compte les indices prosodiques pour accéder au sens* » ; au niveau de la production, « *l'apprenant-locuteur se doit de respecter la structure prosodique s'il veut se faire comprendre* ». Billières (2002) a justifié l'importance de la prosodie d'une L2 pour trois raisons, qui méritent d'être reprises pour mettre en évidence l'importance de son acquisition :

- a) Sur le plan linguistique, cela renvoie à de très nombreuses fonctions importantes que la prosodie joue dans la communication langagière (cf. section 1.1.3) ;
- b) Sur le plan de l'acquisition, l'intonation, voire la prosodie, est souvent l'élément le plus évident que l'apprenant perçoit dès le début d'apprentissage d'une nouvelle

langue étrangère : « *l'intonation* [mais aussi d'autres éléments prosodiques, entre autres, le rythme, la proéminence] *est la première chose que l'élève perçoit, avant même d'identifier des mots isolés ou des groupes de mots. La perception de la parole est un phénomène global : on perçoit par « blocs phoniques » délimités par des montées et des descentes de la voix, des accentuations propres à chaque langue, des pauses réelles ou subjectives. C'est à des étapes ultérieures de traitement que s'opèrent les opérations de reconnaissance d'unités linguistiques plus petites telles que les mots ou les phonèmes. L'intonation permet de découper le continuum sonore en « tranches » d'énoncé qui coïncident souvent avec des groupes de sens ; elle constitue également une trame sur laquelle se réalisent les phonèmes qui, sans elle, n'auraient pas d'existence... et contribue également à leur production correcte : la perception et l'articulation d'une unité sonore dépendent entre autres de sa place à l'intérieur du mouvement **rythmico-intonatif*** ». Malgré cela, c'est un aspect qui demeure « *ignoré de la majorité des enseignants de langues.* » (Billières, 2002 : 42).

c) Sur le plan psychologique, « *parler « sans accent » est toujours valorisant et stimulant, en même temps que cela aide énormément pour un accès au sens plus rapide et moins entaché de risques d'erreurs de compréhension. En outre, une intonation adéquate contribue énormément à l'intégration sociale.* » (ibid.).

Depuis les années 1980s, l'acquisition de la prosodie d'une L2 est considérée comme aussi importante, voire plus importante que l'acquisition des éléments segmentaux (Chun & Levis, 2020). Cependant, la recherche sur l'acquisition de la prosodie ne reçoit toujours pas une attention suffisante (Trouvain & Braun, 2020). Trouvain et Braun (2020) ont proposé l'une des revues de la littérature les plus récentes, qui synthétise de manière globale des études sur l'acquisition des prosodies de différentes L2. Ils ont examiné plusieurs aspects prosodiques, tels que les marquages prosodiques de la structure informationnelle, les modalités intonatives des phrases interrogatives, la segmentation prosodique, l'alignement du pic mélodique, le rythme, la pause, la fluidité, etc., en illustrant chaque sujet par des études existantes. Étant donné que la prosodie est un sujet à la fois vaste et complexe, Trouvain et Braun (2020) n'ont naturellement pu examiner qu'un nombre limité d'études (bien que cela représente déjà beaucoup). Ils ont également proposé quelques pistes pour

des recherches futures dans ce domaine⁸⁸. En plus de la revue de Trouvain et Braun (2020), Jongman et Tremblay (2020) ont réalisé une revue sur l'acquisition de la prosodie lexicale, portant spécifiquement sur les tons lexicaux et les accents lexicaux. Par ailleurs, Chun et Levis (2020) ont proposé une revue portant sur l'enseignement de la prosodie, dans laquelle beaucoup d'études sur l'acquisition des éléments prosodiques des diverses L2 ont été répertoriées.

Toutes ces études dessinent un aperçu de la situation actuelle et globale dans la recherche sur l'acquisition des prosodies des L2, sans aborder en détail pour l'acquisition d'un élément prosodique particulier dans une ou plusieurs langues spécifiques. Néanmoins, la recherche sur l'acquisition de tout élément prosodique est liée aux enjeux pédagogiques mis en relief par Borrell et Salsignac (2002) et par Billières (2002). Ces enjeux correspondent également aux objectifs de notre étude : tant sur le plan de production (« intelligibilité ») que sur celui de perception (« compréhensibilité »), nous nous interrogeons sur les manières dont des apprenants sinophones segmentent une chaîne prosodique et organisent et regroupent des unités linguistiques plus petites en syntagmes plus grands, pour les intégrer ensuite dans une structure rythmique et mélodique qui leur est peu familière. Dans les sous-sections suivantes, nous mettrons l'accent sur les travaux antérieurs qui ont spécifiquement traité de l'acquisition des éléments prosodiques du français par des apprenants sinophones.

1.4.2 Travaux antérieurs sur l'acquisition de la prosodie du français par des apprenants sinophones

Dans la section précédente, nous avons souligné l'importance de la prosodie d'une L2 dans la communication et sa place un peu marginale dans la recherche actuelle. Dans cette section, nous nous concentrerons sur les problèmes et les difficultés chez

⁸⁸ "For L2 sentence prosody research it remains a challenge to define what is 'correct' or 'acceptable' in the target variety on the one hand, and to widen our knowledge of target and source languages on the other. But there are more challenges for future research. Specifically, characteristics that are traditionally considered segmental, such as reduction, should be examined from a prosodic perspective. There is a need to investigate the interface between prosody and other linguistic levels, such as syntax in L2s (e.g. Zubizarreta and Nava 2011). It is also important to extend L2 prosody research to more natural situations, such as dialogues and other interactional behaviour (e.g. Ward and Gallardo 2017)... Finally, fluency and timing are not treated together with intonation- and pitch-related aspects in L2 teaching, L2 assessment, and L2 testing." (Trouvain & Braun, 2020: 617).

des apprenants sinophones au cours de l'acquisition des éléments prosodiques du français.

Après avoir fait une revue de la littérature sur ce sujet, nous avons remarqué que les études existantes sont relativement limitées et datées. Néanmoins, nous les synthétiserons à la fin de la revue afin d'établir un inventaire des problèmes et difficultés indiqués dans ces études et d'ouvrir des nouvelles perspectives pour notre recherche actuelle.

- **Lee (1987) : sur l'intonation**

À notre connaissance, la thèse de Lee (1987) est la première étude à avoir traité de l'acquisition de la prosodie du français (en particulier, l'intonation, dans son étude) par des apprenants sinophones taiwanais. D'après son travail, l'intonation réalisée par les apprenants était caractérisée comme une production de « *discontinuité* » et de « *monotonie* » (*ibid.* : 202). En se basant sur son analyse, ces caractéristiques se manifestent par des écarts de plusieurs aspects, que nous reportons ici dans le Tableau 1.19 :

Tableau 1.19 : Écarts de la prononciation sur le plan intonatif en français chez des apprenants sinophones (adaptés de Lee (1987)).

Aspects prosodiques	Écarts
Mélodique	Discontinuité mélodique : Un plateau mélodique qui neutralise la variation de hauteur, suivi d'une chute brutale quand le mot prosodique se trouve à la fin de la phrase.
Rythmique	- « <i>L'interaction entre les variations de durée et les structures syntaxiques est quasi inexistante : la syllabe allongée se situe aussi bien à la fin de l'unité syntaxique qu'à l'intérieur ou au début de celle-ci. Les mots grammaticaux sont souvent accompagnés d'un allongement de durée</i> » (<i>ibid.</i> : 200) ; - Discontinuité temporelle : pauses mal positionnées.
Hiérarchisation prosodique	- Absence de hiérarchie de degré de dépendance : « <i>Les Chinois ont la difficulté à réaliser une hiérarchie de degré de cohésion d'une part, et ils ont tendance à considérer un mot lexical comme une unité prosodique d'autre part</i> » ⁸⁹ (<i>ibid.</i> : 206).

⁸⁹ Une explication a été donnée par Lee (1987) et il est important de citer cette explication pour notre analyse dans la partie expérimentale : « ... si en français les variations de hauteur et de durée expriment des relations de dépendance différente au sein de l'énoncé, c'est parce qu'elles sont conditionnées par la structure syntaxique. Alors qu'en chinois, les variations de hauteur et de durée ne sont pas tributaires de la hiérarchie syntaxique. L'autonomie de la variation de hauteur et de durée par rapport à la syntaxe dans la langue explique selon nous en grande partie la raison pour laquelle les Chinois ont du mal à s'y retrouver du point de vue syntaxique dans le degré de cohésion à l'intérieur de l'énoncé. » (Lee, 1987 : 209).

- **Shen (1990) : sur la prosodie en général**

Shen (1990) a réalisé une étude portant sur la production et la perception du contour intonatif interrogatif du français chez des sinophones. Ses expériences ont montré que les sinophones novices en français (sans aucune connaissance préalable de la langue) avaient la capacité de distinguer le contour déclaratif de l'interrogatif (la même conclusion chez Faraco & Cavé, 2008), et que les apprenants étaient aussi capables de les produire, malgré une réalisation plus ou moins déviante.

Comparativement aux francophones natifs, ces écarts (Tableau 1.20) se manifestent principalement chez les apprenants par :

Tableau 1.20 : Écarts de la prononciation prosodique en français chez des apprenants sinophones (adaptés de Shen (1990)).

Aspects prosodiques	Écarts
Mélodique	- Une étendue de variation F0 plus large ; - Un registre mélodique plus élevé sur l'ensemble de l'énoncé ; - Un contour interrogatif déviant ⁹⁰ .
Rythmique	- Un patron rythmique plutôt « <i>staccato</i> » au lieu de « <i>legato</i> » comme chez les francophones natifs ; - Allonger la syllabe pénultième.

- **Landron (2007) ; Landron et al. (2016) : sur la prosodie en général**

Landron (2007) a d'abord fait son mémoire de master portant sur l'acquisition de la prosodie française par des locuteurs taiwanais. Dans son mémoire, plusieurs modalités prosodiques ont été examinées : la déclaration, l'interrogation totale vs partielle, l'impératif, la prosodie expressive, etc. Les écarts suivants ont été identifiés et résumés dans le Tableau 1.21 :

⁹⁰ "The most striking foreignness in uttering French rising terminal by Chinese speakers resides on the articulation of the final syllable; either they rise by hypercorrection the pitch of the final syllable higher than speakers of French L1; or they lower first by articulatory habits the pitch of the final syllable, then rise it finally, producing thus a falling-rising terminal instead of a simple rising terminal." (Shen, 1990: 130).

Tableau 1.21 : Écarts de la prononciation prosodique en français chez des apprenants sinophones (adaptés de Landron (2007)).

Aspects prosodiques	Écarts
Mélodique	<ul style="list-style-type: none"> - Ajout d'un ton haut pour les pentes descendantes ou d'un ton bas pour les pentes montantes sur l'avant-dernière syllabe : cela a tendance à rendre la chute ou la montée mélodique très abrupte ; - Une mélodie exagérée ou une mélodie monotone (moins graduelle), manque de marquage des groupes syntaxiques ; - Modalités interrogatives : contours montants bien réalisés mais sans distinction de types de question.
Rythmique	<ul style="list-style-type: none"> - Débit haché, avec la production de petites coupes entre les mots ou les syllabes : cela brise la fluidité ou la continuité du discours ; - Frontières prosodiques mal placées⁹¹.

Il convient de souligner que l'analyse de Landron (2007) concernant les écarts au niveau mélodique identifiés chez des apprenants sinophones, coïncide avec nos hypothèses (bien que celles-ci nécessitent une vérification dans la partie expérimentale). Les problèmes au niveau mélodique indiqués dans Landron (2007) comme la mélodie exagérée (donc grande étendue et contour abrupte) et l'absence de mouvement mélodique graduel en français seraient dus au transfert du mouvement des tons lexicaux vers le français : « *...l'existence des tons dans la prosodie de leur langue maternelle. Ces tons imposent un certain niveau de hauteur à chaque syllabe... et ce niveau de hauteur est fixé dans le lexique (du chinois mandarin). Ainsi, d'éventuelles variations générales de l'intonation (du français) restent en partie dépendantes de ces tons* » (Landron, 2007 : 169). Une expression utilisée dans Landron (2007) décrit de manière très concrète la réalisation habituelle des apprenants sinophones, qui est à la fois la plus fréquente et, selon nous, la plus problématique : ces apprenants ont tendance à « *syllaber l'énoncé* » (Landron, 2007 : 166). Lee (1987 : 222) a aussi donné une description similaire : « *les Chinois poussés par les habitudes constantes de leur langue maternelle, produisent en français une intonation de mot et non de phrase. Pour eux l'unité prosodique ne s'étend pas au-delà du mot lexical* ».

⁹¹ Un problème susceptible d'être étudié davantage mais peu révélé dans le domaine : « *Comme tout apprenant, ceux que nous avons étudiés semblent apprécier les groupes syntaxiques courts, de une [d'une (corrigé par nous)], deux ou trois syllabes. C'est d'ailleurs ce qui pose parfois le problème. Ils peuvent ainsi parfois couper des groupes en fin de mot avec un marquage intonatif simplement pour avoir des groupes plus courts. Il s'agit pour eux de groupes rythmiques avant d'être des unités significatives, au moins en apparence. Pourtant, la production de groupe syntaxique est liée à la compréhension de ce qui est dit.* » (Landron, 2007 : 161). Selon nous, des apprenants produisent souvent avec un débit relativement lent de sorte qu'ils doivent couper à des positions inappropriées pour la raison physiologique.

Dans Landron *et al.* (2016), plusieurs problèmes sur la prononciation du français en tant que L2 chez des sinophones ont été identifiés, la plupart se concentrant sur le niveau segmental. Pour la prosodie, seulement trois problèmes ont été indiqués, sans exemples ou explications détaillées : « *les montées de continuation, les allongements finaux des groupes prosodiques, les descentes finales des phrases déclaratives qui sont réalisées de façon très abrupte sur la ou les dernières syllabes d'une phrase* » (*ibid.* : 199). Landron *et al.* (2016) représente une contribution récente et significative par rapport aux études antérieures dans ce domaine. Pour enrichir cette recherche, il serait intéressant d'explorer davantage certains résultats. Cela pourrait inclure, d'une part une approche quantitative sur les indices acoustiques afin de mieux comprendre les aspects que les apprenants trouvent plus difficiles à maîtriser, et d'autre part, une approche qualitative pour comparer les configurations mélodiques entre locuteurs natifs et apprenants.

- Kockaert et Li (2008) : sur la prosodie en général

Le travail de Kockaert et Li (2008) présente une étude portant sur les écarts de prononciation aux niveaux segmental et suprasegmental et sur des pistes et exercices de remédiation. Même si une grande partie de ce travail était consacrée au niveau segmental, ils ont indiqué quelques écarts concernant la durée syllabique et l'intonation. De plus, les auteurs ont identifié certaines causes possibles de ces écarts, que nous avons récapitulées dans le Tableau 1.22 :

Tableau 1.22 : Écarts de la prononciation prosodique en français chez des apprenants sinophones (adaptés de Kockaert & Li (2008)).

Aspects prosodiques	Manifestations et Causes
Rythmique	<p>Manifestation :</p> <ul style="list-style-type: none"> - Les durées des syllabes inaccentuées et accentuées : les apprenants « <i>tendent à allonger en général la dernière syllabe du groupe rythmique... En revanche, ils tronquent les groupes rythmiques en des séquences de deux ou de trois syllabes, c'est-à-dire qu'ils auront consommé le souffle après avoir énoncé deux ou trois syllabes... La rythmicité cible est ainsi fractionnée, ce qui résulte en une répartition différente des unités rythmico-sémantiques du message.</i> » (Kockaert & Li, 2008 : 61). <p>Cause :</p> <ul style="list-style-type: none"> - La rythmicité du français diffère de la gestion temporelle des syllabes en mandarin, « <i>où les syllabes, toutes coiffées d'un ton, sont prononcées avec plus ou moins la même durée sur toute la phrase. En d'autres mots, le mandarin se caractérise par un isosyllabisme tandis que le français comprime ou dilate les syllabes dans le carcan isotemporel du groupe rythmique... Le français se caractérise ainsi d'un équilibre temporel entre groupes rythmiques et non entre syllabes.</i> » (ibid. : 60).
Mélodique	<p>Manifestations :</p> <ul style="list-style-type: none"> - « <i>L'énoncé assertif interlangue n'enregistre pas une courbe phrastique successivement montante et descendante sur les dernières syllabes [accentuées par défaut] des groupes prosodiques, typique des énoncés assertifs cible.</i> » (ibid. : 66). - « <i>Les courbes mélodiques ont un impact sur chaque syllabe individuelle.</i> » (ibid. : 89). <p>Cause :</p> <ul style="list-style-type: none"> - « <i>Le mandarin étant une langue à tons, il est évident que l'enveloppe mélodique s'avère toute différente, notamment parce que toutes les syllabes sont coiffées d'un tonème. Le mandarin présente une succession de syllabes ayant chacune leur ton, ce qui résulte en une fluctuation syllabique du fondamental.</i> » (ibid. : 66).

Les perspectives proposées pour les recherches futures par Kockaert et Li (2008) s'alignent parfaitement avec les objectifs de notre étude, qui vise à examiner l'organisation et la segmentation temporelles et mélodiques : « *par conséquent, nous jugeons utile de nous consacrer non seulement à la durée vocalique des syllabes accentuées à l'intérieur des groupes rythmiques mais surtout à celle des unités rythmiques mêmes. C'est ainsi que les apprenants pourront aboutir à une rythmicité énonciative accouplant le souffle au message à la française.* » (Kockaert & Li, 2008 : 64).

- **Gabriel et al. (2014) ; Xie (2014) ; Pillot-Loiseau et Xie (2018) : sur le rythme**

Concernant l'acquisition du rythme du français chez des apprenants sinophones, nous avons identifié trois études pertinentes. Dans Gabriel et al. (2014), les

productions rythmiques du français chez les apprenants sinophones et germanophones ont été examinées à l'aide des paramètres quantitatifs rythmiques %V et VarcoV. Ils ont conclu que la production des sinophones était plus proche de celle des francophones natifs que de celle des germanophones. La raison invoquée est que la nature isochronique syllabique du chinois mandarin (vs la nature isochronique accentuelle de l'allemand) constituait un transfert positif lors de l'acquisition du rythme du français. Les conclusions de Gabriel *et al.* (2014) doivent être indicatives, et ce pour plusieurs raisons : 1) tout d'abord, parmi la dizaine de paramètres rythmiques, seuls deux ont été examinés ; 2) comme nous l'avons souligné précédemment, un patron rythmique ne se manifeste pas uniquement par l'isochronisme ; la proéminence régulière, qui est également un aspect rythmique important en français, n'a pas été abordée dans cette étude ; 3) enfin, les profils linguistiques des apprenants de Gabriel *et al.* (2014) représentaient des cas assez spécifiques par rapport à la plupart des sinophones apprenant le français. Leur étude se déroulait dans un lycée en Allemagne, où les participants avaient des profils linguistiques variés : certains avaient le chinois mandarin comme langue maternelle et certains d'autres étaient bilingues en mandarin et en allemand depuis l'enfance. Cela diffère considérablement de la grande majorité des apprenants qui ont grandi dans un contexte homoglotte sinophone et qui n'ont que le chinois mandarin comme langue maternelle.

Xie (2014), un mémoire de Master 2, et Pillot-Loiseau et Xie (2018), qui présentent des résultats basés sur ce même mémoire lors d'un congrès, fournissent davantage de détails concernant l'acquisition du rythme, en examinant plus de paramètres quantitatifs. L'étude de Xie (2014) présente une analyse acoustique dans laquelle six paramètres (ΔC , %V, Varco ΔC , Varco ΔV , rPVI-C, nPVI-V) ont été calculés et comparés en se basant sur 10 phrases prononcées en français (par des francophones natifs et apprenants du niveau intermédiaire-avancé) et 10 phrases prononcées en chinois mandarin. En plus des données acoustiques, ce travail présente aussi une expérience perceptive AXB avec des stimuli, dont les informations segmentales ont été délexicalisées et la F0 neutralisée. Les conclusions tirées de ses analyses sont contradictoires : acoustiquement, parmi les six paramètres mesurés, seuls les paramètres %V et nPVI-V ont permis de différencier significativement le patron rythmique français prononcé par des natifs de celui prononcé par des apprenants

(autrement dit, parmi les six paramètres analysés, les valeurs des quatre (ΔC , $\text{Varco}\Delta C$, $\text{Varco}\Delta V$, $r\text{PVI-C}$) étaient plutôt comparables entre francophones natifs et apprenants). D'un point de vue perceptif, les auditeurs natifs du français ont tendance à considérer que le rythme du français produit par des apprenants est plus similaire à celui du chinois mandarin qu'à celui des francophones natifs. Xie (2014) remet donc en question l'utilisation des paramètres quantitatifs rythmiques dans la recherche sur l'acquisition du rythme en langue seconde.

Une des limites de ce travail est que l'analyse acoustique sur la production rythmique dans l'interlangue n'a pas pris en considération la proéminence. De même, la F_0 , qui est un élément important dans la perception rythmique, a été neutralisée dans l'expérience perceptive. Cela soulève des questions sur les résultats qui seraient obtenus si la F_0 serait aussi intégrée dans cette expérience perceptive sur le rythme, tout en excluant les informations segmentales.

- **Xi et Pillot-Loiseau (2019) ; Xi *et al.* (2020) : sur l'intonation et la segmentation prosodique**

En analysant un même corpus, nos études pilotes (Xi & Pillot-Loiseau, 2019 ; Xi *et al.*, 2020) ont examiné la réalisation des paramètres temporels et mélodiques aux frontières prosodiques. Les résultats obtenus ont non seulement confirmé des conclusions tirées d'études précédentes, mais aussi apporté de nouveaux résultats (Tableau 1.23) :

Tableau 1.23 : Écarts de la prononciation prosodique en français chez des apprenants sinophones selon Xi et Pillot-Loiseau (2019) et Xi *et al.* (2020).

Aspects prosodiques	Écarts
Mélodique	<ul style="list-style-type: none"> - Un même contour montant abrupte sans distinction des modalités intonatives différentes (continuation majeure, question polaire avec ou sans marque syntaxique) (<i>idem</i> dans Landron, 2007) ; - Mélodie monotone, suivie par une variation de F0 exagérée sur la dernière syllabe avant la frontière prosodique (<i>idem</i> dans Lee, 1987 ; Landron, 2007) ; - Contours descendants au milieu de la phrase : identifiés notamment chez certains apprenants du niveau débutant-intermédiaire⁹².
Rythmique	<ul style="list-style-type: none"> - Allongement final de la dernière syllabe avant la frontière prosodique bien réalisé (<i>idem</i> dans Kockaert & Li, 2008) ; - Pausés mal positionnées qui fractionnent un groupe rythmique (<i>idem</i> dans Lee, 1987).

1.4.3 Synthèse

Après avoir revu les études limitées dans la littérature sur l'acquisition des éléments prosodiques du français par des apprenants sinophones, nous avons constaté que des problèmes similaires et récurrents apparaissent dans différentes études. Selon notre analyse, ces problèmes d'acquisition se répartissent principalement en deux catégories : ceux liés à l'organisation et au regroupement des éléments prosodiques selon une unité sémantico-syntaxique correcte, et ceux concernant la segmentation et le marquage des frontières prosodiques.

Au niveau de l'organisation : 1) le problème le plus évident, selon nous, est lié au fait de « *syllaber l'énoncé* » (Landron, 2007), ce qui se manifeste par le fait que certains apprenants « *produisent en français une intonation de mot et non de phrase* » (Lee, 1987 : 222). Autrement dit, la production prosodique du français chez les apprenants se caractérise par une parole hachée, et ils manquent de la compétence pour regrouper les syllabes isolées en syntagmes prosodiques plus larges ; 2) que ce soit en raison de contours mélodiques plats (mélodie monotone) ou descendants au milieu de la phrase qu'une étendue, ou d'une variation de la F0 trop large, tout cela témoigne d'un manque de compétence dans l'organisation des unités phonétiques selon un patron mélodique attendu, tel que le prototype « LHiLH* ».

⁹² Nous supposons que cela soit dû au transfert négatif du chinois mandarin : “*Cao's study revealed that declination inside a prosodic phrase and resetting of pitch register at the beginning of the next phrase are particularities of the Chinese Mandarin rhythmic pattern, which play a demarcative role. Yet in French, the NS [native speakers] use a F0 rise or a F0 peak to express continuation and a F0 fall for termination*” (Xi *et al.*, 2020: 898).

En ce qui concerne le marquage des frontières prosodiques, à l'aide d'indices acoustiques attendus, plusieurs problèmes récurrents se manifestent : 1) au niveau mélodique, l'accentuation de la syllabe pénultième est souvent accompagnée d'une montée ou d'une descente abrupte sur la dernière syllabe ; 2) au niveau temporel, des pauses sont insérées à l'intérieur d'un groupe rythmique, probablement en raison de la méconnaissance des unités de sens.

Avec la présente étude, qui se concentre sur l'organisation et la segmentation prosodiques, nous visons à apporter des nouvelles perspectives sur les problèmes et les difficultés rencontrés par les apprenants sinophones tant au niveau mélodique qu'au niveau temporel.

1.5 L'électroencéphalographie (EEG) et son application aux études de la prosodie et de l'acquisition langagière

Dans cette ultime section de la partie théorique, nous explorerons la technique de l'électroencéphalographie (EEG). Nous commencerons par une brève présentation de ce qu'est l'EEG, en soulignant ses avantages et inconvénients par rapport à d'autres techniques de neuroimagerie. Ensuite, nous procéderons par une revue de la littérature, en nous concentrant notamment sur l'application de l'EEG à l'étude des phénomènes prosodiques, ainsi que sur les conclusions et les découvertes récentes dans ce domaine. Enfin, nous examinerons l'utilisation de l'EEG dans l'étude de l'acquisition d'une langue seconde (L2), en mettant en évidence ses bénéfices dans la recherche sur l'acquisition du langage.

1.5.1 L'électroencéphalographie (EEG) et le potentiel évoqué (ERP)

Avant d'explorer en détail l'utilisation de l'EEG pour la recherche sur la prosodie et l'acquisition langagière, ainsi que les composantes les plus utilisées à cet effet, nous ferons d'abord un bref rappel de la méthode EEG. Nous discuterons également de ses avantages et inconvénients dans le contexte de la recherche sur les sciences du langage et notamment en acquisition langagière.

Le cerveau humain contient environ 85 milliards de neurones, dont le rôle est de recevoir, traiter et transmettre des informations électrochimiques. Un neurone est composé du corps cellulaire, qui contient le noyau, de l'axone, qui est un prolongement qui émet et conduit le signal, et des dendrites, qui agissent comme

récepteurs (Purves *et al.*, 2018). Les neurones communiquent entre eux par le biais des synapses, en utilisant des neurotransmetteurs (ou neuromédiateurs) émis par des synapses (Figure 1.19). Lors de la libération des neurotransmetteurs par les neurones postsynaptiques, les interconnexions sont établies et les informations sont transmises, ce qui induit des variations de potentiel électrique faible, de l'ordre du microvolt (μV).

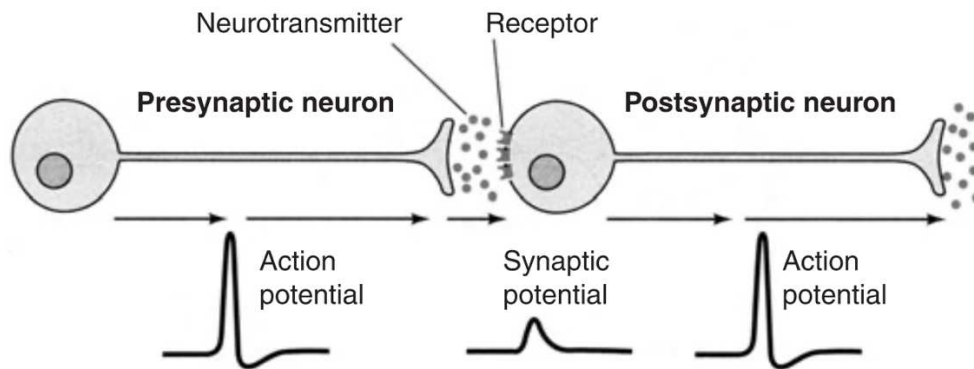


Figure 1.19 : Communication entre neurones conduisant au potentiel électrique (O'Shea, 2005 : 30).

Lors de la communication entre neurones, la technique EEG permet de détecter et d'enregistrer, à travers des électrodes posées sur le scalp, le faible courant électrique qui reflète l'activité cérébrale. Ce signal est ensuite amplifié pour permettre une analyse ultérieure.

Plusieurs études (Kutas & Van Petten, 1994 ; Kaan, 2007 ; Männel, 2008 ; Cohen, 2011 ; Rommers & Federmeier, 2018 ; Leckey & Federmeier, 2019 ; Payne *et al.*, 2020) ont mis en évidence les avantages particuliers de la technique EEG, ainsi que ses inconvénients. Nous mettons en avant ici deux avantages majeurs : 1) Tout d'abord, l'EEG offre une résolution temporelle très élevée, qui permet une précision temporelle de l'ordre de la milliseconde pour l'activité mesurée ; 2) Ensuite, l'expérience EEG peut être réalisée dans une condition passive, sans forcément nécessiter de réponse comportementale. Autrement dit, les participants ne sont pas tenus de réagir aux stimuli, ce qui permet de mesurer le traitement en temps réel (*online processing*) d'une activité langagière sans interruption. Cela est particulièrement avantageux pour étudier l'activité cérébrale chez de très jeunes enfants, qui n'ont pas encore la capacité de réaliser une tâche comportementale, ou chez des participants présentant des pathologies qui les empêchent de les réaliser.

Par ailleurs, la méthode EEG permet de révéler des aspects importants que les méthodes comportementales ne peuvent détecter (Mueller, 2005), notamment dans la recherche sur l'acquisition d'une L2 chez les grands débutants dès le début du processus d'acquisition⁹³.

La méthode EEG présente évidemment des limites. D'une part, sa résolution spatiale est très faible : l'EEG ne fournit qu'une distribution topographique pour connaître la zone approximative d'un traitement neurocognitif donné. Pour obtenir une précision spatiale élevée, il est nécessaire de la combiner avec d'autres méthodes d'imagerie neurologique, comme la fMRI ou la MEG. D'autre part, le signal brut de l'EEG contient beaucoup de bruits (Rommers & Federmeier, 2018), souvent causés par des mouvements oculaires ou corporels pendant l'enregistrement. Ainsi, contrôler un ratio approprié de « signal vs bruit » représente un autre défi lors de la réalisation d'une expérience électroencéphalographique (cf. voir les quatre sources principales de bruit dans Luck (2005)).

Au niveau terminologique, il est important de distinguer « EEG » de « ERP ». Le terme « EEG » peut désigner l'appareil d'électroencéphalographie, mais aussi le signal brut qui est enregistré. Dans une expérience typique utilisant l'EEG, un grand nombre de stimuli est généralement nécessaire pour différencier les conditions (entre 25-60 stimuli dans une expérience EEG sur le langage, selon Leckey & Federmeier, 2019). Cela permet d'extraire le signal cérébral lié à l'événement en question (*ERP effects*⁹⁴, Rommers & Federmeier, 2018), par le biais de la moyenne et de la comparaison de tous les signaux bruts enregistrés des stimuli dans une fenêtre temporelle prédéfinie (Kutas & Van Petten, 1994 ; Kaan, 2007 ; Männel, 2008 ;

⁹³ Un avantage de l'EEG dans la recherche sur l'acquisition d'une L2 chez l'adulte a été mis en évidence dans de nombreuses études, y compris celle de van Hell et Tokowicz (2010): "*ERP studies have also allowed us to conclude that learning has taken place much earlier than would have been thought based on behavioural performance alone. This is because in several studies behavioural performance was at chance when ERP responses demonstrated significant sensitivity to violations of syntax.*" (van Hell & Tokowicz, 2010: 71). Selon Mueller (2005), "*differences in behavioural measures, such as reaction times, are not necessarily the result of the involvement of different neuronal structures, even if they show qualitatively different patterns. Neurophysiological measures can add valuable information about timing and degree of activation of neural networks and therefore provide physiological correlates of behavioural differences. Therefore the question of convergence or divergence of L1 and L2 processing mechanisms is difficult to solve solely on the basis of behavioural data.*" (Mueller, 2005: 153).

⁹⁴ "*Although some components tend to reliably be associated with peaks in the ERP, an individual waveform typically does not allow the researcher to draw conclusions about cognitive processes. Instead, as with most methods, the focus is on differences between conditions, or "ERP effects". An ERP effect is a modulation of an ERP component, or just the difference between two conditions, which, in well-designed studies, isolates particular subprocesses of interest.*" (Rommers & Federmeier, 2018: 249).

Meulman *et al.*, 2016). Le signal ainsi extrait est appelé ERP (*event-related potential*), ou en français « potentiel évoqué », ce qui signifie littéralement un potentiel cérébral évoqué par un événement ou une stimulation extérieure donnée (Kutas & Van Petten, 1994 ; Morgan-Short *et al.*, 2010).

1.5.2 Composantes ERP liées au traitement langagier et prosodique

Selon Steinhauer et Connolly (2008), l'EEG a commencé à être largement utilisée pour la recherche sur les sciences du langage depuis la découverte de la composante N400 par Kutas et Hillyard (1980). Au cours du 20^{ème} siècle, une série de composantes ERP concernant le traitement des différents aspects langagiers ont été découvertes, puis réexaminées et répliquées dans différentes langues. Elles sont désormais considérées comme des paradigmes robustes et fiables pour étudier des phénomènes langagiers. La littérature abonde en revues qui synthétisent les composantes ERP pour l'étude du langage (cf. Kutas & Van Petten, 1994 ; Osterhout & Holcomb, 1995 ; Friederici, 2004 ; Kaan, 2007 ; Männel, 2008 ; Steinhauer & Connolly, 2008 ; Swaab *et al.*, 2012 ; Chapitre 3 de Luck (2014) ; Urbach & Kutas, 2018 ; Leckey & Federmeier, 2019, Payne *et al.*, 2020 pour une revue ; cf. Mueller, 2005 ; Moreno *et al.*, 2008 pour les composantes utilisées dans l'acquisition d'une langue seconde). Dans cette section, nous nous concentrons particulièrement sur les composantes ERP liées au traitement prosodique.

1.5.2.1 *Closure Positive Shift* (CPS) : découverte, évolution et caractéristiques

La composante CPS (*closure positive shift*) a été découverte pour la première fois par Steinhauer *et al.* (1999), qui ont mis en évidence une réponse neurocognitive aux frontières des syntagmes intonatifs en allemand. Cette composante CPS est un potentiel positif, apparaissant bilatéralement dans la zone centro-pariétale, avec une distribution prééminente aux électrodes de la ligne centrale lorsque l'auditeur traite une frontière prosodique⁹⁵ (Steinhauer, 2003). Dans des études ultérieures, la CPS a été démontrée comme une réponse fiable et pure à l'onset des frontières prosodiques, indépendante de la modalité d'input (concernant des onsets qui

⁹⁵ "The component's onset at speech boundaries was likely triggered by the first available acoustic boundary marker, possibly pre-final constituent lengthening." (Steinhauer, 2003: 162).

déclenchent la CPS, cf. voir la discussion dans Bögels *et al.*, 2011). Ainsi, la CPS peut être suscitée par la virgule lors de la lecture silencieuse (CPS visuelle) (Steinhauer & Friederici, 2001 ; Steinhauer, 2003 ; Drury *et al.*, 2016 ; mais cf. Kerkhofs *et al.*, 2008 pour un point de vue différent) et dans la parole délexicalisée ou synthétisée (Steinhauer, 2003 ; Pannekamp *et al.*, 2005 ; Honbolygó *et al.*, 2016). La composante CPS a également été observée lors de la segmentation des séquences continues dans d'autres formes, telles que la musique (Knösche *et al.*, 2005 ; Neuhaus *et al.*, 2006 ; Glushko *et al.*, 2016) et le mouvement corporel (Hilton *et al.*, 2019, 2021). Ces études montrent que lorsque l'auditeur ou le lecteur segmente et structure une suite de signaux, qu'ils soient auditifs ou visuels, langagiers ou musicaux, la CPS tend à être évoquée comme une réponse fiable et robuste.

Par ailleurs, la CPS semble être une réponse graduelle en fonction des niveaux hiérarchiques des frontières prosodiques. Autrement dit, les caractéristiques de la CPS deviennent plus saillantes (e.g., une amplitude plus grande ou une latence plus précoce) quand des indices acoustiques à la position d'une frontière prosodique sont plus saillants ou plus évidents⁹⁶ (Bögels *et al.*, 2011 ; Pauker, 2013 ; Glushko *et al.*, 2016). Cependant, à ce jour, les indices acoustiques précis (allongement final ou ton de frontière⁹⁷) qui déclenchent la CPS à la frontière prosodique restent encore flous (Bögels *et al.*, 2011 ; cf. voir aussi l'étude pilote de Ganga *et al.*, 2024 qui a examiné la relation entre l'effet CPS et chacun des trois indices prosodiques en néerlandais).

Après avoir été démontrée pour l'allemand par Steinhauer *et al.* (1999) (voir aussi Isel *et al.*, 2005), la CPS a été répliquée dans différentes langues : en chinois mandarin (Li *et al.*, 2008 ; Li & Yang, 2009 ; Liu *et al.*, 2010 pour la CPS déclenchée par la ponctuation dans la lecture silencieuse ; Li *et al.*, 2019 pour la CPS déclenchée par les frontières prosodiques précoce vs tardive dans des phrases localement

⁹⁶ "Steinhauer (2003) argues that the CPS amplitude depends on the amount of activation of phonological representations in the brain. He reports a smaller CPS for covert prosodic phrasing... Kerkhofs *et al.* (2008) propose that the size of the CPS depends on the salience of prosodic boundary markers." (Bögels *et al.*, 2011: 430); "The prosodic CPS seems to be modulated by the strength of boundary markers in a graded manner (rather than being an all-or-none response)... For example, it has been shown that boundaries with stronger pre-final lengthening and longer pauses elicit larger CPS amplitudes." (Glushko *et al.*, 2016: 2).

⁹⁷ "However, it remains unclear which aspects of the PB [prosodic breaks] are necessary for the elicitation of a CPS. Since the pause is apparently not necessary, the prefinal lengthening and/or boundary tone remain viable candidates for bringing about the CPS." (Bögels *et al.*, 2011: 428).

ambiguës), en japonais (Wolff *et al.*, 2008), en anglais (Peter *et al.*, 2014), en néerlandais (Bögels *et al.*, 2010), en coréen (Hwang & Steinhauer, 2011) ou en portugais (Batista *et al.*, 2023).

Toutefois, à ce jour, ce résultat neurophysiologique est difficilement généralisable à toutes les langues, car certaines, comme le français, ont été peu étudiées. À notre connaissance, il n'existe qu'une seule étude EEG (Gilbert *et al.*, 2023) portant sur le français canadien, qui a répliqué une CPS en réponse au traitement de la frontière prosodique par des auditeurs francophones.

Étant non seulement une composante utilisée de manière isolée pour étudier le traitement neurocognitif pur de la frontière prosodique, la CPS a été également examinée en combinaison avec d'autres composantes ERP linguistiques, notamment N400 et P600, afin d'étudier l'interaction entre le traitement prosodique et celui de sémantico-syntaxique (e.g., Bögels *et al.*, 2010 ; Itzhak *et al.*, 2010 ; Steinhauer *et al.*, 2010 ; Pauker *et al.*, 2011 ; Roll & Horne, 2011). De plus, elle est utilisée pour étudier l'intégration des diverses informations linguistiques dans la perception et la compréhension de la parole (Kerkhofs *et al.*, 2007 ; Honbolygó *et al.*, 2016).

1.5.2.2 D'autres composantes ERP pour étudier la prosodie de la parole

En plus de la composante CPS, qui permet d'examiner le traitement neurocognitif des frontières prosodiques, d'autres composantes sont utilisées pour étudier d'autres aspects prosodiques : entre autres MMN, N325 et N100.

- Les composantes MMN et N325 :

Ces deux composantes ont été principalement utilisées pour évaluer la perception de la structure métrique de la langue. Avec une polarité négative, la composante MMN apparaît entre 100-250 ms après l'onset du stimulus, avec une distribution marquée aux électrodes des lignes frontale et centrale. Étant sensible au changement d'indice acoustique (comme la F0, la durée, l'intensité ou le timbre, Näätänen & Kreegipuu, 2012), la MMN est en général déclenchée dans le cadre d'une étude « oddball », où deux séries de stimuli répétitifs, et un seul indice acoustique varie entre les deux séries (standards vs déviants). Dans l'étude du langage (cf. Näätänen *et al.*, 2007 pour une revue de littérature), la MMN est souvent utilisée pour tester la discrimination phonétique (Männel, 2008, Winkler *et al.*, 1999) et prosodique. Elle est employée pour examiner divers aspects prosodiques, tels que

l'accentuation, par exemple, en finnois (Ylinen *et al.*, 2009), en français (Aguilera *et al.*, 2014 ; te Rietmolen *et al.*, 2019a) et en d'autres langues (e.g., Honbolygó & Csépe, 2013 ; Honbolygó *et al.*, 2020), ainsi que pour le ton lexical (Kaan *et al.*, 2008). Ces études concernent tant la langue maternelle que l'acquisition d'une L2. La N325, découverte pour la première fois en anglais et en néerlandais (Böcker *et al.*, 1999), est une autre composante pour examiner le traitement de l'accent métrique. Bien que les deux composantes puissent être utilisées pour tester la perception de l'accentuation, la différence entre la N325 et la MMN réside dans le fait que l'attention du participant est requise ou non pendant l'expérience (*ibid.*). Dans la littérature, plusieurs études ont été réalisées en utilisant la composante N325, y compris celle de te Rietmolen *et al.* (2016) pour le français.

- *La composante N100 :*

La N100, une composante neurocognitive sensible aux propriétés physiques (qu'elles soient auditives ou visuelles) est un paradigme récemment utilisé pour étudier la segmentation lexicale ou la différence entre syllabe accentuée et syllabe inaccentuée (Steinhauer & Connolly, 2008). Des études notables incluent celle de Sanders et Neville (2003a), qui explore la relation entre la segmentation de la parole enchaînée et l'intégration d'informations lexicale et syntaxique en anglais ; ainsi que celle de Zhang et Zhang (2019) sur l'incongruence métrique et syntaxique en chinois mandarin.

D'autres études électroencéphalographiques sur le traitement prosodique, mais en recourant à d'autres composantes non spécifiquement liées à la prosodie (notamment P600 et N400 pour examiner l'incongruence entre prosodie et morphosyntaxe), sont également disponibles. Par exemple, Astésano *et al.* (2004) ont utilisé les composantes P800 (une composante sensible à la violation de la modalité déclarative vs interrogative), et N400 pour examiner l'intégration des informations sémantique et prosodique en français. Magne *et al.* (2007) et te Rietmolen *et al.* (2017, 2019b) ont utilisé la composante N400 pour étudier l'influence de la structure métrique congruente vs incongruente du lexique sur l'accès au sens. Llanos *et al.* (2021) ont examiné une nouvelle composante, ARP (*accent related potential*), qui est une marque neurocognitive alignée avec la position de l'accent mélodique du stimulus. Kotz et Paulmann (2007) ont découvert la composante PEP (*prosodic expectancy positivity*), qui est liée avec la violation de la

prosodie émotionnelle. L'étude de Paulmann (2023) est une synthèse récente sur le mécanisme neurophysiologique de la prosodie, dans laquelle nous pourrions trouver plus d'études que nous n'avons pas mentionnées ici.

Bien que plusieurs composantes ERP liées au traitement des différents aspects prosodiques soient disponibles, dans la partie expérimentale de cette présente étude, nous examinerons principalement la composante CPS pour étudier le traitement neurocognitif des frontières prosodiques en français.

1.5.3 Études sur l'acquisition d'une L2 avec l'EEG

Dans cette section, nous nous concentrerons sur l'utilisation de l'EEG dans la recherche sur l'acquisition d'une L2. Après une revue de la littérature, nous présenterons le modèle ERP de convergence L1-L2, proposé par Steinhauer *et al.* (2009). Ce modèle mettra en évidence les avantages spécifiques de l'EEG pour étudier le processus d'acquisition d'une L2.

Comme indiqué dans la section 1.5.1, l'EEG offre non seulement une très haute résolution temporelle, permettant de déterminer le moment précis (*when*) d'une activité cérébrale enregistrée, mais elle peut également révéler comment (*how*) un traitement neurocognitif donné se déroule. Cela est possible grâce aux caractéristiques des composantes, sans nécessiter que le participant réalise une tâche comportementale spécifique (Osterhout *et al.*, 2004 ; Steinhauer *et al.*, 2009 ; Leckey & Federmeier, 2019⁹⁸). Grâce à ces capacités, l'EEG a été appliquée pendant des décennies à l'étude de l'acquisition de la langue maternelle ou première chez les très jeunes enfants. Dans les travaux de Friederici (2005), ainsi que Steinhauer et Connolly (2008), des expériences EEG ont permis d'établir un ordre d'émergence des composantes ERP (Figure 1.20). Cela permet de déterminer à quel moment précis un enfant acquiert certains aspects linguistiques, apportant ainsi de nouvelles compréhensions sur le processus d'acquisition de la L1. Prenons l'exemple de

⁹⁸ "In conclusion, ERPs are a method particularly well-suited to the study of language. They provide a continuous measure, and their excellent temporal resolution allows for the assessment of all aspects of linguistic information down to the millisecond level. The lack of need for an overt task allows measures to be taken in a more ecologically valid way, and also allows researchers to extend their work to populations, including infants and some clinical populations, who cannot follow instructions or give behavioral responses. The ability to make measurements without an overt task also means that ERPs can shed light on aspects of processing that cannot be tapped using behavioral methods – for example, revealing evidence for implicit learning of language regularities in participants acquiring a second language (McLaughlin *et al.*, 2004; Tokowicz & MacWhinney, 2005)." (Leckey & Federmeier, 2019: 65).

l'acquisition de la prosodie de la L1 : les études de Männel et Friederici (2009, 2011) et Männel *et al.* (2013) ont exploré le moment où la composante CPS émerge et se développe pleinement, ainsi que les indices acoustiques privilégiés chez les très jeunes enfants allemands lors de la perception de la prosodie de leur langue maternelle.

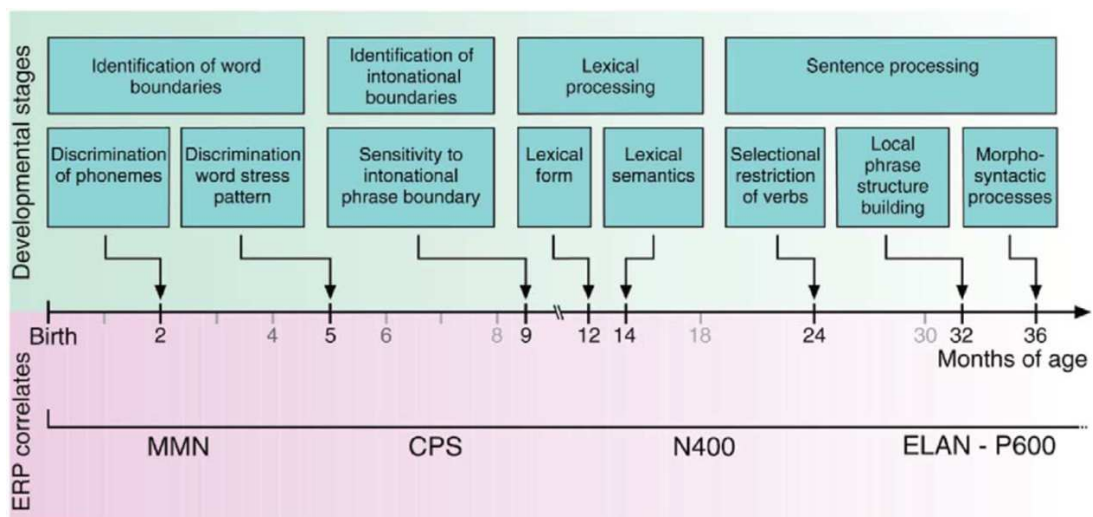


Figure 1.20 : Ordre d'émergence des composantes ERP au cours d'acquisition de la langue maternelle (Friederici, 2005 : 483).

Bien que l'EEG ne soit pas aussi performante que d'autres méthodes de neuroimagerie, telles que la fMRI ou la MEG, en terme de résolution spatiale, elle peut néanmoins fournir une estimation approximative du signal cérébral (*where*), avec la méthode LORETA (*Low Resolution Electromagnetic Tomography*) (Pascual-Marqui *et al.*, 1994, 2002 ; Pascual-Marqui, 1999). Grâce à cette méthode, certaines études ont essayé de décrire le changement spatial des composantes afin d'appréhender l'évolution des régions cérébrales impliquées dans le traitement neurocognitif de certains aspects langagiers au cours d'acquisition d'une L2 (e.g., Osterhout *et al.*, 2008).

Depuis le travail d'Ardal *et al.* (1990), considéré comme la première étude électroencéphalographique sur l'acquisition d'une L2 (Morgan-Short & Tanner, 2014), l'application EEG dans ce domaine a connu un essor considérable au cours des trente dernières années. Les potentiels méthodologiques de l'EEG pour appréhender les phénomènes d'acquisition langagière restent encore largement inexploités (Morgan-Short, 2014). Dans la littérature, une grande quantité d'études

sur l'acquisition de la langue maternelle utilisant la technique EEG est disponible (e.g., Friederici & Thierry, 2008 ; Männel & Friederici, 2008).

À l'instar des études sur l'acquisition de la L1 menées avec l'EEG, nous pensons que les avantages de cette technique peuvent être pleinement transposés à la recherche sur l'acquisition d'une L2. L'EEG se révèle être un outil assez performant pour étudier le processus d'acquisition et le traitement langagier chez l'individu (e.g., Osterhout *et al.*, 2004⁹⁹ ; Tanner *et al.*, 2013 ; Tanner & Van Hell, 2014). Elle permet notamment d'examiner des performances particulières via des composantes spécifiques tout au long de l'acquisition (Royle & Steinhauer, 2014, Di Liberto *et al.*, 2021).

En effet, Steinhauer *et al.* (2009) ont indiqué que les propriétés des composantes ERP chez les apprenants tardifs d'une L2 peuvent refléter le changement de performances et/ou le processus d'acquisition, allant du niveau débutant vers le niveau avancé ou natif (e.g., études de Weber-Fox & Neville (1996) ; White *et al.*, (2012)). Selon ces études, il a été démontré que certains aspects langagiers, tels que la sémantique, la morphosyntaxe ou la grammaire, peuvent être complètement acquis, et même atteindre le niveau natif chez des apprenants tardifs. Cela se manifeste par des similarités neurocognitives des composantes ERP entre les locuteurs natifs et les apprenants¹⁰⁰ (Clahsen & Felser, 2006).

L'apport de la méthode EEG, éventuellement combinée avec d'autres techniques d'imagerie cérébrale, remet en question l'hypothèse de la période critique dans la recherche sur l'acquisition du langage. En se basant sur des découvertes neurophysiologiques récentes, un modèle ERP de convergence L1-L2 a été établi par Steinhauer *et al.* (2009) pour mieux éclairer les phases d'acquisition d'une L2. Ce modèle consistait initialement à comparer les latences, les amplitudes et les distributions topographiques des composantes ERP liées aux traitements morphosyntaxiques entre les locuteurs natifs et les apprenants de différents niveaux (cf. Tableau 1.24).

⁹⁹ "In principle, ERPs provide the temporal resolution and sensitivity needed to reveal how (and potentially where) the L2 is processed, how similar this processing is to L1 processing, and how L2 processing changes as a function of increasing L2 experience." (Osterhout *et al.*, 2004: 287).

¹⁰⁰ "... At higher levels of proficiency, L2 learners' ERPs consist of a biphasic sequence of components that is indistinguishable from that in native speakers." (Steinhauer *et al.*, 2009: 23).

Tableau 1.24 : Différentes phases d'acquisition morphosyntaxique révélées par composantes ERP chez les apprenants tardifs (Steinhauer *et al.*, 2009).

Table 2 ERPs reflecting different stages of morpho-syntactic proficiency in late L2

Stage	Proficiency level	ERP pattern	Cognitive processes	References
1	Novice	No difference	Indifferent perception; performance at or near chance level	Friederici, 2002; Mueller, <i>et al.</i> , 2005 (untrained controls); Rossi <i>et al.</i> , 2006 (2 conditions)
2	Very low proficiency	N400 or right-lateralized/posterior negativities	Difficulties during lexical access and integration; compensatory processing strategy, likely relying on semantic plausibility and pragmatics	Weber-Fox and Neville, 1996; Hahne, 2001; Osterhout <i>et al.</i> , 2006 (1 month); White <i>et al.</i> , 2007 (regarding Broca's aphasics, see also Hagoort <i>et al.</i> , 2003b)
3	Low to intermediate	Small/delayed P600 (possibly preceded by N400s)	Beginning grammaticalization/proceduralization	Osterhout <i>et al.</i> , 2006 (4 months); Steinhauer <i>et al.</i> , 2006;
4	Intermediate	Larger/earlier P600	Late structural reanalysis/repair approaches native-like mechanisms	Osterhout <i>et al.</i> , 2006 (9 months);
5	Intermediate to high/near native-like	Bilateral AN + P600	Near-native processing; early automatic + late controlled processing	Friederici 2002; Bowden <i>et al.</i> , 2007; Morgan-Short <i>et al.</i> , 2007
6	very high/native-like	Lateralized LAN + P600	Native-like processing; early automatic + late controlled processing	Pakulak and Neville, 2004; Steinhauer <i>et al.</i> , 2006; Rossi <i>et al.</i> , 2006; Bowden, 2007 <i>et al.</i> ; Morgan-Short, 2007 <i>et al.</i>

Après une revue de la littérature détaillée, nous avons remarqué que très peu d'études EEG n'ont été faites jusqu'à présent sur l'acquisition de la prosodie d'une L2. Concernant son application à la recherche sur l'acquisition d'une L2, la plupart des références se concentrent actuellement sur les aspects morphosyntaxiques ou grammaticaux (cf. Osterhout *et al.*, 2004, 2006 ; Moreno *et al.*, 2008 ; Sabourin *et al.*, 2013 ; Morgan-Short, 2014 ; Steinhauer, 2014a ; Morgan-Short *et al.*, 2015 pour des revues de la littérature sur tous les aspects langagiers, bien que celles-ci soient incomplètes en raison de l'essor dans le domaine pendant ces dernières années). De manière générale, quel que soit l'aspect langagier examiné, les analyses de la plupart des études EEG sur l'acquisition d'une L2 se sont concentrées sur les trois propriétés fondamentales des composantes ERP : la latence, l'amplitude et la distribution topographique. Parmi ces études, nous avons constaté des conclusions similaires :

- 1) d'une part, la composante ERP apparaît chez les apprenants dès le niveau (grand) débutant (Morgan-Short, 2014), et la latence est l'indice le plus

évident pour différencier les traitements chez les apprenants de niveaux différents (débutant vs intermédiaire vs avancé) et les locuteurs natifs (e.g., Weber-Fox & Neville, 1996 ; McLaughlin *et al.*, 2010¹⁰¹, pour une revue de la littérature sur les manifestations des composantes ERP au cours de l'acquisition grammaticale chez les apprenants d'une L2, mais voir aussi Meulman *et al.*, 2014, qui ont montré le contraire). Il semble que chez les apprenants, notamment ceux du niveau débutant, la présence ou l'absence d'un effet ERP comparable à celui des natifs dépend d'une série de facteurs¹⁰². Les deux les plus importants sont la similarité L1 vs L2 (cette idée a été révélée dans de nombreuses études, e.g., Osterhout *et al.*, 2004 ; Ojima *et al.*, 2005 ; Tokowicz & MacWhinney, 2005 ; Chen *et al.*, 2007 ; Sabourin & Stowe, 2008 ; van Hell & Tokowicz, 2010 ; Foucart & Frenck-Mestre, 2011 ; Tolentino & Tokowicz, 2011¹⁰³ ; Zawiszewski *et al.*, 2011 ; Bowden *et al.*, 2013 ; Meulman *et al.*, 2014¹⁰⁴ ; Díaz *et al.*, 2016, Mickan & Lemhöfer, 2020, etc. Voir aussi une discussion détaillée sur la relation entre similarité L1-L2 et effet ERP dans Morgan-Short, 2014¹⁰⁵) et le niveau de compétence en L2 (Geyer *et*

¹⁰¹ "The most striking finding across all of the studies reviewed here was the clear and consistent change in learners' brain responses to violations of L2 grammatical rules across the first year of instruction." (McLaughlin *et al.*, 2010: 143).

¹⁰² Selon Perani et Abutalebi (2005) : âge d'acquisition, niveau de L2, durée d'immersion, fréquence d'usage, etc. et Steinhauer (2014b) : langue maternelle, environnement d'apprentissage (salle de classe vs immersion) et variants psychométriques (e.g., motivation). Cf. Reiterer (2010), Morgan-Short (2014) et Sabourin *et al.* (2013) pour une revue et une discussion sur les facteurs ci-dessus dans des études d'EEG et/ou cognitives sur l'acquisition d'une L2.

¹⁰³ Les remarques sur la distance L1 vs L2 dans l'acquisition morphosyntaxique semblent intéressantes : "... Thus, results showed that, generally, both early and late AoA (Age of Acquisition), higher and lower proficiency participants exhibited the same components as in the L1 (or in comparison to native speakers) during the processing of crosslinguistically similar constructions. Dissimilar or unique constructions, however, resulted in conflicting findings... The processing of similar constructions was thus generally associated with a comparable degree of activation of common brain areas involved in (morpho)syntactic processing in the case of fMRI studies and with the presence of comparable (morpho)syntactic components in the case of ERP studies. The observed findings also suggest that (morpho)syntactic structures that exist in both languages but that differ in their specific implementation (i.e. dissimilar constructions) suffered from some degree of crosslinguistic competition and negative transfer." (Tolentino & Tokowicz, 2011: 118-119). Cela pourrait être sans doute transposé sur l'acquisition des structures prosodiques d'une L2.

¹⁰⁴ "Failure to achieve native-like processing has also been linked to dissimilarity between L1 and L2, as well as characteristics of the target language." (Meulman *et al.*, 2014: 14).

¹⁰⁵ Une remarque conclusive plus précise : "The results from this body of research reflect a converging but complex picture regarding the influence of L1 features on L2 processing. It appears that L2 learners can evidence natively-like processing of L2 features when they are unique to the L2, although quantitative differences can be evident between L2 and L1 speakers and within learners for features that are either present or absent in L1. Natively-like processing patterns also appear to be evidenced when the feature is expressed similarly in L1 and L2. Interestingly, it seems that the most difficulties for processing L2 features come about when the feature is present in the L1 but is expressed according to a different set

al., 2011 ; Xue *et al.*, 2013¹⁰⁶ ; Kasparian & Steinhauer, 2016 ; Fromont *et al.*, 2020 ; voir aussi la discussion de Kotz (2009) sur la relation entre le niveau de L2, la similarité L1 vs L2 et l'âge dans l'acquisition syntaxique d'une L2).

- 2) D'autre part, en comparant les effets ERP entre natifs et apprenants, il semble que les apprenants d'une L2, même tardifs (adultes), peuvent atteindre un niveau comparable à celui des natifs dans leur acquisition (e.g., Steinhauer *et al.*, 2009 ; Morgan-Short *et al.*, 2012a, b).

Concernant l'acquisition de la prosodie d'une L2 à l'aide de l'EEG, il existe un certain nombre d'études portant sur l'acquisition de l'accentuation ou du ton lexical. Cependant, notre revue de la littérature révèle que la composante CPS reste peu explorée dans l'acquisition prosodique d'une L2. Nous n'avons trouvé que quatre études disponibles sur ce sujet. Nous récapitulons les contextes et les conclusions de ces quatre études dans le Tableau 1.25 :

of rules in the L2." (Morgan-Short, 2014: 25). Ce point doit être pris en considération quant à notre expérience EEG dans la partie expérimentale de cette étude.

¹⁰⁶ "L2 proficiency and L1-L2 similarity interact in complex ways in shaping patterns of neural responses in L2 grammaticalization." (Xue *et al.*, 2013: 262).

Tableau 1.25 : Les quatre études existantes dans la littérature sur l'utilisation de la CPS pour étudier l'acquisition de la prosodie d'une L2.

Étude	L1	L2	Conclusions principales
Gilbert <i>et al.</i> (2023)	Anglais	Français	Les CPS du français et de l'anglais ont été observées tant chez les locuteurs natifs que chez les locuteurs de la L2. Cependant, l'effet CPS apparaît plus tardivement chez les apprenants francophones face aux frontières prosodiques de l'anglais.
	Français	Anglais	
Nickels & Steinhauer (2018)	Chinois mandarin ou Allemand	Anglais	La CPS a été observée tant chez les anglophones natifs que chez les apprenants sinophones et germanophones, avec une plus grande amplitude chez les locuteurs de niveaux avancé et natif, dans les régions antérieure, centrale et postérieure.
Nickels <i>et al.</i> (2013)	Allemand	Anglais	La CPS a été observée tant chez les natifs que chez les apprenants, au niveau centro-pariétal.
Naito-Billen (2012)	Anglais	Japonais	La CPS a été observée tant chez les natifs que chez les apprenants, au niveau centro-frontal.

À partir de ces quatre études, nous avons tiré quelques discussions et conclusions intéressantes.

Tout d'abord, concernant les travaux de Nickels et de ses collaborateurs (2013, 2018). Nickels *et al.* (2013) ont étudié des germanophones natifs apprenant l'anglais, tandis que Nickels et Steinhauer (2018) ont étudié des germanophones et sinophones natifs apprenant l'anglais. Ils ont montré que le traitement prosodique chez les apprenants L2 peut atteindre le niveau natif, tout comme l'acquisition de la morphosyntaxe. En s'appuyant sur l'étude de Steinhauer *et al.* (2009), Nickels *et al.* (2013) ainsi que Nickels et Steinhauer (2018) ont aussi remis en question l'hypothèse de la période critique (Lenneberg, 1967) : selon leurs résultats, l'effet CPS apparaît chez tous les apprenants tardifs d'une L2, même au niveau débutant. D'autre part, ils ont constaté que le niveau en langue étrangère est un facteur plus

déterminant que l'âge d'acquisition¹⁰⁷ (e.g., Ojima *et al.*, 2005 ; Dowens *et al.*, 2010, 2011 ; White *et al.*, 2012).

La discussion sur la présence de la CPS dans ces deux études semble intéressante. Dans Nickels *et al.* (2013) (germanophones natifs apprenant l'anglais), il est indiqué que même si les locuteurs natifs et non-natifs utilisent des indices acoustiques partiellement différents, la CPS a été déclenchée chez tous les apprenants, ce qui contredit l'idée de beaucoup d'études dont Morgan-Short (2014). Dans Nickels et Steinhauer (2018) (germanophones et sinophones natifs apprenant l'anglais), il suppose que la présence de la CPS chez tous les participants (natifs et non-natifs, tous niveaux confondus) pourrait être due à un mécanisme commun dans le traitement de la musique, comme indiqué par Glushko *et al.* (2016)¹⁰⁸. Une autre explication possible de la présence de la CPS, tant chez les locuteurs natifs que non natifs, pourrait aussi être attribuée aux universaux prosodiques : les langues du monde utilisent certains indices acoustiques communs comme marqueurs des frontières prosodiques. Cependant, cette question complexe ne peut pas être expliquée uniquement par les universaux prosodiques. En effet, parmi les indices utilisés pour la segmentation prosodique dans les langues du monde, certains, comme l'allongement final (Vaissière, 1983), seraient véritablement universels, tandis que d'autres varient selon les langues (cf. *supra*^{97 et 105}).

Sur ce point, l'étude de Gilbert *et al.* (2023), qui a examiné la CPS chez les apprenants anglophones du français et inversement chez les apprenants francophones de l'anglais, offre des perspectives intéressantes. Après avoir analysé la CPS dans les deux langues, Gilbert *et al.* (2023) ont conclu que les anglophones natifs pouvaient entièrement percevoir et assimiler l'allongement final du français comme un indice pour détecter la frontière prosodique. Ainsi, la CPS chez les anglophones était tout à

¹⁰⁷ "Despite smaller amplitude differences between the groups at frontal electrodes, the most striking result was the strong main effect of boundary at all electrode levels. We further found that the CPS increases with increasing proficiency. Importantly, this increase does not start from an absent CPS at low levels of proficiency; instead, even at low levels, the CPS is already present and only increases further with higher proficiency... Taken together, the present data are in line with previous studies suggesting that ERP differences in language processing between late L2 learners and native speakers are largely driven by proficiency differences, not by differences in their age of language acquisition." (Nickels & Steinhauer, 2018: 24, 28).

¹⁰⁸ "The fact that prosodic boundaries can routinely be processed – independent of participants' L1 background or L2 proficiency level – is most likely due to the underlying neurocognitive mechanisms being shared with other cognitive domains such as music (Glushko *et al.*, 2016) and because prosodic boundaries are features of all three L1 languages tested." (Nickels & Steinhauer, 2018: 24).

fait comparable à celle chez les francophones natifs quand ils étaient tous confrontés aux frontières prosodiques du français. En revanche, face à la prosodie de l'anglais, la latence de la CPS chez les francophones était relativement tardive par rapport aux anglophones natifs. Cela s'explique par le fait que, pour les francophones, la durée syllabique de l'allongement final en anglais est généralement moins longue et moins saillante que celle en français.

1.5.4 Synthèse

Dans cette dernière section de la partie théorique, nous nous sommes concentrés sur la méthode d'électroencéphalographie (EEG) et son application à la recherche sur les sciences du langage et sur l'acquisition d'une L2. Nous avons commencé par expliquer l'origine de l'activité électrique dans le cerveau et par différencier les termes « EEG » et « ERP ». Les avantages de l'EEG dans la recherche sur l'acquisition langagière et les composantes ERP liées à la prosodie ont été présentés par la suite. Une attention particulière a été accordée à une composante spécifiquement dédiée au traitement neurocognitif des frontières prosodiques, que nous étudierons dans la partie expérimentale : la CPS (*closure positive shift*). Nous avons examiné de près sa découverte, ses caractéristiques électrophysiologiques, l'état actuel de la recherche sur cette composante, ainsi que quelques études sur son application dans la recherche sur l'acquisition prosodique d'une L2.

Dans la partie expérimentale, nous mènerons une expérience électroencéphalographique et examiner la composante CPS pour étudier en particulier le traitement neurocognitif des frontières prosodiques du français chez les auditeurs francophones natifs et les apprenants sinophones du français en tant que L2. Notre analyse portera sur les différentes caractéristiques des composantes CPS chez ces deux groupes d'auditeurs, ce qui nous permettra de comprendre comment la perception des frontières prosodiques se développe chez les apprenants, comparativement aux natifs, au cours de l'acquisition du français comme L2. Dans les sections 1.2 et 1.3, nous avons comparé les différents indices acoustiques utilisés pour signaler les frontières prosodiques en français et en chinois mandarin : de façon simplifiée, la segmentation prosodique en français repose sur la pause, l'allongement final et la montée du contour F0, tandis qu'en chinois mandarin, elle repose sur la pause, l'allongement final mais le réhaussement de registre. Ces

différences dans les indices acoustiques utilisés aux frontières prosodiques des deux langues rejoignent l'idée de Morgan-Short (2014) et de beaucoup d'autres études (cf. *supra*¹⁰⁵). Grâce à notre expérience électroencéphalographique, nous pourrions, à lumière de Gilbert *et al.* (2023), explorer si l'utilisation de ces différents indices mélodiques dans les deux langues influence spécifiquement le traitement neurocognitif des frontières prosodiques du français par des apprenants sinophones de niveaux différents.

Partie expérimentale

La partie expérimentale, composée de trois chapitres, présente les trois expériences menées pour examiner les performances de segmentation et d'organisation de la prosodie du français chez les apprenants sinophones, comparés aux francophones natifs.

Dans les deux premiers chapitres, nous réaliserons deux expériences distinctes. La première expérience, présentée au chapitre 2, portera sur la production. Nous y examinons comment les apprenants sinophones, de différents niveaux, intègrent les éléments mélodiques et temporels de la prosodie pour organiser et segmenter des éléments sonores dans la chaîne parlée. Nous comparerons ces résultats avec ceux des locuteurs francophones natifs afin d'identifier les similitudes et les différences observées chez les apprenants. La deuxième expérience, présentée au chapitre 3, se concentrera sur la perception. Nous y analyserons principalement comment les indices prosodiques en français sont utilisés pour désambiguïser des phrases localement ambiguës, en comparant les performances des auditeurs francophones avec celles des apprenants sinophones.

Dans le chapitre 4, nous réaliserons une expérience d'électroencéphalographie (EEG) afin d'examiner le traitement neurocognitif de la segmentation prosodique. Notre attention se portera plus particulièrement sur le potentiel évoqué CPS. Nous analyserons et comparerons les propriétés neurophysiologiques de cette composante chez des auditeurs francophones natifs et des apprenants sinophones percevant, là aussi, des paires de phrases localement ambiguës.

Chapitre 2 Étude acoustique

2.1 Introduction et questions de recherche

Pour étudier les performances d'organisation et de segmentation prosodiques en français chez les francophones natifs et les apprenants sinophones, nous avons mené deux analyses acoustiques : l'une axée sur la segmentation prosodique et l'autre sur l'organisation prosodique. Ces analyses ont été effectuées sur une série de paires de phrases localement ambiguës.

Concernant la segmentation prosodique, nous avons examiné les indices acoustiques liés à la segmentation en français, en nous concentrant principalement sur l'allongement final et la montée de F0 à la position des dernières syllabes précédant certaines frontières prosodiques. En ce qui concerne l'organisation prosodique, nous avons analysé et comparé les valeurs de neuf paramètres quantitatifs rythmiques, afin d'évaluer les patrons temporels des locuteurs francophones natifs et des apprenants. L'objectif de ces analyses acoustiques est de nous aider à répondre aux questions de recherche suivantes :

- 1) Au niveau de la segmentation prosodique, quels indices acoustiques les locuteurs francophones natifs utilisent-ils pour marquer les frontières prosodiques dans des phrases localement ambiguës ? Quant aux apprenants sinophones, sont-ils capables de segmenter les frontières prosodiques en utilisant les indices mélodiques et temporels attendus en français ?
- 2) Au niveau de l'organisation prosodique, comment les patrons temporels se manifestent-ils chez les francophones natifs et chez les apprenants ? Comment ces patrons temporels évoluent-ils en fonction des niveaux de maîtrise du français ?

En nous basant sur les différences prosodiques entre les deux langues, ainsi que sur les résultats des recherches antérieures, nous formulons les hypothèses suivantes en lien avec ces questions de recherche :

- 1) Pour les francophones natifs, il est raisonnable de supposer qu'ils seront en mesure de marquer les frontières prosodiques au niveau de la segmentation, pour désambiguïser les sens des phrases. Ils utiliseront pour cela les deux indices acoustiques (l'allongement final et la montée mélodique) sur les dernières syllabes précédant les frontières prosodiques.

- 2) Concernant les apprenants, nous supposons qu'ils pourraient également réaliser l'allongement final pour marquer les frontières prosodiques, et ce quel que soit leur niveau en français. En effet, cet indice acoustique est universel et utilisé dans la majorité des langues, y compris en chinois mandarin. Pour ce qui est de la montée de F0 avant les frontières prosodiques, nous supposons également que cela ne poserait pas de difficulté réelle pour les apprenants, qui signaleraient les frontières prosodiques par la montée mélodique sur les dernières syllabes.
- 3) En ce qui concerne l'organisation temporelle, étant donné que le français et le mandarin sont tous deux des langues à isochronie syllabique, les patrons rythmiques des francophones natifs tendraient naturellement vers cette isochronie. Selon des études antérieures sur l'acquisition du rythme en langue seconde (L2), les apprenants présentent souvent des patrons rythmiques marqués par de fortes variations, en raison de débits de parole variables. Nous prévoyons un phénomène similaire chez les apprenants sinophones. Cela étant dit, nous supposons également que plus leur niveau de français est avancé, plus leurs patrons rythmiques se rapprocheront de l'isochronie syllabique caractéristique du français.

2.2 Méthodes

2.2.1 Locuteurs et corpus

Dix locuteurs francophones natifs et 21 apprenants sinophones du français, répartis en trois niveaux de compétence (débutant, intermédiaire et avancé), ont participé à cette expérience. Avant le déroulement de l'expérience, tous les participants ont rempli le questionnaire LEAP-Q (Marian *et al.*, 2007 ; Kaushanskaya *et al.*, 2020) (cf. sections 2.5.1 et 2.5.2 pour les questionnaires destinés aux locuteurs francophones natifs et aux apprenants sinophones). Ce questionnaire a permis de recueillir des informations personnelles et des profils linguistiques détaillés, notamment pour répartir les apprenants en trois groupes en fonction de leur niveau de français. Les locuteurs francophones natifs avaient tous le français pour seule langue maternelle. Les apprenants sinophones étaient étudiants en Île-de-France, et vivaient, au moment de la collecte des données, dans un environnement francophone, où ils utilisaient le français dans leur vie quotidienne. Par ailleurs, aucun des participants,

francophone natif ou apprenant, n'a déclaré souffrir de trouble de la parole ou d'apprentissage. Le Tableau 2.1 ci-dessous récapitule les informations importantes concernant les participants, ainsi que la répartition des apprenants sinophones en fonction de leur niveau :

Tableau 2.1 : Informations (sexe et âge moyen) sur les 31 sujets ayant participé à l'expérience acoustique.

Locuteurs		Nombre (H : Homme ; F : Femme)	Âge Moyen (ET)
Francophones natifs		10 (5H, 5F)	31,3 (8,37)
Apprenants sinophones	Débutants	7 (2H, 5F)	26,29 (4,15)
	Intermédiaires	7 (2H, 5F)	27,14 (2,91)
	Avancés	7 (3H, 4F)	27 (2,83)

Inspirés de Pauker (2013), nous avons établi un corpus contenant cinq paires de phrases localement ambiguës. Chaque paire de phrases a été construite de sorte à avoir les frontières prosodiques situées dans deux positions différentes, l'une avec une frontière qui apparaît relativement plus tôt, appelée « clôture précoce » (CP) et l'autre avec une frontière qui apparaît plus tard, appelée « clôture tardive » (CT) (Tableau 2.2).

Tableau 2.2 : Les cinq paires de phrases utilisées dans l'expérience acoustique.

Conditions prosodiques	Phrase en haut : CP (Clôture Précoce)	Phrase en bas : CT (Clôture Tardive)
	Chaque fois que le serpent mangeait, le rat se cachait.	Chaque fois que le serpent mangeait le rat, le lapin se cachait.
	Chaque fois que le chef cuisinait, le steak sentait très bon.	Chaque fois que le chef cuisinait le steak, le restaurant sentait très bon.
	Chaque fois que l'étudiant travaillait, les notes du cours devenaient plus claires.	Chaque fois que l'étudiant travaillait les notes, les cours devenaient plus claires.
	Chaque fois que le comédien fredonnait, le morceau était joli à l'oreille.	Chaque fois que le comédien fredonnait le morceau, la mélodie était jolie à l'oreille.
	Chaque fois que le professeur enseignait, la théorie littéraire devenait moins obscure.	Chaque fois que le professeur enseignait la théorie littéraire, son raisonnement devenait moins obscur.

Lors de la production de ces phrases, les locuteurs doivent normalement indiquer correctement les positions des frontières prosodiques (CP ou CT) afin que les auditeurs puissent établir la relation hiérarchique syntaxique appropriée entre les constituants prosodiques et ainsi accéder aux sens corrects. Les objectifs de la création de ce corpus sont doubles : d'une part, il s'agit d'observer comment les francophones natifs et les apprenants sinophones désambiguïseront la syntaxe par la réalisation des frontières prosodiques, ce qui nous permettra d'évaluer leurs capacités en segmentation prosodique ; d'autre part, nous pourrons analyser leurs compétences en organisation prosodique en calculant les valeurs des paramètres rythmiques sur l'ensemble de ces phrases.

2.2.2 Traitement et analyses des données

Les locuteurs ont été enregistrés séparément dans la chambre sourde du Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle). Les phrases du corpus étaient imprimées dans un ordre aléatoire sur une feuille, et les locuteurs ont été invités à lire le corpus à deux reprises. Au total, 20 phrases (2 phrases par paire × 5 paires × 2 répétitions) ont été enregistrées, digitalisées à une fréquence de 16,000 Hz et à 16 bits. Au total, nous avons enregistré 620 phrases (20 phrases × 31

sujets (10 locuteurs francophones natifs + 21 apprenants sinophones)). Ces 620 phrases ont été segmentées automatiquement à l'aide de la boîte à outils SPASS (Bigi, 2015, 2016 ; Bigi & Meunier, 2018) dans le logiciel Praat (version 6.4.06) (Boersma & Weenink, 2024). Par la suite, les segmentations et annotations ont été vérifiées et corrigées manuellement si nécessaire. La Figure 2.1 présente un exemple de segmentation pour la phrase « *Chaque fois que le serpent mangeait, le rat se cachait* », produite par une locutrice francophone native.

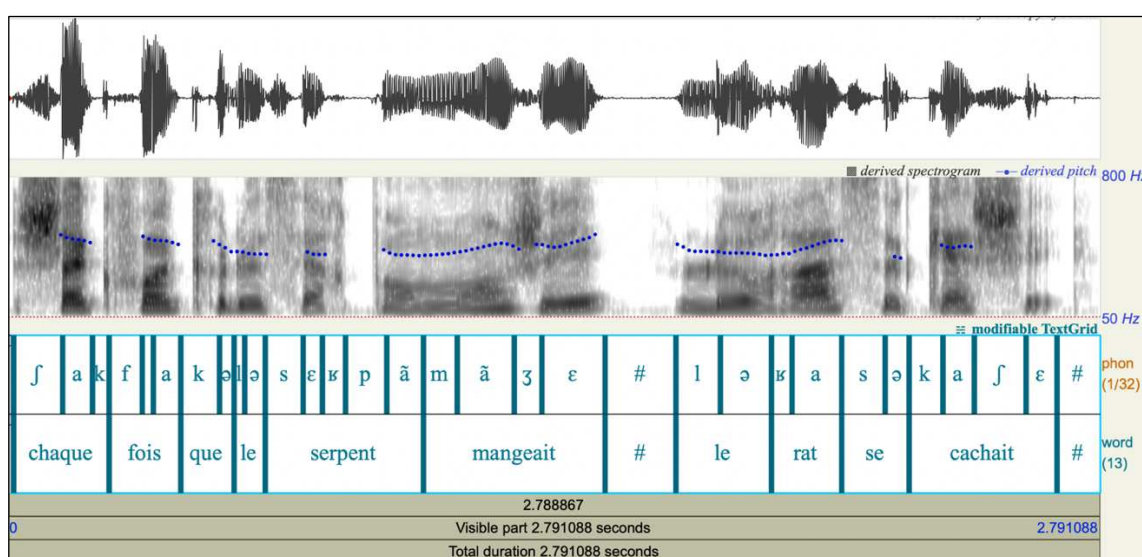


Figure 2.1 : Exemple de segmentation pour la phrase « *Chaque fois que le serpent mangeait, le rat se cachait* ».

Nous avons extrait une série de paramètres acoustiques mélodiques et temporels à l'aide d'un script de Praat (Boersma & Weenink, 2024). Sur le plan mélodique, nous avons comparé les mouvements mélodiques des mêmes voyelles dans les deux phrases de chaque paire. Par exemple, dans la paire de phrases suivante (Tableau 2.3), nous avons réalisé une première comparaison entre les deux voyelles /ɛ/ (en rouge) : l'une précède la frontière CP (notée V-FrontièreCP), tandis que l'autre ne la précède pas (notée V-NonFrontièreCP). Ensuite, nous avons comparé les deux voyelles /a/ (en vert) : l'une précède la frontière CT (notée V-FrontièreCT), et l'autre ne la précède pas (notée V-NonFrontièreCT). Ces comparaisons sont récapitulées dans le Tableau 2.4 pour une meilleure lisibilité.

Tableau 2.3 : Exemple d'une paire de phrases localement ambiguës dans l'expérience de production de la parole (« # » représente la frontière prosodique).

Condition prosodique	Exemple
Clôture précoce (CP)	Chaque fois que le serpent mangeait, le rat se cachait.
Clôture tardive (CT)	Chaque fois que le serpent mangeait le rat, le lapin se cachait.

Tableau 2.4 : Significations des abréviations utilisées dans l'étude de production de la parole.

Dénomination	Signification
V-FrontièreCP	Voyelle qui précède la frontière de clôture précoce
V-NonFrontièreCP	Voyelle qui ne précède pas la frontière de clôture précoce
V-FrontièreCT	Voyelle qui précède la frontière de clôture tardive
V-NonFrontièreCT	Voyelle qui ne précède pas la frontière de clôture tardive

Pour chaque voyelle analysée, nous avons extrait deux valeurs de F0, correspondant respectivement aux positions de 20 % et de 90 % de chaque voyelle. Nous avons ensuite calculé la différence entre ces deux valeurs (c'est-à-dire F0 à 90% - F0 à 20%), ce qui nous a permis de déterminer s'il s'agissait d'une montée ou d'une descente mélodique. Par la suite, comme nous venons de l'indiquer, nous avons comparé les mêmes voyelles dans les deux conditions prosodiques CP et CT (i.e., V-FrontièreCP vs. V-NonFrontièreCP et V-FrontièreCT vs. V-NonFrontièreCT).

Au niveau de la segmentation temporelle, nous avons extrait les durées des voyelles analysées (celles en rouge et en vert dans le Tableau 2.3). Comme pour l'analyse mélodique, nous avons effectué les mêmes comparaisons : V-FrontièreCP vs V-NonFrontièreCP et V-FrontièreCT vs V-NonFrontièreCT. Étant donné que la frontière prosodique en français est marquée, d'un point de vue temporel, par l'allongement de la dernière voyelle, ces mesures nous ont permis d'évaluer si un allongement de durée était réalisé par les locuteurs francophones natifs et les apprenants sinophones sur les voyelles finales précédant les frontières prosodiques CP et CT.

Au niveau de l'organisation temporelle, nous avons calculé, à l'aide du logiciel Correlatore (Mairano & Romano, 2010), les sept paramètres quantitatifs rythmiques les plus étudiés dans la littérature (à savoir %V ; ΔC ; ΔV ; VarcoV ; VarcoC ; nPVI-V ; rPVI-C), ainsi que le paramètre CCI (*Control/Compensation Index*) en tant qu'indicateur complémentaire, proposé par Bertinetto et Bertini (2008). Le Tableau 2.5 rappelle les significations de ces paramètres.

Tableau 2.5 : Significations des paramètres rythmiques mesurés dans cette étude.

Paramètre	Signification
%V	La proportion des intervalles vocaliques au sein de la phrase, calculée comme la somme des durées des intervalles vocaliques divisée par la durée totale de tous les intervalles.
ΔV	L'écart-type des durées d'intervalles vocaliques au sein de la phrase.
ΔC	L'écart-type des durées d'intervalles consonantiques au sein de la phrase.
VarcoV VarcoC	ΔV et ΔC normalisés
rPVI-C	La mesure des variabilités brutes des intervalles consonantiques successifs. $rPVI = \left[\sum_{k=1}^{m-1} d_k - d_{k+1} / (m - 1) \right]$
nPVI-V	rPVI normalisé $nPVI = 100 \left[\sum_{k=1}^{m-1} \left \frac{d_k - d_{k+1}}{(d_k + d_{k+1})/2} \right / (m - 1) \right]$
CCI-V	Le calcul de l'index du niveau I du modèle CCI est modifié sur la base du paramètre PVI. L'index CCI-V ajoute le nombre de segments dans un même intervalle vocalique (Bertinetto & Bertini, 2010), ce qui permet de prendre en considération la complexité de la structure phonotactique. $CI = \frac{100}{m - 1} \sum_{k=1}^{m-1} \left \frac{d_k}{n_k} - \frac{d_{k+1}}{n_{k+1}} \right $
CCI-C	L'index CCI-C ajoute le nombre de segments dans un même intervalle consonantique (Bertinetto & Bertini, 2010).

2.3 Résultats

Dans les sections suivantes, nous présenterons les résultats de nos analyses acoustiques. La section 2.3.1 sera consacrée aux résultats concernant la segmentation mélodique. La section 2.3.2 traitera des aspects temporels : nous y présenterons les résultats de la segmentation temporelle dans la section 2.3.2.1 et ceux de l'organisation temporelle dans la section 2.3.2.2.

2.3.1 Segmentation mélodique

Nous avons extrait et comparé un total de 1240 voyelles. Pour chaque condition prosodique, deux types de comparaisons ont été effectués : une comparaison intra-groupe, qui consistait à comparer les valeurs de F0 entre les voyelles précédant les frontières prosodiques et celles ne les précédant pas, au sein de chacun des quatre

groupes de locuteurs. L'objectif était de déterminer si les locuteurs de chaque groupe utilisaient la montée de F0 en français pour signaler les frontières prosodiques. Les différences de F0 ont été analysées à l'aide d'un modèle linéaire mixte pour chaque condition prosodique, en utilisant la boîte à outils lme4 (Bates *et al.*, 2024) dans RStudio (Version 2024.09). Ce modèle testait l'effet de deux facteurs : la position de la voyelle (deux niveaux : précédant la frontière vs ne précédant pas la frontière), et le niveau de français (quatre niveaux : natif, avancé, intermédiaire, débutant). Les locuteurs ont été inclus comme facteur à effet aléatoire. Le modèle utilisé était le suivant : [Différences de F0 ~ Positions de la voyelle + Niveaux + (1 | Locuteurs)]. Pour approfondir l'analyse des différences de F0 en fonction des positions prosodiques au sein de chaque groupe de locuteurs, nous avons également réalisé une série de tests post-hoc de Tukey à la suite du modèle linéaire mixte.

Pour la comparaison inter-groupe, nous avons examiné les valeurs de F0 des voyelles précédant la même frontière prosodique entre les quatre groupes de locuteurs. Statistiquement, ces valeurs ont été analysées à l'aide d'un test ANOVA, suivi d'une série de tests post-hoc de Tukey.

Nous présenterons les résultats de la segmentation mélodique dans deux sections distinctes. La section 2.3.1.1 sera dédiée aux comparaisons de la condition CP, tandis que la section 2.3.1.2 se concentrera sur les comparaisons de la condition CT. Dans chaque section, nous rapporterons d'abord les résultats de la comparaison « intra-groupe », puis ceux de la comparaison « inter-groupe ».

2.3.1.1 Comparaisons « intra-groupe » et « inter-groupe » pour la condition CP

- Comparaison « intra-groupe »

Le Tableau 2.6 présente les valeurs moyennes (en Hz) des différences de F0 (90% - 20%) ainsi que les écarts-types pour les quatre groupes de locuteurs, pour V-FrontièreCP et V-NonFrontièreCP. Le Tableau 2.6 indique aussi le pourcentage de différence entre les deux conditions, calculé comme suit : $F0 \text{ de V-FrontièreCP} - F0 \text{ de V-NonFrontièreCP} / F0 \text{ de V-FrontièreCP}$.

Tableau 2.6 : Statistiques descriptives des différences de F0 (en Hz) pour V-FrontièreCP et V-NonFrontièreCP pour les quatre groupes de locuteurs.

Niveau / Positions des voyelles	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
V-FrontièreCP	Moyenne : +55,36 ET : 54,04	Moyenne : +63,36 ET : 41,70	Moyenne : +31,69 ET : 27,70	Moyenne : +41,53 ET : 52,82
V-NonFrontièreCP	Moyenne : -10,28 ET : 12,06	Moyenne : +10,23 ET : 18,62	Moyenne : +7,25 ET : 12,45	Moyenne : +27,90 ET : 79,47
V-FrontièreCP vs V-NonFrontièreCP	119% de montée	84% de montée	77% de montée	33% de montée

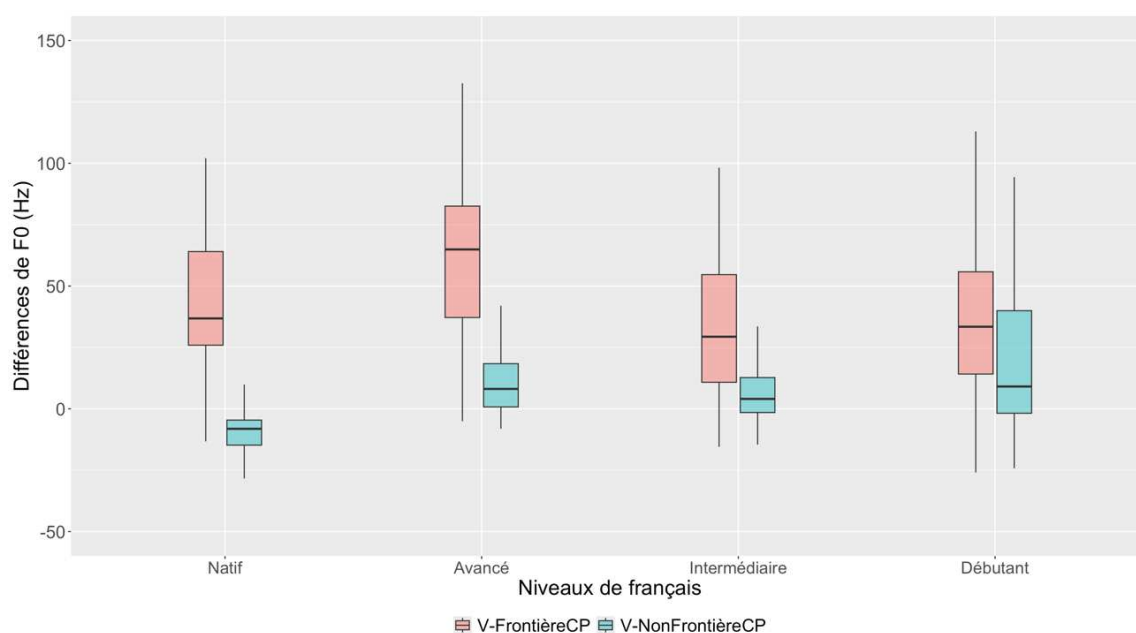


Figure 2.2 : Comparaisons des différences de F0 (en Hz) pour V-FrontièreCP et V-NonFrontièreCP pour les quatre groupes de locuteurs.

Comme l'indiquent les statistiques descriptives (Tableau 2.6) et la Figure 2.2, les différences de F0 des V-FrontièreCP chez tous les groupes de locuteurs sont plus prononcées que celles des V-NonFrontièreCP. Les montées dans les V-FrontièreCP sont respectivement 119%, 84%, 77% et 33% plus marquées par rapport aux V-NonFrontièreCP chez les francophones natifs, ainsi que chez les apprenants avancés, intermédiaires et débutants.

Le modèle linéaire mixte a montré un effet significatif des positions prosodiques des voyelles, tandis que l'effet des niveaux de français n'est pas significatif (voir Tableau 2.7). En d'autres termes, dans les phrases comportant des structures CP, les différences de F0 entre V-FrontièreCP et V-NonFrontièreCP sont significativement distinctes, tant chez les locuteurs francophones natifs que chez les apprenants.

Tableau 2.7 : Résultats du modèle linéaire mixte pour les différences de F0 entre V-FrontièreCP vs V-NonFrontièreCP en fonction des facteurs « positions des voyelles » et « niveaux de français ».

Effet	Positions des voyelles	Niveaux de français
Résultat	b=20,87 ; SE=1,64 ; t=12,70 ; p<0,001	b=8,43 ; SE=6,19 ; t=1,36 ; p=0,24

Pour examiner chaque groupe plus précisément, nous avons réalisé des tests post-hoc de Tukey après le modèle linéaire mixte. Ces tests ont montré des différences significatives chez les locuteurs francophones natifs, ainsi que chez les apprenants avancés et intermédiaires, tandis qu'aucune différence significative n'a été trouvée chez les apprenants débutants (Tableau 2.8).

Tableau 2.8 : Résultats des tests post-hoc de Tukey pour les différences de F0 entre V-FrontièreCP vs V-NonFrontièreCP dans chaque groupe de locuteurs.

Groupe		Résultat des tests post-hoc de Tukey
Francophones natifs		SE=5,54 ; t=11,85 ; p<0,001
Apprenants	Avancés	SE=5,46 ; t=9,73 ; p<0,001
	Intermédiaires	SE=3,64 ; t=6,73 ; p<0,001
	Débutants	SE=11,41 ; t=1,20 ; p=0,23

En résumé, bien que les différences de F0 de V-FrontièreCP soient plus marquées que celles de V-NonFrontièreCP dans tous les groupes, elles ne sont pas significativement différentes chez les apprenants débutants.

- *Comparaison « inter-groupe »*

Les valeurs de V-FrontièreCP pour les quatre groupes de locuteurs ont déjà été indiqués dans la Figure 2.2. Les résultats du test ANOVA (Tableau 2.9) montrent que les différences de F0 des V-FrontièreCP varient significativement entre les quatre groupes.

Tableau 2.9 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de F0 des V-FrontièreCP pour les quatre groupes de locuteurs.

Condition	Résultat du test ANOVA	
V-FrontièreCP	df=3, F=6,69, p<0,001, $\eta^2=0,062$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=7,22, t=1,11, p=0,69
	Natif vs Intermédiaire	SE=7,25, t= -3,27, p=0,007
	Natif vs Débutant	SE=7,22, t= -1,92, p=0,22
	Avancé vs Intermédiaire	SE=7,85, t=4,03, p<0,001
	Avancé vs Débutant	SE=7,83, t=2,79, p=0,03
	Intermédiaire vs Débutant	SE=7,85, t=1,25, p=0,59

Les tests post-hoc de Tukey, également présentés dans le Tableau 2.9, révèlent que les différences de F0 des apprenants avancés sont comparables à celles des locuteurs natifs, mais diffèrent significativement de celles des niveaux intermédiaires et débutants. Cela suggère que les apprenants avancés produisent une montée de F0 proche de celle des natifs dans la condition CP, tandis que les montées de F0 des apprenants intermédiaires et débutants sont similaires. Cependant, un résultat inattendu émerge : aucune différence significative n'est observée entre natifs et débutants. Cela ne signifie pas pour autant que les apprenants débutants utilisent la montée de F0 pour désambiguïser les phrases localement ambiguës. En effet, comme le montre la comparaison intra-groupe, les montées de F0 entre V-FrontièreCP et V-NonFrontièreCP chez les débutants sont similaires, contrairement aux attentes (voir Tableau 2.8).

2.3.1.2 Comparaisons pour la condition CT

- Comparaison « intra-groupe »

Comme le montrent le Tableau 2.10 et la Figure 2.3, les locuteurs francophones natifs, ainsi que les apprenants avancés et intermédiaires ont réalisé en moyenne une plus grande montée de F0 pour V-FrontièreCT comparée à V-NonFrontièreCT. En revanche, les apprenants débutants ont produit une montée de F0 plus grande pour V-NonFrontièreCT.

Tableau 2.10 : Statistiques descriptives des différences de F0 (en Hz) pour V-FrontièreCT vs V-NonFrontièreCT pour les quatre groupes de locuteurs.

Niveau Positions des voyelles	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
V-FrontièreCT	Moyenne : +42,46 ET : 40,26	Moyenne : +54,98 ET : 46,47	Moyenne : +21,28 ET : 33,59	Moyenne : +29,77 ET : 42,53
V-NonFrontièreCT	Moyenne : +7,75 ET : 22,75	Moyenne : +32,72 ET : 55,20	Moyenne : +14,01 ET : 44,53	Moyenne : +42,14 ET : 76,37
V-FrontièreCT vs V-NonFrontièreCT	82% de montée	40% de montée	34% de montée	-42%

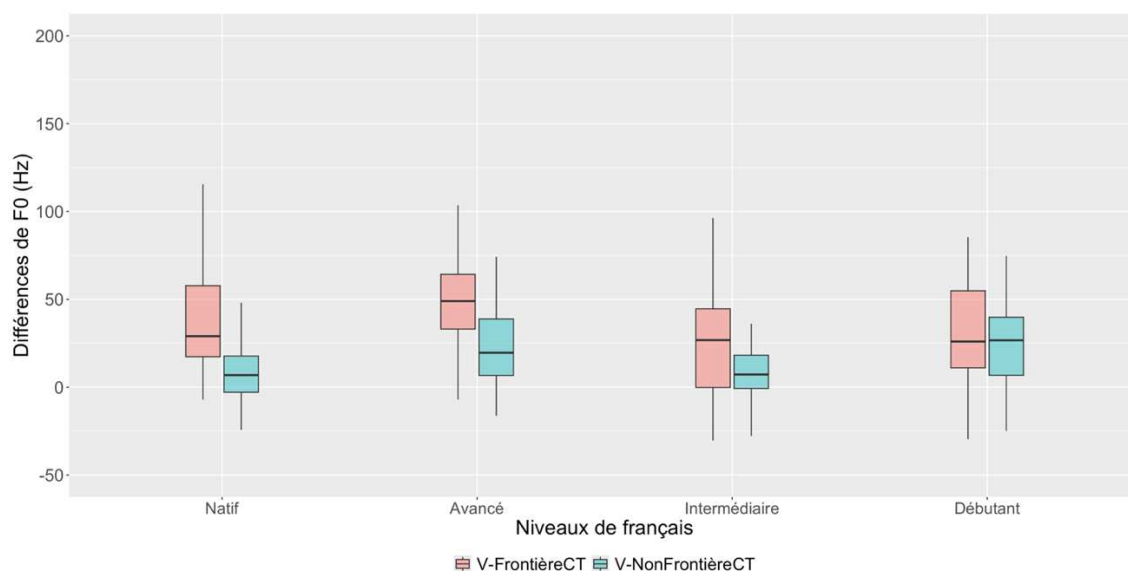


Figure 2.3 : Comparaisons des différences de F0 (en Hz) pour V-FrontièreCT vs V-NonFrontièreCT pour les quatre groupes de locuteurs.

Le modèle linéaire mixte a montré un effet significatif pour les deux facteurs : positions des voyelles et niveaux de français (Tableau 2.11). Cela signifie que, dans les phrases avec des structures CT, les variations de F0 entre V-FrontièreCT et V-NonFrontièreCT sont significativement différentes. Cependant, lorsque l'on considère chaque groupe, les tests post-hoc de Tukey ont révélé des différences significatives chez les locuteurs natifs, ainsi que chez les apprenants avancés et intermédiaires, mais pas chez les apprenants débutants (Tableau 2.12).

Tableau 2.11 : Résultats du modèle linéaire mixte pour les différences de F0 entre V-FrontièreCT vs V-NonFrontièreCT en fonction des facteurs « positions des voyelles » et « niveaux de français ».

Effet Résultat	Positions des voyelles	Niveaux de français
	b=7,53 ; SE=1,79 ; t=4,20 ; p<0,001	b=13,21 ; SE=5,61 ; t=2,35 ; p=0,03

Tableau 2.12 : Résultats des tests post-hoc de Tukey pour les différences de F0 entre V-FrontièreCT vs V-NonFrontièreCT dans chaque groupe de locuteurs.

Groupe		Résultat des tests post-hoc de Tukey
Francophones natifs		SE=4,62 ; t=7,51 ; p<0,001
Apprenants	Avancés	SE=8,62 ; t=2,58 ; p=0,01
	Intermédiaires	SE=6,67 ; t=1,09 ; p=0,02
	Débutants	SE=6,82 ; t=0,17 ; p=0,86

En résumé, les locuteurs natifs, avancés et intermédiaires ont signalé les frontières prosodiques en utilisant l'indice mélodique attendu, c'est-à-dire la montée de F0 pour les V-FrontièreCT par rapport aux V-NonFrontièreCT. En revanche, une telle différence n'a pas été observée chez les apprenants débutants.

- *Comparaison « inter-groupe »*

Les valeurs pour V-FrontièreCT pour les quatre groupes de locuteurs ont déjà été indiqués dans la Figure 2.3. Les résultats du test ANOVA (Tableau 2.13) montrent que les différences de F0 pour V-FrontièreCT sont significativement distinctes entre les quatre groupes. Les résultats des tests post-hoc de Tukey montrent que, pour le niveau avancé, les différences de F0 sont comparables à celles du niveau natif, mais diffèrent significativement de celles des niveaux intermédiaires et débutants. Cela suggère que les apprenants avancés présentent une montée de F0 proche de celle des locuteurs natifs. En revanche, aucune différence significative n'a été observée entre intermédiaires et débutants.

Par ailleurs, les tests post-hoc révèlent qu'il n'y a pas de différence non plus entre les niveaux natif et débutant. Comme pour la condition CP, cela n'indique pas que les apprenants débutants utilisent la montée de F0 pour marquer la frontière CT. En effet, comme le montre notre comparaison intra-groupe (voir Tableau 2.12), les montées de F0 pour V-FrontièreCT et V-NonFrontièreCT chez les apprenants débutants étaient comparables, et ils ont même réalisé une montée de F0 plus élevée pour V-NonFrontièreCT comparée à V-FrontièreCT.

Tableau 2.13 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de F0 des V-FrontièreCT pour les quatre groupes de locuteurs.

Condition	Résultat du test ANOVA	
V-FrontièreCT	df=3, F=9,25, p<0,001, $\eta^2=0,083$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=6,38, t=1,96, p=0,20
	Natif vs Intermédiaire	SE=6,38, t= -3,32, p=0,006
	Natif vs Débutant	SE=6,38, t= -1,99, p=0,19
	Avancé vs Intermédiaire	SE=6,92, t=4,87, p<0,001
	Avancé vs Débutant	SE=6,92, t=3,65, p=0,002
	Intermédiaire vs Débutant	SE=6,92, t=1,23, p=0,61

2.3.2 Aspects temporels

2.3.2.1 Segmentation temporelle

Nous avons analysé et comparé les durées des mêmes voyelles que celles analysées pour la segmentation mélodique. Comme pour les analyses mélodiques, deux types de comparaisons ont été effectués pour chaque condition prosodique. La comparaison « intra-groupe » visait à déterminer si les locuteurs de chaque groupe utilisaient l’allongement final en français pour signaler les frontières prosodiques. Les durées ont été analysées statistiquement à l’aide d’un modèle linéaire mixte : [Valeurs de durée ~ Positions des voyelles + Niveaux + (1 | Locuteurs)]. Des tests post-hoc de Tukey ont ensuite été réalisés pour examiner la situation spécifique de chaque groupe. Pour la comparaison « inter-groupe », nous avons comparé les durées des voyelles précédant la même frontière prosodique entre les quatre groupes. Ces valeurs ont été analysées à l’aide d’un test ANOVA, suivi de tests post-hoc de Tukey.

Les résultats de la segmentation temporelle seront présentés en deux sections distinctes : la section 2.3.2.1.1 pour les comparaisons de la condition CP, et la section 2.3.2.1.2 pour celles de la condition CT. Dans chaque section, nous rapporterons d’abord les résultats de la comparaison « intra-groupe », puis ceux de la comparaison « inter-groupe ».

2.3.2.1.1 Comparaisons pour la condition CP

- Comparaison « intra-groupe »

Comme l'indiquent le Tableau 2.14 et la Figure 2.4, les durées des V-FrontièreCP chez les francophones natifs et les apprenants intermédiaires et avancés sont supérieures à celles des V-NonFrontièreCP, ce qui n'est pas le cas pour les débutants. Pour les francophones natifs, les durées des V-FrontièreCP sont 54% plus longues que celles des V-NonFrontièreCP. Pour les apprenants avancés et intermédiaires, les V-FrontièreCP sont respectivement 42% et 22% plus longues par rapport aux V-NonFrontièreCP. Les débutants n'ont pas présenté les mêmes différences (-6%).

Tableau 2.14 : Statistiques descriptives des durées des voyelles (en ms) pour V-FrontièreCP et V-NonFrontièreCP pour les quatre groupes de locuteurs.

Niveau / Positions des voyelles	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
V-FrontièreCP	Moyenne : 171,12 ET : 32,18	Moyenne : 191,91 ET : 48,04	Moyenne : 190,85 ET : 50,77	Moyenne : 191,87 ET : 43,91
V-NonFrontièreCP	Moyenne : 78,80 ET : 32,62	Moyenne : 111,54 ET : 37,35	Moyenne : 149,06 ET : 70,00	Moyenne : 203,75 ET : 87,85
V-FrontièreCP vs V-NonFrontièreCP	54% plus longue	42% plus longue	22% plus longue	-6%

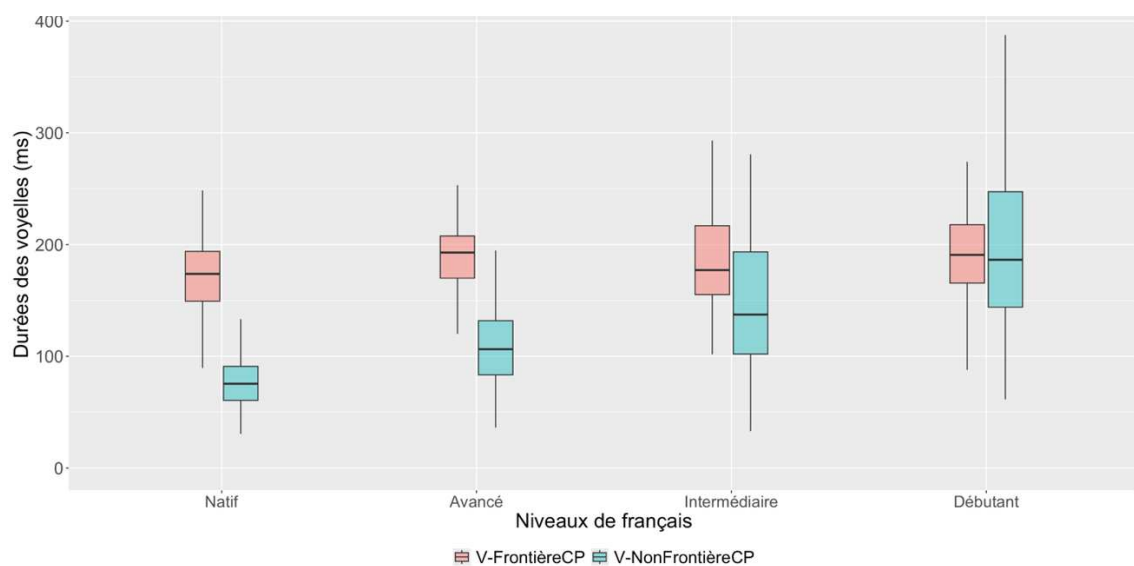


Figure 2.4 : Comparaisons des durées des voyelles (en ms) pour V-FrontièreCP vs V-NonFrontièreCP pour les quatre groupes de locuteurs.

Le modèle linéaire mixte a mis en évidence un effet significatif à la fois pour les facteurs « positions des voyelles » et « niveaux de français » (Tableau 2.15). Cela indique que les durées des voyelles diffèrent de manière significative entre les V-FrontièreCP et les V-NonFrontièreCP, et qu'il existe également une différence notable entre les quatre groupes de locuteurs.

Tableau 2.15 : Résultats du modèle linéaire mixte pour les durées entre V-FrontièreCP vs V-NonFrontièreCP en fonction des facteurs « positions des voyelles » et « niveaux de français ».

Effet Résultat	Positions des voyelles	Niveaux de français
	b=27,28 ; SE=2,04 ; t=13,39 ; p<0,001	b=36,68 ; SE=8,08 ; t=4,54 ; p<0,001

Les tests post-hoc de Tukey ont montré une différence significative entre les durées des V-FrontièreCP et des V-NonFrontièreCP chez les locuteurs francophones natifs, ainsi que chez les apprenants avancés et intermédiaires. Cependant, aucune différence significative n'a été observée chez les apprenants débutants (Tableau 2.16).

Tableau 2.16 : Résultats des tests post-hoc de Tukey pour les durées entre V-FrontièreCP vs V-NonFrontièreCP dans chaque groupe de locuteurs.

Groupe		Résultat des tests post-hoc de Tukey
Francophones natifs		SE=4,58 ; t=20,15 ; p<0,001
Apprenants	Avancés	SE=7,27 ; t=11,05 ; p<0,001
	Intermédiaires	SE=10,36 ; t=4,03 ; p<0,001
	Débutants	SE=11,73 ; t= -1,01 ; p=0,31

En d'autres termes, les locuteurs natifs, avancés et intermédiaires ont marqué les frontières prosodiques CP par un allongement final. En revanche, chez les apprenants débutants, les durées des voyelles dans les deux positions (V-FrontièreCP et V-NonFrontièreCP) étaient comparables, ce qui montre qu'ils n'ont pas pu utiliser la durée vocalique pour distinguer la présence ou l'absence des frontières prosodiques CP.

- *Comparaison « inter-groupe »*

Les résultats du test ANOVA (Tableau 2.17) indiquent que les durées des V-FrontièreCP diffèrent significativement entre les quatre groupes. Les tests post-hoc

de Tukey, également présentés dans le même tableau, montrent que la durée des voyelles pour les V-FrontièreCP chez les francophones natifs est significativement différente de celle des apprenants des trois niveaux. En revanche, les durées des V-FrontièreCP sont similaires entre les trois groupes d'apprenants.

Tableau 2.17 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de durées des V-FrontièreCP pour les quatre groupes de locuteurs.

Condition	Résultat du test ANOVA	
V-FrontièreCP	df=3, F=5,04, p=0,002, $\eta^2=0,047$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=6,74, t=3,08, p=0,01
	Natif vs Intermédiaire	SE=6,74, t=2,93, p=0,02
	Natif vs Débutant	SE=6,74, t=3,08, p=0,01
	Avancé vs Intermédiaire	SE=7,31, t=0,15, p=0,99
	Avancé vs Débutant	SE=7,31, t=0,006, p=1
	Intermédiaire vs Débutant	SE=7,31, t=0,14, p=0,99

Ce constat est corroboré par les statistiques descriptives (voir Tableau 2.14), qui révèlent que les moyennes des durées des V-FrontièreCP sont d'environ 191 ms chez les apprenants (probablement à cause d'un débit plus long), contre environ 170 ms chez les locuteurs natifs. Il est toutefois important de noter que, selon la comparaison intra-groupe, les locuteurs natifs ainsi que les apprenants intermédiaires et avancés ont tous été capables de signaler la condition CP par un allongement final significatif pour les V-FrontièreCP. En revanche, les apprenants débutants ont produit des durées comparables entre les V-FrontièreCP et les V-NonFrontièreCP, ne marquant ainsi pas, comme attendu, la présence des frontières prosodiques CP.

2.3.2.1.2 Comparaisons pour la condition CT

- Comparaison « intra-groupe »

Nous avons effectué les mêmes analyses pour examiner l'allongement final des voyelles précédant (ou pas) les frontières CT. Le Tableau 2.18 récapitule les valeurs moyennes des durées vocaliques et les écarts-types pour les quatre groupes de locuteurs. Les résultats montrent que les natifs produisent des durées de V-FrontièreCT 34% plus longues par rapport à celles de V-NonFrontièreCT. Pour les

apprenants avancés, intermédiaires et débutants, les durées de V-FrontièreCT sont respectivement de 19% et 10% plus longues par rapport à celles de V-NonFrontièreCT. Les débutants n'ont pas produit les mêmes différences (-9%).

Tableau 2.18 : Statistiques descriptives des durées des voyelles (en ms) pour V-FrontièreCT et V-NonFrontièreCT pour les quatre groupes de locuteurs.

Niveau / Positions des voyelles	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
V-FrontièreCT	Moyenne : 147,68 ET : 40,34	Moyenne : 162,07 ET : 40,04	Moyenne : 169,48 ET : 51,12	Moyenne : 185,07 ET : 44,83
V-NonFrontièreCT	Moyenne : 97,99 ET : 34,49	Moyenne : 130,58 ET : 38,56	Moyenne : 152,74 ET : 54,52	Moyenne : 200,87 ET : 59,43
V-FrontièreCT vs V-NonFrontièreCT	34% plus longue	19% plus longue	10% plus longue	-9%

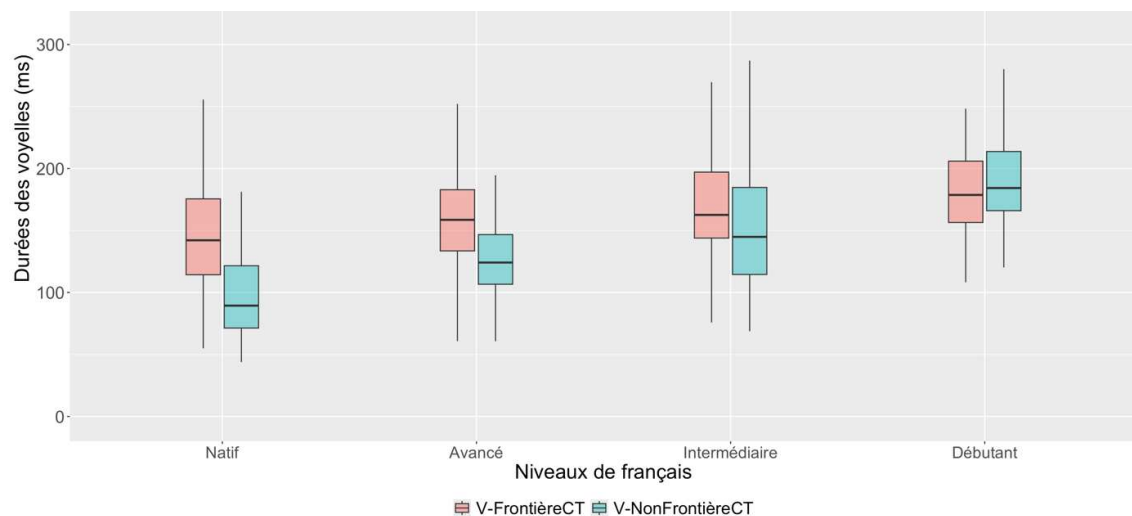


Figure 2.5 : Comparaisons des durées des voyelles (en ms) pour V-FrontièreCT vs V-NonFrontièreCT pour les quatre groupes de locuteurs.

Le modèle linéaire mixte a mis en évidence des effets significatifs tant pour les « positions des voyelles » que pour les « niveaux de français » (Tableau 2.19). Ces résultats confirment que les durées des voyelles diffèrent de manière significative entre les deux positions prosodiques, et qu'il existe également une différence significative entre les groupes de locuteurs. Les résultats des tests post-hoc (Tableau 2.20) montrent que les locuteurs natifs, avancés et intermédiaires ont réalisé un

allongement final significativement plus long pour V-FrontièreCT par rapport à V-NonFrontièreCT. Pour les apprenants débutants, les durées des V-FrontièreCT n'étaient pas significativement différentes de celles des V-NonFrontièreCT.

Tableau 2.19 : Résultats du modèle linéaire mixte pour les durées entre V-FrontièreCT vs V-NonFrontièreCT en fonction des facteurs « positions des voyelles » et « niveaux de français ».

Effet Résultat	Positions des voyelles	Niveaux de français
	b=11,69 ; SE=1,74 ; t=6,74 ; p<0,001	b=37,12 ; SE=6,53 ; t=5,68 ; p<0,001

Tableau 2.20 : Résultats des tests post-hoc de Tukey pour les durées entre V-FrontièreCT vs V-NonFrontièreCT dans chaque groupe de locuteurs.

Groupe		Résultat des tests post-hoc de Tukey
Francophones natifs		SE=5,31 ; t=9,36 ; p<0,001
Apprenants	Avancés	SE=6,64 ; t=4,74 ; p<0,001
	Intermédiaires	SE=8,93 ; t=1,87 ; p=0,04
	Débutants	SE=8,90 ; t= -1,78 ; p=0,08

- *Comparaison « inter-groupe »*

Les durées moyennes de V-FrontièreCT pour les 4 groupes de locuteurs ont déjà été illustrées dans la Figure 2.5. Comme le montre cette figure, plus le niveau de français est bas, plus la durée des V-FrontièreCT est longue, suggérant un débit plus lent. Selon les résultats du test ANOVA (Tableau 2.21), les durées des voyelles pour V-FrontièreCT diffèrent significativement entre les quatre groupes. Les résultats des tests post-hoc (Tableau 2.21) indiquent que ces durées chez les locuteurs natifs sont plutôt comparables à celles des apprenants avancés, tandis qu'elles diffèrent significativement de celles des apprenants intermédiaires et débutants.

Tableau 2.21 : Résultats des tests ANOVA et post-hoc de Tukey pour les différences de durées des V-FrontièreCT pour les quatre groupes de locuteurs.

Condition	Résultat du test ANOVA	
V-FrontièreCT	df=3, F=10,37, p<0,001, $\eta^2=0,092$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=6,85, t=2,10, p=0,16
	Natif vs Intermédiaire	SE=6,85, t=3,18, p=0,009
	Natif vs Débutant	SE=6,85, t=5,46, p<0,001
	Avancé vs Intermédiaire	SE=7,43, t= -1,00, p=0,75
	Avancé vs Débutant	SE=7,43, t= -3,10, p=0,01
	Intermédiaire vs Débutant	SE=7,43, t=2,10, p=0,16

2.3.2.2 Organisation temporelle

Pour analyser l'organisation temporelle des productions rythmiques chez les francophones natifs et les apprenants sinophones, nous avons calculé les valeurs des paramètres rythmiques quantitatifs suivants : %V, ΔC et ΔV , VarcoV et VarcoC, nPVI-V et rPVI-C, ainsi que CCI-C et CCI-V (voir Tableau 2.5 pour leurs définitions). Pour chacun de ces paramètres, les valeurs obtenues ont été analysées à l'aide d'un test ANOVA, en considérant le facteur « niveaux de français » comme un effet fixe. Ensuite, une série de tests post-hoc de Tukey a été réalisée afin de comparer les données « inter-groupes ». Nous présenterons ci-dessous les résultats obtenus pour chacun des neuf paramètres étudiés.

- Paramètres %V :

Les statistiques descriptives (Tableau 2.22) et les résultats du test ANOVA (Tableau 2.23) indiquent une différence significative entre les quatre groupes, avec la valeur de %V la plus élevée chez les apprenants débutants et la plus faible chez les francophones natifs (voir aussi la Figure 2.6).

Tableau 2.22 : Statistiques descriptives des valeurs %V pour les quatre groupes de locuteurs.

Locuteurs Paramètre	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
%V	Moyenne : 45,41 ET : 4,46	Moyenne : 46,92 ET : 5,14	Moyenne : 47,65 ET : 6,02	Moyenne : 49,89 ET : 5,47

Tableau 2.23 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de %V pour les quatre groupes de locuteurs.

Paramètre	Résultat du test ANOVA	
%V	df=3, F=20,59, p<0,001, $\eta^2=0,091$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=0,58, t=2,61, p=0,04
	Natif vs Intermédiaire	SE=0,58, t=3,89, p<0,001
	Natif vs Débutant	SE=0,58, t=7,77, p<0,001
	Avancé vs Intermédiaire	SE=0,63, t= -1,18, p=0,64
	Avancé vs Débutant	SE=0,63, t= -4,76, p<0,001
	Intermédiaire vs Débutant	SE=0,63, t= -3,58, p=0,002

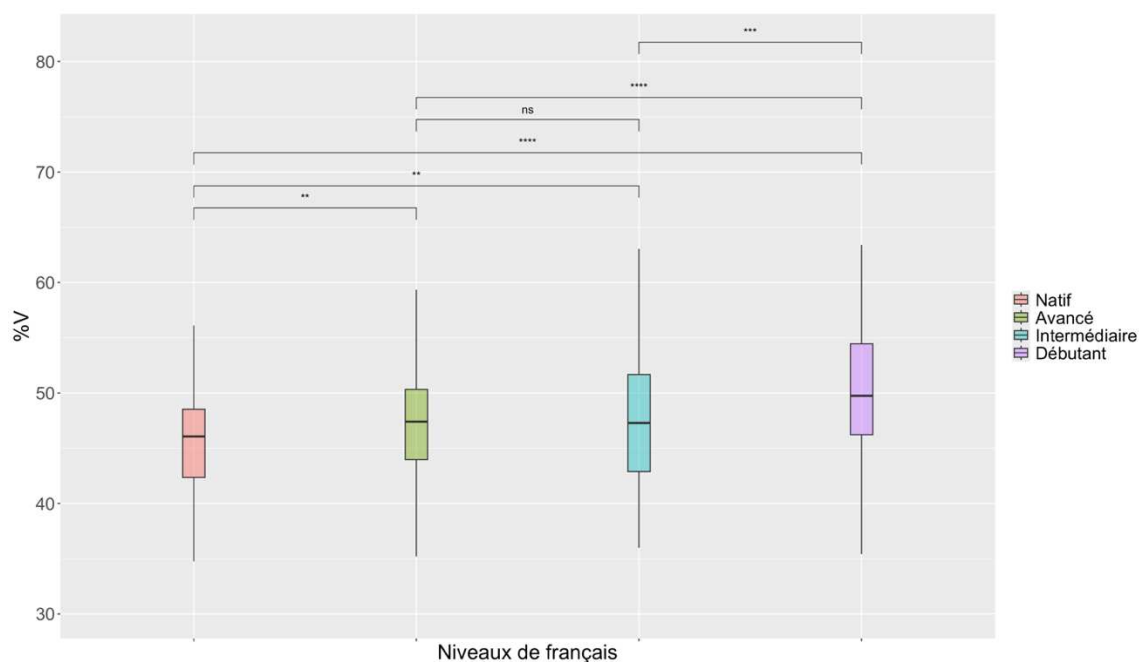


Figure 2.6 : Comparaison des valeurs de %V pour les quatre groupes de locuteurs.

Les tests post-hoc de Tukey montrent également que les différences de %V sont significatives dans les comparaisons inter-groupes, à l'exception de celles entre les apprenants avancés et intermédiaires, dont les valeurs de %V sont comparables. Ces résultats suggèrent que l'acquisition du rythme en français est un processus graduel. À mesure que le niveau de français progresse, la proportion vocalique dans la production rythmique diminue, se rapprochant ainsi d'un patron natif.

- Paramètres ΔV et ΔC :

Nous présentons les statistiques descriptives pour les deux paramètres bruts ΔV et ΔC dans le Tableau 2.24, ainsi que leurs distributions dans la Figure 2.7.

Tableau 2.24 : Statistiques descriptives des valeurs ΔV et ΔC pour les quatre groupes de locuteurs.

Locuteurs Paramètre	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
ΔV	Moyenne : 36,33 ET : 9,19	Moyenne : 49,08 ET : 13,91	Moyenne : 56,65 ET : 19,34	Moyenne : 73,06 ET : 22,69
ΔC	Moyenne : 48,06 ET : 10,65	Moyenne : 63,28 ET : 14,26	Moyenne : 66,67 ET : 14,65	Moyenne : 83,93 ET : 30,36

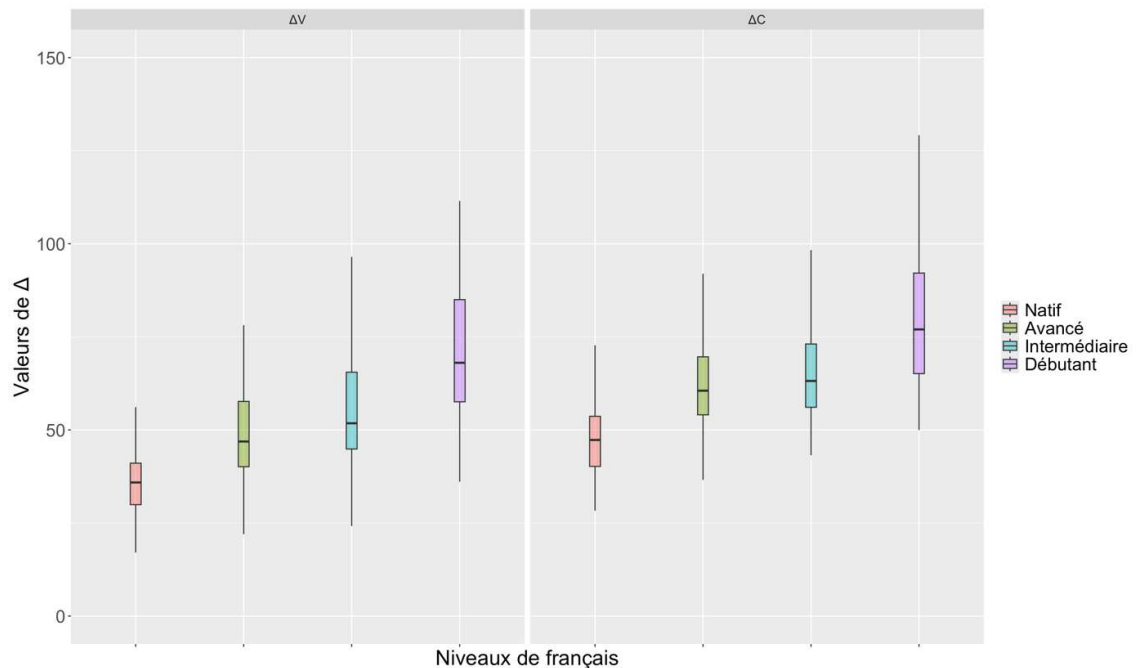


Figure 2.7 : Distribution des valeurs de ΔV (gauche) et ΔC (droite) pour les quatre groupes de locuteurs.

Comme indiqué dans le Tableau 2.24 et illustré sur la Figure 2.7, les valeurs de ΔV et ΔC diminuent à mesure que le niveau de français progresse. Cela signifie qu'au niveau natif, les variations des durées vocaliques et consonantiques sont les plus stables, avec une moindre variation. En revanche, au niveau débutant, ces durées présentent les variations les plus importantes, reflétant un patron rythmico-temporel instable.

En complément de la Figure 2.7, nous avons comparé ces deux paramètres entre les quatre groupes en s'inspirant de la célèbre figure de Ramus (1999) (Figure 2.8). De manière cohérente avec les observations précédentes, cette figure montre également que les durées vocaliques et consonantiques sont les plus stables chez les locuteurs natifs et les plus variables chez les apprenants débutants.

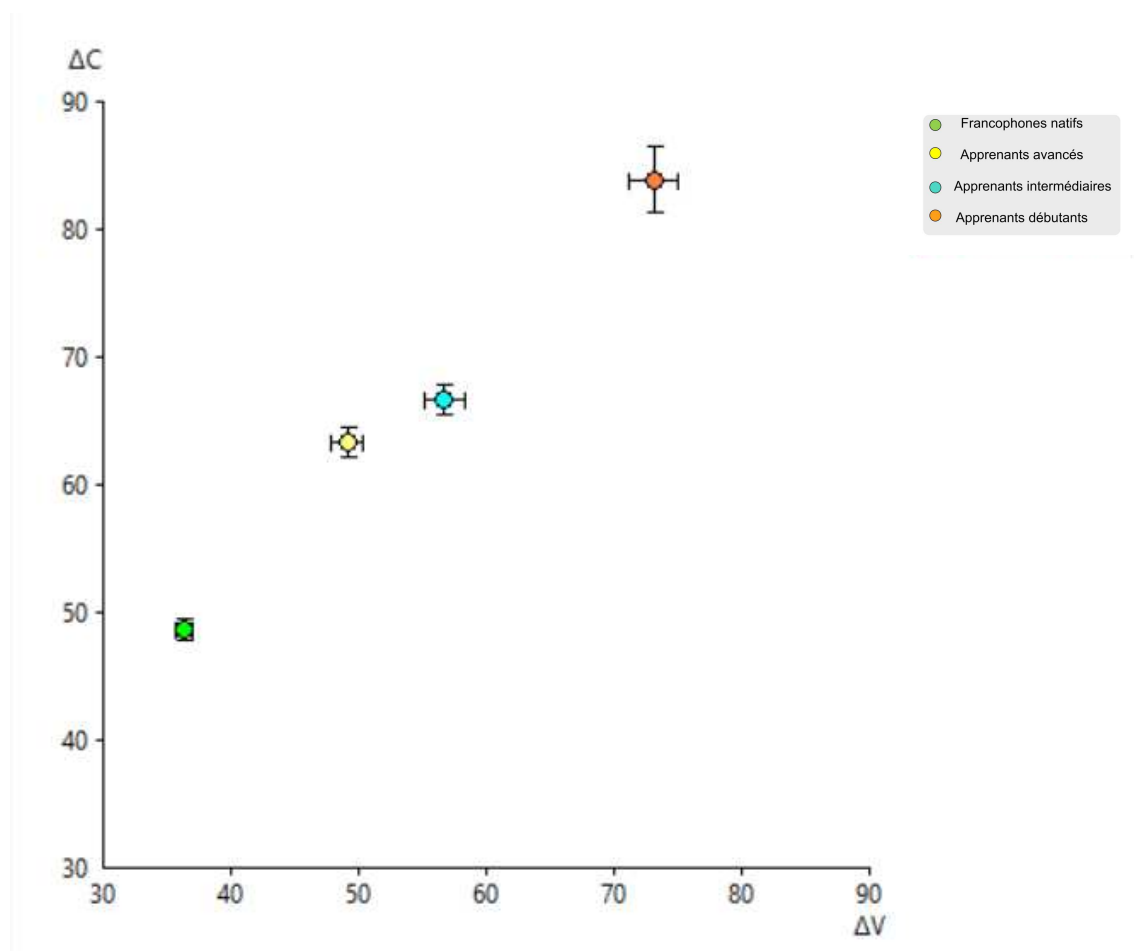


Figure 2.8 : Comparaison des valeurs de ΔV et ΔC pour les quatre groupes de locuteurs.

Statistiquement, tant pour le paramètre ΔV que pour celui de ΔC , les tests ANOVA et post-hoc de Tukey (Tableau 2.25) ont révélé les différences significatives entre les quatre groupes (sauf celles de ΔC entre apprenants avancés et intermédiaires).

Tableau 2.25 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de ΔV et ΔC pour les quatre groupes de locuteurs.

Paramètre	Résultat du test ANOVA	
ΔV	df=3, F=141,71, p<0,001, $\eta^2=0,41$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=1,82, t=7,02, p<0,001
	Natif vs Intermédiaire	SE=1,82, t=11,19, p<0,001
	Natif vs Débutant	SE=1,82, t=20,23, p<0,001
	Avancé vs Intermédiaire	SE=1,97, t= -3,85, p<0,001
	Avancé vs Débutant	SE=1,97, t= -12,18, p<0,001
	Intermédiaire vs Débutant	SE=1,97, t= -8,33, p<0,001
ΔC	df=3, F=105,69, p<0,001, $\eta^2=0,34$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=2,03, t=7,50, p<0,001
	Natif vs Intermédiaire	SE=2,03, t=9,17, p<0,001
	Natif vs Débutant	SE=2,03, t=17,68, p<0,001
	Avancé vs Intermédiaire	SE=2,20, t= -1,54, p=0,41
	Avancé vs Débutant	SE=2,20, t= -9,39, p<0,001
	Intermédiaire vs Débutant	SE=2,20, t= -7,84, p<0,001

- Paramètres *VarcoV* et *VarcoC* :

Nous reporterons ci-dessous les résultats pour les deux paramètres normalisés *VarcoV* et *VarcoC*. Dans le Tableau 2.26, nous présentons les statistiques descriptives.

Tableau 2.26 : Statistiques descriptives des valeurs *VarcoV* et *VarcoC* pour les quatre groupes de locuteurs.

Locuteurs Paramètre	Francophones	Apprenants		
	natifs	Avancés	Intermédiaires	Débutants
<i>VarcoV</i>	Moyenne :	Moyenne :	Moyenne :	Moyenne :
	44,82	47,30	47,62	45,57
<i>VarcoC</i>	ET : 9,11	ET : 9,86	ET : 9,57	ET : 9,07
	Moyenne :	Moyenne :	Moyenne :	Moyenne :
	52,02	56,57	54,09	54,14
	ET : 8,24	ET : 8,87	ET : 8,19	ET : 12,95

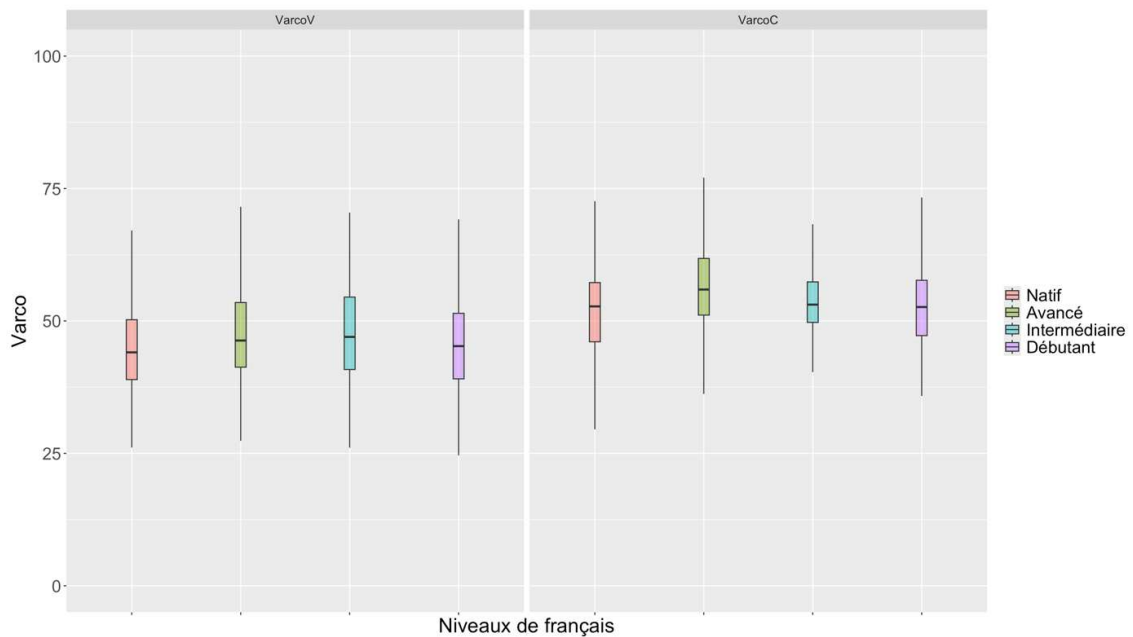


Figure 2.9 : Distribution des valeurs de VarcoV (gauche) et VarcoC (droite) pour les quatre groupes de locuteurs.

Comme le montrent le Tableau 2.26 et la Figure 2.9, les valeurs des deux paramètres normalisés VarcoV et VarcoC ne présentent pas la même tendance évidente que celles des paramètres bruts ΔV et ΔC . Les statistiques descriptives révèlent des moyennes très proches entre les quatre groupes pour ces deux paramètres. Nous avons aussi comparé ces deux paramètres entre les quatre groupes et les résultats, illustrés dans la Figure 2.10, montrent que les valeurs de VarcoV sont plutôt similaires entre les locuteurs natifs et les apprenants, tandis que pour VarcoC, elles sont les plus faibles chez les locuteurs natifs et les plus élevées chez les apprenants avancés. Cela est statistiquement confirmé par les tests ANOVA et les tests post-hoc de Tukey (Tableau 2.27). Pour le paramètre VarcoV, les résultats ne montrent aucune différence significative entre les groupes. En ce qui concerne le paramètre VarcoC, une seule différence significative a été observée entre les niveaux natif et avancé.

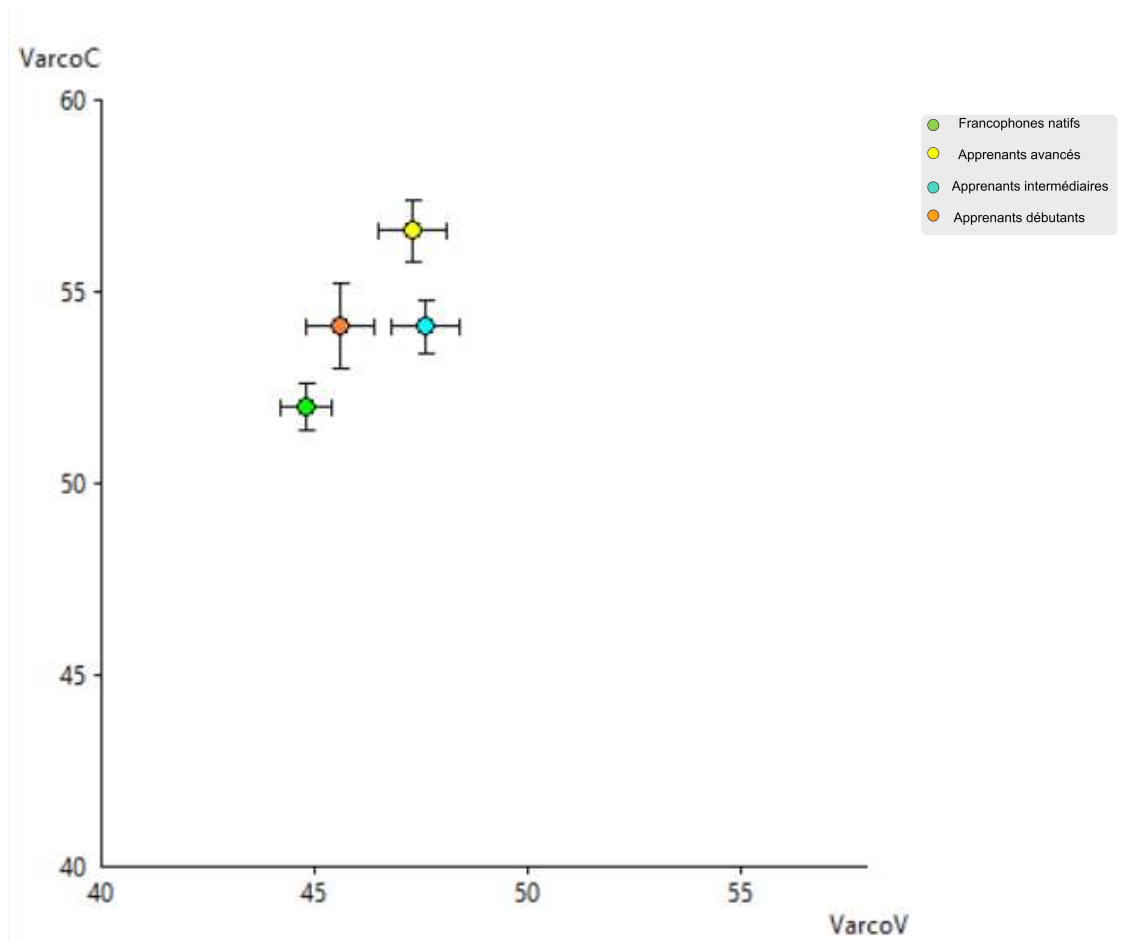


Figure 2.10 : Comparaison des valeurs de VarcoV et VarcoC pour les quatre groupes de locuteurs.

Pour rappel, VarcoV et VarcoC sont des valeurs normalisées. Cela montre qu'une fois le débit de parole normalisé, les patrons rythmiques chez les francophones natifs et les apprenants deviennent similaires, tant pour les intervalles vocaliques que pour les intervalles consonantiques.

Tableau 2.27 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de VarcoV et VarcoC pour les quatre groupes de locuteurs.

Paramètre	Résultat du test ANOVA	
VarcoV	df=3, F=3,37, p=0,018, $\eta^2=0,016$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=1,03, t=2,41, p=0,08
	Natif vs Intermédiaire	SE=1,03, t=2,71, p=0,04
	Natif vs Débutant	SE=1,03, t=0,73, p=0,89
	Avancé vs Intermédiaire	SE=1,12, t= -0,28, p=0,99
	Avancé vs Débutant	SE=1,12, t=1,54, p=0,41
	Intermédiaire vs Débutant	SE=1,12, t=1,82, p=0,26
VarcoC	df=3, F=6,18, p<0,001, $\eta^2=0,029$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=1,06, t=4,30, p<0,001
	Natif vs Intermédiaire	SE=1,06, t=1,96, p=0,21
	Natif vs Débutant	SE=1,06, t=2,00, p=0,19
	Avancé vs Intermédiaire	SE=1,15, t=2,16, p=0,14
	Avancé vs Débutant	SE=1,15, t=2,12, p=0,15
	Intermédiaire vs Débutant	SE=1,15, t= -0,04, p=1

- Paramètres nPVI-V et rPVI-C :

Nous récapitulons ci-dessous les statistiques descriptives pour nPVI-V et rPVI-C (Tableau 2.28). Les Figures 2.11 et 2.12 illustrent la distribution et la comparaison de ces deux paramètres entre les 4 groupes.

Tableau 2.28 : Statistiques descriptives des valeurs nPVI-V et rPVI-C pour les quatre groupes de locuteurs.

Locuteurs Paramètre	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
nPVI-V	Moyenne :	Moyenne :	Moyenne :	Moyenne :
	44,11	52,56	53,33	51,64
	ET : 8,35	ET : 10,52	ET : 10,15	ET : 11,53
rPVI-C	Moyenne :	Moyenne :	Moyenne :	Moyenne :
	55,60	71,03	75,73	92,78
	ET : 15,43	ET : 18,41	ET : 17,99	ET : 26,51

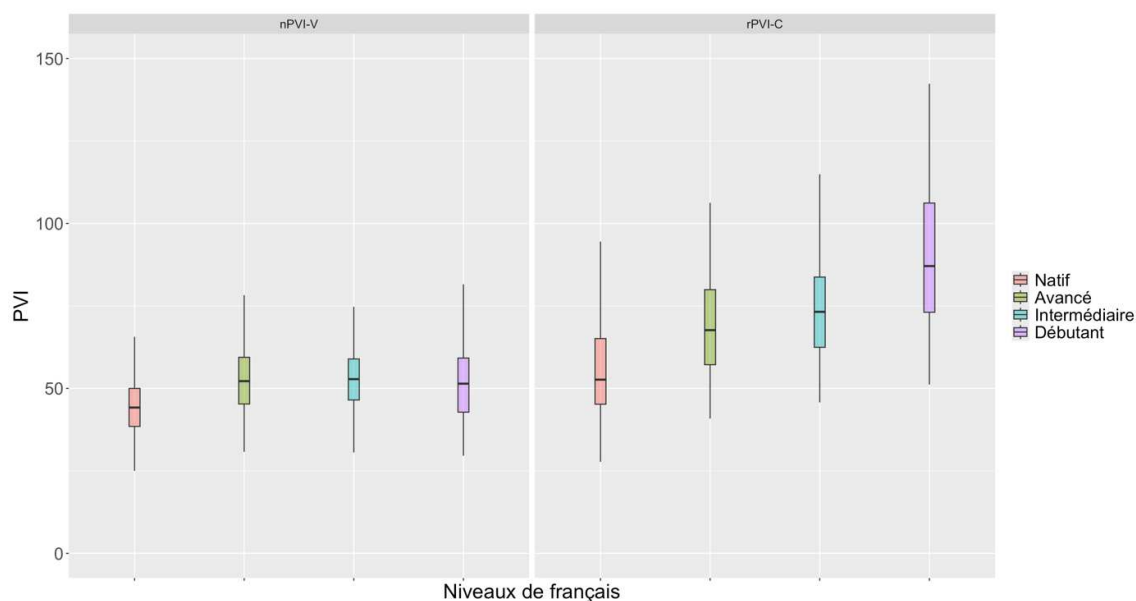


Figure 2.11 : Distribution des valeurs de nPVI-V et rPVI-C chez les quatre groupes de locuteurs.

Tableau 2.29 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de nPVI-V et rPVI-C pour les quatre groupes de locuteurs.

Paramètre	Résultat du test ANOVA	
nPVI-V	df=3, F=32,29, p<0,001, $\eta^2=0,14$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=1,11, t=7,64, p<0,001
	Natif vs Intermédiaire	SE=1,11, t=8,34, p<0,001
	Natif vs Débutant	SE=1,11, t=6,81, p<0,001
	Avancé vs Intermédiaire	SE=1,20, t= -0,64, p=0,92
	Avancé vs Débutant	SE=1,20, t=0,77, p=0,87
	Intermédiaire vs Débutant	SE=1,20, t=1,41, p=0,50
rPVI-C	df=3, F=100,71, p<0,001, $\eta^2=0,33$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=2,16, t=7,14, p<0,001
	Natif vs Intermédiaire	SE=2,16, t=9,31, p<0,001
	Natif vs Débutant	SE=2,16, t=17,20, p<0,001
	Avancé vs Intermédiaire	SE=2,35, t= -2,00, p=0,19
	Avancé vs Débutant	SE=2,35, t= -9,27, p<0,001
	Intermédiaire vs Débutant	SE=2,35, t= -7,27, p<0,001

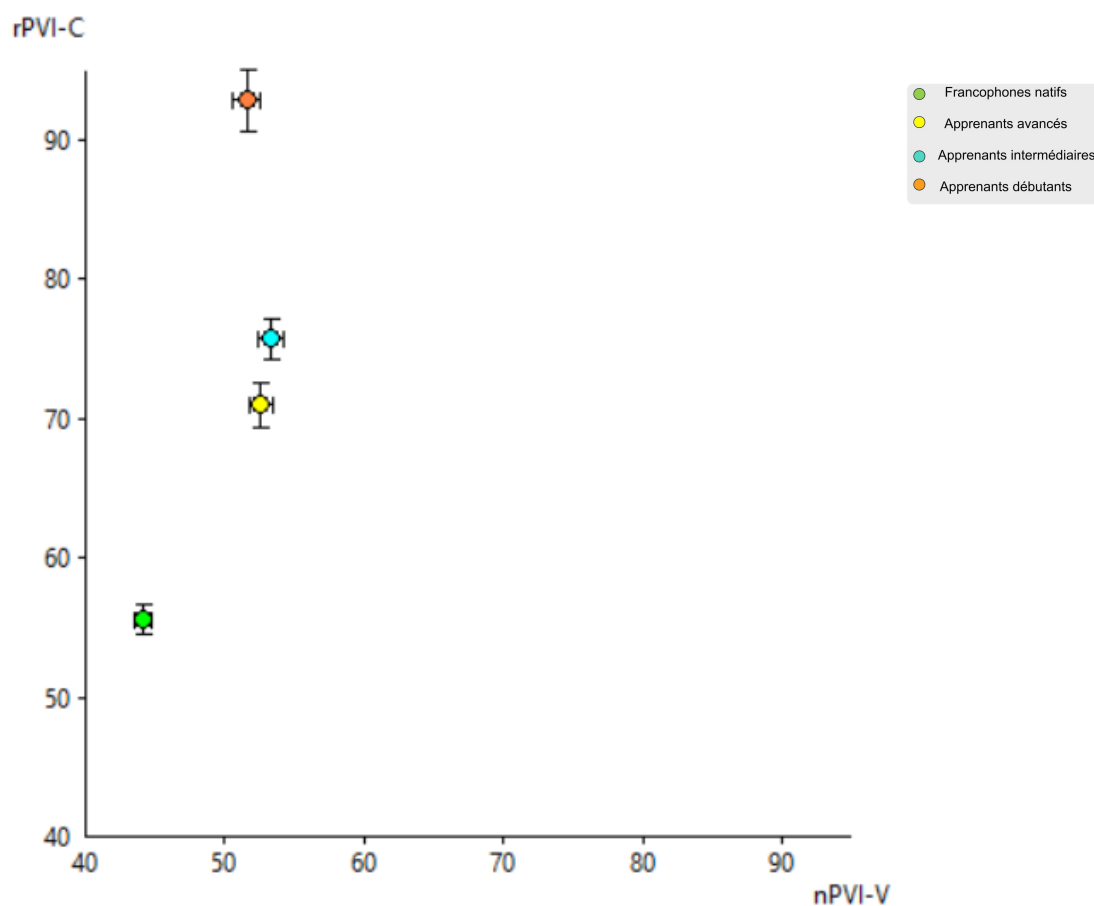


Figure 2.12 : Comparaison des valeurs de nPVI-V et rPVI-C pour les quatre groupes de locuteurs.

Pour les deux paramètres nPVI-V et rPVI-C, les valeurs les plus faibles sont observées chez les locuteurs natifs, ce qui suggère une plus grande stabilité des intervalles consonantiques et vocaliques successifs dans leur patron rythmique. Chez les apprenants, les valeurs de nPVI-V sont très similaires entre les trois niveaux. Concernant rPVI-C, les différences entre les groupes suivent la même tendance que celles observées pour ΔC : comme le montrent les statistiques descriptives (Tableau 2.28) et les Figures 2.11 et 2.12, plus le niveau est faible, plus la valeur est élevée. Les tests ANOVA et post-hoc de Tukey (Tableau 2.29) confirment ces observations. Pour nPVI-V, les valeurs des locuteurs natifs diffèrent significativement de celles des apprenants, tandis qu'elles restent comparables entre les trois niveaux d'apprenants. Cela indique que, même après normalisation du débit de parole, les productions des apprenants pour les intervalles vocaliques successifs ne sont pas encore comparables à celles des locuteurs natifs. En ce qui concerne rPVI-C, les valeurs

différent significativement entre les groupes, à l'exception des apprenants avancés et intermédiaires, dont les valeurs sont comparables.

- Paramètres CCI-C et CCI-V

Nous présenterons ci-dessous les statistiques descriptives des valeurs CCI-C et CCI-V pour les quatre groupes de locuteurs (Tableau 2.30), et illustrer leurs valeurs sur la Figure 2.13 afin de les mettre en évidence.

Tableau 2.30 : Statistiques descriptives des valeurs CCI-V et CCI-C chez les quatre groupes de locuteurs.

Locuteurs Paramètre	Francophones natifs	Apprenants		
		Avancés	Intermédiaires	Débutants
CCI-V	Moyenne : 33,58 ET : 8,13	Moyenne : 51,65 ET : 14,54	Moyenne : 60,66 ET : 21,59	Moyenne : 79,56 ET : 26,29
CCI-C	Moyenne : 38,47 ET : 9,40	Moyenne : 41,39 ET : 12,28	Moyenne : 44,42 ET : 13,33	Moyenne : 63,55 ET : 25,45

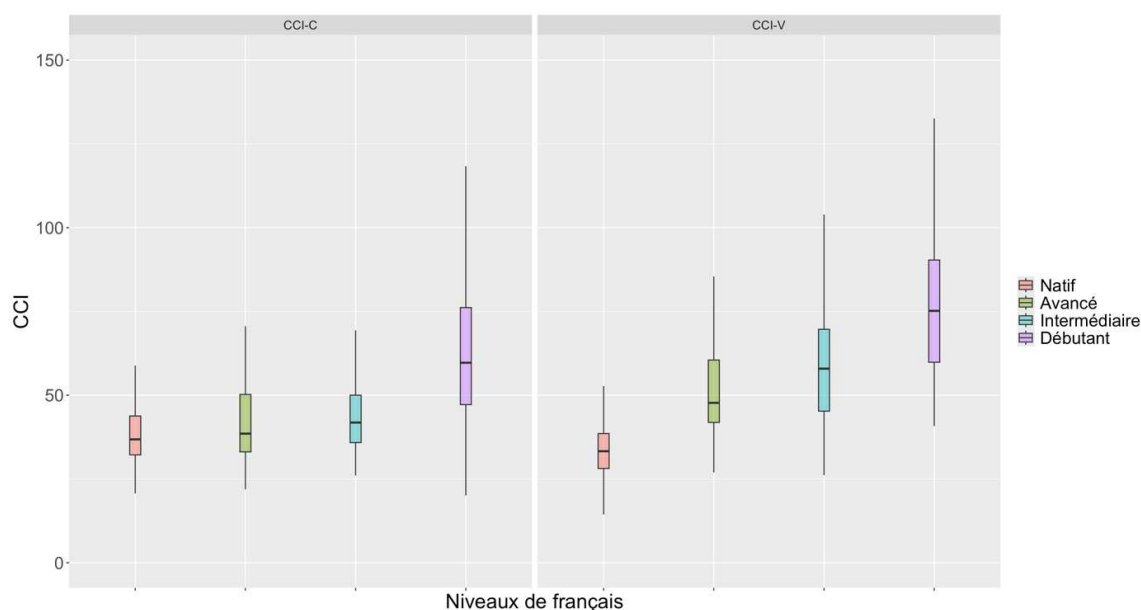


Figure 2.13 : Distribution des valeurs de CCI-C (gauche) et CCI-V (droite) pour les quatre groupes de locuteurs.

Étant donné que les paramètres CCI sont sensibles à la variation du débit de parole, les valeurs de CCI-C et CCI-V, à l'instar d'autres paramètres bruts, sont les plus élevées chez les apprenants débutants, suivies des intermédiaires et des avancés, et

elles sont les plus faibles chez les francophones natifs. Nous avons réalisé des tests ANOVA et des tests post-hoc de Tukey, dont les résultats sont présentés dans le Tableau 2.31. Pour le paramètre CCI-V, toutes les comparaisons post-hoc de Tukey ont révélé des différences significatives entre les groupes (Tableau 2.31). En revanche, pour le paramètre CCI-C, des patrons comparables ont été observés entre les locuteurs natifs et les apprenants avancés, ainsi qu'entre les apprenants avancés et intermédiaires.

Tableau 2.31 : Résultats des tests statistiques ANOVA et post-hoc de Tukey pour les valeurs de CCI-V et CCI-C pour les quatre groupes de locuteurs.

Paramètre	Résultat du test ANOVA	
CCI-V	df=3, F=183,51, p<0,001, $\eta^2=0,47$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=2,00, t=9,03, p<0,001
	Natif vs Intermédiaire	SE=2,00, t=13,52, p<0,001
	Natif vs Débutant	SE=2,00, t=22,96, p<0,001
	Avancé vs Intermédiaire	SE=2,17, t= -4,15, p<0,001
	Avancé vs Débutant	SE=2,17, t= -12,85, p<0,001
	Intermédiaire vs Débutant	SE=2,17, t= -8,70, p<0,001
CCI-C	df=3, F=77,11, p<0,001, $\eta^2=0,16$	
	Tests post-hoc de Tukey	
	Natif vs Avancé	SE=1,74, t=1,68, p=0,34
	Natif vs Intermédiaire	SE=1,74, t=3,42, p=0,004
	Natif vs Débutant	SE=1,74, t=14,43, p<0,001
	Avancé vs Intermédiaire	SE=1,89, t= -1,61, p=0,37
	Avancé vs Débutant	SE=1,89, t= -11,76, p<0,001
	Intermédiaire vs Débutant	SE=1,89, t= -10,15, p<0,001

Les études de Bertinotti et Bertini (2010), ainsi que de Bertinetto *et al.* (2012), qui ont examiné les patrons rythmiques à l'aide des paramètres CCI, ont classé le français et le chinois mandarin comme des langues plutôt contrôlantes au niveau I (cf. Tableau 1.13). En nous appuyant sur les travaux de Bertinetto et Bertini (2010), nous avons représenté les valeurs des paramètres CCI-C et CCI-V des quatre groupes en reproduisant la Figure 2.14.

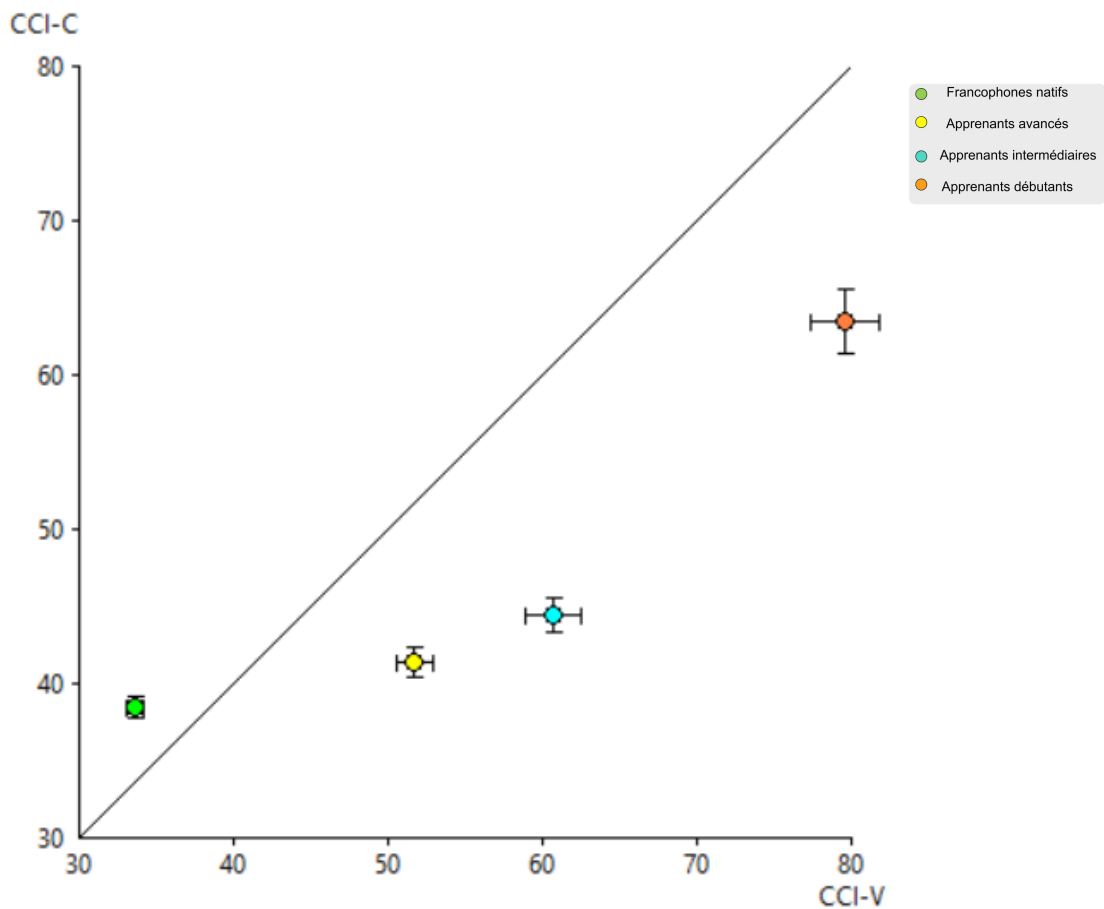


Figure 2.14 : Comparaison des valeurs de CCI-V et CCI-C pour les quatre groupes de locuteurs.

Comme illustré sur la Figure 2.14, les valeurs de CCI-C et CCI-V des francophones natifs s'alignent près de la ligne bisectrice, ce qui reflète une production contrôlée (ou syllabique). Cela indique une fluctuation quasi identique entre les intervalles consonantiques et vocaliques chez les locuteurs natifs. Cette observation est confirmée quantitativement par les valeurs assez proches de CCI-V (33,58) et CCI-C (38,47) chez les natifs.

En revanche, chez les apprenants, les valeurs de CCI-C et CCI-V sont plus élevées que celles des natifs et ne sont pas identiques entre elles, s'éloignant ainsi de la ligne bisectrice. Cela suggère une plus grande variabilité dans les mouvements articulatoires des apprenants, traduisant des fluctuations segmentales plus variables par rapport à la stabilité observée chez les natifs. Cependant, à mesure que le niveau des apprenants progresse vers celui des natifs, les valeurs des CCI et leurs écarts-types diminuent progressivement (Tableau 2.30, Figure 2.14). De plus, l'écart entre CCI-C et CCI-V se réduit : chez les débutants, la différence est de 14,67 ; chez

les intermédiaires, de 12,14 ; chez les avancés, de 6,38 ; et chez les natifs, de 4,37 (cf. Tableau 2.30).

Ces résultats permettent de conclure que, au fil de leur apprentissage, les apprenants produisent des fluctuations de plus en plus comparables et proches entre les intervalles vocaliques et consonantiques. Cela reflète une évolution vers un patron rythmique davantage contrôlant et syllabique, similaire à celui des locuteurs natifs.

2.4 Discussion

Notre étude a permis d'identifier un ensemble de paramètres acoustiques qui caractérisent les stratégies de segmentation et d'organisation prosodique mises en œuvre par les locuteurs natifs et les apprenants lors de la production de paires de phrases localement ambiguës. Nous discuterons de ces paramètres dans les deux sections suivantes : la section 2.4.1 sera consacrée à la segmentation prosodique, tandis que la section 2.4.2 traitera de l'organisation prosodique.

2.4.1 Segmentation prosodique

Nous tenterons de répondre dans cette section aux deux questions de recherche posées au début de notre étude, en nous appuyant sur les résultats obtenus :

1. Quels indices acoustiques les locuteurs natifs francophones utilisent-ils pour marquer les frontières prosodiques dans des phrases localement ambiguës ?
2. Les apprenants sinophones réussissent-ils à segmenter ces frontières en utilisant les indices mélodiques et temporels typiques du français ?

Nous commencerons par récapituler les résultats concernant la segmentation mélodique et ensuite temporelle, en mettant en lumière les corrélats acoustiques identifiés chez les quatre groupes de locuteurs. Nous discuterons notamment des performances des apprenants, en explorant les variations liées à leur niveau de maîtrise du français et les facteurs qui pourraient expliquer ces différences.

- Segmentation mélodique

Pour chaque paire de phrases localement ambiguës, nous avons mesuré les différences de F0 entre les voyelles identiques situées dans deux positions prosodiques distinctes : l'une précédant immédiatement la frontière prosodique et

l'autre ne la précédant pas. Cette analyse nous a permis de déterminer si les locuteurs marquaient les frontières prosodiques par des montées de F0, et comment ces variations étaient influencées par leur niveau de maîtrise du français.

Les résultats ont montré que les locuteurs natifs produisaient une montée de F0 significative sur les dernières voyelles précédant les frontières prosodiques, par rapport à celles qui ne les précédaient pas. Ce phénomène a été observé dans les deux conditions, CP et CT. Parmi les groupes d'apprenants sinophones, les apprenants de niveaux avancé et intermédiaire ont également montré des montées de F0 plus importantes sur les voyelles précédant les frontières dans les conditions CP et CT. En revanche, les apprenants débutants n'ont pas présenté de différences significatives entre voyelles précédant ou pas les frontières, et ce indépendamment de la condition prosodique.

En résumé, la montée de F0 a été efficacement utilisée par les locuteurs natifs ainsi que par les apprenants avancés et intermédiaires comme un indice mélodique permettant de signaler les différentes frontières prosodiques (CP et CT) dans des phrases localement ambiguës. En revanche, les apprenants débutants n'ont pas encore développé la capacité d'utiliser cet indice pour distinguer ces frontières prosodiques.

- *Segmentation temporelle*

Pour évaluer si l'allongement final était utilisé comme indice temporel dans la segmentation prosodique, nous avons analysé les durées des mêmes voyelles examinées lors de l'analyse de la segmentation mélodique. Les résultats montrent que les locuteurs natifs, ainsi que les apprenants de niveaux intermédiaire et avancé, ont produit une durée significativement plus longue sur les voyelles situées immédiatement avant les frontières prosodiques, tant dans la condition CP que CT. En revanche, chez les apprenants débutants, les durées des voyelles précédant ou non les frontières prosodiques étaient comparables. Autrement dit, ces derniers n'étaient pas encore capables de signaler les frontières prosodiques à l'aide de l'allongement final.

Notre première hypothèse, formulée au début de cette étude sur les locuteurs natifs (« Les francophones natifs marqueront les frontières prosodiques à l'aide des deux indices acoustiques : l'allongement final et la montée mélodique. »), a été confirmée.

Les locuteurs natifs ont utilisé ces deux indices simultanément pour signaler les frontières prosodiques CP et CT. Cependant, notre deuxième hypothèse, concernant les apprenants sinophones (« Les apprenants pourraient utiliser l’allongement final et la montée de F0 sur les dernières syllabes pour marquer les frontières prosodiques, quel que soit leur niveau en français. »), n’a été que partiellement validée. En effet, seuls les apprenants de niveaux avancé et intermédiaire ont été capables d’utiliser ces deux indices pour marquer les frontières prosodiques.

Les résultats correspondent donc globalement à nos attentes, à l’exception des apprenants débutants qui n’ont exploité ni la montée de F0 ni même l’allongement final, pourtant un indice commun au français et au mandarin. Nous proposons trois explications à ces difficultés.

La première explication est liée à la compréhension linguistique limitée des débutants. En raison d’un apprentissage récent, les débutants maîtrisaient mal le vocabulaire et la syntaxe des phrases de notre corpus. Malgré une phase de préparation avant la lecture, ils peinaient probablement à comprendre le sens global de ces phrases, et donc à détecter leur ambiguïté locale. Comme le montrent les études d’Allbritton et al. (1996) et Ferreira & Dell (2000), la production d’indices prosodiques dépend principalement de la conscience des locuteurs de la présence de l’ambiguïté dans une phrase. Or, dans notre protocole, les phrases étaient présentées en ordre aléatoire, sans regroupement par paires, pour éviter que les locuteurs (notamment natifs) ne prennent conscience de cette ambiguïté et modifient artificiellement leurs productions. Si les apprenants avancés et les natifs ont pu s’appuyer sur leur compréhension et probablement aussi sur la ponctuation (e.g., virgules) pour segmenter les phrases, les débutants, confrontés à des difficultés lexico-syntaxiques, n’ont pas pu accéder aux nuances sémantiques nécessaires pour une bonne segmentation.

Une deuxième explication possible est l’incapacité des apprenants débutants à organiser les mots en groupes rythmiques, une compétence clé en prosodie. La prosodie implique à la fois de segmenter la parole et de l’organiser en unités cohérentes. Or, les débutants produisaient les phrases souvent mot à mot, avec des pauses fréquentes après chaque substantif. Ce phénomène, qui est assez fréquent chez les débutants, est illustré dans les deux figures 2.15 et 2.16 ci-dessous. Cette parole hachée empêchait toute hiérarchie prosodique naturelle. Plutôt que de lier

les mots en groupes rythmiques (e.g., syntagmes), les apprenants, notamment débutants, isolaient chaque élément, rendant impossible l'émergence de contours intonatifs ou temporels significatifs.

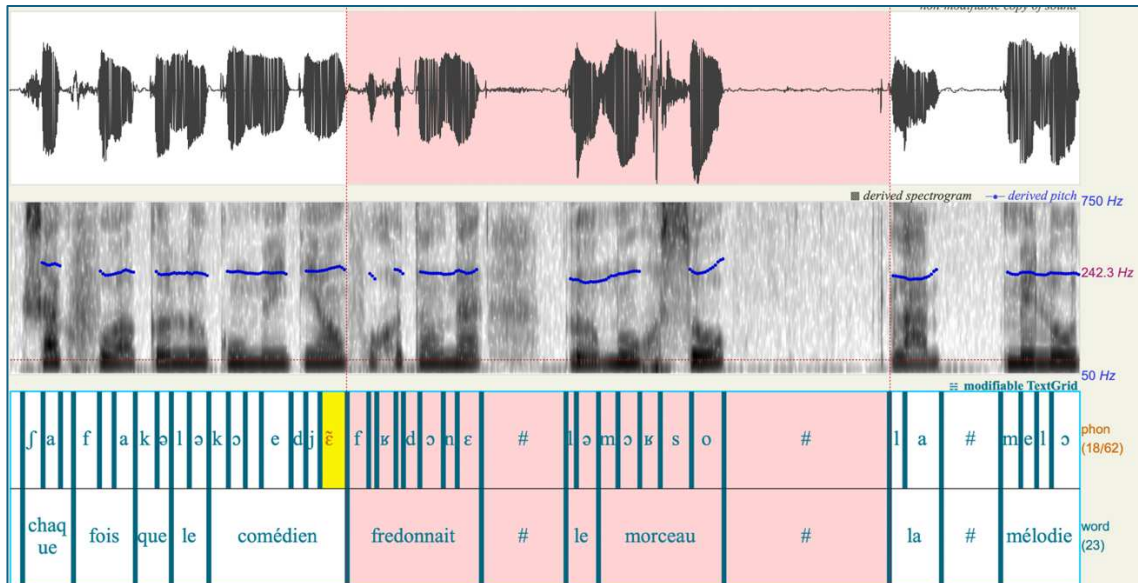


Figure 2.15 : La production de la phrase « Chaque fois que le comédien fredonnait le morceau, la mélodie était jolie à l'oreille. » dans la condition CT, par un apprenant débutant.

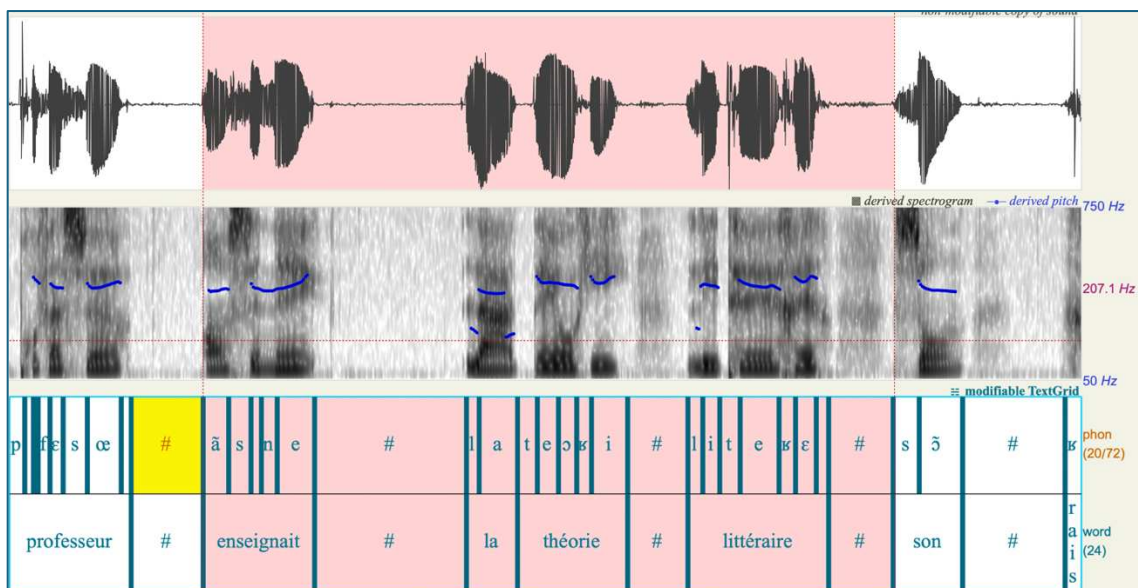


Figure 2.16 : La production de la phrase « Chaque fois que le professeur enseignait la théorie littéraire, son raisonnement devenait moins obscur. » dans la condition CT, par un apprenant débutant.

La troisième explication est liée à l'impact du débit de parole lent sur la production des débutants. Chez les natifs, l'allongement final des voyelles avant une frontière prosodique CP ou CT émerge grâce à un débit régulier et une quasi-isochronie

syllabique, impliquant des durées similaires, sauf pour la dernière syllabe. Comme nous l'avons montré plus haut, le débit de parole est lent chez les débutants, une caractéristique typique des apprenants de L2, selon Polyanskaya *et al.* (2017) (voir aussi Ordin & Polysanskaya (2015) ; Polyanskaya & Ordin (2019)). Ce débit lent a uniformisé les durées des voyelles, même celles qui précèdent immédiatement une frontière prosodique. Combiné à des pauses fréquentes, cela a eu pour effet d'atténuer fortement l'allongement final. En revanche, dès le niveau intermédiaire, l'amélioration du débit et de la fluidité a permis de distinguer acoustiquement les voyelles aux frontières prosodiques, confirmant que cet indice devient opérationnel avec la maîtrise progressive de la langue.

Pour résumer, les difficultés des débutants relèvent d'un ensemble de facteurs liés à la compréhension, au rythme, et au débit de parole. Ces difficultés seront par la suite résorbées à mesure que l'acquisition des apprenants progresse.

2.4.2 Organisation prosodique

Dans cette section, nous abordons l'autre question posée au début de notre étude : comment les patrons temporels prosodiques se manifestent-ils chez les francophones natifs et les apprenants sinophones, et comment évoluent-ils en fonction des niveaux en français ? Pour répondre à cette question, nous avons analysé neuf paramètres rythmiques (%V, ΔC , ΔV , VarcoV, VarcoC, nPVI-V, rPVI-C, CCI-C et CCI-V) afin de comparer les productions des locuteurs natifs et des apprenants de niveaux débutant, intermédiaire et avancé.

Les résultats obtenus montrent que les paramètres rythmiques reflètent assez clairement l'évolution de l'acquisition prosodique en français. Le français et le mandarin, bien qu'ils soient tous les deux des langues à isochronie syllabique, présentent quelques différences. Les valeurs des paramètres %V, ΔV et ΔC sont systématiquement plus élevées en mandarin qu'en français, comme l'ont aussi montré Romano et Mairano (2010). Nous avons observé la même tendance chez les apprenants : les débutants affichent des valeurs proches du mandarin, tandis que les apprenants avancés convergent vers celles des natifs. Ces résultats suggèrent qu'au début de l'apprentissage, les apprenants produisent des patrons rythmiques similaires au mandarin, mais que ces patrons évoluent progressivement vers ceux du français avec l'augmentation du niveau de maîtrise.

Pour les paramètres normalisés (VarcoV, VarcoC et nPVI-V), nos résultats indiquent que l'organisation temporelle des trois groupes d'apprenants est comparable à celle des locuteurs natifs, une fois le facteur de débit de parole neutralisé. Cette proximité entre natifs et apprenants s'explique par deux éléments. Premièrement, le transfert positif du mandarin au français est favorisé par leur typologie commune basée sur l'isochronie syllabique. Deuxièmement, les deux langues partagent une structure syllabique assez similaire, notamment avec une prédominance des syllabes ouvertes, ce qui constitue un avantage pour les apprenants et facilite leur acquisition des patrons rythmiques du français.

Cependant, pour d'autres paramètres rythmiques, notamment rPVI-C, CCI-C et CCI-V, des différences significatives apparaissent entre les groupes. Ces paramètres bruts montrent une tendance similaire à celle des valeurs de ΔV et ΔC : plus le niveau d'apprentissage est élevé, plus les valeurs se rapprochent de celles des natifs. Le paramètre rPVI-C, qui mesure la stabilité des intervalles consonantiques, révèle la difficulté des apprenants à produire les groupes consonantiques complexes du français. Contrairement au mandarin, dont la structure syllabique prototypique (CGVX) est simple et relativement stable, le français autorise des groupes consonantiques plus complexes (plusieurs clusters CC dans nos données). Ces structures plus variées peuvent constituer un défi, notamment pour les débutants, entraînant une instabilité dans leurs productions. Pour autant, nos analyses montrent une progression graduelle : à mesure que le niveau en français augmente, les apprenants parviennent à stabiliser leurs productions et à se rapprocher là aussi des valeurs des locuteurs natifs.

Les résultats des paramètres CCI (CCI-C et CCI-V) confirment cette tendance. Chez les locuteurs natifs, les valeurs CCI-C et CCI-V sont quasi égales, reflétant un patron rythmique « contrôlant », caractérisé par des fluctuations comparables entre intervalles vocaliques et consonantiques. À l'inverse, chez les apprenants, plus le niveau est bas, plus la production est instable, avec des différences marquées entre CCI-C et CCI-V. Ces écarts diminuent progressivement du niveau débutant au niveau avancé, indiquant une acquisition graduelle des structures rythmiques complexes du français.

En résumé, ces analyses mettent en évidence un double processus dans l'acquisition prosodique des apprenants sinophones. D'un côté, des similarités structurales entre

le mandarin et le français, telles que la prédominance des syllabes ouvertes, facilitent un transfert rythmique positif. D'un autre côté, la complexité des groupes consonantiques français constitue un obstacle, mais qui est surmonté progressivement au fil de l'apprentissage.

Les résultats de nos analyses permettent de tirer plusieurs conclusions importantes concernant l'acquisition du rythme en français par des apprenants sinophones.

Premièrement, en ce qui concerne l'étude du rythme langagier, bien que les paramètres quantitatifs rythmiques aient été critiqués dans des travaux antérieurs, les neuf paramètres utilisés dans notre analyse se sont révélés homogènes et pertinents pour décrire le processus d'acquisition rythmique d'une langue seconde (L2). Les paramètres bruts, non normalisés pour le débit, montrent des valeurs plus élevées chez les apprenants débutants et plus faibles chez les locuteurs natifs, reflétant une progression claire. En revanche, les paramètres normalisés pour le débit indiquent des valeurs comparables entre les natifs et les apprenants, quel que soit leur niveau, ce qui souligne l'importance de neutraliser l'effet du débit dans l'analyse des patrons rythmiques.

Deuxièmement, nous avons introduit deux paramètres moins étudiés dans la recherche sur l'acquisition du rythme en L2 : les indices CCI-V et CCI-C. Ces paramètres, qui mesurent les fluctuations dans les intervalles vocaliques et consonantiques, ont apporté de nouvelles perspectives sur les patrons rythmiques des locuteurs. Nos résultats montrent que le patron rythmique des natifs est de type « contrôlant », caractérisé par une fluctuation équilibrée entre les intervalles vocaliques et consonantiques, comme l'ont déjà révélé des études antérieures (Bertinetto & Bertini, 2010 ; Bertinetto *et al.*, 2012). En revanche, les apprenants présentent des fluctuations plus variables, reflétant leur progression dans l'acquisition du rythme. Ces paramètres, utilisés pour la première fois dans une étude sur l'acquisition du rythme en L2, ont permis de capturer des caractéristiques rythmiques inédites chez les apprenants. Cependant, notre analyse s'est limitée au niveau I, et il serait pertinent d'inclure le niveau II dans des recherches futures pour approfondir ces observations.

Troisièmement, le débit de parole s'est révélé être un facteur clé dans l'acquisition des aspects temporels. D'une part, il influence la fiabilité des paramètres non normalisés, rendant difficile la comparaison des patrons rythmiques entre natifs et

apprenants. D'autre part, un débit lent, typique des apprenants débutants, réduit la prééminence temporelle, notamment l'allongement final, dans leurs productions rythmiques. Ce problème macrorythmique constitue un obstacle majeur à l'acquisition d'un rythme proche de celui des natifs.

Notre étude représente une contribution significative à la recherche sur l'acquisition de la prosodie en L2, en combinant l'analyse des aspects mélodiques et temporels, ce qui est rarement fait dans la littérature. En effet, selon Trouvain & Braun (2020 : 605) : « *Research combining L2 melodic and timing aspects is still sparse, as noted by Mennen and de Leeuw (2014)* ». Cette étude constitue également une avancée dans l'étude de l'acquisition de la prosodie française par des apprenants sinophones. Contrairement aux travaux antérieurs, souvent fragmentés et limités à un seul aspect prosodique ou à un seul niveau d'apprenants, notre étude a analysé simultanément les aspects mélodiques et temporels auprès d'apprenants de trois niveaux (débutants, intermédiaires et avancés).

Pour finir, cette étude s'est concentrée sur une analyse acoustique de la production des phrases localement ambiguës en français, afin d'examiner comment les natifs et les apprenants désambigüisent ces phrases par la prosodie. Une étude complémentaire de perception, basée sur une tâche comportementale, sera présentée dans le chapitre suivant. Elle explorera comment les natifs et les apprenants traitent en temps réel les structures prosodiques ambiguës CP et CT.

Si vous avez déjà passé une ou plusieurs épreuve(s) suivante(s), merci de nous renseigner le résultat de la dernière session :

Quel type d'épreuve	Quand ? (mois/année)	Dans quel pays ?	Niveau équivalent du CECRL
TCF/TEF			Cliquez ici
DELFF			Cliquez ici
DALF			Cliquez ici
Autre (à préciser) :			Cliquez ici

Je n'ai jamais passé d'épreuve sur mon niveau de compétence du français.

Globalement, comment vous vous évaluez votre niveau de compétence du français actuel :

bilingue avancé intermédiaire débutant

Sur une échelle de zéro à dix, choisissez dans le menu votre niveau de compétence à l'oral et en compréhensions du français :

Oral	Cliquez ici	Compréhension de la langue orale	Cliquez ici	Lecture	Cliquez ici
------	-------------	----------------------------------	-------------	---------	-------------

Sur une échelle de zéro à dix, évaluez-vous dans quelle mesure vous êtes actuellement exposé(e) au français dans les contextes suivants :

Interagir avec vos amis	Cliquez ici	Interagir avec vos collègues (à l'école ou au travail)	Cliquez ici
Interagir avec la famille	Cliquez ici	Regarder la télévision / Écouter la radio	Cliquez ici
Lecture (romans, journaux, etc.)	Cliquez ici	Auto-apprentissage / Auto-perfectionnement	Cliquez ici

Parlez-vous d'autre(s) langue(s) étrangère(s) excepté le français ? Oui Non

(**Attention** : « Parler » ici sous-entend que vous pouvez vous exprimer dans cette langue de façon claire et détaillée avec un degré d'aisance.)

Si oui, précisez laquelle (lesquelles) :

Âge d'apprentissage de cette(ces) langue(s) étrangère(s) : ans.

Avez-vous déjà eu des problèmes de vision , problèmes d'audition , troubles du langage , ou troubles d'apprentissage ? (Cochez toutes les cases applicables).

Seriez-vous d'accord que nous vous contacterons pour vous inviter à participer à nos expériences prochaines (avec une rémunération) ? Oui Non Peut-être

Merci beaucoup de votre temps !

Chapitre 3 Étude perceptive

3.1 Introduction

L'objectif de ce chapitre est d'examiner comment les francophones natifs et les apprenants sinophones perçoivent les frontières prosodiques en français. Plus précisément, comment ils traitent les frontières prosodiques à différentes positions pour désambiguïser les structures syntaxiques ambiguës en français.

La plupart des études qui ont examiné comment la prosodie aide à désambiguïser des phrases syntaxiquement ambiguës ont porté sur le traitement « *off-line* », en s'intéressant à la désambiguïstation syntaxique une fois que l'auditeur a écouté la phrase entière (Beach, 1991 ; Warren *et al.*, 1995). Comparativement, peu de travaux ont examiné comment la prosodie aide à désambiguïser la syntaxe en temps réel (le traitement « *on-line* »). Pour le français, il y a notamment eu les travaux de Millotte *et al.* (2007, 2008) qui ont examiné la désambiguïstation au niveau du syntagme phonologique, avec les paires de phrases comme :

(a.) [[Le petit chien]_{PP} [mord la laisse]_{PP} [qui le retient]_{PP}]_{IPH}

(b.) [[Le petit chien mort]_{PP} [sera enterré demain]_{PP}]_{IPH}

(« PP » pour « syntagme phonologique » (*phonological phrase*) ; « IPH » pour « syntagme intonatif » (*intonational phrase*))

Leurs études de production et de perception ont montré que les francophones natifs ont spontanément produit les indices prosodiques (montée de F0 et allongement final) qui différencient ce type de phrases ambiguës, sans même se rendre compte de l'ambiguïté. D'autre part, les francophones natifs ont pu recourir à ces mêmes indices prosodiques pour prédire la structure syntaxique et résoudre l'ambiguïté syntaxique dans la perception. Dans Millotte *et al.* (2008), ils ont aussi montré que les auditeurs avaient un temps de traitement plus court pour le mot ambigu qui pouvait être attaché directement au syntagme phonologique précédent.

Dans ce chapitre, nous avons réalisé une expérience de perception afin d'examiner comment les francophones natifs et les apprenants sinophones du français traitent en temps réel les indices prosodiques du français quand ils sont confrontés à des phrases ambiguës. Nous avons examiné, à l'aide de paires de phrases localement ambiguës telles que (c.) et (d.) ci-dessous, si les auditeurs, en se basant sur les

indices prosodiques présents aux positions de « IPH1 », traitent correctement le syntagme nominal « *le rat* ». Dans la phrase (c.), « *le rat* » est « sujet », tandis que dans la phrase (d.), il fonctionne comme « complément d'objet direct ».

(c.) [Chaque fois que le serpent mangeait]_{IPH1} [**le rat** se cachait]_{IPH2}

(d.) [Chaque fois que le serpent mangeait **le rat**]_{IPH1} [le lapin se cachait]_{IPH2}

Inspirée de Millotte *et al.* (2007, 2008), notre étude utilise le même protocole expérimental, avec comme principal objectif de répondre aux questions suivantes :

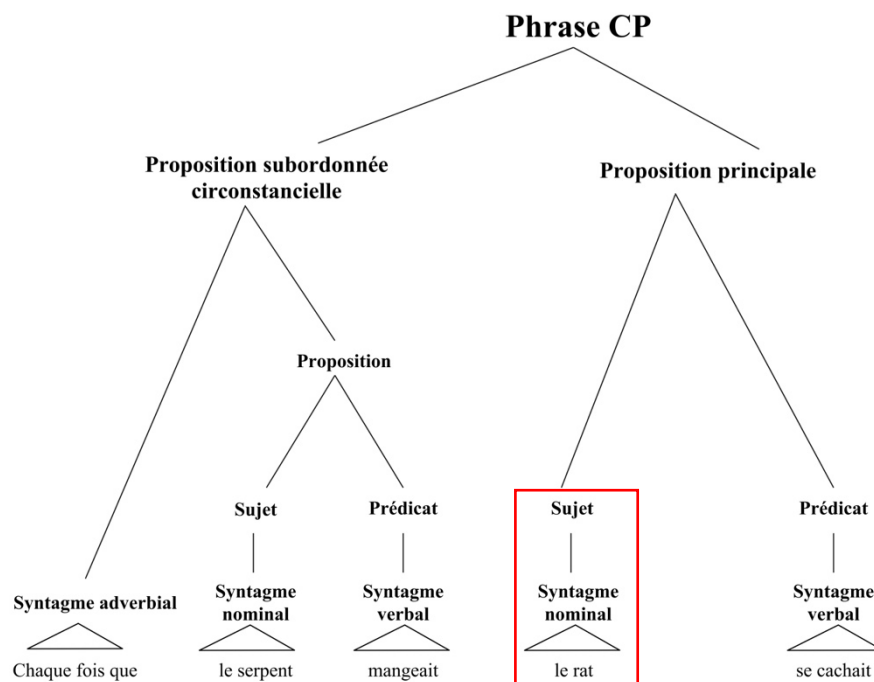
- 1) Les auditeurs francophones et sinophones sont-ils capables, en se basant uniquement sur les indices prosodiques, de déterminer les fonctions syntaxiques d'un ensemble de syntagmes nominaux inclus dans des phrases ambiguës en français ?
- 2) Si les auditeurs parviennent à distinguer les syntagmes ambigus grâce à la prosodie, leurs performances varient-elles en fonction de leur niveau en français et/ou des structures prosodiques ambiguës ?
- 3) Enfin, la désambiguïsation syntaxique par la prosodie est-elle influencée par la transitivité des verbes ?

Sur la base des conclusions de Millotte *et al.* (2007, 2008), nous pouvons légitimement supposer que les auditeurs francophones natifs seront parfaitement capables de distinguer les fonctions syntaxiques à l'aide des indices prosodiques. Par ailleurs, étant donné que l'allongement final et la pause sont également utilisés en chinois mandarin pour la segmentation prosodique, nous émettons l'hypothèse que les apprenants sinophones seront également en mesure de désambiguïser les différentes structures syntaxiques en se basant sur ces mêmes indices. Enfin, nous supposons que la transitivité des verbes influencera l'attribution des fonctions syntaxiques des syntagmes nominaux (e.g., « *le rat* » dans les phrases (c.) et (d.)). Ainsi, avec un verbe principalement transitif, le syntagme nominal aura tendance à être traité comme un complément d'objet direct, tandis qu'avec un verbe qui est utilisé essentiellement comme intransitif, le syntagme nominal sera plus souvent interprété comme un sujet.

3.2 Méthodologie

3.2.1 Les stimuli

Nous avons utilisé les mêmes types de phrases que celles utilisées dans notre étude de production de la parole. Nous avons sélectionné 20 paires de phrases localement ambiguës (cf. section 3.5 pour la liste de ces 20 paires de phrases), où chaque paire a été construite de sorte à avoir les frontières prosodiques situées dans deux positions différentes : l'une précédant le syntagme nominal, appelée « clôture précoce (CP) » et l'autre suivant le syntagme nominal, appelée « clôture tardive (CT) ». Un exemple de ces paires de phrases est présenté dans le Tableau 3.1. Le même syntagme nominal dans chaque paire (ici « *le rat* ») a deux fonctions syntaxiques différentes : « sujet » dans la proposition principale de la condition CP (carré en rouge dans la Figure 3.1) et « complément d'objet direct » dans la proposition subordonnée de la condition CT (carré en rouge dans la Figure 3.1).



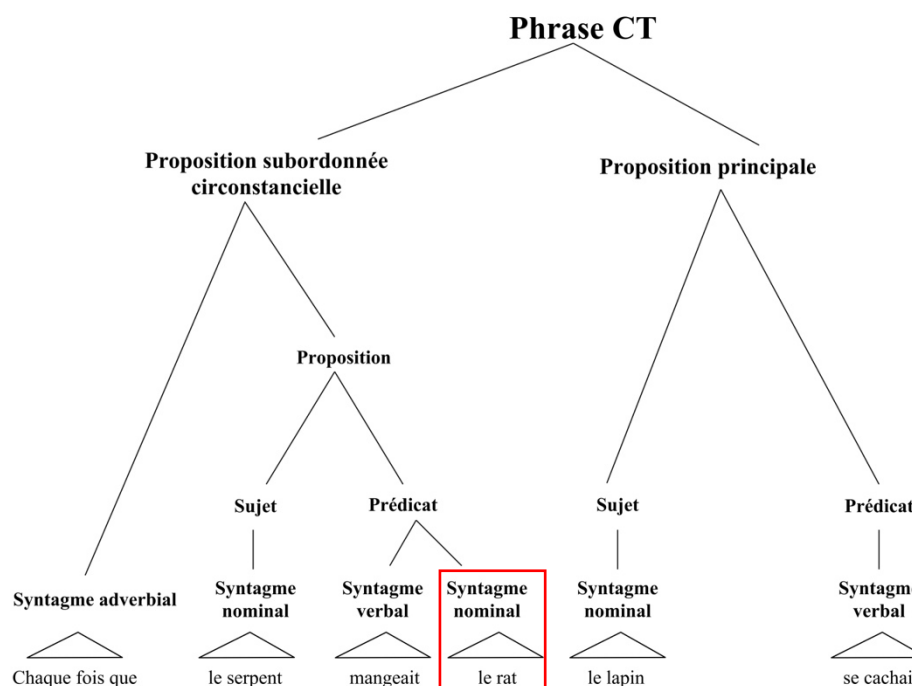


Figure 3.1 : Structure syntaxique arborescente pour une phrase avec la frontière CP (haut) CT (bas).

En plus de ces 40 phrases (20 paires de phrases), 5 autres phrases, ne présentant aucune ambiguïté, ont été ajoutées pour servir de distracteurs (par exemple, « *Chaque fois que j'allais à la piscine, il y avait du monde.* »).

Tableau 3.1 : Trois types de phrases utilisées dans l'expérience perceptive.

Condition	Exemple
Clôture précoce (CP) (<i>Early Closure</i>)	Chaque fois que le serpent mangeait, le rat se cachait.
Clôture tardive (CT) (<i>Late Closure</i>)	Chaque fois que le serpent mangeait le rat, le lapin se cachait.
Phrase distractive	Chaque fois que j'allais à la piscine, il y avait du monde.

Les 45 phrases, imprimées sur une feuille, ont été lues dans un ordre aléatoire par une locutrice native du français dans la chambre sourde du Laboratoire de Phonétique et Phonologie (CNRS & Université Sorbonne Nouvelle). Les phrases enregistrées ont été digitalisées à 16,000 Hz et à 16 Bits. Des analyses acoustiques ont été effectuées pour déterminer les différences de F0 et de durée des voyelles finales du syntagme verbal (en rouge dans le Tableau 3.1) et du syntagme nominal (en vert) dans les deux conditions CP et CT. Les valeurs de F0 obtenues renvoient à la différence entre la F0 à 90 % et à 20 % de la même voyelle pour chaque condition (e.g., la différence entre la F0 à 90 % et à 20 % de /a/ de « rat » dans la CP). Nous

avons aussi calculé la différence entre la CP et la CT (indiquée en pourcentage dans le Tableau 3.2 et représentée dans la Figure 3.2). La même procédure a été utilisée pour obtenir les valeurs de durée (Tableau 3.2, Figure 3.3).

Tableau 3.2 : Différences de F0 et de durée de la même voyelle dans les conditions prosodiques CP et CT.

Syntagme	Condition prosodique	Différence de F0		Différence de durée	
		CT	CP	CT	CP
Syntagme nominal (SN)	CT	+42,1 Hz	49% de plus de montée (t(18) = -2,2, p < 0,05).	119 ms	27,9% plus longue (t(19) = -4,6, p < 0,001).
	CP	+21,45 Hz		86 ms	
Syntagme verbal (SV)	CT	+74 Hz	109% de plus de montée (t(19) = 6,9, p < 0,001).	133 ms	46,1% plus longue (t(19) = 11, p < 0,001).
	CP	-6,7 Hz		72 ms	

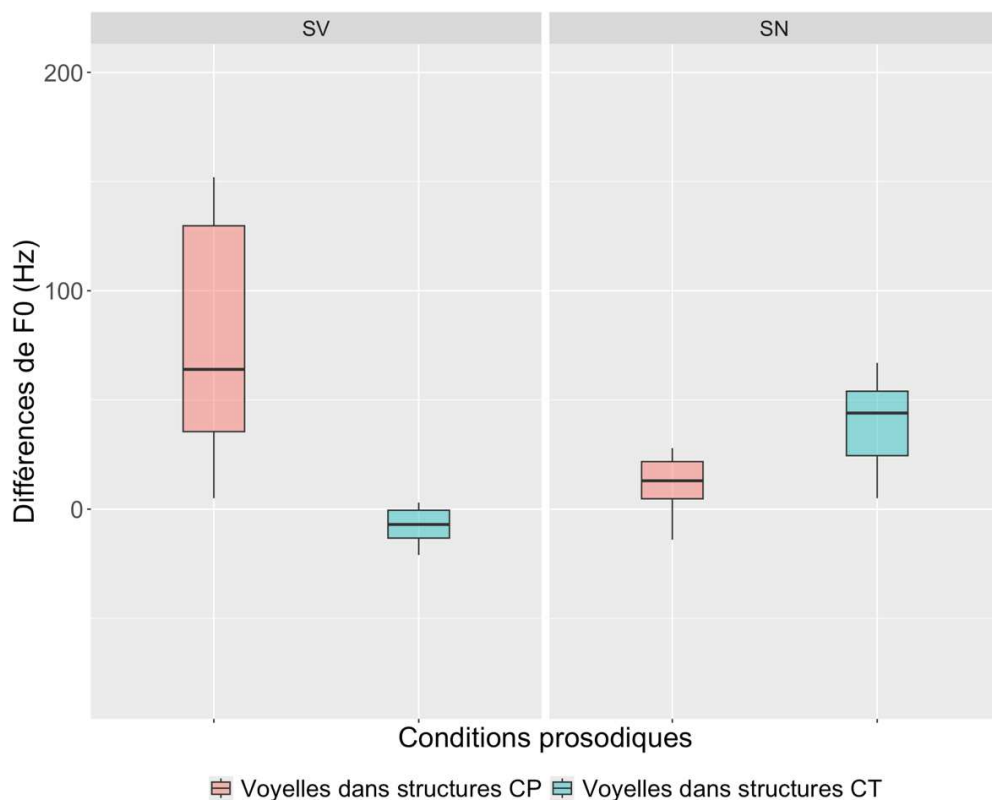


Figure 3.2 : Comparisons des différences de F0 (Hz) des voyelles en fonction des conditions prosodiques dans les syntagmes verbaux (SV, gauche) et nominaux (SN, droite).

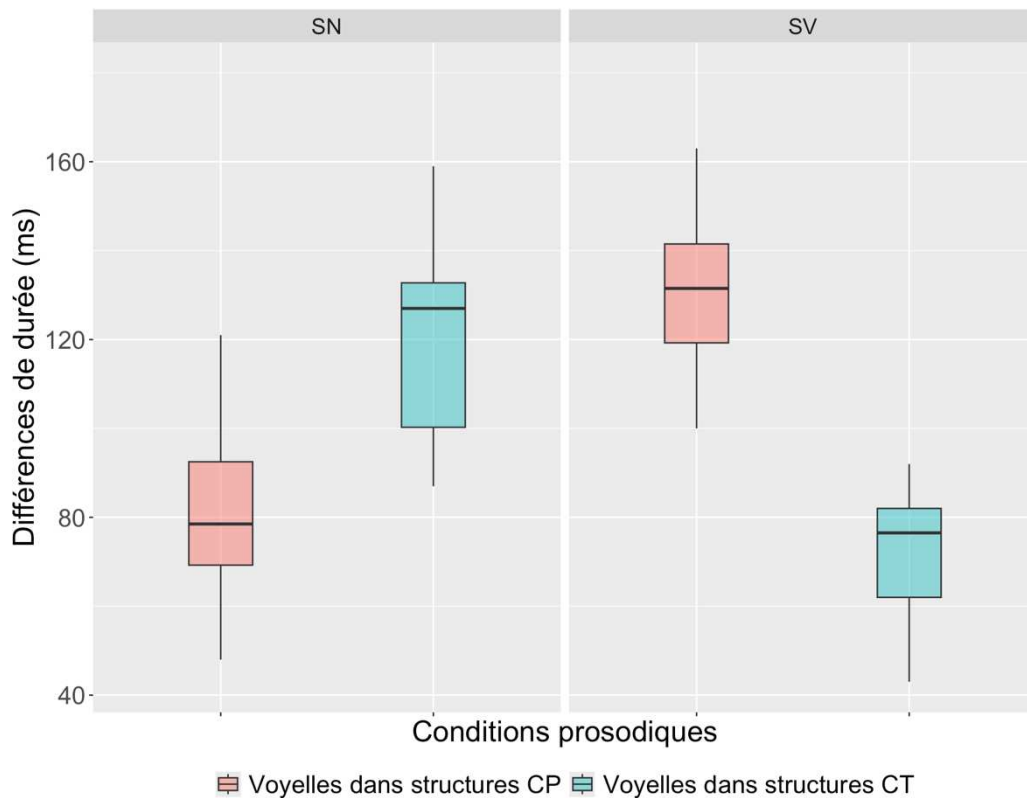


Figure 3.3 : Comparaisons des différences de durée (ms) des voyelles en fonction des conditions prosodiques dans les syntagmes nominaux (SN, gauche) et verbaux (SV, droite).

Les T tests ont révélé des effets significatifs pour les conditions prosodiques (CP vs CT) sur chacune des variables dépendantes mesurées : la F0 et la durée, tant pour les syntagmes nominaux que verbaux (Tableau 3.2). La condition CT est signalée par une plus grande montée de F0 de la voyelle finale du syntagme nominal (49 % de plus comparée à la condition CP). La condition CP est signalée par une montée encore plus forte de F0 de la voyelle finale du syntagme verbal (109 % de plus comparée à la même voyelle dans la condition CT). Les différences de durée pour les conditions CP et CT vont dans la même direction : la durée de la voyelle finale du syntagme nominal est 27,9 % plus longue par rapport à la même voyelle dans la condition CP. De même, la voyelle finale du syntagme verbal est 46,1 % plus longue dans la condition CP comparée à la même voyelle dans la condition CT. Par ailleurs, une pause est systématiquement marquée à la frontière du syntagme intonatif de chaque phrase, comme l'illustrent les exemples donnés dans la Figure 3.4. Prises ensemble, les analyses acoustiques montrent que la locutrice enregistrée a bien réalisé les indices acoustiques attendus pour marquer les frontières du syntagme intonatif : la montée de F0, l'allongement final et la pause.

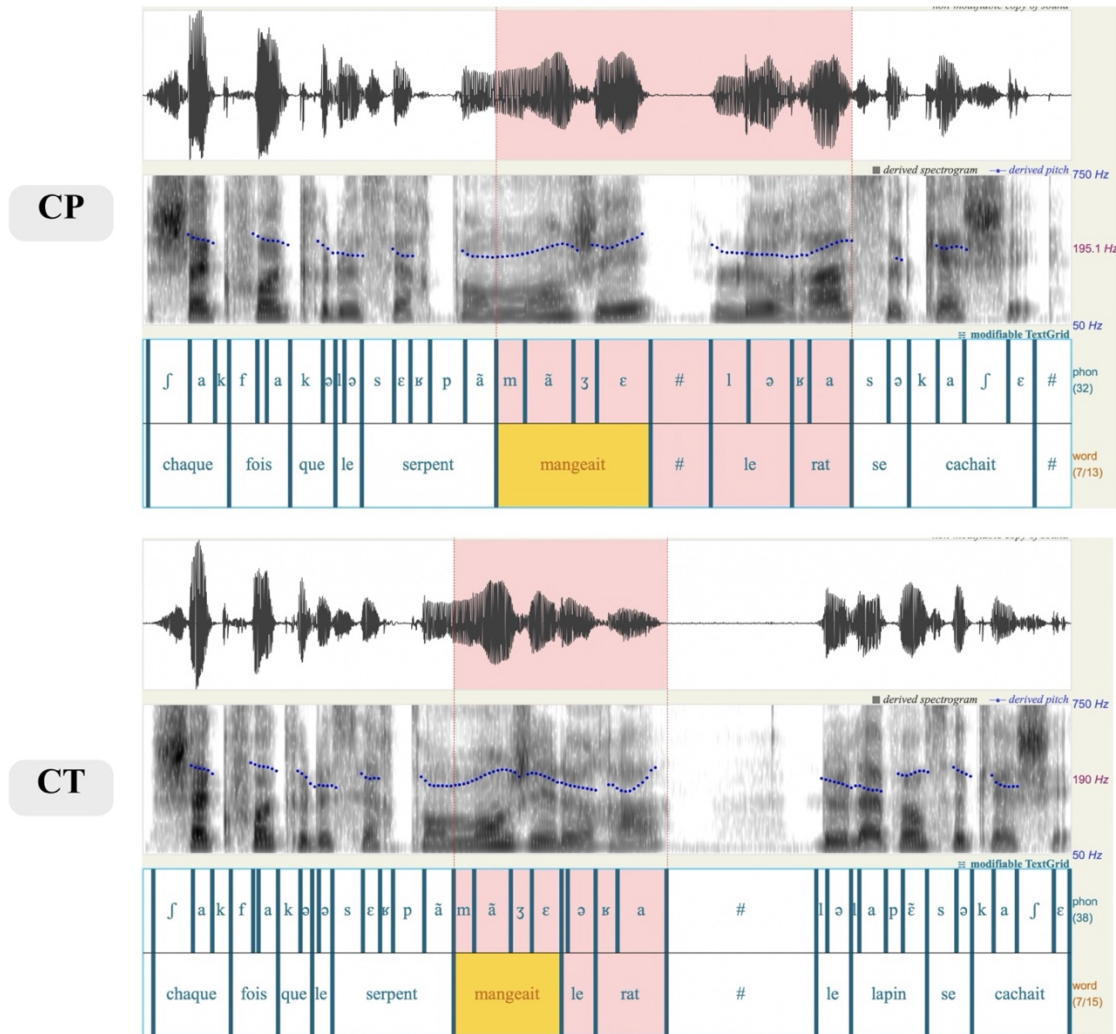


Figure 3.4 : Signaux acoustiques et contours mélodiques d'un exemple représentant les conditions CP (haut) et CT (bas).

Nous avons ensuite procédé à une troncation des phrases enregistrées à l'offset de la dernière voyelle du syntagme nominal pour constituer les stimuli de l'expérience perceptive (e.g., « *Chaque fois que le serpent mangeait, le rat se cachait.* » vs « *Chaque fois que le serpent mangeait le rat, le lapin se cachait.* »). Cette procédure avait pour objectif d'avoir des stimuli identiques, qui ne se distinguaient que sur le plan prosodique. Les 40 stimuli ainsi obtenus ont été répartis dans deux blocs pour que deux stimuli issus de la même paire de phrases n'apparaissent jamais dans le même bloc. Nous avons ajouté à ces deux blocs les 5 stimuli distracteurs, coupés à une position autre que la frontière du syntagme intonatif. Ainsi, notre expérience perceptive contenait deux blocs, chacun contenant 25 stimuli (Figure 3.5).

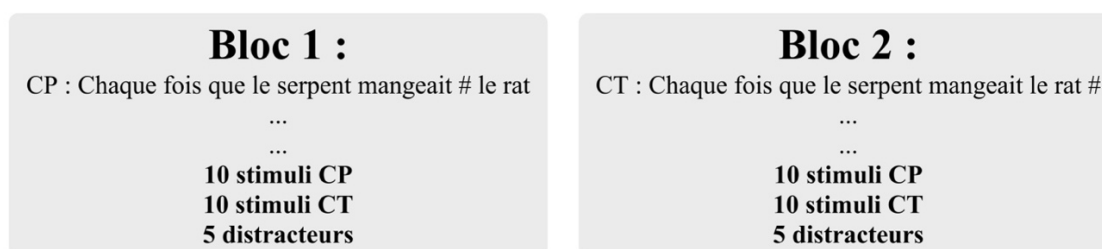


Figure 3.5 : Deux blocs de stimuli utilisés dans l'étude de perception.

3.2.2 Participants et procédure

Vingt locuteurs francophones natifs et quarante apprenants sinophones (Tableau 3.3) ont pris part à cette expérience de perception. Aucun des participants ne présente de trouble langagier ou auditif. Comme pour l'expérience acoustique, chaque apprenant sinophone avait, préalablement à l'expérience, rempli le questionnaire LEAP-Q (Marian *et al.*, 2007) pour pouvoir évaluer son profil linguistique et son niveau en langue française. Cela nous a permis de diviser les participants en deux groupes : niveau intermédiaire et niveau avancé (Tableau 3.3). Contrairement à l'expérience acoustique, nous n'avons pas inclus d'apprenants débutants pour cette expérience, car à ce stade de leur apprentissage, il est peu probable qu'ils puissent comprendre correctement les stimuli et utiliser les indices prosodiques pour résoudre les ambiguïtés syntaxiques.

Tableau 3.3 : Informations (sexe et âge) des participants dans l'expérience perceptive.

Participants		Nombre et sexe	Âge moyen (ans) ; Écart-type
Francophones natifs		4H ; 16F	23,1 ; 4,4
Apprenants sinophones	Niveau avancé	2H ; 18F	24,2 ; 2,4
	Niveau intermédiaire	5H ; 15F	21,95 ; 2,7

Les participants ont signé un consentement de recherche avant de commencer l'expérience et chacun a reçu une indemnisation de dix euros en espèces pour leur participation. Ils ont été testés individuellement dans une salle calme, en se servant d'un ordinateur, d'un casque et d'une feuille pour écrire leurs réponses. La consigne suivante a été donnée à chaque participant oralement en français et en mandarin : « *Cette expérience est une tâche de complétion de phrases, pendant laquelle vous devez, après chaque stimuli écouté, le compléter à l'écrit par n'importe quel contenu qui vous vient à l'esprit* ». La seule condition a été que le stimulus complété forme une phrase

grammaticalement correcte. Après l'explication des consignes, l'expérimentateur a donné un exemple en complétant une phrase qui n'apparaissait pas dans les stimuli. L'expérience a été lancée une fois que le participant avait confirmé la bonne compréhension des consignes. Les stimuli ont été diffusés un par un sur ordinateur et projetés sur des diapositives ; le participant devant, après chaque stimulus, écrire le contenu complété sans recopier toute la phrase. Les participants avaient la possibilité d'écouter chaque stimulus autant de fois qu'ils le souhaitaient avant de répondre. Pour les vingt participants dans chacun des trois groupes, dix participants ont été testés avec le Bloc 1 et dix autres participants ont été testés avec le Bloc 2. Nous avons recueilli un total de 1500 stimuli complétés (25 stimuli (10 CP + 10 CT + 5 distracteurs) pour chacun des 60 participants). Pour analyser les données, nous nous sommes limités aux stimuli des conditions CP et CT : nous avons examiné si le mot cible (i.e., le syntagme nominal) était correctement traité comme « sujet » dans la condition CP et comme « complément d'objet direct » dans la condition CT. La présentation détaillée des réponses correctes et incorrectes fournies par chaque participant des trois groupes se trouve dans l'Annexe 3.5.1 et 3.5.2 (Tableaux 3.8 et 3.9).

3.3 Résultats

3.3.1 Résultats généraux chez les trois groupes de participants

Parmi les réponses données par les participants sinophones, neuf réponses ont été éliminées (indiquées par « \ » dans les Tableaux 3.8 et 3.9) car les mots cibles n'ont été interprétés ni comme « sujet » ni comme « complément d'objet direct ». Nous commencerons par analyser ces résultats de manière globale, en tenant compte de l'ensemble des participants et en comparant les conditions prosodiques CP et CT. Nous avons recensé le nombre de réponses correctes pour chaque bloc, chaque condition prosodique et chaque groupe de participants. Ensuite, nous avons calculé les moyennes, ainsi que les nombres maximum et minimum des réponses correctes. Les résultats obtenus sont détaillés dans la Figure 3.6 détaillés et dans le Tableau 3.4. Ensuite, nous examinerons les résultats spécifiques à chaque groupe. Enfin, nous nous intéresserons à l'impact éventuel de la nature transitive des verbes sur la désambiguïsation syntaxique.

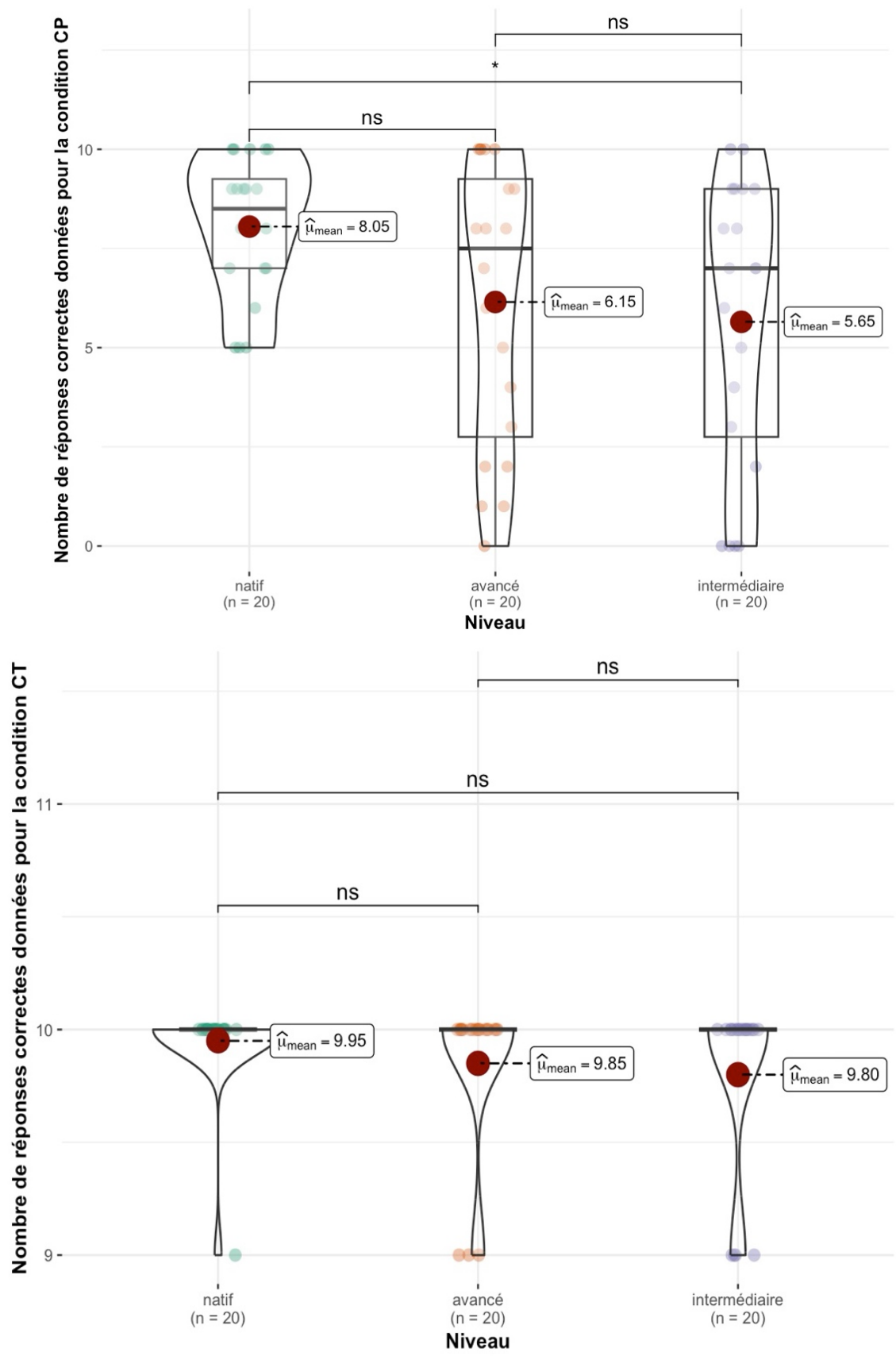


Figure 3.6 : Comparaison de nombre de réponses correctes données par les trois groupes (natifs vs avancés vs intermédiaires) pour la condition prosodique CP (haut) et la condition prosodique CT (bas).

Tableau 3.4 : Statistiques descriptives des résultats obtenus dans la tâche de complétion en fonction des conditions prosodiques (CP vs CT) pour les trois niveaux de participants (natifs vs avancés vs intermédiaires).

Item	Natif			Avancé			Intermédiaire		
	Total	Bloc 1	Bloc 2	Total	Bloc 1	Bloc 2	Total	Bloc 1	Bloc 2
	Condition Prosodique CP								
Moyenne de nombre de réponses correctes	8,05	7,7	8,4	6,15	5,6	6,7	5,65	5,7	5,6
Écart-type	1,76	1,7	1,84	3,57	2,88	4,24	3,63	3,77	3,69
Réponses correctes minimum	5	5	5	0	2	0	0	0	0
Réponses correctes maximum	10	10	10	10	10	10	10	10	10
Condition Prosodique CT									
Moyenne de nombre de réponses correctes	9,95	9,9	10	9,85	9,7	10	9,80	9,8	9,8
Écart-type	0,22	0,32	0	0,37	0,48	0	0,41	0,42	0,42
Réponses correctes minimum	9	9	10	9	9	10	9	9	9
Réponses correctes maximum	10	10	10	10	10	10	10	10	10

Tableau 3.5 : Résultats des tests post-hoc de Tukey pour les interactions entre chaque condition prosodique et chaque niveau de groupe de participants. (Les interactions significatives sont mises en gras.).

Niveau	Interaction entre conditions prosodiques	t	P _{Tukey}	Condition prosodique	Interaction entre deux niveaux	t	P _{Tukey}
Natif	CP * CT	-2,71	0,08	CP	Natif * Avancé	-2,71	0,08
					Natif * Intermédiaire	-3,42	0,011 *
Avancé	CP * CT	-5,28	< 0,001 **	CT	Avancé * Intermédiaire	0,71	0,98
					Natif * Avancé	-0,14	1
Intermédiaire	CP * CT	-5,92	< 0,001 **	CT	Natif * Intermédiaire	-0,21	1
					Avancé * Intermédiaire	0,07	1

- *Résultats globaux :*

Nous avons utilisé une analyse de covariance (ANCOVA) pour comparer le nombre de réponses correctes des trois groupes de participants dans les conditions prosodiques CP et CT, en tenant compte du bloc (1 ou 2) comme covariable. Les résultats montrent un effet significatif pour la « condition prosodique » ($F(1, 113) = 316,88, p < 0,001, \eta^2 = 0,34$), tandis que le bloc n'est pas un facteur significatif. De manière générale, comme l'illustre la Figure 3.6, les participants, qu'ils soient natifs, apprenants avancés ou intermédiaires, ont correctement identifié les syntagmes nominaux cibles comme des « compléments d'objet direct » dans la condition prosodique CT (Figure 3.6, bas). Seuls les stimuli suivants ont été considérés comme CP : {Chaque fois que ma sœur dansait le ballet, ...} par un natif ; {Chaque fois que l'étudiant travaillait les notes, ...} par trois avancés et un intermédiaire ; {Chaque fois que le garçon jouait le jeu, ...} et {Chaque fois que l'on partageait la pizza, ...} par un intermédiaire. Nous considérons que ces cas isolés relèvent davantage du hasard ou d'un manque d'attention, et non d'une incapacité des participants à traiter correctement les syntagmes inclus dans ces stimuli. Ces exceptions ne remettent donc pas en cause le fait que les auditeurs, qu'ils soient natifs ou apprenants, ont pu s'appuyer efficacement sur les indices acoustiques pour reconnaître les structures CT. En revanche, les structures CP ont posé davantage de difficultés. Dans cette condition, les syntagmes nominaux ont été interprétés de manière variable : tantôt comme des « sujets » tantôt comme des « compléments d'objet direct », et ce chez tous les auditeurs.

L'analyse ANCOVA a révélé un effet significatif du facteur « niveau » ($F(2, 113) = 18, p = 0,03, \eta^2 = 0,04$), indiquant que les performances des trois groupes de participants diffèrent de manière significative. Pour mieux comprendre ces différences, ainsi que l'interaction entre les conditions prosodiques et les niveaux de maîtrise du français, nous avons effectué des tests post-hoc de Tukey. Les résultats détaillés de ces analyses sont présentés dans le Tableau 3.5, et seront également décrits ci-dessous en fonction des niveaux de français.

- *Résultats par groupe de participants :*

Pour les participants natifs, les stimuli avec la structure CP (Figure 3.6, haut), ont été correctement identifiés avec une moyenne de 8,05 réponses correctes sur 10

(Tableau 3.4), ce qui indique qu'ils ont, malgré quelques erreurs, bien réussi à identifier les structures CP dans l'ensemble. Concernant les stimuli avec la structure CT, les natifs ont obtenu une moyenne de 9,95 réponses correctes sur 10. Ce résultat souligne encore davantage qu'ils n'ont rencontré aucune difficulté pour cette condition. En outre, notre test post-hoc de Tukey n'a pas révélé d'interaction significative entre les conditions CP et CT (Tableau 3.5). Cela signifie que les participants natifs ont traité les deux conditions de manière statistiquement comparable, et qu'ils ont su utiliser efficacement les indices prosodiques pour désambiguïser les structures syntaxiques.

Chez les apprenants avancés, nous avons observé une tendance similaire pour la condition CT, avec une moyenne de 9,85 réponses correctes sur 10 stimuli. En revanche, en ce qui concerne les stimuli CP, les apprenants avancés ont affiché des performances inférieures par rapport aux natifs, avec en moyenne seulement 6,15 réponses correctes sur 10 stimuli en moyenne. Pour un apprenant avancé (S10, Tableau 3.9), aucun des 10 stimuli CP n'a été correctement identifié. Le test post-hoc de Tukey a révélé une interaction significative entre les conditions CP vs CT ($t = -5,28$, $p < 0,001$) (Tableau 3.5) pour ce groupe. Comme l'indiquent les statistiques descriptives (Tableau 3.4), les apprenants avancés ont pu identifier les frontières CT en se basant sur les indices prosodiques, mais ils n'ont pas encore parfaitement réussi à bien distinguer les frontières CT de celles de CP.

Face aux stimuli CP et CT, les apprenants intermédiaires ont montré des performances similaires à celles des apprenants avancés. Pour les stimuli CT, ils ont obtenu une moyenne de 9,8 réponses correctes sur 10, ce qui montre qu'ils n'ont rencontré aucune difficulté particulière pour reconnaître cette structure. En revanche, pour la condition CP, leur performance a été nettement inférieure, avec une moyenne de 5,65 réponses correctes sur 10 stimuli (Tableau 3.4). Quatre participants intermédiaires (S3 et S10 dans le Tableau 3.8 ; S1 et S4 dans le Tableau 3.9) n'ont pas réussi à assigner correctement aucun des 10 stimuli CP. Le test post-hoc de Tukey a révélé une interaction significative entre les conditions CP vs CT ($t = -5,92$, $p < 0,001$) (Tableau 3.5). Comme le montrent les statistiques descriptives (Tableau 3.4), les apprenants intermédiaires ont pu identifier les frontières CT, mais ils étaient encore incapables de distinguer les frontières CT de celles de CP.

- *Comparaison entre les trois groupes de participants : les performances chez les trois groupes de participants varient-elles en fonction de leur niveau en français et/ou des structures prosodiques ambiguës ?*

Afin de répondre à cette question de recherche, il est important d'examiner également les différences des traitements entre les trois groupes de participants pour une même condition prosodique.

Comme le montrent les comparaisons par paires (Tableau 3.5), les tests post-hoc de Tukey montrent qu'aucune différence significative n'est observée pour les réponses aux stimuli CT. Cela est vrai pour les groupes « natifs vs avancés » ($t = -0,14$, $p = 1$), pour « natifs vs intermédiaires » ($t = -0,21$, $p = 1$), et « avancés vs intermédiaires » ($t = 0,07$, $p = 1$). Ces résultats montrent que les apprenants, quel que soit leur niveau, se comportent comme les natifs, et ne rencontrent aucune difficulté à percevoir les frontières CT en se basant uniquement sur les indices prosodiques.

En revanche, pour les stimuli de structure CP, les tests post-hoc de Tukey (Tableau 3.5) ont révélé une différence significative entre les groupes « natifs vs intermédiaires » ($t = -3,42$, $p = 0,01$). Cependant, aucune différence significative n'est observée pour les comparaisons entre groupes « natifs vs avancés » ($t = -2,71$, $p = 0,08$), et entre groupes « avancés vs intermédiaires » ($t = 0,71$, $p = 0,98$). Ces résultats suggèrent que les apprenants intermédiaires traitent les frontières CP de manière différente par rapport aux natifs. Tandis que les performances des apprenants avancés sont comparables à celles des natifs, et à la fois à celles des intermédiaires. Cela suggère que les apprenants avancés n'ont pas encore atteint un niveau de maîtrise nettement supérieur à celui des intermédiaires.

Ces observations fournissent des arguments en faveur de l'acquisition progressive de la prosodie sur le plan perceptif. En particulier, l'acquisition des structures prosodiques plus complexes comme CP, est un processus graduel. À mesure que le niveau des apprenants se rapproche de celui des natifs, ils deviennent progressivement capables de percevoir les frontières prosodiques CP, témoignant ainsi d'une amélioration dans leur traitement des indices prosodiques.

3.3.2 Analyse des résultats en fonction des transitivités des verbes

La disparité dans les traitements entre les conditions CP vs CT chez les auditeurs, notamment natifs, nous poussent à explorer l'éventuelle influence de la transitivité

des verbes sur les réponses des auditeurs, en plus des indices prosodiques disponibles. Pour rappel, la tâche de complétion réalisée par les participants consistait à déterminer la fonction syntaxique du syntagme nominal. Dans la condition CP, ce syntagme joue le rôle du sujet de la proposition principale, tandis que dans la condition CT, il sert de complément d'objet direct de la proposition subordonnée circonstancielle (voir la Figure 3.1). La question est donc de déterminer si les auditeurs pourraient attribuer au syntagme nominal la fonction de « sujet », lorsque le verbe qui le précède est perçu comme plutôt intransitif, et la fonction de « complément d'objet direct » lorsque le verbe qui le précède est perçu comme plutôt transitif.

Afin de répondre à cette question, nous avons réalisé une analyse supplémentaire des réponses données par les auditeurs, en classant les verbes utilisés dans les stimuli tronqués en fonction de leur transitivité. Cette catégorisation a été effectuée à l'aide de la base de données *LVF (Les Verbes français)* (Dubois, 1991 ; François *et al.*, 2007). Cette base répertorie, pour chaque verbe de la langue française, l'ensemble des structures syntaxiques attestées, telles que les structures transitives, intransitives, ou pronominales. Ces structures sont ensuite classées et hiérarchisées en fonction de leur taux d'acceptation dans l'usage.

En nous appuyant sur cette base, nous avons examiné chaque verbe utilisé dans les stimuli et noté toutes les structures qui y sont répertoriées. Par la suite, nous avons classé ces verbes en deux groupes : un groupe contenant les verbes qui peuvent être employés à la fois de manière transitive et intransitive (n=13) ; et un groupe contenant les verbes qui s'emploient principalement de manière transitive (n=5) (Tableau 3.6).

Tableau 3.6 : Catégorisation des verbes en deux groupes en fonction de leur transitivité.

Groupe	Verbes
Transitif et Intransitif	'appeler, chanter, chasser, cuisiner, danser, dessiner, enseigner, fermer, jouer ¹⁰⁹ , lire, manger, souffler, travailler'
Principalement transitif	'conduire, partager, regarder, rénover, visiter'

¹⁰⁹ Dans l'expérience perceptive, ce verbe a été utilisé dans trois stimuli différents. Ainsi, dans le Tableau 3.6, nous avons seulement 18 verbes.

Lors des présentations des résultats obtenus auprès des trois groupes de participants dans les sections précédentes, nous avons montré que les auditeurs natifs et les apprenants rencontraient des difficultés particulières à traiter correctement les stimuli de structure CP. En nous appuyant sur cette catégorisation, et si les transitivités des verbes influencent l'attribution des fonctions syntaxiques des syntagmes nominaux, nous formulons les hypothèses suivantes :

- 1) Les verbes principalement transitifs (Tableau 3.6), que les auditeurs ont l'habitude d'associer à un complément d'objet direct, pourraient induire une erreur d'interprétation. Lorsqu'ils sont intégrés dans des structures CP, les syntagmes nominaux qui suivent ces verbes risquent d'être perçus à tort comme compléments d'objet direct, au lieu d'être identifiés comme sujets de la proposition suivante.
- 2) Les auditeurs natifs seraient principalement impactés par la transitivité des verbes dans le traitement des stimuli. Cela s'explique par le fait qu'il ne s'agit pas d'un traitement purement prosodique ou phonétique, mais d'une interaction entre syntaxe et prosodie. Les apprenants avancés et intermédiaires, dans leur phase actuelle d'acquisition du français, ne seraient probablement pas encore en mesure de traiter simultanément les informations prosodiques et morphosyntaxiques. Par ailleurs, cette interaction syntaxico-prosodique est d'autant plus compliquée que les stimuli utilisés dans notre étude étaient tronqués ;
- 3) Pour les verbes qui peuvent s'employer à la fois de manière transitive et intransitive, leur transitivité ne devrait pas constituer un biais dans la perception des stimuli localement ambigus et tronqués.

Nous avons analysé les réponses données par les trois groupes de participants, en nous concentrant particulièrement sur les verbes qui avaient donné lieu à des réponses incorrectes. Nous récapitulons dans le Tableau 3.7 ces verbes en fonction des conditions prosodiques et des groupes de participants. Pour chaque verbe, nous avons noté le nombre de réponses incorrectes entre parenthèses. Par exemple, « *jouer (7)* » signifie que des francophones natifs ont, de manière incorrecte, considéré ce verbe comme transitif à 7 reprises dans la condition CP.

Tableau 3.7 : Verbes donnés lieu à des réponses incorrectes chez les trois groupes de participants.

Niveau de français	Condition prosodique	Verbes dans les réponses incorrectes
Francophones natifs	CP	jouer (7), lire (6), enseigner (6), partager (4), danser (2), regarder (2), visiter (2), chanter (2), travailler (1), conduire (1), souffler (1), appeler (1), chasser (1), dessiner (1), rénover (1), fermer (1)
	CT	danser (1)
Apprenants avancés	CP	jouer (13), rénover (6), souffler (5), enseigner (5), chanter (4), dessiner (4), partager (4), danser (3), lire (3), travailler (3), visiter (3), appeler (3), chasser (3), conduire (2), manger (2), regarder (2), fermer (2), cuisiner (1)
	CT	travailler (3)
Apprenants intermédiaires	CP	jouer (15), danser (6), enseigner (6), regarder (6), visiter (6), souffler (4), travailler (4), rénover (4), partager (4), conduire (3), cuisiner (3), fermer (3), lire (3), manger (3), appeler (3), chanter (3), chasser (3), dessiner (3)
	CT	travailler (1), jouer (1), partager (1)

Dans chacun des trois groupes d'auditeurs, nous avons constaté que le verbe « jouer » a engendré le plus grand nombre de réponses incorrectes dans la condition CP. Cela s'explique en grande partie par le fait que ce verbe a été utilisé dans trois stimuli différents :

- *Chaque fois que l'acteur {jouait, la scène amusait les spectateurs. / jouait la scène, son costume amusait les spectateurs.}*
- *Chaque fois que le garçon {jouait, le jeu durait des heures. / jouait le jeu, le dîner durait des heures.}*
- *Chaque fois que les parents {jouaient, le jeu semblait facile. / jouaient le jeu, les règles semblaient faciles.}*

Outre le verbe « jouer », les verbes ayant engendré le plus grand nombre de réponses incorrectes dans la condition CP sont « lire », « enseigner » et « partager ». Parmi ces verbes, seul « partager » s'emploie de manière principalement transitive, tandis que « jouer », « lire », « enseigner » peuvent être employés à la fois de manière transitive et intransitive. De plus, la base de données LVF (Dubois, 1991 ; François *et al.*, 2007) établit un classement hiérarchique des structures syntaxiques possibles, en fonction de leur fréquence d'acceptation dans l'usage. Pour le verbe « jouer », la structure intransitive occupe la première place, surpassant toutes les autres structures

syntaxiques. Si la transitivité d'un verbe influençait la façon dont un auditeur interprète le syntagme nominal suivant un verbe comme « jouer », qui privilégie fortement une structure intransitive, nous n'aurions pas dû observer un si grand nombre de réponses incorrectes pour ce verbe dans la condition CP. En ce qui concerne les deux autres verbes « lire » et « enseigner », étant donné qu'ils acceptent à la fois des structures transitives et intransitives, les syntagmes nominaux intégrés dans les stimuli CP ne devraient pas être quasi systématiquement interprétés comme des compléments d'objet direct.

Chez les deux groupes d'apprenants, les stimuli CP ont révélé que les verbes posant le plus de difficulté étaient « jouer », « enseigner », « souffler » et « rénover ». Là encore, si la transitivité jouait un rôle prépondérant, les syntagmes nominaux intégrés dans les stimuli CP ne devraient pas être quasi systématiquement interprétés comme des compléments d'objet direct. Par ailleurs, nos résultats, détaillés dans l'annexe 3.5.2, montrent que certains apprenants intermédiaires n'ont donné aucune réponse correcte pour la complétion des stimuli CP, et ce quelle que soit la transitivité des verbes. Cela suggère, comme mentionné précédemment, que dans leur stade actuel d'apprentissage, les apprenants ne parviennent probablement pas encore à totalement intégrer simultanément les informations prosodiques et morphosyntaxiques lors du traitement des phrases.

Nos observations, aussi bien pour les francophones natifs que pour les apprenants, montrent que la transitivité des verbes n'a pas influencé le traitement des stimuli CP vs CT. Ainsi, notre hypothèse 1 (« Les verbes principalement transitifs intégrés dans des structures CP pourraient induire une erreur d'interprétation. ») n'a pas été confirmée : les verbes ayant généré le plus de réponses incorrectes s'emploient à la fois de manière transitive et intransitive, contrairement à ce que nous avons anticipé. Les résultats pour les apprenants tendent à confirmer l'hypothèse 2 : « les apprenants avancés et intermédiaires ne seraient probablement pas encore en mesure de traiter simultanément les informations prosodiques et morphosyntaxiques ». Leur capacité à interpréter les syntagmes nominaux en fonction de la transitivité des verbes est encore limitée. Comme indiqué dans les Tableaux 3.8 et 3.9, les 18 verbes testés dans les stimuli CP ont tous donné lieu à des réponses incorrectes, quelle que soit leur transitivité. Notamment, 4 apprenants sur

40 n'ont donné aucune réponse correcte pour la condition CP, soulignant leurs difficultés à traiter les indices acoustiques présents dans les stimuli CP.

Enfin, l'hypothèse 3 (« les verbes à la fois transitif et intransitif ne devraient pas constituer un biais dans la perception des stimuli localement ambigus et tronqués ») n'a pas été confirmée. Ces verbes ont généré un grand nombre de réponses incorrectes, tant chez les francophones natifs que chez les apprenants. Si la transitivité des verbes jouait un rôle déterminant, ces verbes auraient dû entraîner moins de réponses incorrectes, voire aucune.

Il est important de considérer avec prudence notre analyse des données basée sur la transitivité des verbes, car plusieurs limitations réduisent la portée de nos observations. Tout d'abord, cette analyse ayant été réalisée après le recueil des données, la répartition des verbes entre les deux catégories n'est pas équilibrée : les verbes à la fois transitifs et intransitifs (13 verbes) sont bien plus nombreux que les verbes principalement transitifs (5 verbes). Ensuite, malgré nos recherches, nous n'avons pas trouvé de base de données fournissant des indications précises sur la fréquence ou l'occurrence des verbes selon leur usage transitif ou intransitif. Cette absence d'indicateurs objectifs a conduit à une catégorisation inévitablement subjective de la transitivité des verbes.

3.4 Discussion

Cette étude de perception a utilisé des syntagmes nominaux inclus dans des phrases localement ambiguës pour examiner si les francophones natifs et les apprenants sinophones (intermédiaires et avancés) pouvaient prédire correctement leurs fonctions syntaxiques (sujet ou complément d'objet direct), sur la base des indices prosodiques. Les résultats indiquent qu'en condition CT, les participants, quel que soit leur niveau, ont identifié correctement les syntagmes nominaux comme « sujets », grâce aux indices acoustiques. En revanche, la condition CP a présenté des difficultés, notamment pour les apprenants sinophones, qui ont interprété ces syntagmes comme « sujets » mais aussi comme « compléments d'objet direct ».

Bien que les francophones natifs aient enregistré moins de réponses correctes en CP par rapport à CT, l'analyse statistique n'a pas montré de différence significative entre ces deux conditions. Les locuteurs natifs du français sont donc capables de désambigüiser les structures syntaxiques en s'appuyant sur les indices prosodiques,

un résultat qui corrobore ceux de Millotte *et al.* (2007, 2008). En revanche, les apprenants ont obtenu significativement moins de bonnes réponses pour CP, signalant une difficulté spécifique avec cette structure. Par ailleurs, les performances des apprenants avancés sont légèrement supérieures à celles des intermédiaires, ce qui suggère que l'acquisition des structures prosodiques complexes, comme CP, est un processus graduel, permettant aux apprenants de mieux percevoir les frontières prosodiques au fur et à mesure de leur progression.

Nos résultats illustrent parfaitement la préférence pour la structure CT dans le traitement syntaxique en temps réel. Cette tendance est cohérente avec la prédiction du *Late Closure Preference* (Frazier, 1979), confirmée par plusieurs études antérieures (e.g., Altmann *et al.*, 1998 ; Igoa *et al.*, 1998 ; Maia *et al.*, 2007 ; et Zagar *et al.* (1997) pour le français). Selon Frazier (1979), le traitement de la structure CT est un processus initial et par défaut dans le traitement de la parole : lorsque la grammaire le permet, le lecteur ou l'auditeur associe systématiquement tout nouvel élément à la clause en cours de traitement. Un exemple adapté de Paape et Vasishth (2021) illustre bien cette idée :

(e.) Since Jay always jogs a mile seems like a very short distance to him

Lors du traitement de la phrase (e.), le locuteur ou le lecteur suit un processus linéaire, attachant a priori automatiquement le syntagme nominal « *a mile* » comme complément d'objet direct de la clause « *Since Jay always jogs...* ». Cependant, lorsque le verbe « *seems* » entre dans le traitement linéaire en cours, le locuteur/lecteur se rend compte que le syntagme nominal « *a mile* » ne peut pas être un complément d'objet, mais qu'il débute une nouvelle clause en tant que sujet. Cette préférence pour la condition CT peut s'expliquer par la stratégie de l'attachement minimal (*Minimal Attachment Strategy*), qui vise à réduire la charge cognitive et à optimiser l'utilisation de la mémoire de travail (Kimball, 1973 ; Frazier & Fodor, 1978 ; Frazier, 1987 ; Bergmann *et al.*, 2008). Selon cette stratégie, le traitement syntaxique privilégie l'interprétation la plus linéaire et la plus simple, en cherchant à minimiser le nombre de clôtures syntaxiques, sauf lorsque celles-ci deviennent indispensables pour la compréhension. Ainsi, pour la phrase (e.), le cerveau interprète d'abord le syntagme nominal « *a mile* » comme un complément d'objet direct, car cette lecture demande moins d'efforts cognitifs et permet d'éviter des

réanalyses coûteuses. À l'inverse, dans la condition CP, où le syntagme nominal doit être interprété comme sujet, une structure syntaxique plus complexe est requise, ce qui pourrait alourdir le traitement cognitif, surtout dans le contexte de stimuli tronqués et localement ambigus.

Les auditeurs de notre expérience ont affiché un processus de traitement similaire, en attachant (quasi)systématiquement le syntagme nominal ambigu à la proposition précédente. Un tel traitement est à la fois grammatical et cognitivement moins coûteux. En revanche, cette stratégie par défaut a conduit à des erreurs pour certains stimuli tronqués de structures CP, interprétés à tort comme des structures CT, notamment chez les apprenants, voire parfois chez les natifs.

Une autre explication possible de la différence de traitement entre structures CP et CT repose sur l'*Informative Boundary Hypothesis* (Clifton Jr. *et al.*, 2002 ; Watson & Gibson, 2005 ; Carlson *et al.*, 2009a, 2009b), qui met en avance l'importance du contexte prosodique environnant. Selon cette hypothèse, l'effet d'une frontière prosodique ne dépend pas uniquement des indices acoustiques absolus, mais également des autres frontières apparues avant ET après celle-ci dans la phrase. Lors de la perception linéaire et continue de la parole, le traitement ne se limite pas à une frontière isolée, mais s'étend progressivement à une structure syntaxio-prosodique plus large et globale (Lee & Garnsey, 2012). Cela permet à l'auditeur d'établir et d'ajuster la hiérarchie prosodique et les relations syntaxiques entre syntagmes. Afin d'illustrer cette idée, considérons nos deux formes syntaxiques issues d'un même stimulus :

(f.) *Chaque fois que le serpent mangeait # le rat*

(g.) *Chaque fois que le serpent mangeait # le rat se cachait #*

Dans les stimuli tronqués de notre expérience, bien que la frontière prosodique CP dans (f.) (indiquée par « # ») soit acoustiquement marquée, son effet maximal n'est atteint que dans un contexte complet, comme en (g.). Dans cette phrase complète, après le syntagme nominal « *le rat* », une frontière mineure apparaît entre « *le rat* » et « *se cachait* », suivie d'indices acoustiques marquant la fin de phrase sur « *se cachait* ». Ces frontières successives, mineure et finale, renforcent l'effet de la frontière prosodique CP après « *mangeait* », permettant ainsi aux auditeurs d'interpréter correctement « *le rat* » comme le sujet de la proposition suivante.

Dans notre expérience, les stimuli étaient coupés juste après le syntagme nominal et avaient une durée relativement courte (3-4 secondes). Cette configuration obligeait les auditeurs à procéder à un traitement incrémental et local. En l'absence d'une structure prosodique complète et contextuelle, les indices tels que la pause, l'allongement final et la montée de F0, n'étaient pas suffisamment saillants et informatifs, notamment pour les apprenants (et aussi pour certains natifs).

Les travaux de Kjelgaard et Speer (1999), ainsi que de Millotte *et al.* (2008) soutiennent cette idée. Ils ont examiné la désambiguïsation syntaxique avec une prosodie neutre ou incongruente, et ont constaté que la structure CP était traitée plus lentement comparée à la structure CT. Cela suggère que le traitement de la structure CP nécessite plus de ressources temporelles ou d'efforts cognitifs lorsque les indices prosodiques ne sont pas suffisamment informatifs.

Carlson *et al.* (2001) ont aussi montré que le manque d'indices prosodiques clairs dans les phrases ambiguës favorise le traitement CT chez l'auditeur, conformément à la prédiction *Late Closure Preference*. Dans notre étude, les frontières prosodiques des stimuli CP, marquées notamment par une pause avant le syntagme nominal, n'ont pas suffi à empêcher ce processus par défaut. Ainsi, rattacher le syntagme nominal au verbe précédent, même dans la condition CP, reste compatible avec un traitement par défaut tout en étant grammaticalement correct.

Nos résultats, en comparant les performances des trois groupes, apportent des informations sur le processus de transfert prosodique. Face aux stimuli CP, les apprenants avancés se situaient entre les groupes natifs et intermédiaires, bien que leurs performances soient plus proches de celles des intermédiaires. Cela reflète une tendance chez les apprenants avancés à s'éloigner progressivement du traitement par défaut favorisant la structure CT, tout en se rapprochant du niveau des natifs. Les analyses statistiques ont en effet révélé une différence significative entre les intermédiaires et les natifs dans le traitement de la structure CP, mais aucune différence significative entre les natifs et les avancés. Cela suggère que la capacité à désambiguïser les structures syntaxiques complexes comme CP à l'aide de la prosodie n'est pas encore acquise chez les apprenants intermédiaires, mais commence à émerger chez les avancés, sans toutefois être totalement maîtrisée.

Du point de vue de l'acquisition de la prosodie d'une L2, les études d'Ip et Cutler (2018, 2022) apportent des éclairages sur le transfert prosodique. Bien que certains

indices acoustiques, comme l'allongement final ou la pause, soient universellement utilisés pour la segmentation prosodique, leur hiérarchie d'importance varie entre les langues. Ip et Cutler (2022) ont comparé la stratégie de désambiguïsation par la prosodie en production et perception en anglais et en mandarin. Ils ont montré que pour les locuteurs sinophones, la pause est l'indice prédominant, voire suffisant, pour désambiguïser (e.g., voir Fon *et al.* (2010) ; voir aussi le Tableau 1.3), tandis que les anglophones privilégient les indices mélodiques. En anglais, l'allongement final est préféré à la pause, alors qu'en mandarin, la pause prime sur l'allongement final.

Ip et Cutler (2018) suggère que l'acquisition de la prosodie en L2 subit aussi le processus de transfert : les apprenants doivent s'adapter à une nouvelle hiérarchie d'indices prosodiques. Ainsi, même si les indices acoustiques sont communs entre la L1 et la L2, leur traitement en L2 n'est ni automatique ni systématique. Les apprenants doivent s'adapter aux nouveaux poids des indices prosodiques utilisés dans leur langue maternelle mais exploités de manière différente en L2. Ce nouvel ordre devient progressivement ancré comme une nouvelle catégorie phonologique. Dans notre expérience perceptive, les stimuli CP comportaient une pause suivie d'un seul syntagme nominal, sans contexte prosodique supplémentaire. Cette pause, bien que présente, perdait en saillance pour les apprenants, rendant difficile la distinction CP et CT. D'autres indices acoustiques, comme l'allongement final et la montée de F0, n'ayant qu'un faible poids en mandarin, n'étaient pas suffisamment informatifs pour être combinés à la pause en une stratégie efficace de désambiguïsation.

Pour conclure, il est important de souligner que l'utilisation de stimuli tronqués dans notre expérience perceptive répondait à une nécessité méthodologique. Cela signifie toutefois que les performances des participants ne reflètent pas pleinement la manière dont les locuteurs natifs et les apprenants traitent des structures syntaxiques complètes. Afin d'examiner ce traitement dans des phrases non tronquées, nous avons mené une autre étude, basée sur l'électroencéphalographie (EEG), qui sera présentée dans le chapitre suivant. Cette approche nous permettra d'explorer en temps réel le traitement des structures ambiguës CP et CT, par les natifs et les apprenants, cette fois à l'aide de phrases complètes.

3.5 Annexes

3.5.1 Les stimuli utilisés dans le premier bloc de l'expérience perceptive, ainsi que les réponses données respectivement par les trois groupes de participants pour les deux conditions prosodiques CP et CT dans la tâche de complétion

P1 : Chaque fois que le garçon visitait le château, le donjon était très désert.

P2 : Chaque fois que mon père conduisait la voiture, le phare tombait en panne.

P3 : Chaque fois que l'acteur jouait, la scène amusait les spectateurs.

P4 : Chaque fois que le chef cuisinait le steak, le restaurant sentait très bon.

P5 : Chaque fois que l'auteur lisait son roman, le personnage devenait vivant.

P6 : Chaque fois que le serpent mangeait le rat, le lapin se cachait.

P7 : Chaque fois que l'employé appelait, le patron répondait.

P8 : Chaque fois que ma sœur dansait le ballet, son profil semblait très beau.

P9 : Chaque fois que le chanteur chantait, son album devenait un succès.

P10 : Chaque fois que Marie soufflait la bougie, le feu s'éteignait immédiatement.

P11 : Chaque fois que le chien chassait, le chat s'enfuyait.

P12 : Chaque fois que l'artiste dessinait, le portrait devenait plus coloré.

P13 : Chaque fois que la fille regardait le film, l'histoire l'attirait pendant quelques heures.

P14 : Chaque fois que le professeur enseignait, la théorie littéraire devenait moins obscure.

P15 : Chaque fois que le vendeur fermait, le supermarché avait besoin d'un nettoyage.

P16 : Chaque fois que le garçon jouait le jeu, le dîner durait des heures.

P17 : Chaque fois que le voisin rénouvait, la maison était en désordre.

P18 : Chaque fois que l'étudiant travaillait les notes, les cours devenaient plus claires.

P19 : Chaque fois que les parents jouaient, le jeu semblait facile.

P20 : Chaque fois que l'on partageait, la pizza finissait très vite.

Tableau 3.8 : Réponses correctes (cases vides) et incorrectes (cases en rouge) pour la tâche de complétion des 20 phrases (P1-P20) dans le premier bloc pour les 10 sujets (S1 à S10) des trois groupes.

Groupe de participants	Numéro de stimuli	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20	
	Condition prosodique du stimulus	CT	CT	CP	CT	CT	CT	CP	CT	CP	CT	CP	CP	CT	CP	CP	CT	CP	CT	CP	CP	
Francophones natifs	S1																					
	S2																					
	S3																					
	S4																					
	S5																					
	S6																					
	S7																					
	S8																					
	S9																					
	S10																					
Apprenants avancés	S1																					
	S2																					
	S3																					
	S4																					
	S5																					
	S6																					
	S7																					
	S8																					
	S9																					
	S10																					
Apprenants intermédiaires	S1																					
	S2																					
	S3																					
	S4																					
	S5																					
	S6																					
	S7																					
	S8																					
	S9																					
	S10																					

(« \ » représente la réponse éliminée de notre analyse.)

3.5.2 Les stimuli utilisés dans le deuxième bloc de l'expérience perceptive, ainsi que les réponses données respectivement par les trois groupes de participants pour les deux conditions prosodiques CP et CT dans la tâche de complétion

P1 : Chaque fois que le chien chassait le chat, les souris s'enfuyaient.

P2 : Chaque fois que le voisin rénouvait la maison, l'allée était en désordre.

P3 : Chaque fois que le chanteur chantait son album, le produit devenait un succès.

P4 : Chaque fois que les parents jouaient le jeu, les règles semblaient faciles.

P5 : Chaque fois que l'employé appelait le patron, la secrétaire répondait.

P6 : Chaque fois que mon père conduisait, la voiture tombait en panne.

P7 : Chaque fois que le chef cuisinait, le steak sentait très bon.

P8 : Chaque fois que l'on partageait la pizza, le dîner finissait très vite.

P9 : Chaque fois que ma sœur dansait, le ballet semblait très beau.

P10 : Chaque fois que le garçon jouait, le jeu durait des heures.

P11 : Chaque fois que l'auteur lisait à voix haute, son roman devenait vivant.

P12 : Chaque fois que le professeur enseignait la théorie littéraire, son raisonnement devenait moins obscur.

P13 : Chaque fois que le serpent mangeait, le rat se cachait.

P14 : Chaque fois que l'étudiant travaillait, les notes du cours devenaient plus claires.

P15 : Chaque fois que l'artiste dessinait le portrait, la toile devenait plus colorée.

P16 : Chaque fois que le garçon visitait, le château était très désert.

P17 : Chaque fois que le vendeur fermait le supermarché, le rayon avait besoin d'un nettoyage.

P18 : Chaque fois que l'acteur jouait la scène, son costume amusait les spectateurs.

P19 : Chaque fois que Marie soufflait, la bougie s'éteignait immédiatement.

P20 : Chaque fois que la fille regardait, le film l'attirait pendant quelques heures.

Tableau 3.9 : Réponses correctes (cases vides) et incorrectes (cases en rouge) pour la tâche de complétion des 20 phrases (P1 à P20) dans **le deuxième bloc** pour les 10 sujets (S1 à S10) des trois groupes.

Groupe de participants	Numéro de stimuli	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	P11	P12	P13	P14	P15	P16	P17	P18	P19	P20		
		CT	CT	CT	CT	CT	CP	CP	CT	CP	CP	CP	CT	CP	CP	CT	CP	CT	CT	CT	CP	CP	
Francophones natifs	S1																						
	S2																						
	S3																						
	S4																						
	S5																						
	S6																						
	S7																						
	S8																						
	S9																						
	S10																						
Apprenants avancés	S1																						
	S2																						
	S3																						
	S4																						
	S5																						
	S6																						
	S7																						
	S8																						
	S9																						
	S10																						
Apprenants intermédiaires	S1																						
	S2																						
	S3																						
	S4																						
	S5																						
	S6																						
	S7																						
	S8																						
	S9																						
	S10																						

(« \ » représente la réponse éliminée de notre analyse.)

Chapitre 4 Étude électroencéphalographique

4.1 Introduction et questions de recherche

Ce chapitre présente les résultats de l'expérience électroencéphalographique (EEG) menée pour étudier le traitement neurocognitif de la frontière prosodique du français chez des francophones natifs et des apprenants sinophones intermédiaires et avancés. Cette expérience a notamment pour objectif d'examiner la composante CPS afin de répondre aux questions suivantes :

- 1) La CPS serait-elle déclenchée lorsque les auditeurs francophones natifs perçoivent des frontières prosodiques différentes (CP vs CT) dans des phrases localement ambiguës ?
- 2) La CPS serait-elle déclenchée lorsque les apprenants sinophones perçoivent les mêmes phrases ambiguës ? Comment le niveau de français modulerait-il la composante CPS ?
- 3) Notre expérience perceptive a montré que les auditeurs francophones natifs et les apprenants avaient particulièrement du mal à traiter la structure CP tandis que la structure CT ne posait aucun problème. Comment cette différence comportementale se manifesterait-elle à travers les réponses CPS ?

Sur la base de notre revue de littérature sur la composante CPS et sur l'utilisation de l'EEG dans les recherches sur l'acquisition d'une L2 (cf. section 1.5.3), nous formulons les prédictions suivantes pour les trois questions de recherche :

- 1) Les frontières des structures prosodiques CP et CT déclencheront la CPS chez les auditeurs francophones natifs. En effet, la CPS a été observée dans plusieurs langues, et compte tenu de son caractère universel, nous postulons que la CPS pourrait également refléter le traitement neurocognitif des frontières prosodiques du français.
- 2) Les effets neurophysiologiques sont parfois observés dès le tout début de l'apprentissage d'une langue seconde, même si aucun effet n'est observé au niveau comportemental. Ainsi, nous prédisons que les frontières prosodiques CP et CT déclencheront également la CPS chez les apprenants sinophones. Toutefois, selon le niveau de maîtrise de la L2, des différences qualitatives pourraient être observées entre les deux groupes d'apprenants bilingues. L'amplitude de la CPS pourrait être plus petite et/ou la latence du pic de la

CPS pourrait être longue pour les apprenants intermédiaires. Cependant, comme l'indique le modèle de convergence L1-L2 proposé par Steinhauer *et al.* (2009) (Figure 1.20), il y aurait peut-être des différences concernant les amplitudes et/ou les latences des CPS entre les francophones natifs et les apprenants sinophones, en particulier pour le niveau intermédiaire.

- 3) Nous avons interprété les résultats de l'expérience perceptive à la lumière des prédictions *Late Closure Preference* (Frazier, 1979), *Minimal Attachment Strategy* (Kimball, 1973 ; Frazier & Fodor, 1978 ; Frazier, 1987 ; Bergmann *et al.*, 2008) et *Informative Boundary Hypothesis* (Clifton Jr. *et al.*, 2002 ; Watson & Gibson, 2005 ; Carlson *et al.*, 2009a, 2009b) (cf. section 3.4). Nous avons insisté sur l'importance du contexte prosodique complet pour l'établissement de la bonne relation hiérarchique entre les constituants syntaxico-prosodiques. Avec un contexte prosodique complet, nous prédisons qu'il n'y aurait pas de différence dans les CPS concernant les traitements neurocognitifs CP et CT.

4.2 Méthode

4.2.1 Participants

Afin de répondre à nos questions de recherche, nous avons recruté 60 locuteurs pour participer à l'expérience EEG, dont 20 participants francophones natifs, ayant le français comme seule langue maternelle, et 40 sinophones natifs apprenants du français avec le niveau intermédiaire ou avancé en français¹¹⁰ (Tableau 4.1).

Tableau 4.1 : Informations sur les trois groupes de participants de l'expérience EEG.

Groupe de participants	Nombre	Sexe (H : Homme, F : Femme)	Âge moyen (Écart-type)
Francophones natifs	20	6H, 14F	22,2 (3,30)
Apprenants sinophones avancés	20	3H, 17F	24,3 (2,00)
Apprenants sinophones intermédiaires	20	5H, 15F	24,25 (3,11)

¹¹⁰ La méthode de classification des apprenants en groupes intermédiaire et avancé était identique à celle utilisée dans les études acoustique et perceptive, basée sur leur niveau autodéclaré dans le questionnaire LEAP-Q (Marian *et al.*, 2007) prérempli avant l'expérience EEG.

Tous les participants ont signé un consentement de recherche (cf. section 4.6.1) avant de commencer l'expérience et chacun a reçu une indemnisation de vingt à trente euros en espèces pour sa participation. Ils ont aussi rempli un questionnaire sur leurs informations personnelles et sur leur profil d'apprentissage des langues (cf. section 4.6.2 pour les questionnaires utilisés dans l'expérience EEG). Tous les participants sont droitiers (Oldfield, 1971) et ont une vision normale ou corrigée. Aucun d'entre eux n'a déclaré présenter de trouble langagier, auditif, neurologique ou psychiatrique. Par ailleurs, aucun des 60 participants n'a pris part à l'expérience acoustique ou perceptive, afin de garantir qu'ils n'avaient jamais été exposés aux stimuli utilisés dans l'expérience EEG.

4.2.2 Stimuli

Les mêmes phrases, enregistrées par la même locutrice native du français que celle de l'expérience perceptive, ont été utilisées comme stimuli dans l'expérience EEG. Les stimuli dans l'expérience EEG étaient constitués de 150 phrases, dont 50 paires de phrases ambiguës avec les structures prosodiques CP et CT et 50 phrases distractrices qui ne présentent aucune ambiguïté (cf. section 4.6.3 pour les 150 stimuli utilisés dans l'expérience EEG).

Pour chaque phrase, deux marqueurs d'événement (déclencheurs, ang. *triggers*) ont été placés à l'onset de la dernière syllabe du syntagme verbal et du syntagme nominal. Ainsi, pour l'ensemble des 150 stimuli, nous avons 4 déclencheurs pour les quatre conditions prosodiques (Figure 4.1, Tableau 4.2).

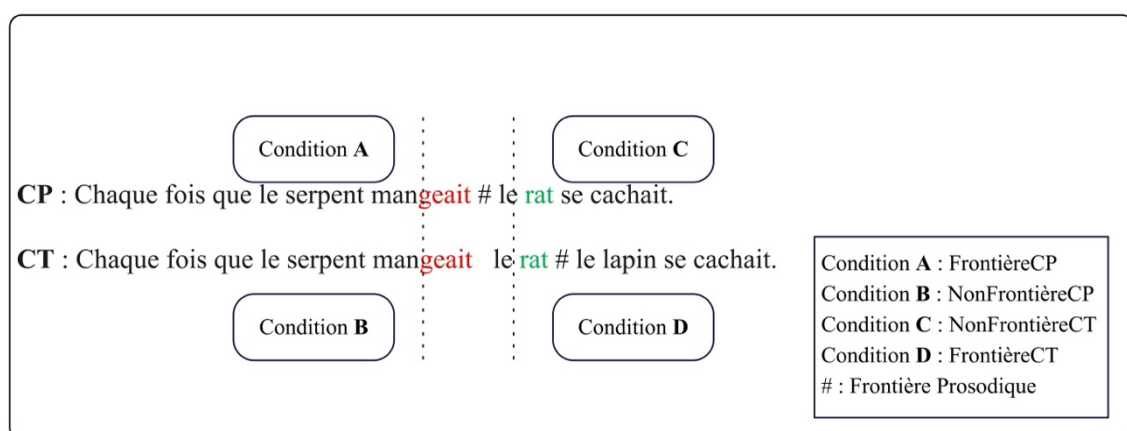


Figure 4.1 : Exemple de positions des quatre déclencheurs intégrés dans une paire de stimuli ambiguës.

Tableau 4.2 : Les dénominations et les significations des quatre conditions prosodiques de l'expérience EEG.

Condition	Dénomination	Signification
A	<i>FrontièreCP</i>	Présence d'une frontière de clôture précoce
B	<i>NonFrontièreCP</i>	Absence d'une frontière de clôture précoce
C	<i>NonFrontièreCT</i>	Absence d'une frontière de clôture tardive
D	<i>FrontièreCT</i>	Présence d'une frontière de clôture tardive

Cela nous a permis de comparer dans un premier temps les réponses CPS en fonction de la présence vs absence d'une frontière, et aussi de la position précoce vs tardive d'une frontière prosodique.

4.2.3 Procédure

Le protocole expérimental a été examiné et approuvé par le Comité d'Éthique de la Recherche de la Sorbonne Nouvelle (Avis CER-USN-01-2023) (cf. section 4.6.4) et par le Service de Protection des Données du CNRS (2-22087) (cf. section 4.6.5). L'expérience s'est déroulée au Laboratoire MoDyCo (UMR 7114, CNRS & Université Paris Nanterre). Les 150 stimuli ont été aléatoirement répartis en 5 blocs, avec 30 stimuli dans chaque bloc. Une courte pause a été proposée entre chaque bloc. Les participants ont été testés individuellement dans une pièce isolée sur les plans électrique et acoustique. Ils étaient confortablement installés sur un fauteuil, face à un écran d'ordinateur placé à 80 cm de leurs yeux, et un haut-parleur. L'expérimentateur a donné la consigne suivante : « *Vous allez écouter 150 phrases en français. Après chaque phrase, un mot s'affiche sur l'écran. Si ce mot apparaît dans la phrase que vous venez d'entendre, vous appuyez, avec la main droite, sur la touche « J » du clavier pour le choix « oui » ; s'il n'apparaît pas dans la phrase, vous appuyez sur la touche « F » du clavier pour le choix « non » (tâche de vérification lexicale). Il faut rester le plus immobile possible pendant l'expérience.* ». Après l'explication des consignes, un bloc d'entraînement constitué de 9 stimuli (qui ne font pas partie des 150 stimuli) a été proposé aux participants afin qu'ils se familiarisent avec la tâche expérimentale. L'expérience débutait lorsque le participant confirmait avoir bien compris les consignes. Les stimuli ont été diffusés un par un par le logiciel *Presentation* (version 24.0), avec le mot cible projeté sur l'écran à la fin de chaque stimulus (Figure 4.2). Entre chaque bloc (tous les 30 stimuli), une pause d'un temps illimité a été insérée pour que le participant puisse se reposer.

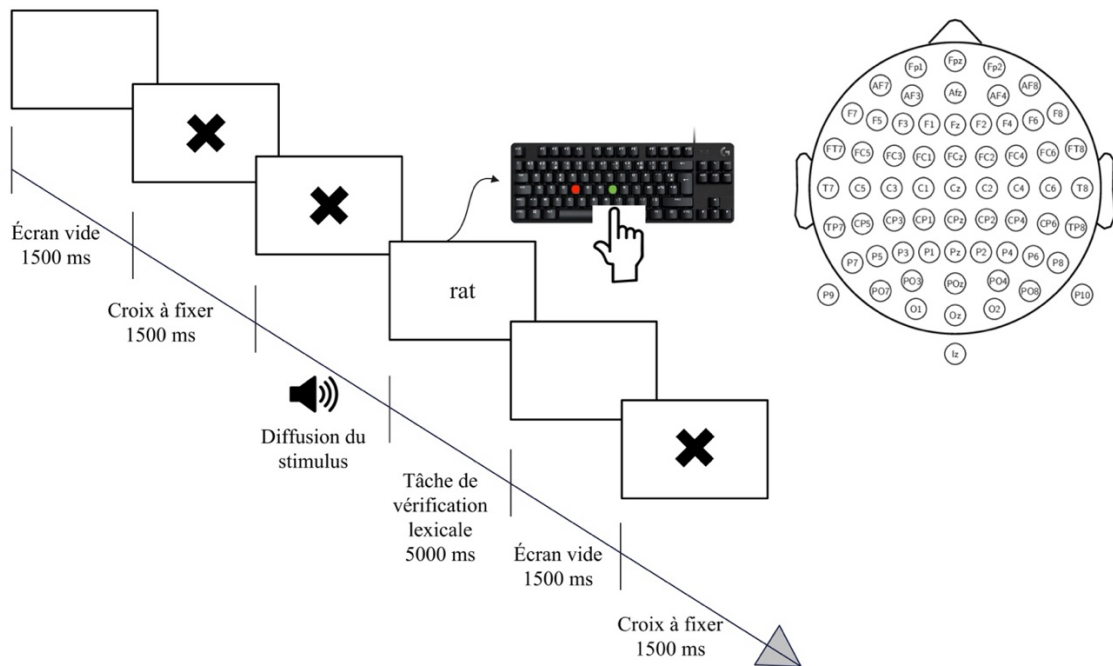


Figure 4.2 : Exemple du déroulé temporel d'un stimulus (partie gauche) et positions des 64 électrodes du système Biosemi ActiveTwo (partie droite).

Les signaux EEG ont été enregistrés en continu à l'aide de 64 électrodes (Ag-AgCl électrodes ; Biosemi ActiveTwo system, Amsterdam, Pays-Bas) placées sur un bonnet (10/20 international system), avec deux électrodes supplémentaires au niveau des mastoïdes. Afin de détecter les mouvements oculaires, quatre électrodes EOG (électrooculographie) ont été utilisées, une à gauche et une à droite de chaque œil, et une au-dessus et une au-dessous de l'œil gauche. Les signaux EEG étaient numérisés à une fréquence d'échantillonnage de 512 Hz. Un filtrage en temps réel a été appliqué avec une bande passante entre 0,05 Hz et 100 Hz.

4.3 Prétraitement et analyses

4.3.1 Prétraitement des données

Les données ont été prétraitées hors ligne à l'aide des logiciels Matlab (R2024a) et EEGLAB (Delorme & Makeig, 2004). Pour chaque participant, le signal EEG continu a été re-référencé en utilisant les deux électrodes au niveau des mastoïdes gauche et droite, ré-échantillonné à 256 Hz et filtré dans une bande passante de 0,5 à 30 Hz. Ensuite, les données ont été visuellement inspectées pour éliminer les segments affectés par d'importants artefacts liés aux mouvements de la tête et pour remplacer les canaux de signal de mauvaise qualité par des données d'interpolation spatiale. Les artefacts, comprenant les mouvements oculaires et clignements des yeux, ont

été supprimés à l'aide d'une Analyse en Composantes Indépendantes (ICA). Des segments de données (*epochs*) correspondant aux quatre déclencheurs ont été extraits dans ERPLAB (Lopez-Calderon & Luck, 2014) : le segment temporel a été choisi entre 200 ms avant le déclencheur (ligne de base) et 1500 ms après (i.e., fenêtre temporelle d'analyse (FT) de [-200, 1500]). Tout segment présentant un voltage supérieur à 100 μV a été supprimé, indépendamment de l'électrode considérée. Le taux de rejet pour chaque participant était compris entre 0% et 30%, et de 16,53% en moyenne (écart-type = 11,03). Les signaux restants, exempts d'artéfacts, ont été inclus dans la procédure statistique du grand moyennage.

4.3.2 Analyses des données

En nous appuyant sur des études antérieures (e.g., Drury *et al.*, 2016 ; Gilbert *et al.*, 2023), qui ont utilisé la CPS comme marqueur de frontières prosodiques, nous avons défini trois régions d'intérêt (ROIs) pour l'analyse de nos données : frontale, centrale et pariétale, avec 14 électrodes dans chaque ROI (Figure 4.3) :

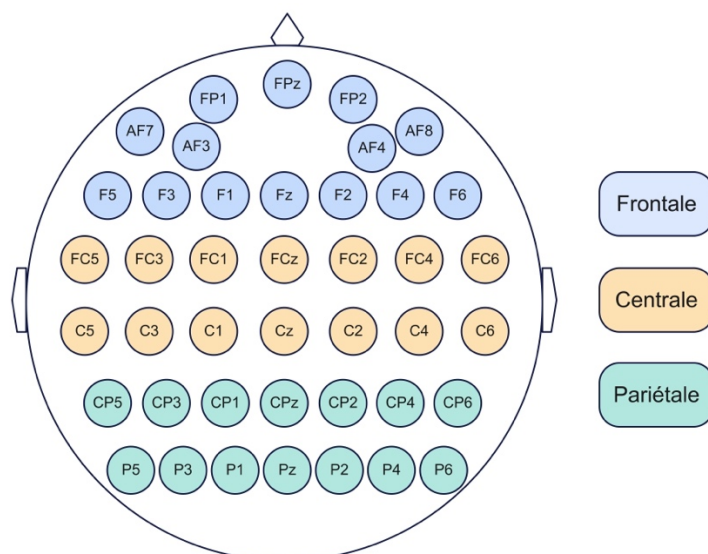


Figure 4.3 : Électrodes dans les trois régions d'intérêt (ROIs) prédéfinies dans notre étude EEG.

Pour définir les fenêtres temporelles d'analyse (FTs), nous n'avons pas procédé à l'inspection visuelle, qui risque d'introduire un biais de l'observateur (Keil *et al.*, 2014 ; Luck & Gaspelin, 2017 ; Fields & Kuperberg, 2020). À la place, nous avons adopté une approche plus objective, qui permet de sélectionner les FTs en fonction des latences des zones fractionnées (Luck, 2014). Cette approche repose sur

l'analyse du décours temporel de la différence des ondes en fractionnant en un pourcentage prédéfini sur une fenêtre de mesure entière. Par exemple, dans la Figure 4.4, la FT [200-700] ms est définie pour l'analyse d'une composante donnée, et l'approche de latence de la zone fractionnée permet de trouver la position à 50%, qui découpe l'ensemble de la FT mesurée en deux zones équivalentes.

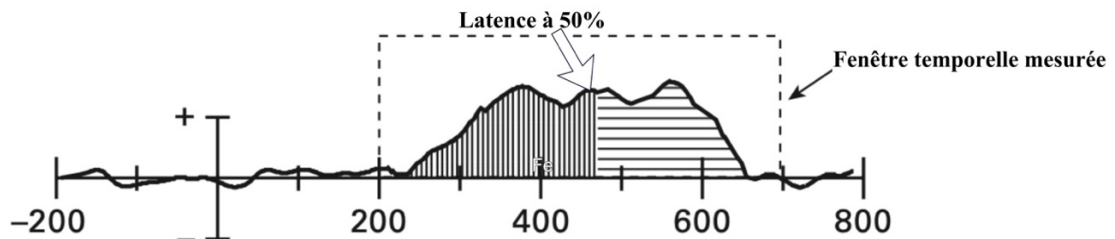


Figure 4.4 : Illustration de l'approche de latence d'une zone fractionnée (adaptée de Luck (2014 : 298)).

De manière similaire, en utilisant la même approche, nous avons déterminé non seulement la position à 50%, considérée comme le pic de la composante, mais aussi la position à 15%, considérée comme le début de la FT de la composante (onset), et la position à 85%, considérée comme la fin de la FT de la composante (offset). Pour ce faire, avec le signal du grand moyennage de chaque groupe, nous avons d'abord extrait la différence entre les conditions prosodiques¹¹¹ dans ERPLAB (Lopez-Calderon & Luck, 2014). En nous basant sur les signaux de différence créés pour chacun des trois groupes, nous avons ensuite extrait les latences précises aux positions 15%, 50% et 85%, où les valeurs d'amplitude deviennent positives (« *rectified area (negative values become positive)* ») sur les 42 électrodes dans les trois ROIs au sein de la FT [0-1500] ms. Les latences moyennes des 42 électrodes dans les trois positions sont notées comme onset, pic et offset des effets CPS pour les trois groupes de participants (Tableau 4.3) :

¹¹¹ Pour chaque différence des signaux, il s'agit toujours de « Frontière – NonFrontière » (i.e. « Signal de la condition *FrontièreCP* – Signal de la condition *NonFrontièreCP* » et « Signal de la condition *FrontièreCT* – Signal de la condition *NonFrontièreCT* »). Étant donné que la composante CPS représente un effet positif, nous nous attendons à une valeur positive de la différence, ce qui implique l'émergence de la CPS.

Tableau 4.3 : Onsets, pics et offsets (identifiés par l'approche de latence de la zone fractionnée [0-1500] ms) du résidu positif de la différence « Frontière – NonFrontière » pour chaque condition prosodique de chaque groupe de participants.

Groupe	Condition Prosodique	Onset (ms)	Pic (ms)	Offset (ms)
Francophones natifs	<i>FrontièreCP vs NonFrontièreCP</i>	455	591	720
	<i>FrontièreCT vs NonFrontièreCT</i>	562	726	924
Apprenants avancés	<i>FrontièreCP vs NonFrontièreCP</i>	610	702	797
	<i>FrontièreCT vs NonFrontièreCT</i>	634	843	1062
Apprenants intermédiaires	<i>FrontièreCP vs NonFrontièreCP</i>	572	724	843
	<i>FrontièreCT vs NonFrontièreCT</i>	596	886	1096

Les latences de l'onset et de l'offset sont donc définies comme le début et la fin de la FT dans laquelle la CPS est censée être déclenchée. Nous avons ensuite défini les trois FTs pour chaque condition prosodique de chaque groupe de participants (Tableau 4.4) :

Tableau 4.4 : Trois FTs de la composante CPS pour chaque condition prosodique et chaque groupe de participants.

Groupe	Condition Prosodique	1^{ère} FT (ms)	2^{ème} FT (ms)	3^{ème} FT (ms)
Francophones natifs	<i>FrontièreCP vs NonFrontièreCP</i>	[0, 455]	[455, 720]	[720, 1500]
	<i>FrontièreCT vs NonFrontièreCT</i>	[0, 562]	[562, 924]	[924, 1500]
Apprenants avancés	<i>FrontièreCP vs NonFrontièreCP</i>	[0, 610]	[610, 797]	[797, 1500]
	<i>FrontièreCT vs NonFrontièreCT</i>	[0, 634]	[634, 1062]	[1062, 1500]
Apprenants intermédiaires	<i>FrontièreCP vs NonFrontièreCP</i>	[0, 572]	[572, 843]	[843, 1500]
	<i>FrontièreCT vs NonFrontièreCT</i>	[0, 596]	[596, 1096]	[1096, 1500]

Les valeurs de l'amplitude moyenne des électrodes ont ensuite été extraites à l'aide du logiciel EPRLAB (Lopez-Calderon & Luck, 2014) dans les trois FTs et pour toutes les conditions prosodiques. Les amplitudes moyennes ont été analysées à l'aide d'un modèle linéaire mixte pour chaque FT. Cela implique de tester dans un premier temps l'effet de la frontière prosodique (deux niveaux : Frontière vs NonFrontière) et l'effet de la ROI (trois niveaux : frontale vs centrale vs pariétal) et de répondre aux questions de recherche 1) et 2). Les participants ont été considérés comme facteur à effet aléatoire. Le modèle établi était donc le suivant : [Amplitude moyenne ~ Condition Prosodique * ROI + (1 | Participant)] et puis analysé dans JASP (Version 0.19.3).

Outre l'approche « traditionnelle »¹¹² faisant appel au modèle linéaire mixte, nous avons mis en œuvre des analyses univariées de masse (*Mass Univariate*) (Groppe *et al.*, 2011a, b) afin d'enrichir l'analyse de nos données CPS. L'approche d'analyse univariée de masse procède par tests statistiques à chaque point temporel sur l'ensemble des données, tout en évitant de définir au préalable une ou plusieurs FT(s) et/ou ROI(s). Par ailleurs, cette approche permet de réduire la possibilité d'avoir des erreurs de Type I¹¹³ (Luck & Gaspelin, 2017 ; Fields & Kuperberg, 2020).

Ainsi, pour analyser les données CPS chez les francophones natifs et les apprenants sinophones, et en nous inspirant des travaux de Fields et Kuperberg (2020), ainsi que Brooks *et al.* (2017), nous avons adopté dans un premier temps l'approche traditionnelle et puis compléter les résultats par l'analyse univariée de masse. Nous avons effectué ces analyses à l'aide de la boîte à outils *Factorial Mass Univariate ERP Toolbox* (FMUT) (Fields, 2017 ; Fields & Kuperberg, 2020), à travers une série de

¹¹² L'approche traditionnelle est caractérisée par un moyennage temporel et spatial : au niveau temporel, on mesure la valeur moyenne d'une ou plusieurs propriété(s) neurophysiologique(s) (souvent amplitude moyenne, ou latence, pic, etc.) d'un effet ERP dans une FT prédéfinie à travers l'inspection visuelle (Luck, 2014) ; au niveau spatial, en suivant des études similaires, une série d'électrodes sont sélectionnées comme représentatives (définie sous terme de ROI (*region of interest* dans la littérature), dont les valeurs moyennes dans la FT prédéfinie sont extraites et comparées par des analyses statistiques (Fields & Kuperberg, 2020). Cette approche, étant beaucoup critiquée (voir Kilner, 2013 ; Luck & Gaspelin, 2017 ; Fields & Kuperberg, 2020 ; Zhang & Kappenman, 2024), reste toutefois la méthodologie la plus adoptée dans la majorité d'études électroencéphalographiques.

¹¹³ "...the mass univariate approach, in which a separate *t* test (or related statistic) is computed at every time point for every electrode site, and some kind of correction for multiple comparisons is applied to control the overall Type I error rate. The traditional Bonferroni correction is usually unreasonably conservative, but a variety of other correction factors are now available and are implemented in free, open-source analysis packages, such as the Mass Univariate Toolbox and FieldTrip. These approaches are still fairly conservative, but they may be the best option when no a priori information is available to guide the choice of latency windows and electrode sites." (Luck & Gaspelin, 2017: 150).

tests de clusters basés sur les permutations¹¹⁴. En général, lorsqu'un potentiel électrique lié à un événement (ERP) est déclenché, celui-ci ne se manifeste pas de manière aléatoire sur le cuir chevelu, mais plutôt sous la forme de plusieurs régions et points temporels adjacents, créant ainsi des « clusters ». Le test basé sur les clusters¹¹⁵ permet d'identifier plus efficacement un effet ERP, tout en le distinguant du bruit, qui est souvent aléatoire et qui peut être confondu avec un véritable effet ERP¹¹⁶ (Groppe *et al.*, 2011a, Fields & Kuperberg, 2020). Selon les recommandations du tutoriel¹¹⁷, l'analyse univariée de masse commence généralement par une première analyse sur une FT restreinte, où un effet ERP est supposé être observé. Si cette première analyse n'aboutit à aucun résultat significatif, une seconde analyse est alors effectuée sur la FT entière afin de vérifier la présence d'effets ERP en dehors de la FT prédéfinie. Il faut néanmoins noter qu'un test sur une FT trop large ou entière pourrait réduire la significativité d'un effet ERP.

Dans l'analyse univariée de masse de notre étude, nous avons adopté donc la procédure suivante : pour chaque condition prosodique de chaque groupe de participants, nous avons effectué un test de clusters basé sur les permutations. Nous avons testé d'abord la FT restreinte (définie par l'approche de latences des zones fractionnées, Tableau 4.4), où la CPS est supposée être observée. Si aucun résultat significatif n'est obtenu, nous avons procédé ensuite à l'analyse de la FT entière [0-1500] ms. Après chaque analyse effectuée par la boîte à outils FMUT, un rapport de

¹¹⁴ Nous avons choisi le test de clusters basé sur les permutations car il s'adapte bien à un effet ERP qui dure relativement longtemps, comme c'est le cas pour la CPS : "*cluster-based permutation tests with weak control of FWER, is much more powerful than permutation-based strong control of FWER for detecting broadly distributed effects that span many time points and sensors.*" (Groppe *et al.*, 2011a: 1723); "*...the cluster-based test is probably the best suited for mass univariate analysis of ERP/ERF data when one is not interested in potentially very focally distributed effects.*" (Groppe *et al.*, 2011b: 1735).

¹¹⁵ Voici une présentation de comment un cluster significatif est calculé : "*First, some threshold for cluster inclusion is defined; often this is the F-value that would be significant with no correction. Spatially and temporally adjacent points that exceed this threshold are considered a cluster. A cluster mass statistic is calculated by summing all the F-values in a cluster. The null distribution for the cluster mass statistic is estimated via a permutation approach, and observed clusters exceeding the 1 - α percentile of the distribution are considered significant*". (Explication extraite du tutoriel de la boîte à outils FMUT (<https://github.com/ericcfields/FMUT/wiki/Mass-Univariate-Statistics-and-Corrections>)).

¹¹⁶ "*The motivation for forming clusters of tests is that true ERP/ERF effects are generally more likely than ERP/ERF noise to occur coherently across multiple adjacent time points and sensors. Consequently, ERP/ERF effects will typically stand out more clearly from noise using cluster-based statistics.*" (Groppe *et al.*, 2011b: 1718). "*...for a cluster statistic (i.e., a statistic representing the size of a cluster of adjacent time points and electrodes showing an effect larger than some prespecified threshold...*" (Fields & Kuperberg, 2020: 5).

¹¹⁷ https://openwetware.org/wiki/Mass_Univariate_ERP_Toolbox:_general_advice

résultats est automatiquement généré indiquant le seuil de significativité et les électrodes du cluster identifié, ainsi que les latences et la position du pic de l'effet. Dans notre étude, les tests de clusters basés sur les permutations ont été menés avec 10 000 itérations, .75 comme distance maximale entre électrodes considérées comme un cluster, 0,05 comme niveau d'alpha¹¹⁸.

4.4 Résultats des effets CPS

Pour étudier la CPS dans les stimuli de notre étude, comme illustré dans la Figure 4.1, nous avons 4 déclencheurs insérés dans les stimuli, pour marquer les conditions *FrontièreCP*, *NonFrontièreCP*, *FrontièreCT*, *NonFrontièreCT* (cf. Tableau 4.2 pour leurs significations). Ainsi, dans nos analyses, nous avons comparé les amplitudes moyennes, d'une part entre *FrontièreCP* et *NonFrontièreCP*, et d'autre part entre *FrontièreCT* et *NonFrontièreCT*. Étant donné que la CPS est une composante ERP positive, nous nous attendons à ce que les amplitudes moyennes soient significativement plus positives ou plus élevées dans la condition *FrontièreCP* par rapport à la condition *NonFrontièreCP*, et dans la condition *FrontièreCT* par rapport à la condition *NonFrontièreCT*. Si cela se confirme, nous pourrions déduire que la frontière prosodique est traitée sur le plan neurocognitif comme l'atteste la présence d'une CPS.

Dans les sections suivantes, nous reporterons d'abord les résultats d'analyse des données CPS pour chacun des trois groupes de participants : la section 4.4.1 pour les francophones natifs, la section 4.4.2 pour les apprenants avancés, et la section 4.4.3 pour les apprenants intermédiaires. Ces premiers résultats par groupes nous permettront de répondre aux questions de recherche 1) et 2), c'est-à-dire d'examiner si les effets CPS sont observables dans les structures CP, puis dans celles

¹¹⁸ Explication sur l'analyse de clusters basée sur les permutations (reprise du tutorial de *Factorial Mass Univariate ERP Toolbox* (<https://github.com/ericcfields/FMUT/wiki/Mass-univariate-statistics-and-corrections>): "This correction is based on testing clusters for significance, rather than individual electrodes/time points. First, some threshold for cluster inclusion is defined; often this is the F-value that would be significant with no correction. Spatially and temporally adjacent points that exceed this threshold are considered a cluster. A cluster mass statistic is calculated by summing all the F-values in a cluster. The null distribution for the cluster mass statistic is estimated via a permutation approach, and observed clusters exceeding the 1 - α percentile of the distribution are considered significant. This provides weak family-wise error control: If no significant effects are present, the probability of a significant cluster is α . However, significance applies only to a cluster as a whole: you cannot reject the null at individual electrodes/time points within a cluster with a specified probability of error or level of confidence."

de CT, respectivement chez les francophones natifs et les apprenants avancés et intermédiaires. Dans chaque section, les signaux ERP seront illustrés avec 9 électrodes représentatives : Fz, F3, F4 du niveau frontal ; Cz, C3, C4 du niveau central ; Pz, P3, P4 du niveau pariétal (voir entre autres Steinhauer *et al.*, 1999 ; Steinhauer, 2003 ; Pauker *et al.*, 2011). Les signaux de toutes les 42 électrodes dans les 3 ROIs sont représentés dans les annexes de ce chapitre (cf. section 4.6.6).

Si les CPS sont déclenchées dans les conditions CP et CT dans chacun des trois groupes de participants, nous avons examiné ensuite les différences dans les traitements neurocognitifs des frontières prosodiques CP vs CT au sein de chaque groupe. Dans notre étude sur la perception, nous avons mis en évidence une différence dans le traitement comportemental entre les frontières CP et CT. Nous avons interprété les résultats de cette étude dans le cadre de trois prédictions. Lorsque le contexte prosodique d'une phrase est incomplet, les auditeurs se voient contraints d'établir des relations hiérarchiques appropriées entre les constituants syntaxiques, favorisant ainsi un traitement de clôture tardive (CT) (*Late Closure Preference*). Ce traitement par défaut est compatible avec la stratégie de l'attachement minimal (*Minimal Attachment Strategy*), qui permet de réduire la charge cognitive et d'optimiser l'utilisation de la mémoire de travail. Lorsque le contexte prosodique d'une phrase est complet et disponible dans la perception de la parole, l'effet d'une frontière prosodique donnée dépend d'autres frontières apparues avant et après dans la phrase (*Informative Boundary Hypothesis*) (cf. section 3.4 pour plus de détails). Cela nous mène à nous interroger : comment les différences observées dans le traitement comportemental des frontières CP vs CT se manifesteraient-elles sur le plan neurocognitif à travers les effets CPS, lorsque le contexte prosodique complet est pleinement disponible dans le stimulus ? Afin de répondre à ce questionnement, pour chaque groupe de participants, nous comparerons les différences d'amplitudes moyennes des 42 électrodes des trois ROIs entre la présence et l'absence d'une frontière prosodique, autrement dit la différence « *FrontièreCP – NonFrontièreCP* » vs la différence « *FrontièreCT – NonFrontièreCT* ». Ces comparaisons nous ont permis de comparer le degré d'intégration neurocognitive des informations prosodiques en fonction des positions précoce et tardive. Nous avons utilisé également l'approche univariée de masse pour affiner les comparaisons et compléter les résultats. Les résultats

obtenus ont permis de répondre à la question de recherche 3), à savoir comparer les traitements neurocognitifs entre CP et CT, et d'examiner s'il existe des différences au niveau du traitement comportemental (ou perceptif) par rapport au traitement neurocognitif.

Et enfin, dans la section 4.4.4, nous avons comparé le traitement d'une même condition prosodique entre les participants, d'une part pour les frontières CP dans les trois groupes, et d'autre part pour les frontières CT dans les trois groupes. Ces comparaisons nous ont permis d'examiner comment les traitements neurocognitifs d'une même condition prosodique varieraient en fonction de niveaux de français.

4.4.1 CPS chez les francophones natifs

4.4.1.1 Conditions *FrontièreCP* vs *NonFrontièreCP*

Les modèles linéaires mixtes ont révélé une différence significative entre *FrontièreCP* et *NonFrontièreCP*, pour toutes les FTs ([0-455], [455-720], [720-1500]) (Tableau 4.5). Toutefois, une amplitude significativement plus positive pour la condition *FrontièreCP* est observée uniquement dans la FT [455-720] ms. En revanche, pour les deux autres FTs [0-455] ms et [720-1500] ms, l'amplitude de l'onde est significativement plus négative dans la condition *FrontièreCP* comparée à la condition *NonFrontièreCP* (Figure 4.5). La positivité unique pour la condition *FrontièreCP* dans la FT [455-720] suggère que la CPS est déclenchée dans cette FT.

Tableau 4.5 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les participants **francophones natifs**.

FT (ms)	Effet de conditions prosodiques	Interaction (Conditions prosodiques * ROI)
[0-455]	b=0,27 ; SE=0,034 ; t= -7,8 ; p<0,001	b= -0,22 ; SE=0,049 ; t= -4,45 ; p<0,001
[455-720]	b=0,58 ; SE=0,054 ; t=10,7 ; p<0,001	b= -0,16 ; SE=0,077 ; t= -2,06 ; p=0,11
[720-1500]	b=0,44 ; SE=0,054 ; t= -8,1 ; p<0,001	b= -0,005 ; SE=0,076 ; t= -0,07 ; p=0,02

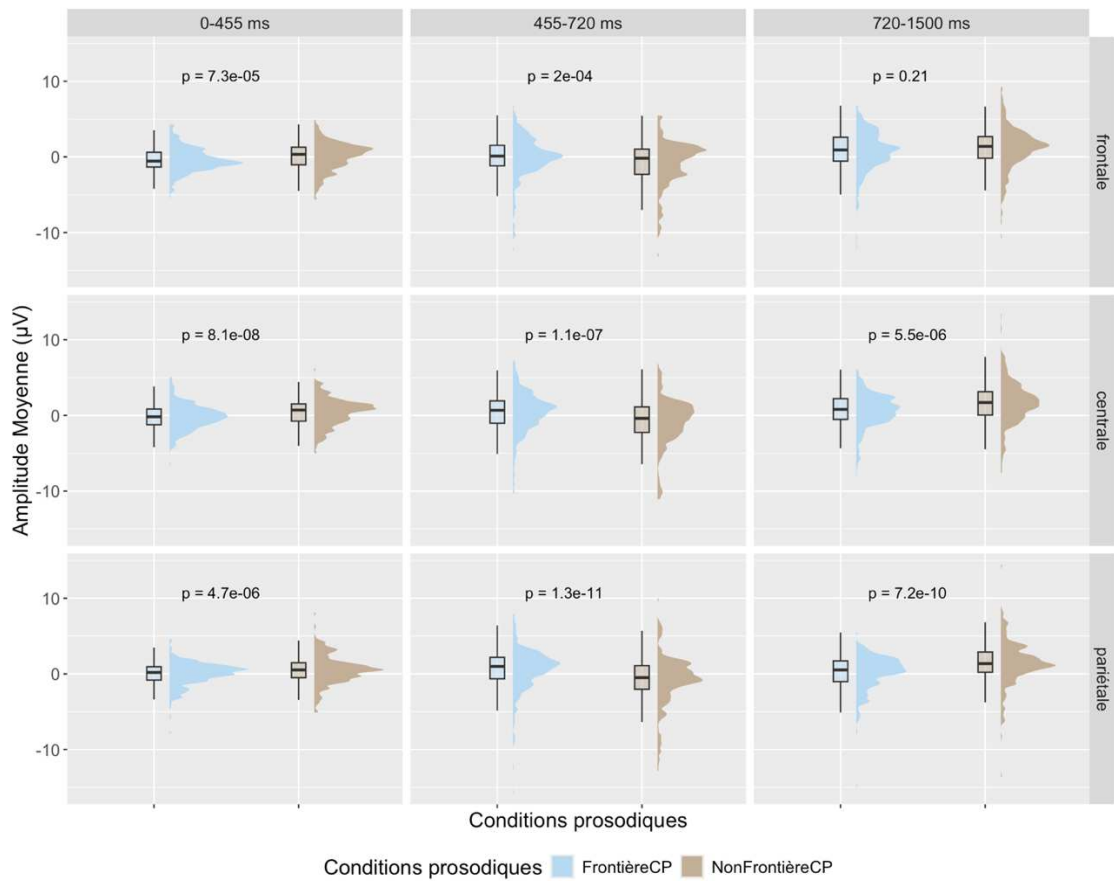


Figure 4.5 : Différences d'amplitudes moyennes entre les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* dans les trois FTs et les trois ROIs chez les participants **francophones natifs**.

Afin de confirmer ce constat, nous avons étudié les signaux ERP (Figure 4.6) et comparé les cartes topographiques des trois FTs (Figure 4.7).

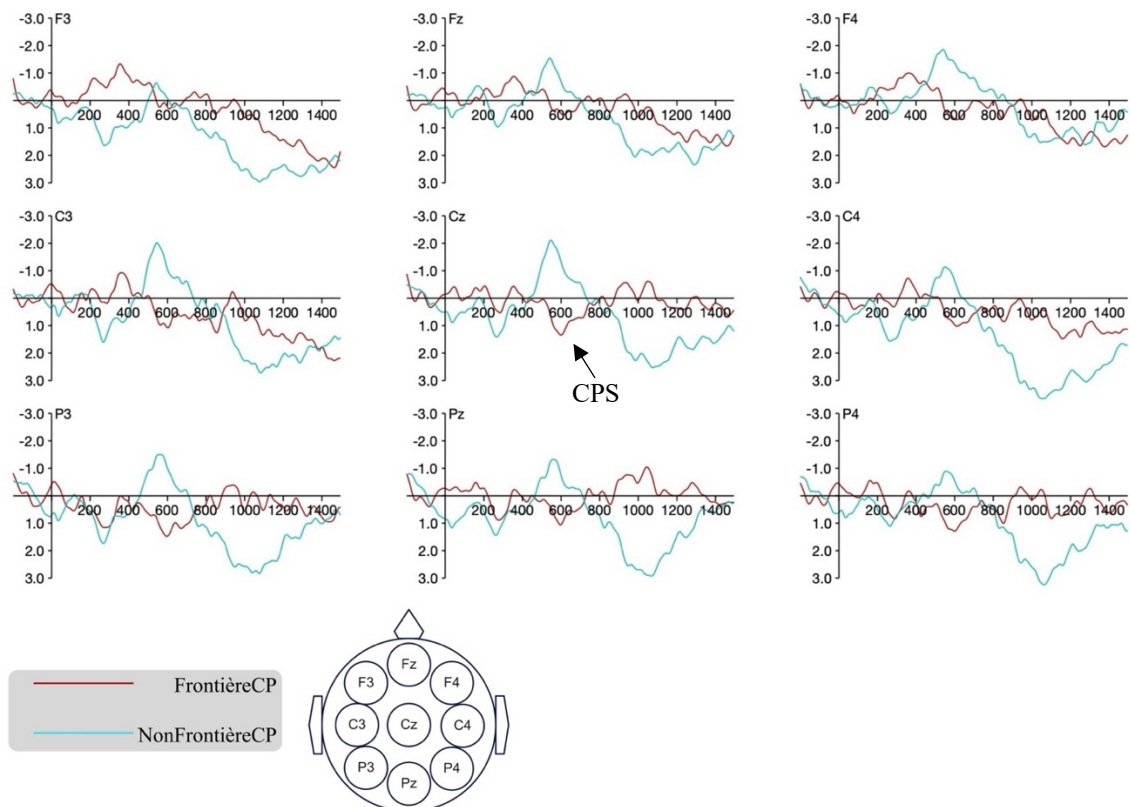


Figure 4.6 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les participants **francophones natifs**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

Comme le montrent les signaux ERP (Figure 4.6), l'effet CPS de la condition *FrontièreCP* chez les francophones natifs s'étend sur toutes les neuf électrodes représentatives¹¹⁹, tant sur la ligne centrale qu'au niveau bilatéral, aux alentours de [450-750] ms. Sur les cartes topographiques des trois FTs (Figure 4.7), la positivité de l'effet CPS pour la condition *FrontièreCP* chez les francophones natifs se concentre principalement au niveau centro-pariétal dans la FT [455-720].

¹¹⁹ Bien que l'effet CPS sur les électrodes frontales soit relativement faible par rapport aux électrodes centrales et pariétales, le résultat du modèle linéaire mixte dans la FT [455-720] ms (Figure 4.5) indique bien une amplitude significativement plus grande dans la condition *FrontièreCP* par rapport à la condition *NonFrontièreCP* sur le niveau frontal. Par ailleurs, les effets CPS sont observés également dans la plupart des électrodes frontales (cf. Figure 4.40, les 42 électrodes pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP*, chez les participants francophones natifs dans l'annexe 4.6.6).

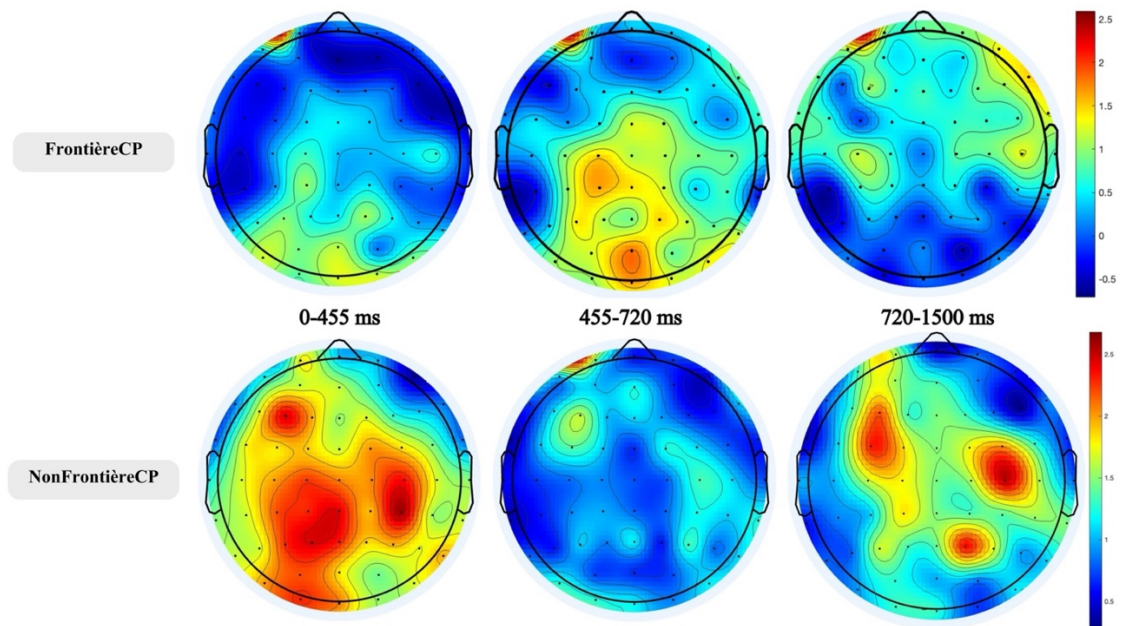


Figure 4.7 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les participants **francophones natifs**.

Par l'analyse univariée de masse, aucun cluster significatif n'a été identifié, ni dans la FT restreinte [455-720] ms (avec toutes les valeurs p égales ou supérieures à 0,08), ni dans la FT entière [0-1500] ms (avec toutes les valeurs p égales ou supérieures à 0,11).

En résumé, les résultats indiquent que les auditeurs francophones perçoivent et traitent les frontières CP. Cela se manifeste par la composante CPS observée sur l'ensemble des électrodes, avec un effet se concentrant aux niveaux central et pariétal dans la FT [455-720], tant sur la ligne centrale que sur les régions bilatérales.

4.4.1.2 Conditions *FrontièreCT* vs *NonFrontièreCT*

Pour les trois FTs [0-562] ms, [562-924] ms et [924-1500] ms des conditions *FrontièreCT* vs *NonFrontièreCT*, les modèles linéaires mixtes ont révélé une différence significative dans toutes les trois FTs (Tableau 4.6). Toutefois, une amplitude significativement plus positive pour la condition *FrontièreCT* est observée seulement dans la deuxième FT ([562-924] ms). En revanche, pour les deux autres FTs [0-562] ms et [924-1500] ms, les amplitudes moyennes sont significativement plus négatives dans la condition *FrontièreCT* comparées à la condition *NonFrontièreCT* (Figure 4.8). Autrement dit, la CPS pour la condition *FrontièreCT* est déclenchée uniquement pendant la FT [562-924] ms.

Tableau 4.6 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les participants **francophones natifs**.

FT (ms)	Effet de conditions prosodiques	Interaction (Conditions prosodiques * ROI)
[0-562]	b=0,45 ; SE=0,04 ; t=10,96 ; p<0,001	b=0,15 ; SE=0,059 ; t=2,62 ; p=0,014
[562-924]	b=0,27 ; SE=0,057 ; t= -4,69 ; p<0,001	b=0,23 ; SE=0,08 ; t=2,80 ; p<0,001
[924-1500]	b=0,29 ; SE=0,06 ; t=4,72 ; p<0,001	b=0,36 ; SE=0,086 ; t=4,22 ; p<0,001

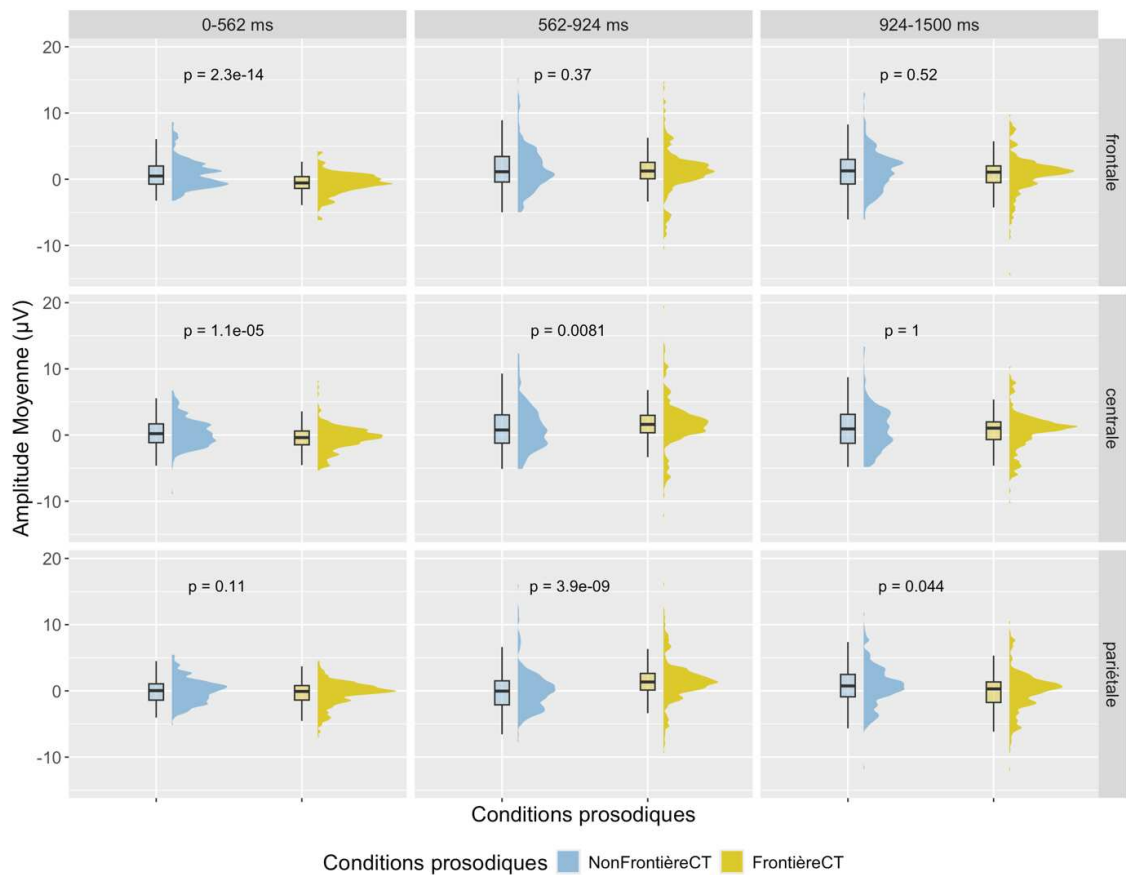


Figure 4.8 : Différences d'amplitudes moyennes entre les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* dans les trois FTs et les trois ROIs chez les participants **francophones natifs**.

Nous avons examiné ensuite les signaux ERP (Figure 4.9) et les chronogrammes des trois FTs (Figure 4.10) pour vérifier le constat des analyses statistiques.

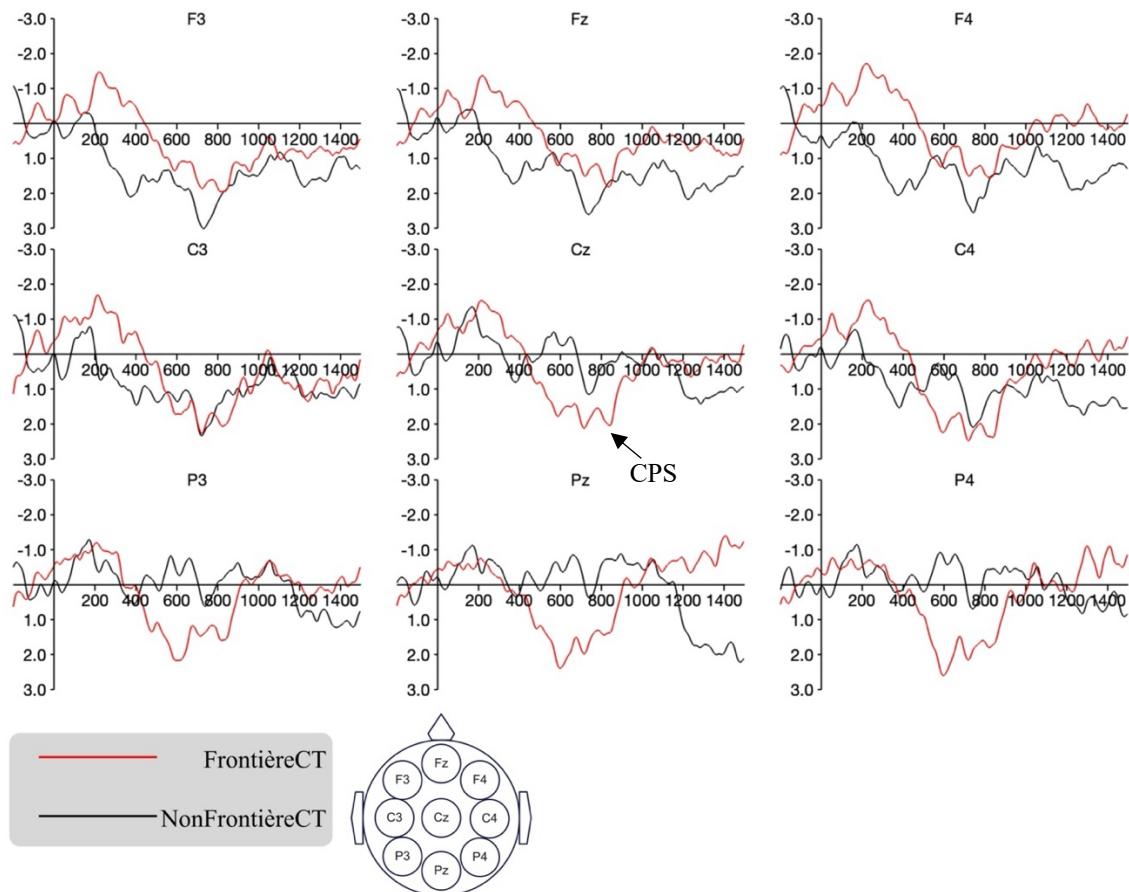


Figure 4.9 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les participants **francophones natifs**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

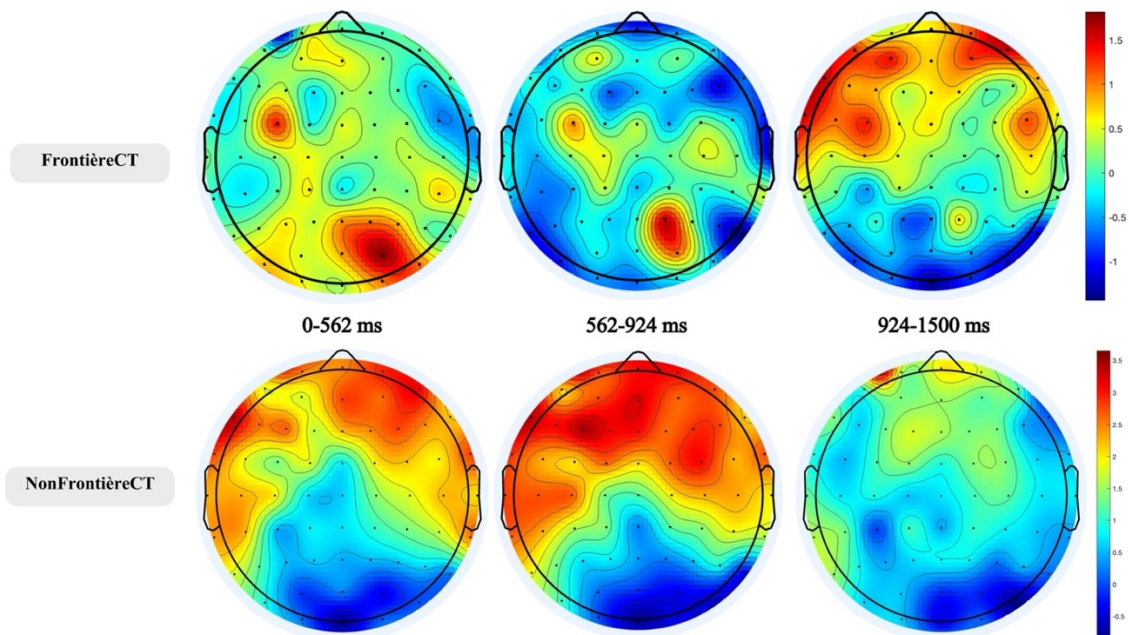


Figure 4.10 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les participants **francophones natifs**.

Comme le montrent les signaux ERP (Figure 4.9) et les cartes topographiques (Figure 4.10), la composante CPS pour *FrontièreCT* chez les francophones natifs émerge au niveau pariétal à partir de 400 ms après la dernière syllabe avant les frontières prosodiques CT. Son effet s'étend au niveau centro-pariétal, avec une latéralisation à droite dans la FT [562-924] ms. Le résultat ci-dessus tiré des signaux ERP et des cartes topographiques correspond aux résultats des analyses statistiques (Figure 4.8) : dans la FT [562-924] ms, les amplitudes moyennes pour *FrontièreCT* sont significativement plus positives comparées à *NonFrontièreCT* uniquement au niveau centro-pariétal tandis qu'au niveau frontal, la différence d'amplitudes moyennes n'est pas significative.

Nous avons effectué ensuite l'analyse univariée de masse : sur la FT restreinte [562-924] ms, un cluster significatif ($p=0,04$) a été identifié. Il se concentre, comme révélée par l'approche traditionnelle, au niveau centro-pariétal (Figure 4.11), avec un pic sur l'électrode Oz à 691 ms après l'onset de la dernière syllabe avant les frontières prosodiques CT.

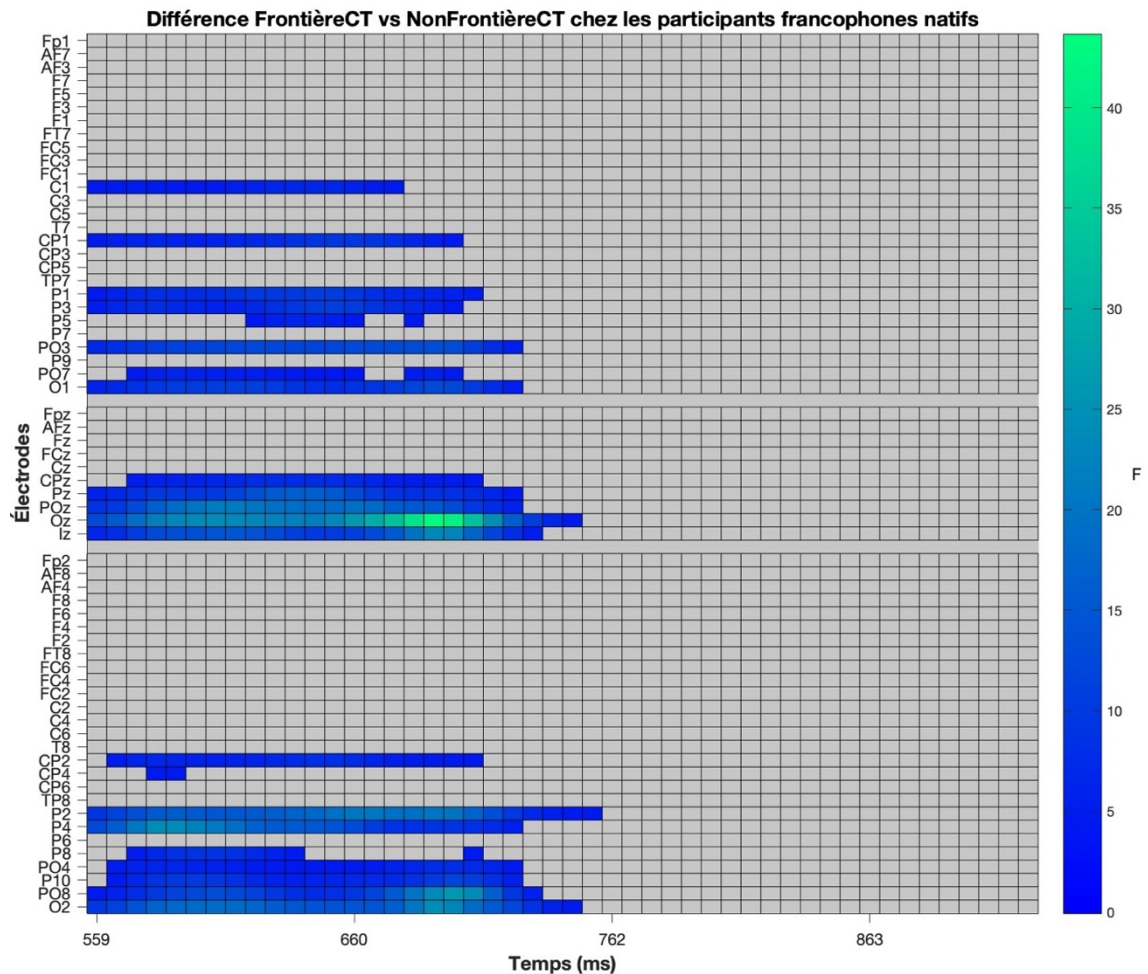


Figure 4.11 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* dans la FT [562-924] ms chez les participants **francophones natifs**.

Pour résumer, les résultats des approches traditionnelle et univariée de masse indiquent que les francophones natifs perçoivent et traitent les frontières CT. Cela se manifeste par la composante CPS observée pendant l'intervalle temporel de [562-924] ms, au niveau centro-pariétal, avec une latéralisation à droite.

4.4.1.3 Comparaison des effets CPS dans les conditions *FrontièreCP* vs *FrontièreCT*

Dans les deux sections précédentes (4.4.1.1 et 4.4.1.2), les résultats de nos analyses montrent que les composantes CPS sont déclenchées lorsque les auditeurs francophones natifs perçoivent et traitent respectivement les frontières prosodiques CP et CT (Figure 4.12).

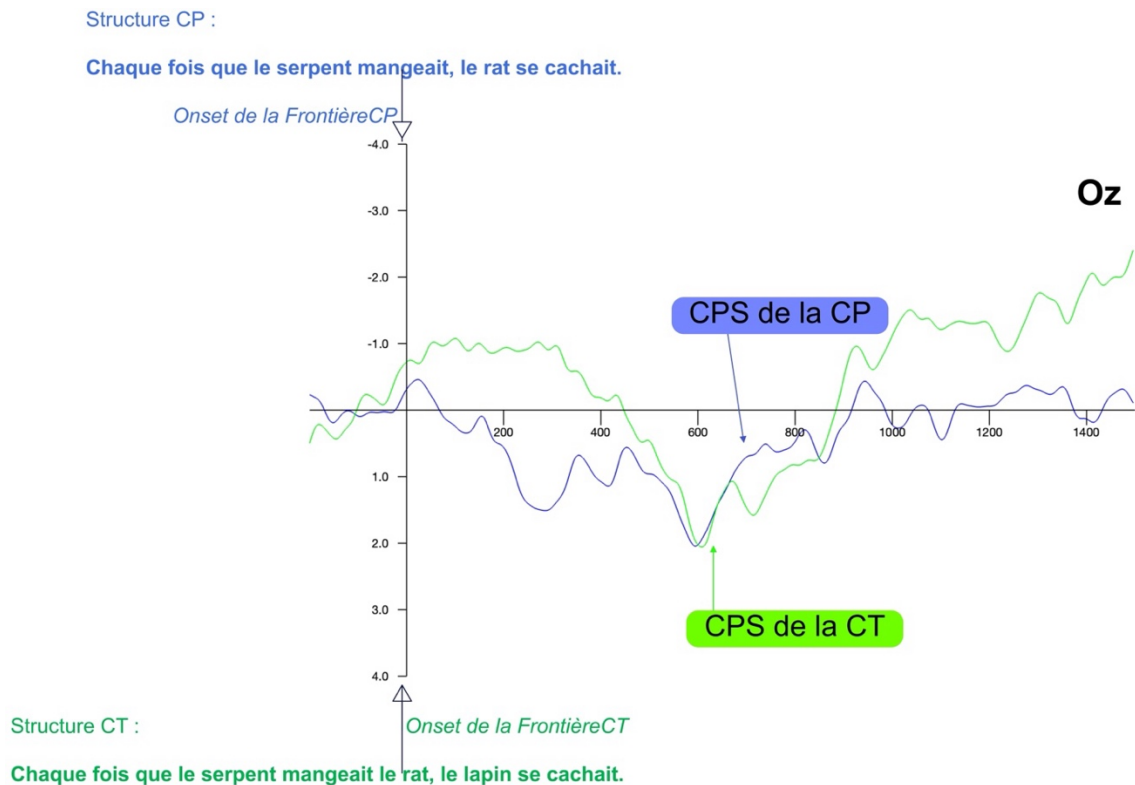


Figure 4.12 : Les composantes CPS pour *FrontièreCP* (haut, bleu) et *FrontièreCT* (bas, vert) sur l'électrode représentative Oz chez les participants **francophones natifs**.

Dans cette section, nous tenterons de répondre à la question de recherche 3 (Comment la différence comportementale dans l'expérience de perception se manifesterait-elle à travers les réponses CPS ?). Nous avons examiné la différence dans les traitements neurocognitifs entre CP et CT chez les francophones natifs. Comme le montre l'électrode représentative Oz (Figure 4.12), il semble qu'il n'y aurait pas de différence entre les deux composantes CPS pour les conditions CP et CT, ce qui suggère qu'il n'y aurait probablement pas de différence dans leur traitement neurocognitif. Pour vérifier cela, nous avons calculé les différences d'amplitudes moyennes des 42 électrodes des trois ROIs entre la présence et l'absence d'une frontière prosodique : d'une part la différence « *FrontièreCP–NonFrontièreCP* » dans la FT où la CPS de la condition CP est observée (soit [455-720] ms) et d'autre part, la différence « *FrontièreCT–NonFrontièreCT* » dans la FT où la CPS de la condition CT est observée (soit [562-924] ms). Ensuite, pour comparer « *FrontièreCP – NonFrontièreCP* » vs

« *FrontièreCT – NonFrontièreCT* », un t-test non-paramétrique¹²⁰ (Test des rangs signés de Wilcoxon) a été effectué. Le résultat du test indique une différence significative ($z=2,54$, $p=0,01$), avec la différence de « *FrontièreCP – NonFrontièreCP* » (moyenne = 1,16; médiane = 1,24) plus grande que celle de « *FrontièreCT – NonFrontièreCT* » (moyenne=0,53 ; médiane=1). Comme le montre la Figure 4.13, la différence du traitement neurocognitif des conditions prosodiques CP vs CT se manifeste plutôt au niveau frontal.

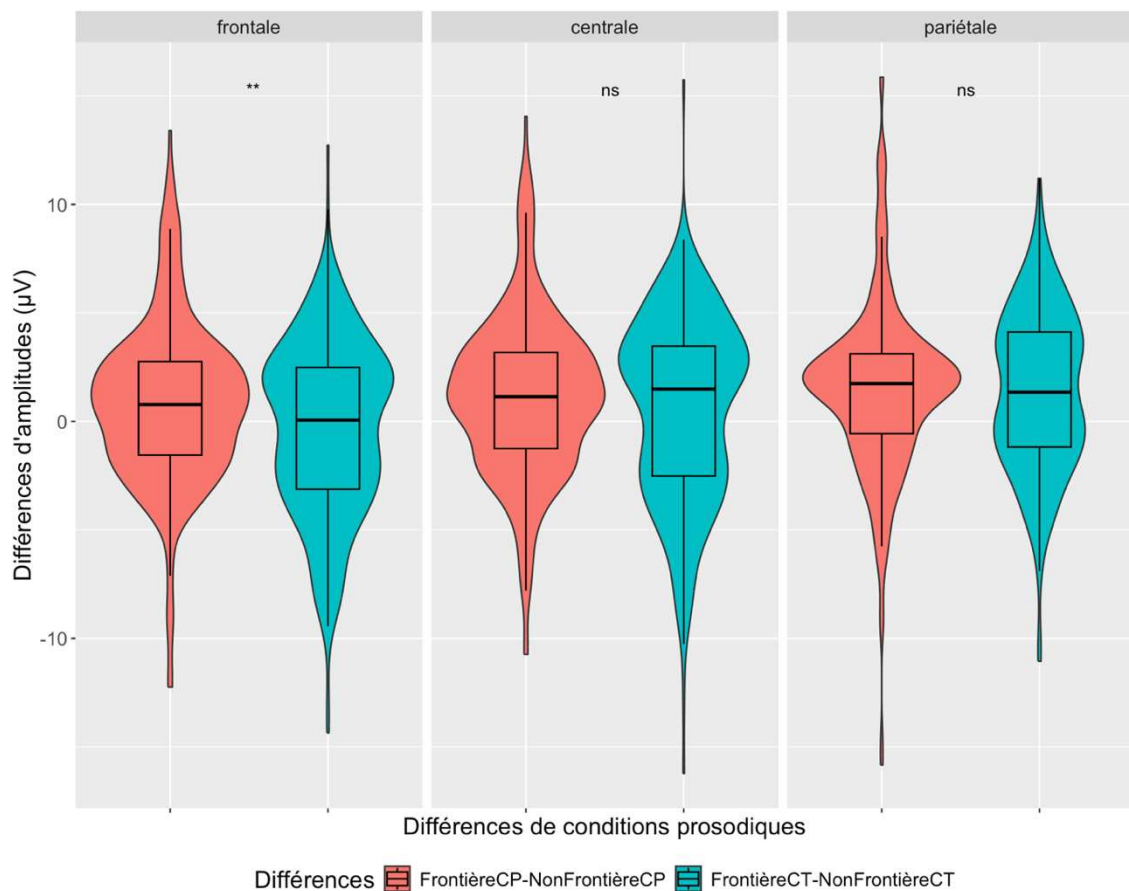


Figure 4.13 : Différences d'amplitudes comparées entre « *FrontièreCP-NonFrontièreCP* » (FT [455-720] ms) vs « *FrontièreCT-NonFrontièreCT* » (FT [562-924] ms) chez les participants **francophones natifs**.

Pour compléter ces résultats, nous avons effectué l'analyse univariée de masse, d'abord dans la FT [450-930] ms, qui couvre les deux FTs où les CPS pour les conditions CP et CT sont observées. Le test de clusters basé sur les permutations n'a pas permis d'identifier de cluster significatif (avec toutes les valeurs p égales ou

¹²⁰Étant donné le résultat du test Shapiro-Wilk significatif ($W=0,97$, $p<0,001$), qui indique une déviation de la normalité dans les données, nous avons procédé à l'alternative non-paramétrique au test de Student.

supérieures à 0,07). Nous avons ensuite élargi la FT à [0-1500] ms, et là aussi aucun cluster significatif n'a été identifié (avec toutes les valeurs p égales ou supérieures à 0,16).

Pour conclure, les comparaisons entre les conditions CP vs CT par l'approche traditionnelle montrent que les francophones natifs traitent les deux types de frontières prosodiques de manière similaire. Les composantes CPS pour les deux conditions se concentrent au niveau centro-pariétal et les différences d'amplitudes entre « *FrontièreCP – NonFrontièreCP* » et « *FrontièreCT – NonFrontièreCT* » sont comparables. Seule une différence a été révélée au niveau frontal, avec « *FrontièreCP – NonFrontièreCP* » légèrement plus grande que « *FrontièreCT – NonFrontièreCT* ». Cela s'explique par l'effet positif observé au niveau frontal dans la condition *FrontièreCP* chez les francophones natifs.

4.4.2 CPS chez les apprenants sinophones avancés

4.4.2.1 Conditions *FrontièreCP* vs *NonFrontièreCP*

Pour chacune des trois FTs ([0-610] ms, [610-797] ms et [797-1500] ms) fixées par l'approche de latence de la zone fractionnée, les modèles linéaires mixtes ont révélé une différence significative (Tableau 4.7). Toutefois, une amplitude significativement plus positive dans la condition *FrontièreCP* est observée seulement dans la FT [610-797] ms. En revanche, dans les deux FTs [0-610] ms et [797-1500] ms, l'amplitude de l'onde est significativement plus négative dans la condition *FrontièreCP* comparée à la condition *NonFrontièreCP* (Figure 4.14). Cela implique que la CPS pour la condition *FrontièreCP* chez les apprenants sinophones avancés est observée dans la FT [610-797] ms.

Tableau 4.7 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants avancés**.

FT (ms)	Effet de conditions prosodiques	Interaction (Conditions prosodiques * ROI)
[0-610]	b=0,55 ; SE=0,071 ; t= -7,8 ; p<0,001	b= -0,05 ; SE=0,1 ; t= -0,48 ; p=0,61
[610-797]	b=0,41 ; SE=0,09 ; t=4,49 ; p<0,001	b=0,06 ; SE=0,13 ; t=0,5 ; p=0,72
[797-1500]	b=0,98 ; SE=0,1 ; t= -9,9 ; p<0,001	b=0,04 ; SE=0,14 ; t=0,3 ; p=0,81

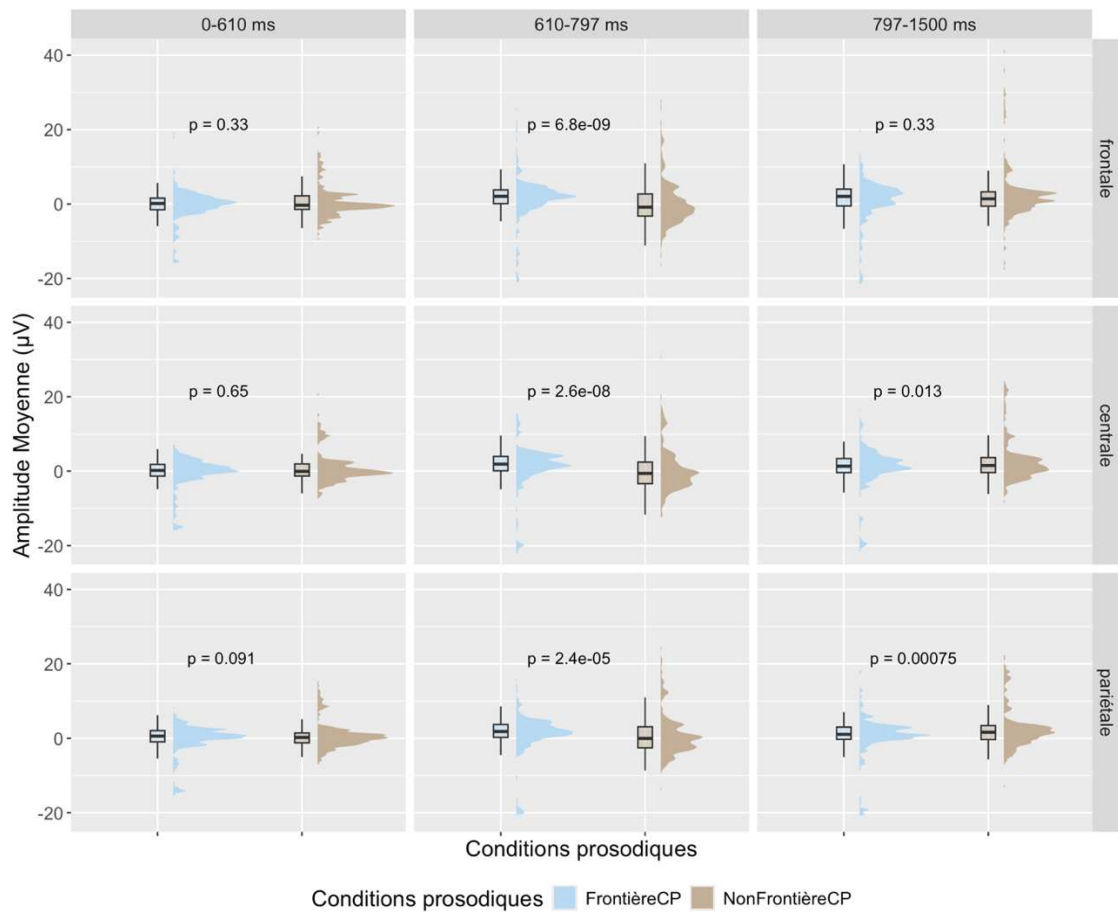


Figure 4.14 : Différences d'amplitudes moyennes entre les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* dans les trois FTs et les trois ROIs chez les **apprenants avancés**.

Pour confirmer ce constat, nous avons examiné ensuite les signaux ERP sur les électrodes représentatives des trois ROIs (Figure 4.15) et les cartes topographiques des trois FTs (Figure 4.16) pour les conditions *FrontièreCP* et *NonFrontièreCP* chez les apprenants avancés.

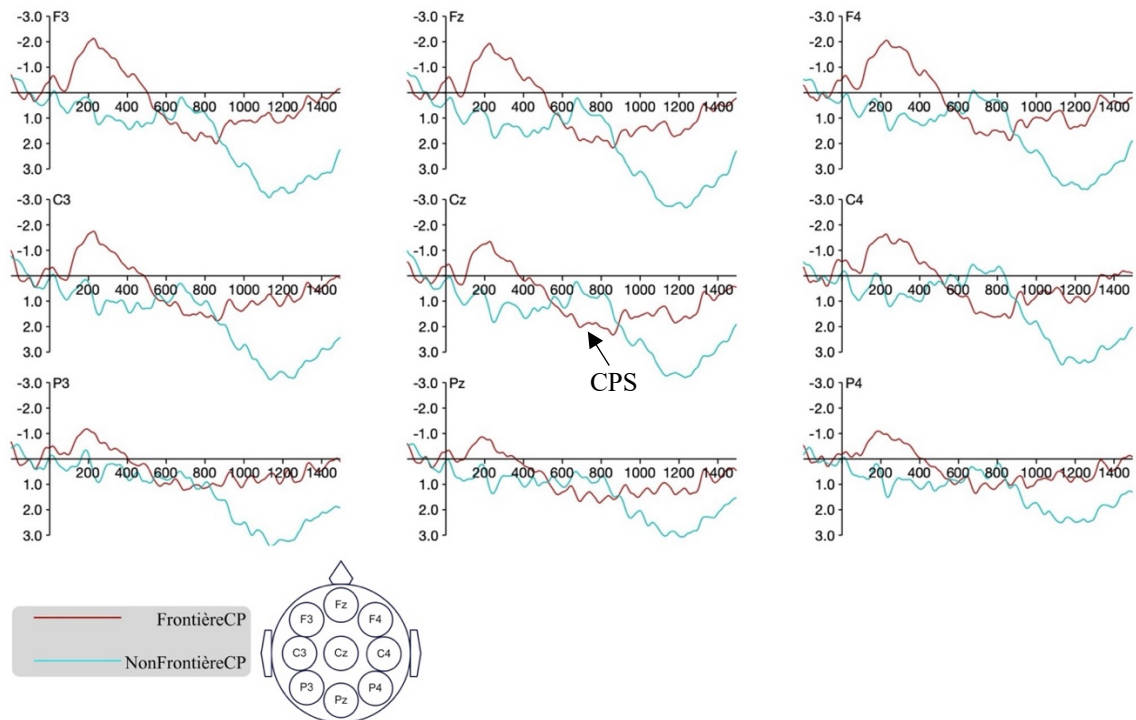


Figure 4.15 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants avancés**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

Comme le montre la Figure 4.15, les effets CPS de la condition *FrontièreCP* (tracé rouge) chez les apprenants avancés se présentent notamment au niveau des électrodes bilatérales de la région centro-frontale, tandis qu'au niveau pariétal, les effets CPS sont très faibles voire invisibles.

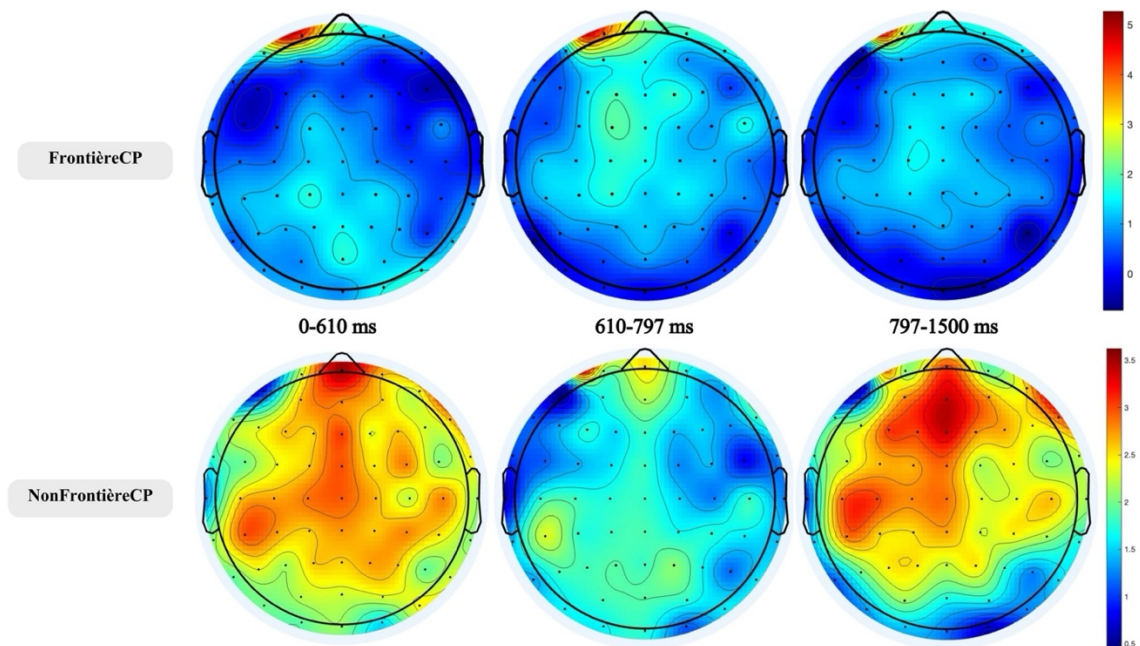


Figure 4.16 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants avancés**.

Les cartes topographiques (Figure 4.16) indiquent la même tendance : pour la condition *FrontièreCP* dans la FT [610-797] ms, la polarité positive pour la condition *FrontièreCP* se concentre au niveau centro-frontal.

Pour affiner ces résultats, nous avons effectué l'analyse univariée de masse. Cependant, aucun cluster significatif n'a été identifié par les tests de clusters basés sur les permutations, ni dans la FT [610-797] ms (avec toutes les valeurs p égales ou supérieures à 0,46), ni dans la FT entière [0-1500] ms (avec toutes les valeurs p égales ou supérieures à 0,29).

En résumé, les résultats obtenus par l'approche traditionnelle indiquent que les apprenants avancés perçoivent et traitent les frontières CP. Cela se manifeste par la composante CPS observée au niveau centro-frontal, tant sur la ligne centrale que sur les régions bilatérales, dans la FT [610-797] ms pour la condition *FrontièreCP*.

4.4.2.2 Conditions *FrontièreCT* vs *NonFrontièreCT*

Pour les trois FTs ([0-634] ms, [634-1062] ms et [1062-1500] ms), les modèles linéaires mixtes ont révélé une différence significative pour les deux FTs [634-1062] ms et [1062-1500] ms : les amplitudes dans ces deux FTs sont significativement plus positives dans la condition *FrontièreCT* comparée à la condition *NonFrontièreCT* (Tableau 4.8, Figure 4.17). Ces résultats statistiques suggèrent que la CPS pour *FrontièreCT* chez les apprenants avancés est déclenchée dans la FT [634-1062] ms et persiste jusqu'à la FT [1062-1500] ms.

Tableau 4.8 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants avancés**.

FT (ms)	Effet de conditions prosodiques	Interaction (Conditions prosodiques * ROI)
[0-634]	b=0,02 ; SE=0,04 ; t=0,5 ; p=0,62	b= -0,05 ; SE=0,06 ; t= -0,72 ; p=0,42
[634-1062]	b= -1,72 ; SE=0,08 ; t= -22,98 ; p<0,001	b=0,07 ; SE=0,11 ; t=0,66 ; p=0,49
[1062-1500]	b= -1,15 ; SE=0,08 ; t= -14,11 ; p<0,001	b= -0,01 ; SE=0,12 ; t= -0,11 ; p=0,86

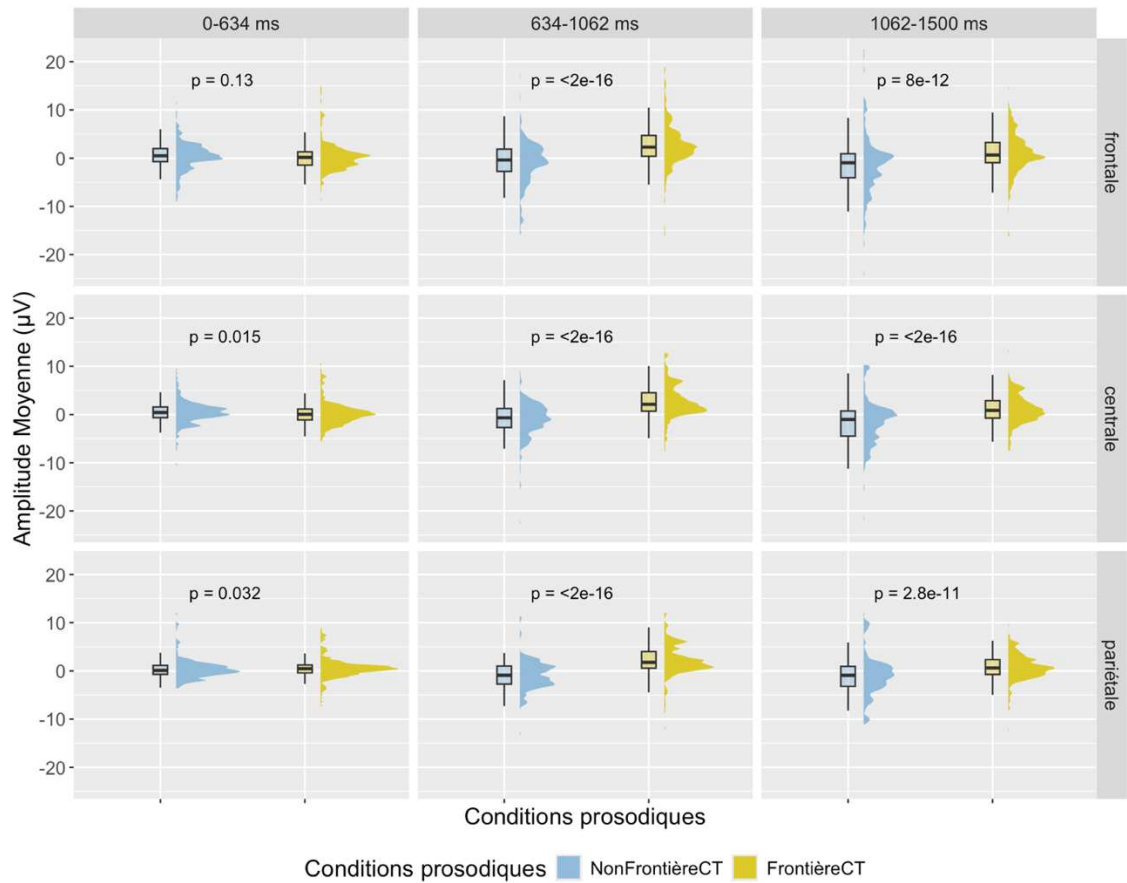


Figure 4.17 : Différences d'amplitudes moyennes entre les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* dans les trois FTs et les trois ROIs chez les **apprenants avancés**.

Nous avons examiné ensuite les signaux CPS sur les 9 électrodes représentatives des trois ROIs pour les conditions *FrontièreCT* et *NonFrontièreCT* chez les apprenants avancés (Figure 4.18), ainsi que les cartes topographiques des trois FTs (Figure 4.19).

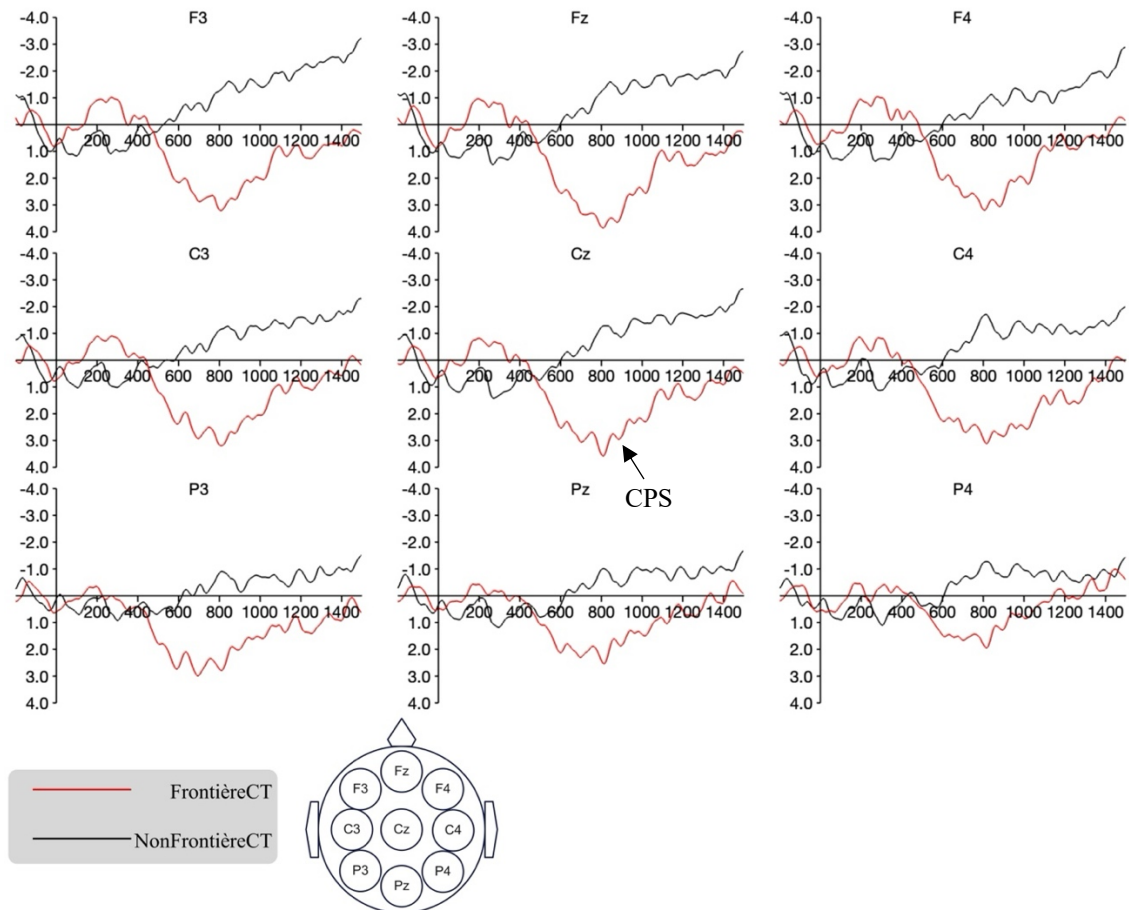


Figure 4.18 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants avancés**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

Comme le montre la Figure 4.18, pour la condition *FrontièreCT* chez les apprenants avancés, l'effet CPS dure relativement long : il émerge à partir de 400 ms après l'onset de la dernière syllabe avant les frontières prosodiques CT et il persiste dans les deux FTs consécutives [634-1062] ms et [1062-1500] ms.

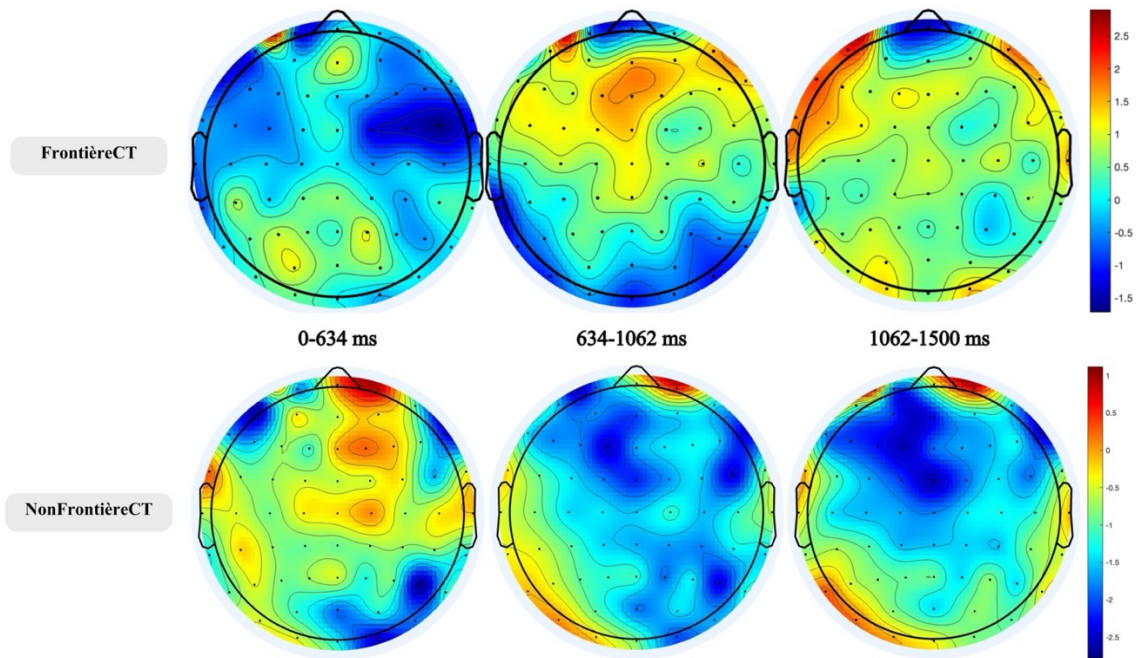


Figure 4.19 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants avancés**.

Sur les chronogrammes topographiques des trois FTs (Figure 4.19), la polarité positive de la CPS se présente d'abord à la région centro-frontale (FT [634-1062] ms, Figure 4.19 milieu haut) et puis s'étend sur l'ensemble du cerveau (FT [1062-1500] ms, Figure 4.19 droite haut).

Afin d'affiner ces résultats sur la CPS de la condition *FrontièreCT* chez les apprenants avancés, nous avons effectué les tests de clusters basés sur les permutations dans la FT [634-1500] ms et un cluster significatif a été identifié ($p=0,006$) (Figure 4.20).

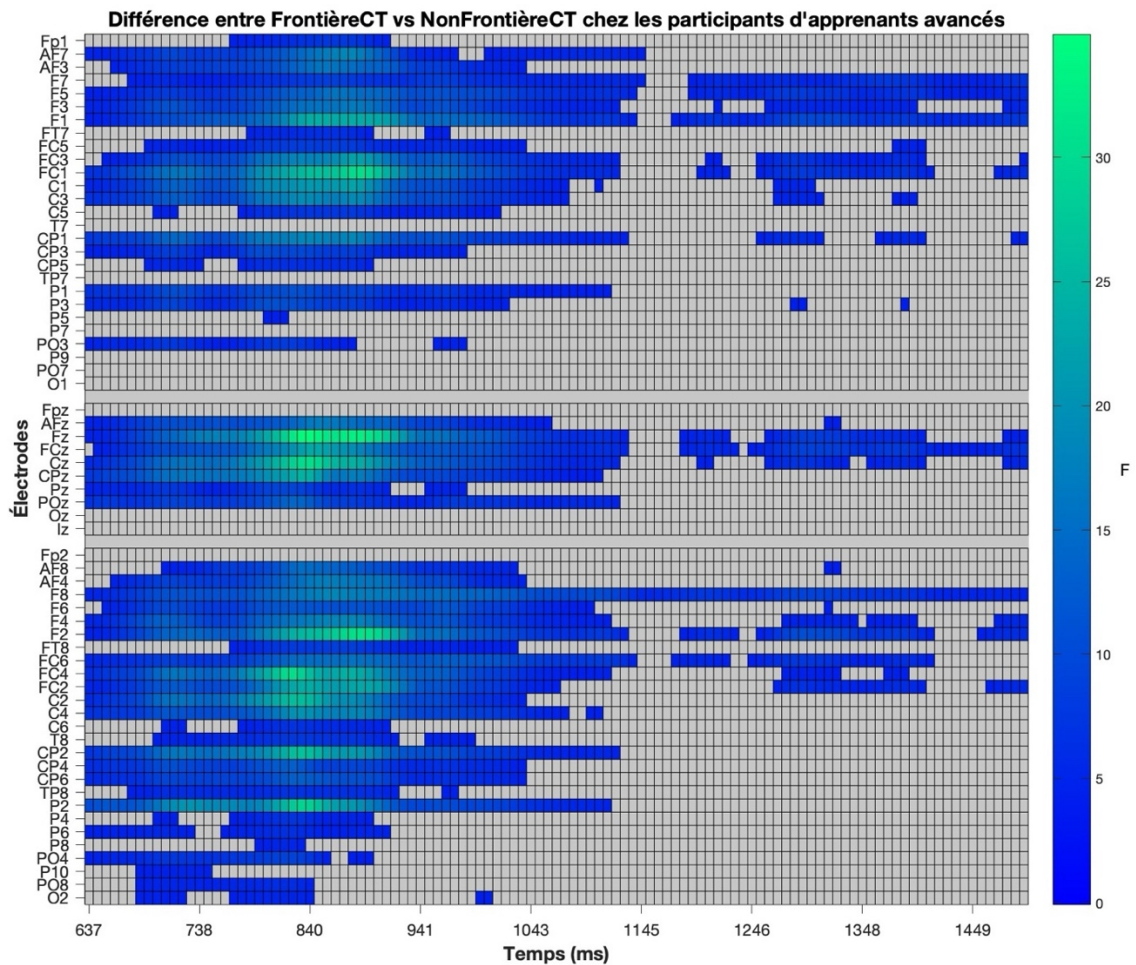


Figure 4.20 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* dans la FT [634-1500] ms chez les **apprenants avancés**.

Le cluster identifié se présente sur l'ensemble du cerveau (Figure 4.20), et d'après le rapport de résultats généré, il est avec un pic sur l'électrode Fz vers 832 ms après les frontières prosodiques CT.

Pour résumer, les résultats obtenus par les approches traditionnelle et univariée de masse montrent que les apprenants avancés perçoivent et traitent sur le plan neurocognitif les frontières prosodiques CT. La composante CPS pour la condition *FrontièreCT* chez ces apprenants est un effet fort. Cela se manifeste par sa durée longue, qui persiste dans la FT [634-1500] ms et par sa distribution topographique large, qui s'étend sur l'ensemble du cerveau.

4.4.2.3 Comparaison des effets CPS dans les conditions *FrontièreCP* vs *FrontièreCT*

Dans les deux sections précédentes (4.4.2.1 et 4.4.2.2), les résultats de nos analyses montrent que les composantes CPS sont déclenchées lorsque les apprenants sinophones avancés perçoivent et traitent respectivement les frontières prosodiques CP et CT (Figure 4.21).

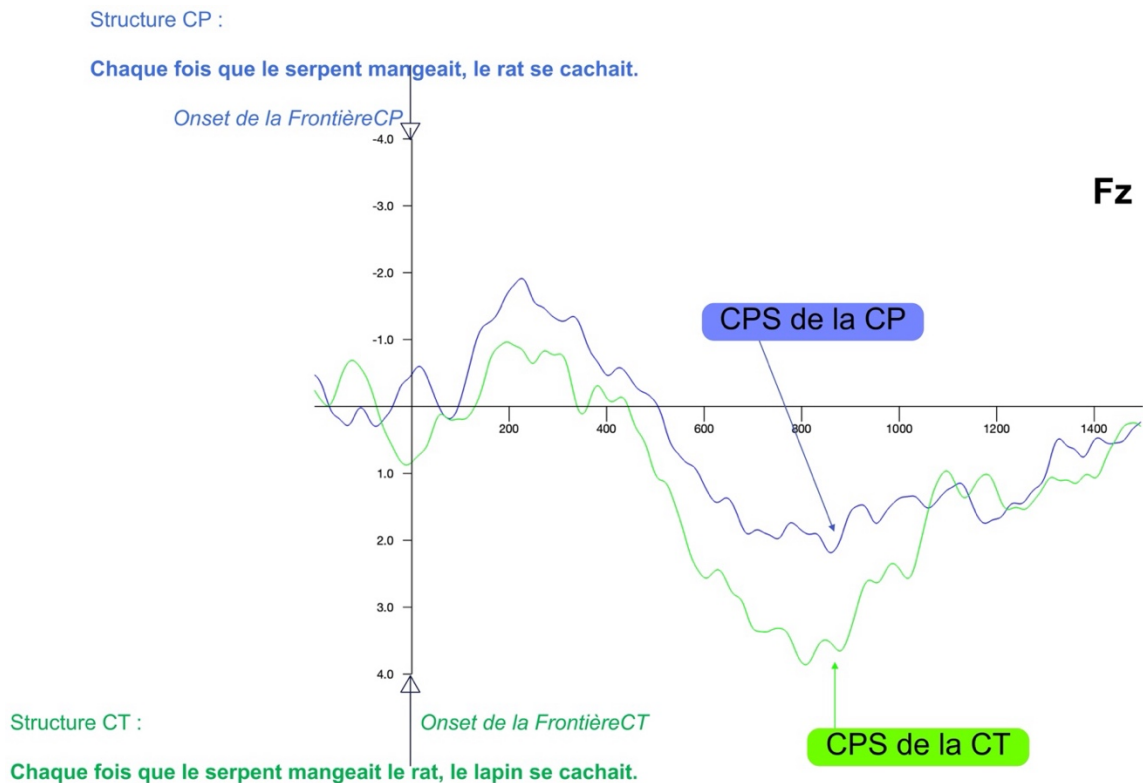


Figure 4.21 : Les composantes CPS pour *FrontièreCP* (haut, bleu) et *FrontièreCT* (bas, vert) sur l'électrode représentative Fz chez les **apprenants avancés**.

En suivant les mêmes démarches de comparaisons que celles effectuées dans la section 4.4.1.3, nous avons examiné la différence dans les traitements neurocognitifs entre CP et CT chez les apprenants avancés afin de répondre à la question de recherche 3 (Comment la différence comportementale dans l'expérience de perception se manifesterait-elle à travers les réponses CPS ?). Comme l'illustre l'électrode représentative Fz (Figure 4.21), l'amplitude de l'effet CPS est plus grande dans la condition CT que dans la CP, ce qui suggère, sur le plan neurocognitif, que les apprenants avancés sont plus sensibles aux frontières prosodiques CT. Pour vérifier ce constat, nous avons comparé la différence entre « *FrontièreCP–NonFrontièreCP* » dans la FT [610-797] ms et « *FrontièreCT–NonFrontièreCT* » dans la FT [634-1500]

ms avec un t-test non-paramétrique¹²¹ (Test des rangs signés de Wilcoxon). Le résultat du test a révélé une différence significative ($z=3,42$, $p<0,001$), avec la différence de « *FrontièreCT-NonFrontièreCT* » (moyenne = 2,81 ; médiane = 2,40) plus grande que celle de « *FrontièreCP-NonFrontièreCP* » (moyenne = 0,81 ; médiane = 2,36). Comme le montre la Figure 4.22, la différence du traitement neurocognitif des conditions prosodiques CP vs CT se manifeste plutôt au niveau pariétal.

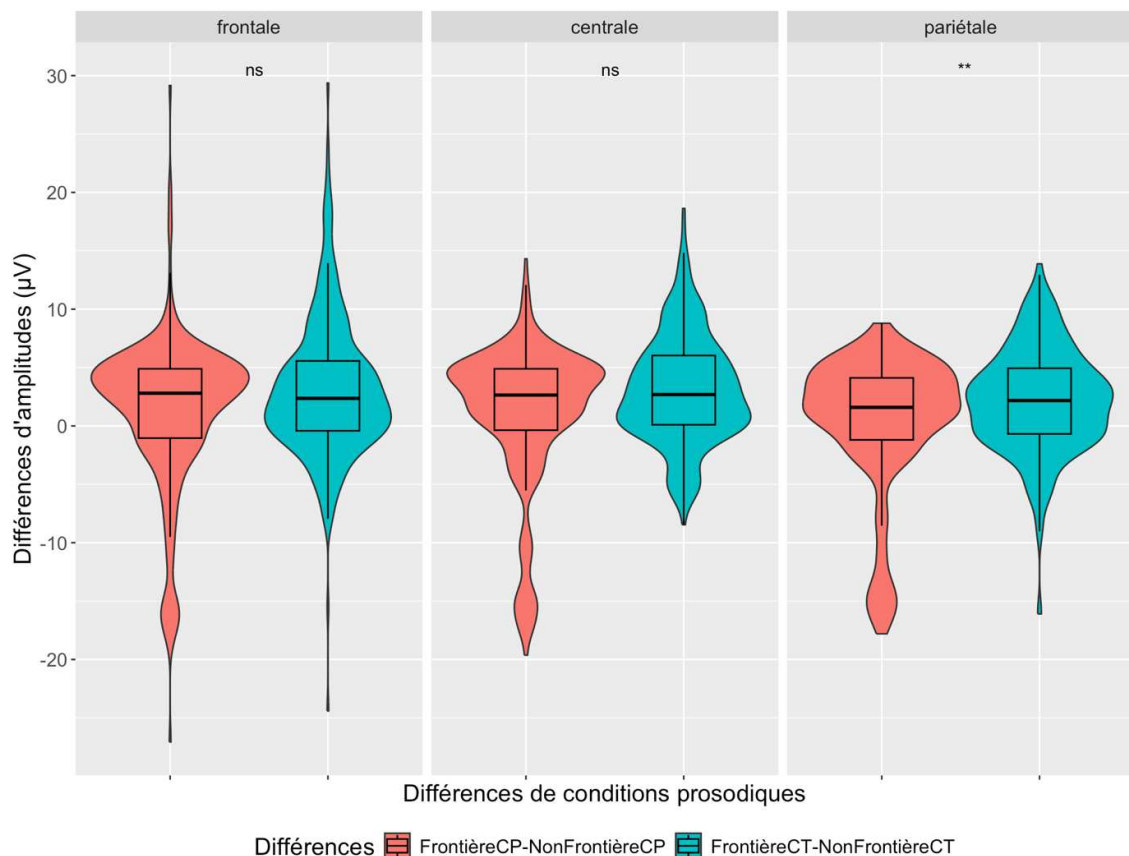


Figure 4.22 : Différences d'amplitudes comparées entre « *FrontièreCP-NonFrontièreCP* » (FT [610-797] ms) vs « *FrontièreCT-NonFrontièreCT* » (FT [634-1500] ms) chez les **apprenants avancés**.

Pour compléter ces résultats, nous avons effectué une analyse univariée de masse, d'abord dans la FT [610-1500] ms, qui couvre les deux FTs où les CPS pour les conditions CP et CT sont observées. Le test de clusters basé sur les permutations n'a pas identifié de cluster significatif (avec toutes les valeurs p égales ou supérieures à

¹²¹Étant donné le résultat du test Shapiro-Wilk significatif ($W=0,91$, $p<0,001$), qui indique une déviation de la normalité dans les données, nous avons procédé à l'alternative non-paramétrique au test de Student.

0,45). Nous avons ensuite élargi la FT à [0-1500] ms et, là non plus, aucun cluster significatif n'a été identifié (avec toutes les valeurs p égales ou supérieures à 0,64). Pour conclure, les différences d'amplitudes entre « *FrontièreCP–NonFrontièreCP* » et « *FrontièreCT–NonFrontièreCT* » révélées par l'approche traditionnelle montrent que les apprenants avancés traitent les deux types de frontières prosodiques de manière similaire. Les composantes CPS pour les deux conditions se concentrent au niveau centro-frontal. Seule une différence est observée au niveau pariétal, avec « *FrontièreCT–NonFrontièreCT* » plus grande que « *FrontièreCP–NonFrontièreCP* ». Cela peut s'expliquer par le fait que l'effet positif est présent au niveau pariétal dans la condition CT chez les apprenants avancés, tandis que dans la condition CP, l'effet CPS ne se présente pas au même niveau.

4.4.3 CPS chez les apprenants sinophones intermédiaires

4.4.3.1 Conditions *FrontièreCP* vs *NonFrontièreCP*

Les modèles linéaires mixtes ont révélé une différence significative pour toutes les trois FTs ([0-572], [572-843] et [843-1500]) (Tableau 4.9). Toutefois, une amplitude significativement plus positive pour la condition *FrontièreCP* est observée uniquement dans la FT [572-843] ms. En revanche, pour les deux autres FTs [0-572] ms et [843-1500] ms, l'amplitude de l'onde est significativement plus négative dans la condition *FrontièreCP* comparée à la condition *NonFrontièreCP* (Figure 4.23). La positivité unique pour la condition *FrontièreCP* dans la FT [572-843] ms suggère que la CPS est déclenchée dans cette FT.

Tableau 4.9 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants intermédiaires**.

FT (ms)	Effet de conditions prosodiques	Interaction (Conditions prosodiques * ROI)
[0-572]	b= -0,27 ; SE=0,06 ; t= -4,65 ; p<0,001	b= -0,12 ; SE=0,08 ; t= -1,45 ; p=0,31
[572-843]	b=0,78 ; SE=0,09 ; t=8,69 ; p<0,001	b=0,08 ; SE=0,13 ; t=0,61 ; p=0,62
[843-1500]	b= -0,3 ; SE=0,09 ; t= -3,19 ; p=0,001	b=0,47 ; SE=0,13 ; t=3,53 ; p<0,001

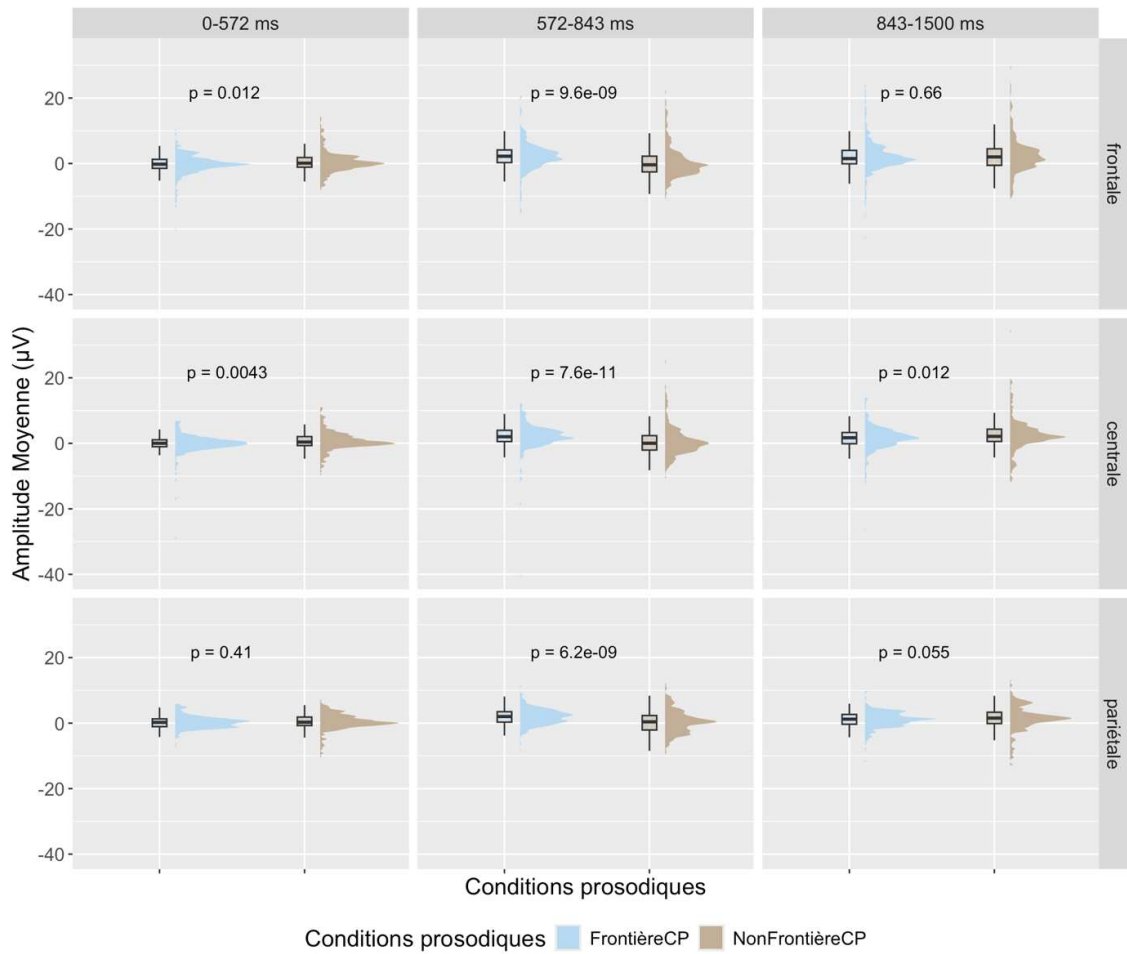


Figure 4.23 : Différences d’amplitudes moyennes entre les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* dans les trois FTs et les trois ROIs chez les **apprenants intermédiaires**.

Afin de confirmer ce constat, nous avons examiné les signaux ERP (Figure 4.24) et comparé les cartes topographiques des trois FTs (Figure 4.25) ci-dessous.

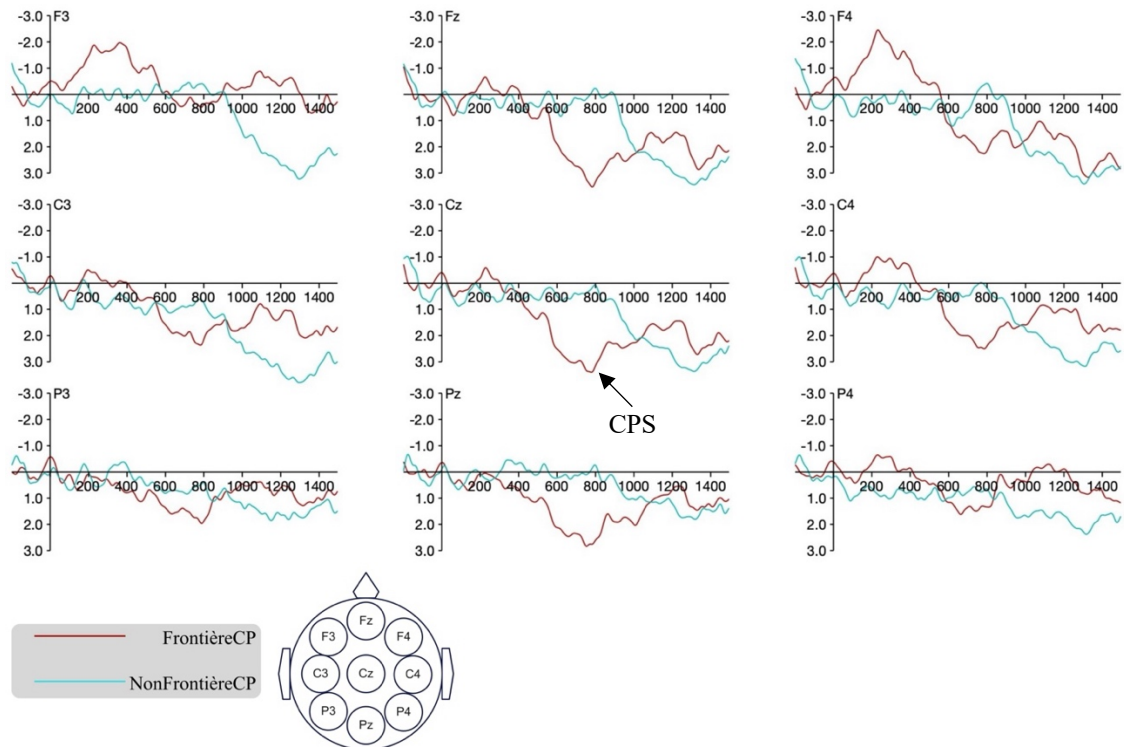


Figure 4.24 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants intermédiaires**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

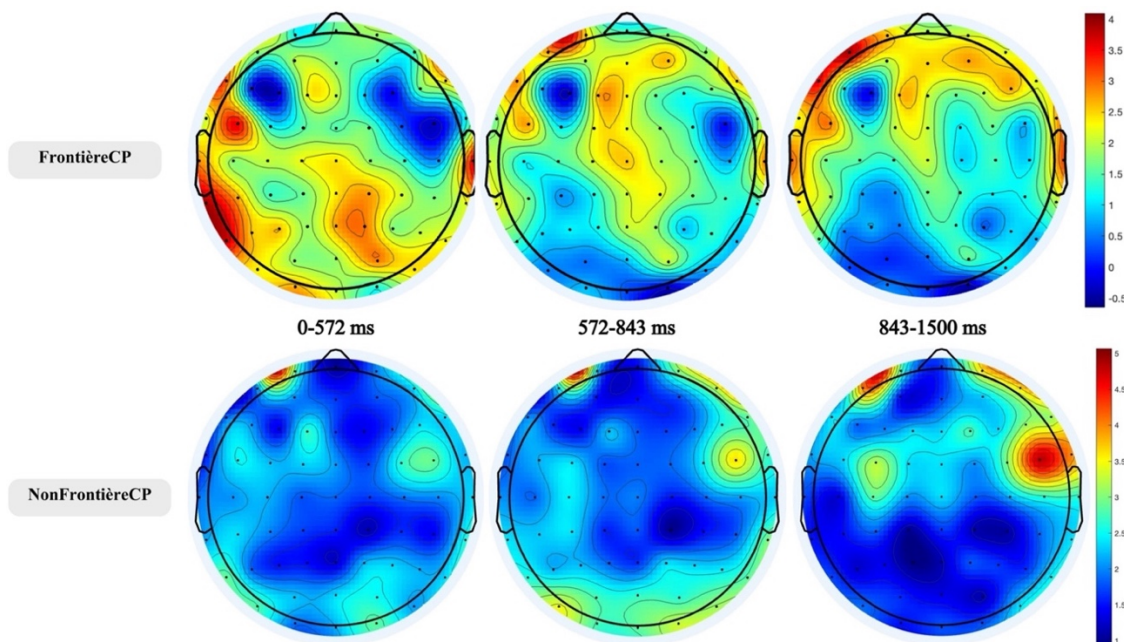


Figure 4.25 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants intermédiaires**.

Comme le montrent les signaux CPS (Figure 4.24), l'effet CPS (tracé rouge) se présente sur toutes les 9 électrodes représentatives vers 600 ms. En comparant les cartes topographiques des trois FTs (Figure 4.25), dans la FT [572-843] ms de la

condition *FrontièreCP*, une polarité positive se présente sur l'ensemble du cerveau, avec une concentration sur la ligne centrale du niveau centro-frontal.

Afin d'affiner ces résultats, nous avons ensuite mené l'analyse univariée de masse. Le test de clusters basés sur les permutations effectué dans la FT [572-843] a permis d'identifier un cluster significatif ($p=0,04$), avec un pic sur l'électrode CP2 vers 793 ms (Figure 4.26) après l'onset de la dernière syllabe avant les frontières prosodiques CP.

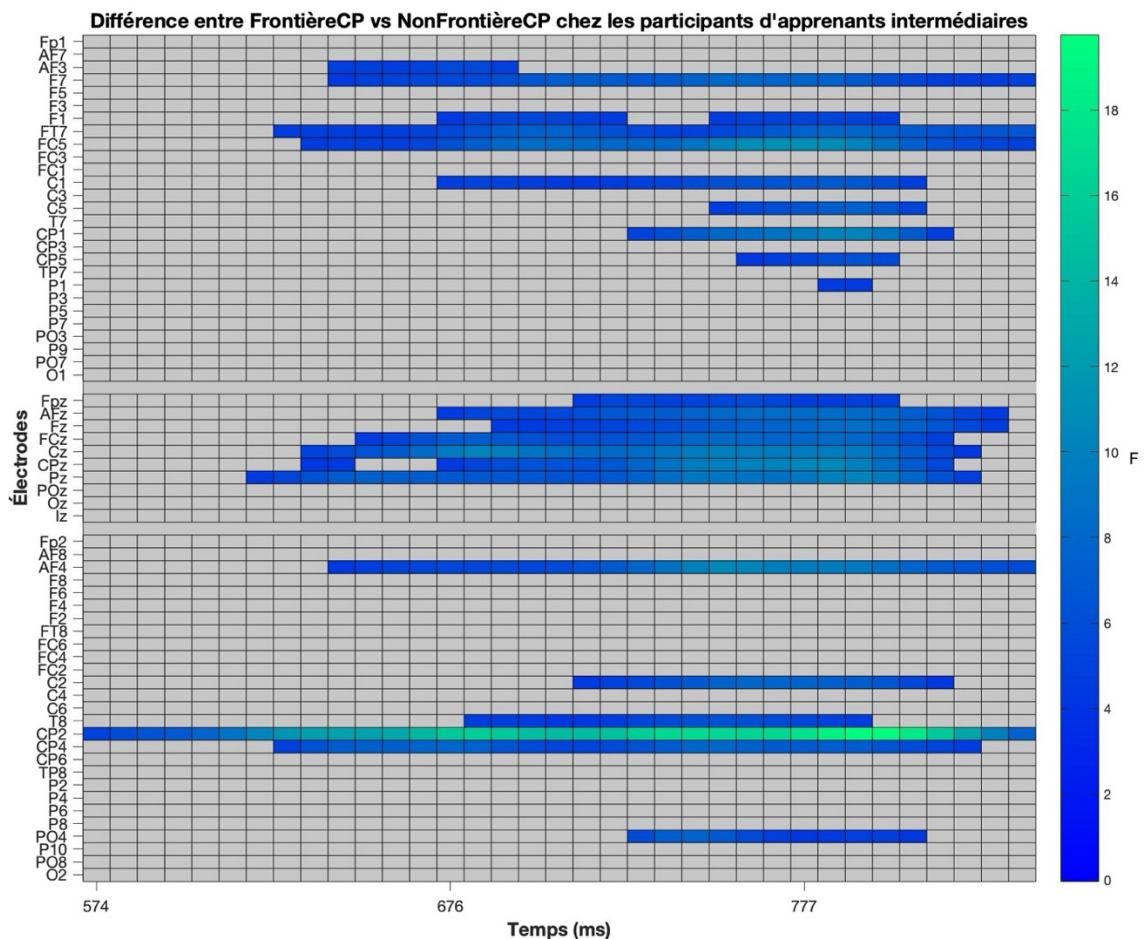


Figure 4.26 : Résultat du test de clusters basés sur les permutations pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* dans la FT [572-843] ms chez les **apprenants intermédiaires**.

En résumé, les analyses traditionnelle et univariée de masse montrent que les apprenants intermédiaires, à l'instar des avancés et des natifs, perçoivent et traitent bien les frontières prosodiques CP. Cela se manifeste par la composante CPS observée à partir de 570 ms après l'onset de la dernière syllabe avant les frontières CP. Cet effet se présente sur l'ensemble du cerveau, avec un effet plus fort sur la ligne centrale du niveau centro-frontal.

4.4.3.2 Conditions *FrontièreCT* vs *NonFrontièreCT*

Les modèles linéaires mixtes ont révélé une différence significative entre *FrontièreCT* et *NonFrontièreCT* pour les FTs [596-1096] et [1096-1500] (Tableau 4.10) : les amplitudes de l'onde sont plus positives dans la condition *FrontièreCT* comparée à la condition *NonFrontièreCT* (Figure 4.27). Cette positivité implique que la CPS de la condition *FrontièreCT* est déclenchée dans la FT [596-1500] ms chez les apprenants intermédiaires.

Tableau 4.10 : Résultats des modèles linéaires mixtes testés dans les trois FTs pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants intermédiaires**.

FT (ms)	Effet de conditions prosodiques	Interaction (Conditions prosodiques * ROI)
[0-596]	b=0,01 ; SE=0,05 ; t=0,20 ; p=0,84	b=0,26 ; SE=0,08 ; t=3,37 ; p=0,002
[596-1096]	b= -0,81 ; SE=0,10 ; t= -8,45 ; p<0,001	b=0,61 ; SE=0,14 ; t=4,49 ; p<0,001
[1096-1500]	b= -0,75 ; SE=0,11 ; t= -6,65 ; p<0,001	b=0,71 ; SE=0,16 ; t=4,45 ; p<0,001

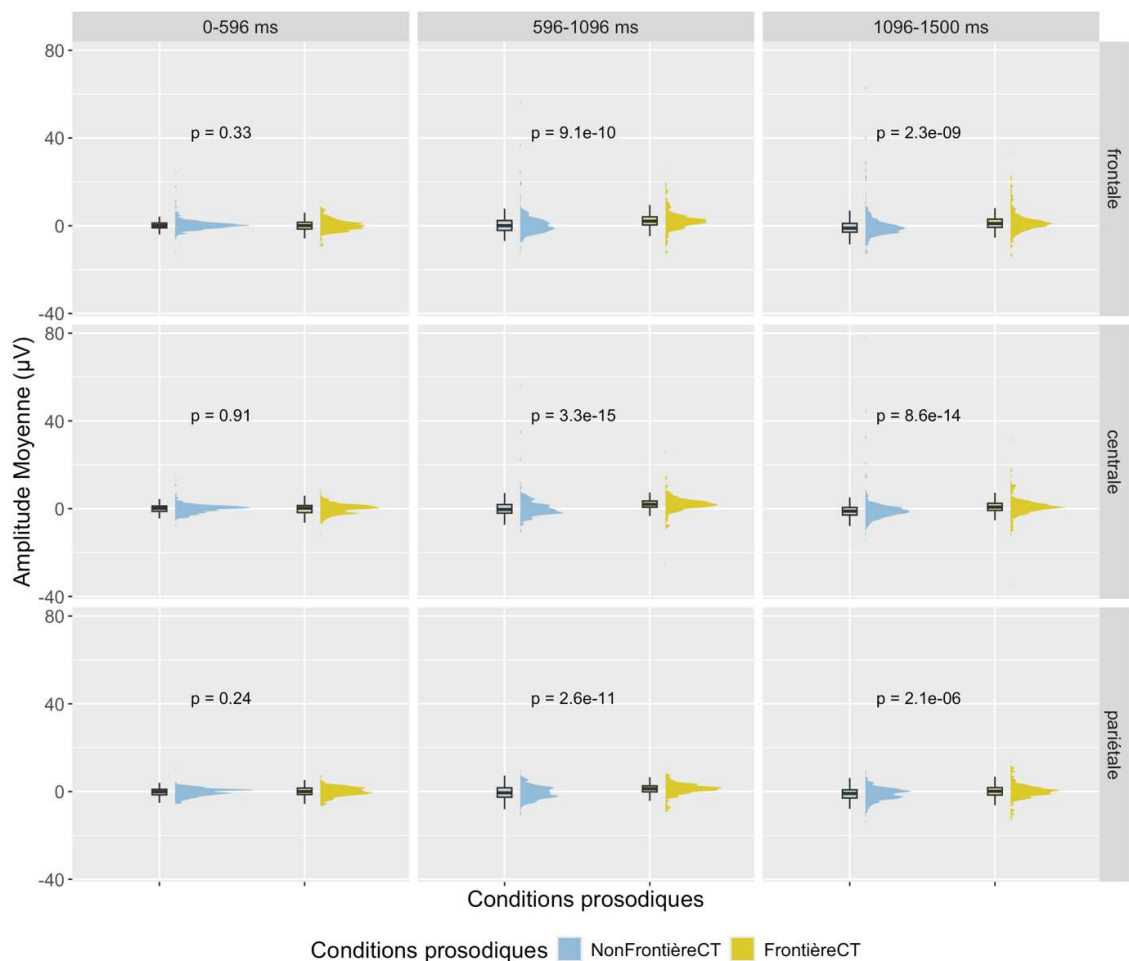


Figure 4.27 : Différences d'amplitudes moyennes entre les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* dans les trois FTs et les trois ROIs chez les **apprenants intermédiaires**.

Afin de vérifier ce constat, nous avons examiné les signaux ERP (Figure 4.28) et comparé les cartes topographiques des trois FTs (Figure 4.29).

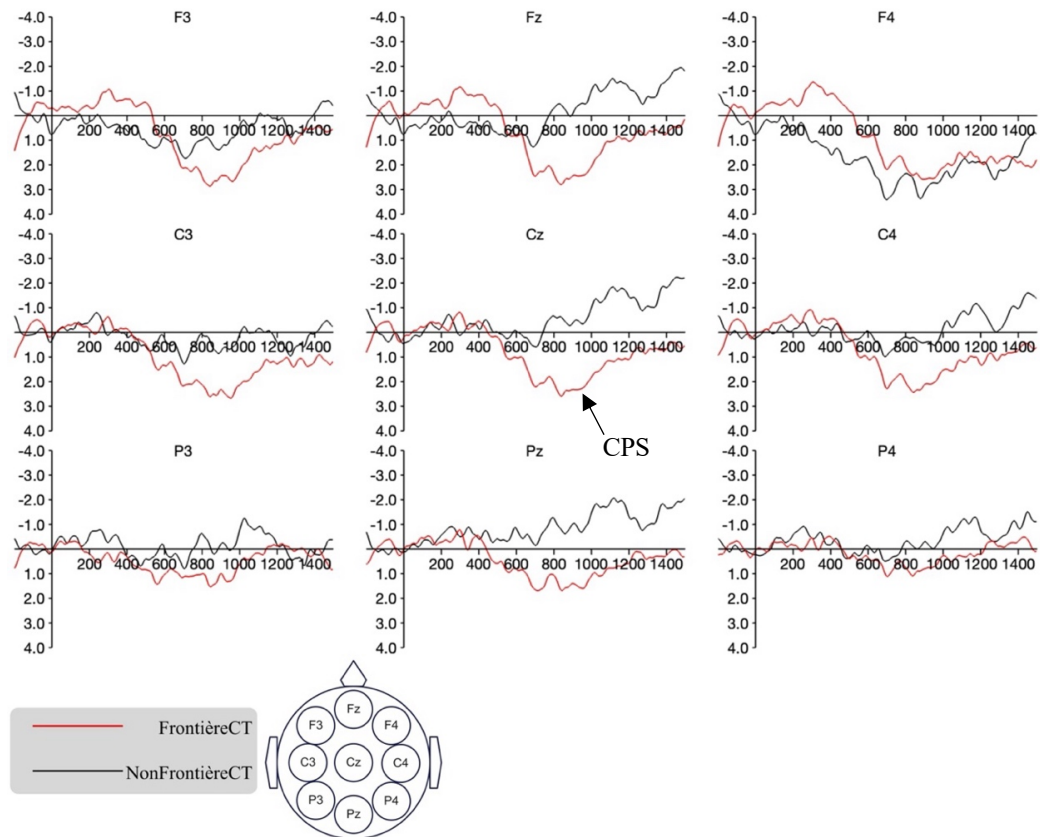


Figure 4.28 : Effets CPS sur 9 électrodes représentatives pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants intermédiaires**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

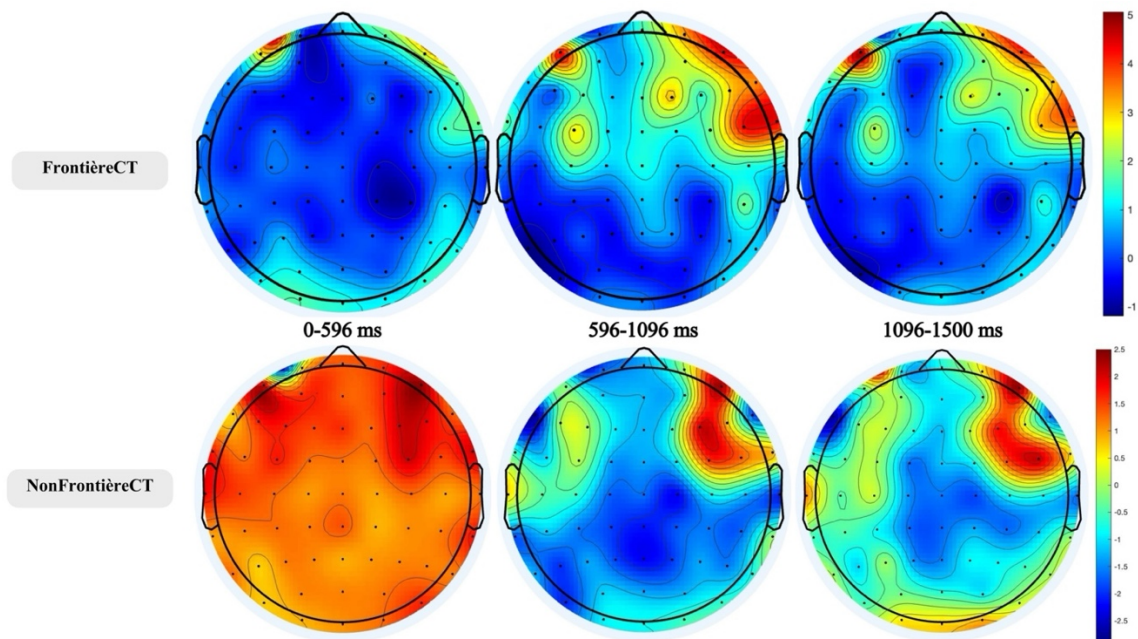


Figure 4.29 : Cartes topographiques des trois FTs pour la comparaison entre conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants intermédiaires**.

par la composante CPS déclenchée à partir de 600 ms après l'onset de la dernière syllabe avant les frontières prosodiques CT et elle dure jusqu'à peu près 1200 ms. L'effet CPS dans cette condition chez ce groupe de participants se concentre principalement au niveau centro-frontal.

4.4.3.3 Comparaison des effets CPS dans les conditions *FrontièreCP* vs *FrontièreCT*

Dans les sections 4.4.3.1 et 4.4.3.2, les résultats de nos analyses montrent que les composantes CPS sont déclenchées lorsque les apprenants sinophones intermédiaires perçoivent et traitent respectivement les frontières prosodiques CP et CT (Figure 4.31).

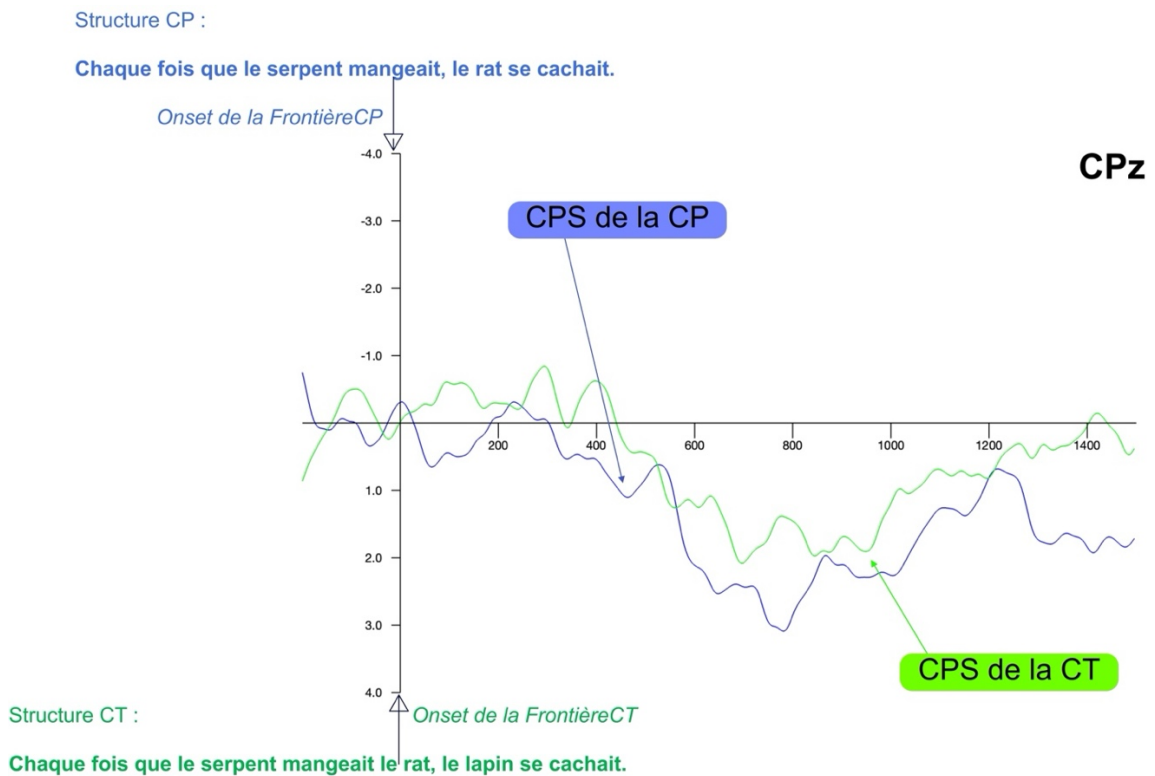


Figure 4.31 : Les composantes CPS pour *FrontièreCP* (haut, bleu) et *FrontièreCT* (bas, vert) sur l'électrode représentative CPz chez les **apprenants intermédiaires**.

Nous avons examiné la différence dans les traitements neurocognitifs entre CP et CT chez les apprenants intermédiaires. Pour cela, nous avons comparé dans un premier temps la différence entre « *FrontièreCP–NonFrontièreCP* » dans la FT [572-843] ms et « *FrontièreCT–NonFrontièreCT* » dans la FT [596-1500] ms avec un t-test non-

paramétrique¹²² (Test des rangs signés de Wilcoxon). Le résultat de ce test n'a pas révélé de différence significative ($z = -1,42$, $p = 0,16$), avec la différence de « *FrontièreCP–NonFrontièreCP* » (moyenne=1,55 ; médiane=1,94) comparable à celle de « *FrontièreCT–NonFrontièreCT* » (moyenne=1,56 ; médiane=1,87). Comme le montre la Figure 4.32, la différence du traitement neurocognitif des conditions prosodiques CP vs CT n'est pas significative dans chacune des trois ROIs.

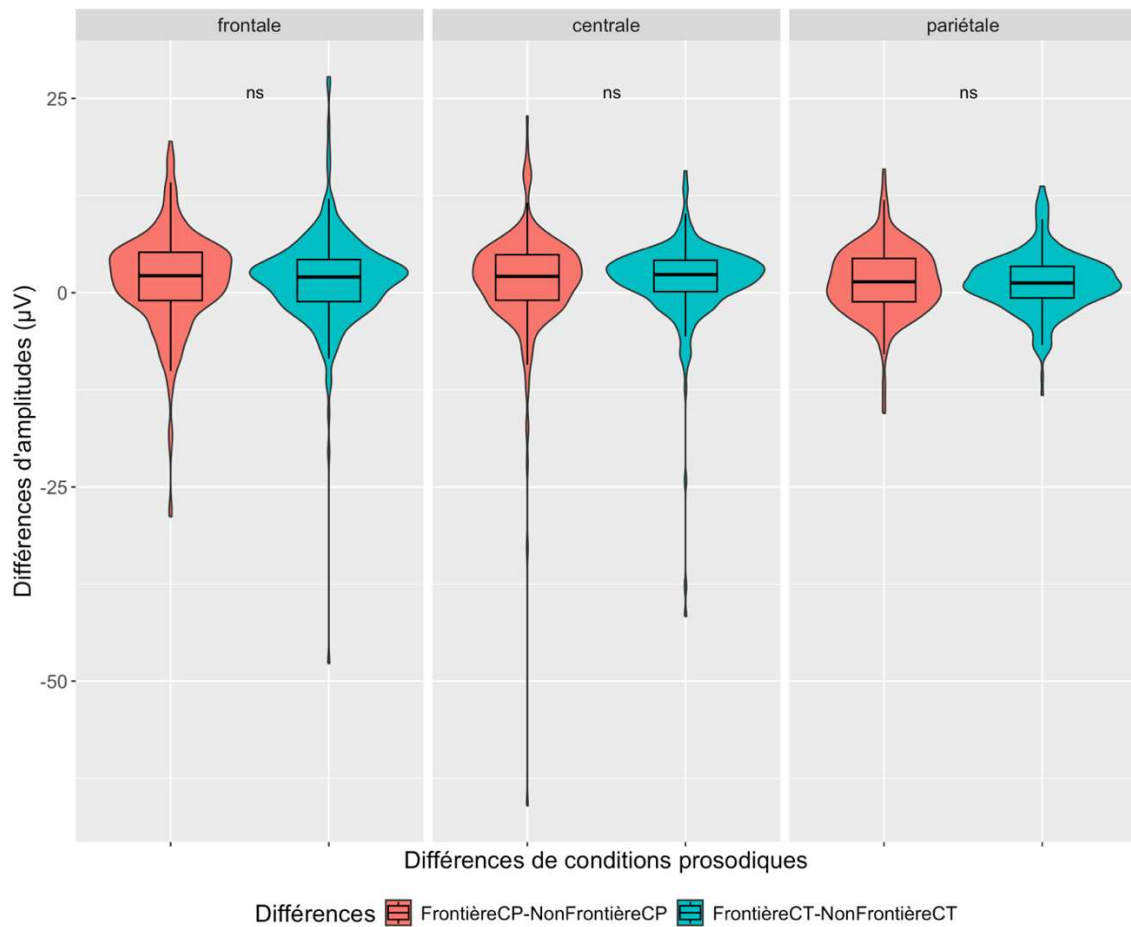


Figure 4.32 : Différences d'amplitudes comparées entre « *FrontièreCP-NonFrontièreCP* » (FT [572-843] ms) vs « *FrontièreCT-NonFrontièreCT* » (FT [596-1500] ms) chez les **apprenants intermédiaires**.

Pour compléter ces résultats, nous avons effectué l'analyse univariée de masse dans la FT [570-1500] ms, qui couvre les deux FTs où les CPS pour les conditions CP et CT sont observées. Le test de clusters basé sur les permutations a identifié un cluster

¹²²Étant donné le résultat du test Shapiro-Wilk significatif ($W=0,91$, $p<0,001$), qui indique que la distribution des données est déviée de la normalité, nous avons procédé à l'alternative non-paramétrique au test de Student.

significatif ($p=0,005$) dans la FT [895-1500] ms (Figure 4.33), avec un pic sur l'électrode CP4 à 1449 ms.

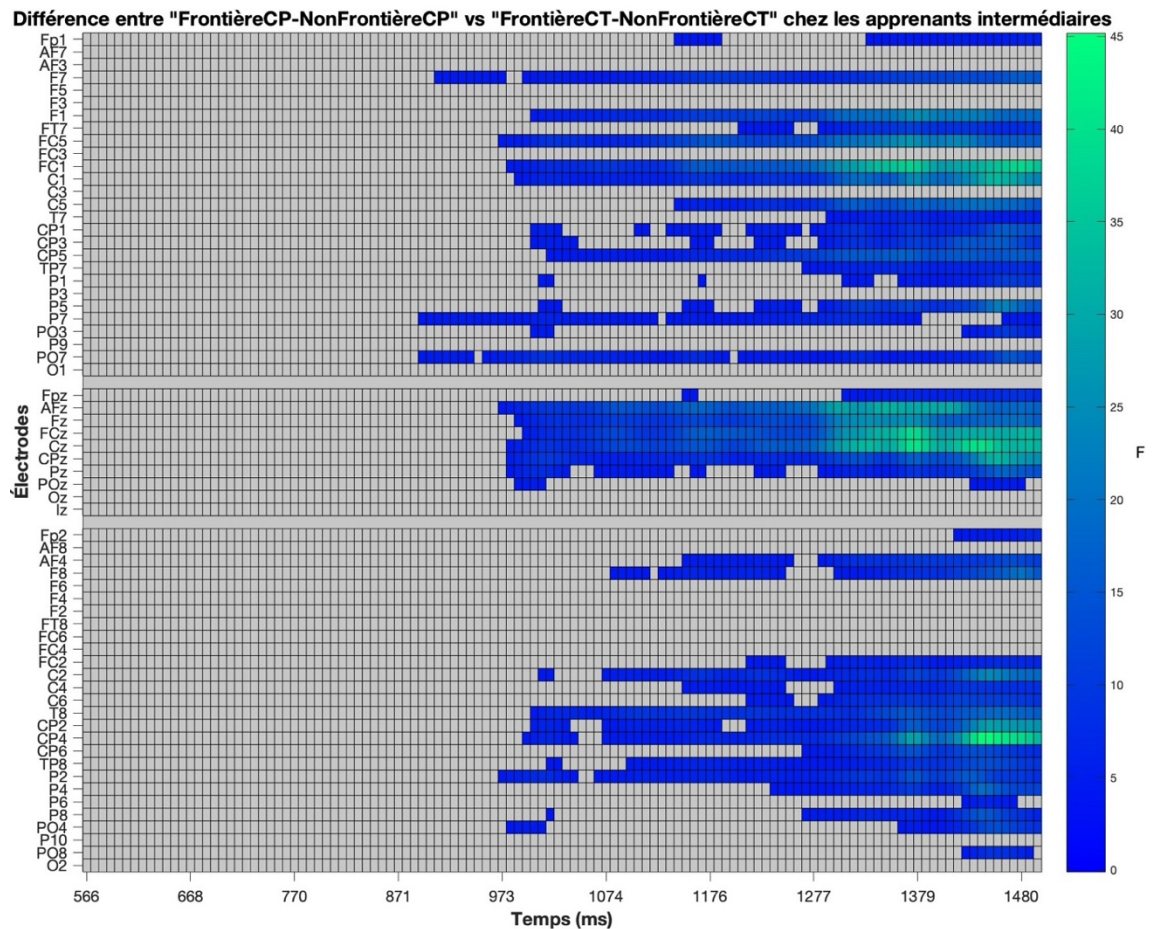


Figure 4.33 : Résultat du test de clusters basés sur les permutations pour la comparaison entre « *FrontièreCP– NonFrontièreCP* » vs « *FrontièreCT–NonFrontièreCT* » dans la FT [570-1500] ms chez les **apprenants intermédiaires**.

Pour conclure, les résultats des comparaisons entre les conditions prosodiques CP vs CT par l'approche traditionnelle montrent que les apprenants intermédiaires traitent les deux types de frontières prosodiques de manière similaire. L'amplitude de « *FrontièreCP– NonFrontièreCP* » ne diffère pas significativement de celle de « *FrontièreCT–NonFrontièreCT* », et c'est le cas pour les trois ROIs. Cela s'explique par les ROIs communes des composantes CPS pour les deux conditions CP et CT : elles se concentrent toutes les deux au niveau centro-frontal. Néanmoins, par l'analyse univariée de masse effectuée dans la FT [570-1500] ms, un cluster significatif a été identifié dans une FT de [895-1500] ms. Cela découle probablement du fait que l'effet CPS de la condition CP se présente dans une FT relativement courte ([572-843]) tandis que celui de la CT dure entre [596-1500] ms.

4.4.4 Comparaison des effets CPS en fonction de niveaux de français

En plus des comparaisons à l'intérieur d'un même groupe pour les conditions CP et CT, nous avons examiné les traitements neurocognitifs d'une même condition prosodique entre les groupes de participants, en fonction de leurs niveaux de français. Nous avons mené une série de tests de clusters dans la boîte à outils *Mass Univariate ERP Toolbox* (Groppe *et al.*, 2011a) afin de faire les comparaisons par paire : francophones natifs vs apprenants avancés, francophones natifs vs apprenants intermédiaires, et apprenants avancés vs intermédiaires. Pour chaque paire, nous avons comparé les différences de « *FrontièreCP–NonFrontièreCP* » et les différences de « *FrontièreCT–NonFrontièreCT* ». Pour les tests de clusters, nous avons utilisé les mêmes paramètres que ceux indiqués dans la section 4.3.2. Nous avons d'abord réalisé le test de clusters sur une FT restreinte, correspondant à l'intervalle temporel couvrant les effets CPS dans les deux groupes. Si aucun résultat significatif n'a été observé, nous avons élargi le test à la FT entière [0-1500] ms.

4.4.4.1 Comparaisons des effets CPS dans la condition *FrontièreCP* pour les trois groupes

Sur la base des latences des CPS obtenues pour la condition *FrontièreCP* chez les trois groupes de participants dans les sections précédentes, nous avons défini les trois FTs suivantes (Tableau 4.11) pour les trois paires de comparaisons :

Tableau 4.11 : FT pour les tests de clusters « *FrontièreCP–NonFrontièreCP* » dans les comparaisons par paire.

Comparaison entre les deux groupes	FT commune (ms)
Francophones natifs vs Apprenants avancés	[450-800]
Francophones natifs vs Apprenants intermédiaires	[450-840]
Apprenants avancés vs Apprenants intermédiaires	[570-840]

a) Comparaison entre les francophones natifs et les apprenants avancés :

Dans la FT [450-800] ms, le test de clusters n'a révélé aucun cluster significatif (avec la valeur $p=0,39$). Nous avons élargi la FT à [0-1500] ms pour identifier d'éventuels clusters significatifs en dehors de la FT initialement définie. Cependant, aucun résultat significatif n'a été trouvé sur l'ensemble de la FT [0-1500] ms, toutes les valeurs p étant égales ou supérieures à 0,84). Cela indique qu'il n'y a pas de

différence significative dans les traitements neurocognitifs des frontières prosodiques CP entre les francophones natifs et les apprenants avancés.

b) *Comparaison entre les francophones natifs et les apprenants intermédiaires :*

L'analyse de clusters réalisée sur la FT [450-840] ms a identifié un cluster significatif ($p=0,03$), localisé sur l'électrode TP7 (Figure 4.34).

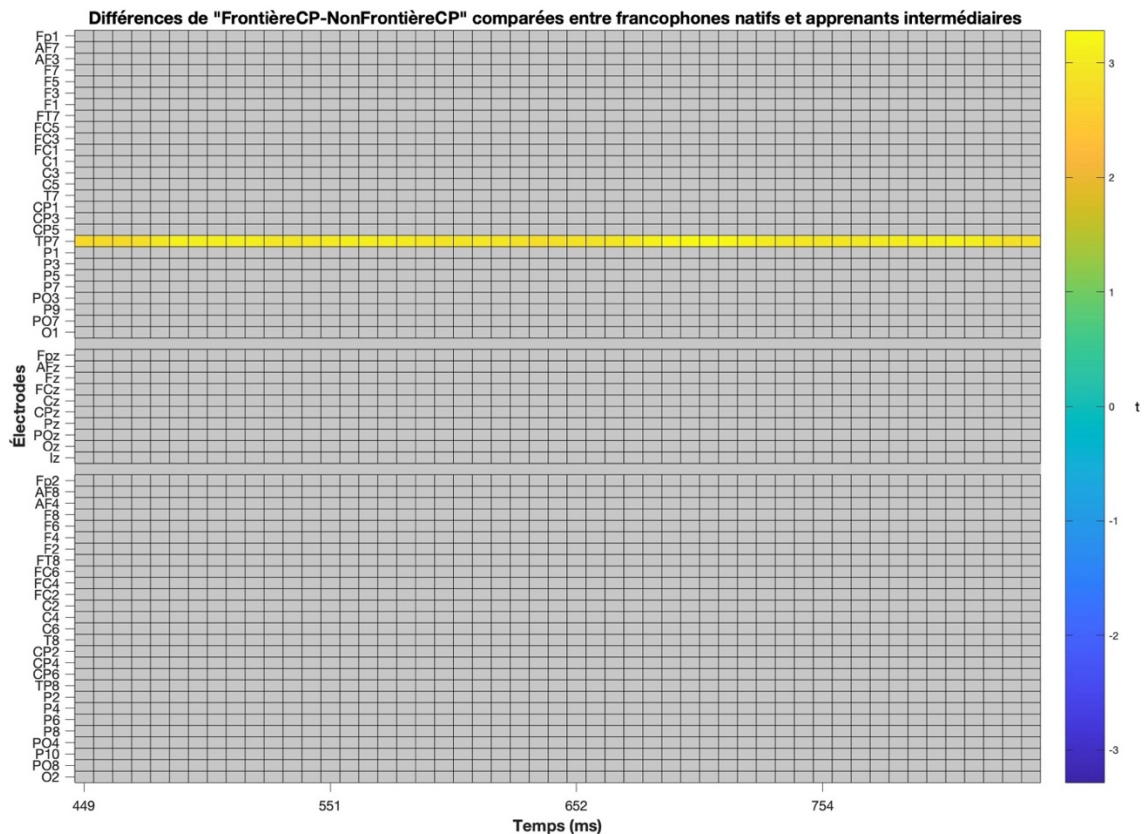


Figure 4.34 : Résultat du test de clusters pour les différences « *FrontièreCP–NonFrontièreCP* » dans la FT [450-840] ms comparées entre les francophones natifs et les apprenants intermédiaires.

Le résultat de ce test indique qu'il existe une différence significative dans les traitements neurocognitifs entre les francophones natifs et les apprenants intermédiaires quand ils perçoivent les frontières prosodiques CP. Le cluster identifié (Figure 4.34) suggère que cette différence se concentre principalement au niveau pariétal.

c) *Comparaison entre les apprenants avancés et les apprenants intermédiaires :*

Les tests de clusters n'ont révélé aucun cluster significatif, ni sur la FT [570-840] ms (avec toutes les valeurs p égales ou supérieures à 2), ni sur la FT plus large [0-1500]

ms (avec toutes les valeurs p égales ou supérieures à 0,64). Ces résultats suggèrent qu'il n'y a pas de différence dans le traitement neurocognitif des frontières prosodiques CP entre les apprenants avancés et intermédiaires.

Pour conclure, les résultats des trois tests de comparaisons par paires suggèrent que, pour les frontières prosodiques CP, les traitements neurocognitifs sont comparables entre les francophones natifs et les apprenants avancés, ainsi qu'entre les apprenants avancés et intermédiaires. Cependant, une différence significative a été observée entre les francophones natifs et les apprenants intermédiaires, cette différence étant principalement localisée au niveau pariétal.

4.4.4.2 Comparaisons des effets CPS dans la condition *FrontièreCT* pour les trois groupes

De la même manière que dans la section 4.4.4.1, nous avons d'abord fixé trois FTs (Tableau 4.12) dans lesquelles les tests de clusters ont été menés afin de comparer les traitements neurocognitifs de la condition CT entre groupes.

Tableau 4.12 : FTs pour les tests de clusters « *FrontièreCT–NonFrontièreCT* » dans les comparaisons par paire.

Comparaison entre les deux groupes	FT commune (ms)
Francophones natifs vs Apprenants avancés	[535-1500]
Francophones natifs vs Apprenants intermédiaires	[559-1500]
Apprenants avancés vs Apprenants intermédiaires	[535-1500]

a). Comparaison entre les francophones natifs et les apprenants avancés :

Sur la FT [535-1500] ms, aucun cluster significatif n'a été identifié par le test de clusters (avec les valeurs p égales ou supérieures à 0,84). Nous avons ensuite élargi la FT et répété le test sur la FT [0-1500] ms, mais là aussi, aucun cluster significatif n'a été révélé (avec les valeurs p égales ou supérieures à 0,98).

b). Comparaison entre les francophones natifs et les apprenants intermédiaires :

Le test de clusters sur la FT [559-1500] ms n'a révélé aucun cluster significatif (avec les valeurs p égales ou supérieures à 0,56). En élargissant la FT testé à [0-1500] ms, aucun résultat significatif n'a été obtenu non plus (avec les valeurs p égales ou supérieures à 0,67).

c). Comparaison entre les apprenants avancés et les apprenants intermédiaires :

Là aussi, les tests de clusters n'ont permis d'identifier aucun cluster significatif, ni sur la FT [535-1500] ms (avec les valeurs p égales ou supérieures à 1,17), ni sur la FT entière [0-1500] ms (avec les valeurs p égales ou supérieures à 1,38).

En conclusion, les résultats de ces trois tests de comparaisons par paires suggèrent que, pour les frontières prosodiques CT, les traitements neurocognitifs des auditeurs francophones natifs et d'apprenants sinophones sont tout à fait comparables.

4.5 Discussion

Dans cette section, nous répondons aux trois questions de recherche formulées en introduction (section 4.1), en nous appuyant sur l'examen des composantes CPS dans les deux conditions prosodiques chez les trois groupes de participants.

1) La CPS serait-elle déclenchée lorsque des auditeurs francophones natifs perçoivent les frontières prosodiques différentes (CP vs CT) dans des phrases localement ambiguës ?

Les résultats de nos analyses montrent que la composante CPS constitue un marqueur neurocognitif robuste associé au traitement des frontières prosodiques en français. Lorsque les francophones natifs détectent ces frontières, qu'elles soient précoces ou tardives, les indices acoustiques qui leur sont associés déclenchent systématiquement cette composante. Bien que la CPS ait déjà été documentée dans plusieurs langues (cf. section 1.5.2.1), cette étude représente, à notre connaissance, la première mise en évidence expérimentale de son activation en français métropolitain.

Par ailleurs, nos résultats corroborent les conclusions des études antérieures menées dans diverses langues, confirmant qu'il s'agit d'un potentiel évoqué bilatéral avec une large distribution (e.g., Steinhauer *et al.*, 1999 ; Kerkhofs *et al.*, 2007 ; Bögels *et al.*, 2010). Chez les francophones natifs, les effets de la CPS se manifestent toujours de manière bilatérale, bien que de légères variations aient été observées au niveau de la distribution topographique. Pour la structure CP, l'effet de la CPS s'étend sur l'ensemble des électrodes, avec une concentration dans la région centro-pariétale, tandis que pour la structure CT, la CPS est principalement activée dans la zone centro-pariétale.

L'analyse comparée des conditions *FrontièreCP* vs *NonFrontièreCP* chez les francophones natifs a mis en évidence d'autres effets liés au traitement syntaxico-prosodique. En particulier, une composante de négativité a été identifiée dans la condition *NonFrontièreCP*, apparaissant au même moment que la CPS observée en *FrontièreCP*. Cette activité neuronale se caractérise par une topographie étendue, se manifestant non seulement de manière évidente sur les électrodes représentatives (Figure 4.35), mais également presque sur la quasi-totalité des électrodes (Figure 4.36).

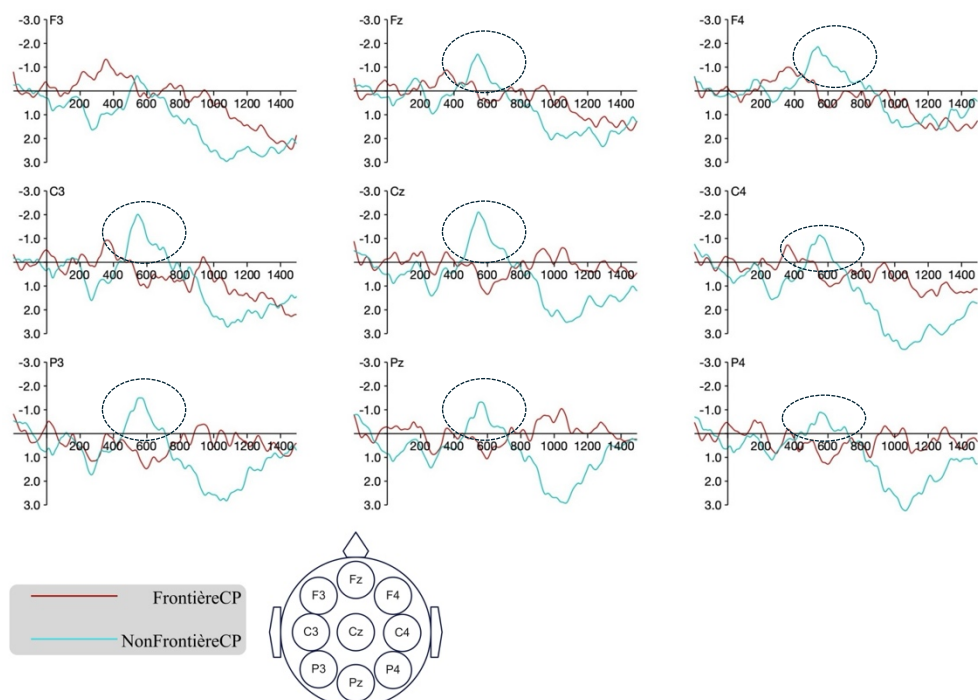


Figure 4.35 : Effets de négativité (entourés en pointillés noirs) dans la condition « *NonFrontièreCP* » sur les 9 électrodes représentatives chez les participants **francophones natifs**.

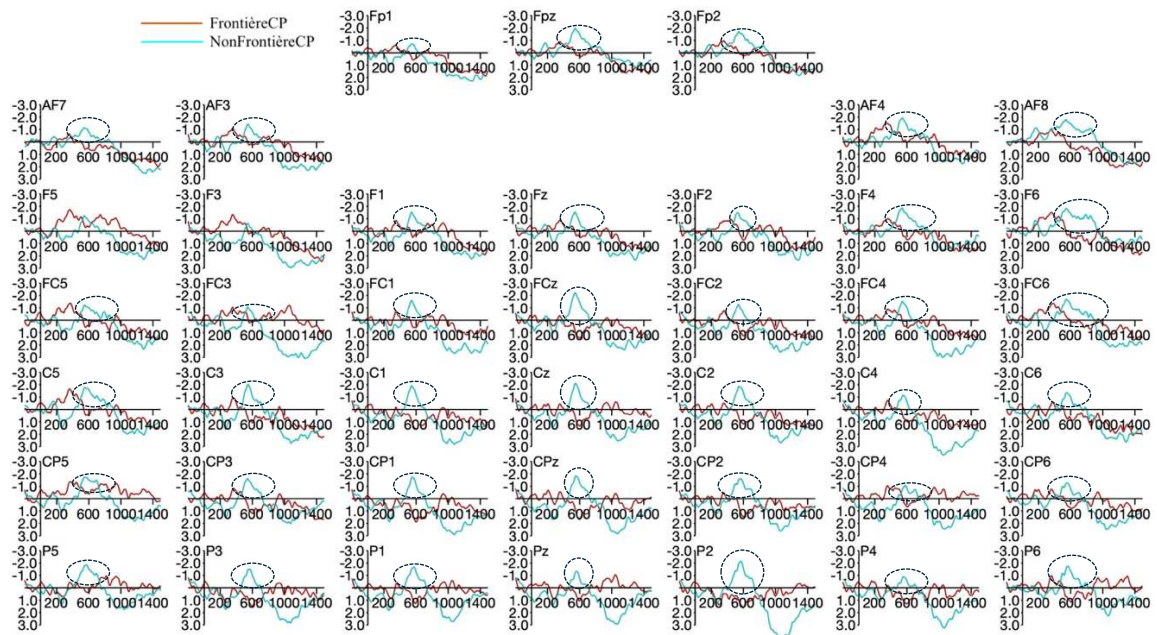


Figure 4.36 : Effets de négativité (entourés en pointillés noirs) dans la condition « *NonFrontièreCP* » sur les 42 électrodes des 3 ROIs chez les participants **francophones natifs**.

Une composante de négativité comparable a été rapporté dans les travaux de Steinhauer (2003), Pauker *et al.* (2011), Bögels *et al.* (2013), ainsi que Nickels et Steinhauer (2018). Dans les études de Nickels et Steinhauer (2018), ainsi que Pauker *et al.* (2011), cet effet, caractérisé par une distribution principalement sur la ligne centrale, a été désigné sous le terme de « *ramp-like negativity* ». Il serait lié à un mécanisme d'expectation cognitive, reflétant l'anticipation d'une frontière tardive lorsque les auditeurs percevaient des stimuli de type CT.

Selon Nickels et Steinhauer (2018), ce mécanisme neurocognitif correspond à un processus descendant (de haut en bas, *top-down process*), illustrant une interaction étroite entre prosodie et syntaxe. Dans notre étude EEG, tout comme dans l'étude de perception, les auditeurs, en écoutant un stimulus, anticipaient des indices acoustiques leur permettant d'établir les relations syntaxiques entre les constituants et de déterminer la fonction du syntagme nominal. Prenons l'exemple d'un stimulus de structure CT :

(a.) Chaque fois que le serpent mangeait (1) le rat (2), ...

Lors de la perception d'un stimulus comme (a.), l'absence d'indices acoustiques signalant une frontière du syntagme intonatif à la position (1) (correspondant à la frontière CP), conduit les auditeurs à anticiper une frontière intonative tardive à la position (2) (correspondant à la frontière CT). Parallèlement, ils interprètent le

syntagme nominal (par exemple, « *le rat* ») comme complément d'objet direct de la proposition en cours de traitement. Cette anticipation se traduit, dans les réponses électroencéphalographiques, par une négativité observée à la position (1), coïncidant avec la CPS déclenchée dans la structure CP. À la position (2), la frontière intonative tardive est marquée par l'apparition naturelle de la CPS. Notons que, dans notre étude, cette négativité présentait une distribution étendue sur l'ensemble des électrodes (Figure 4.36), ce qui diffère des observations de Pauker *et al.* (2011) et de Nickels et Steinhauer (2018), où l'effet était principalement concentré au niveau frontal.

2) *La CPS serait-elle déclenchée lorsque les apprenants sinophones perçoivent les mêmes phrases ambiguës ? Comment le niveau de français modulerait-il la composante CPS ?*

Chez les apprenants sinophones, tant au niveau intermédiaire qu'avancé, la CPS est également observée lors du traitement des frontières prosodiques CP et CT. Ce résultat, cohérent avec nos attentes, confirme l'hypothèse formulée en introduction de ce chapitre. Des études antérieures portant sur l'acquisition de différents aspects d'une L2 ont abouti à une conclusion similaire : même à un niveau débutant, il est possible d'observer la composante électroencéphalographique associée à un aspect linguistique donné, bien qu'elle puisse présenter des différences au niveau de la latence ou de la carte topographique (cf. section 1.5.3). Les études EEG portant plus spécifiquement sur l'acquisition de la prosodie d'une L2 confirment des résultats similaires : dans toutes les études antérieures sur la CPS chez des apprenants de différentes L2 (cf. Tableau 1.25 ; Gilbert *et al.*, 2023 ; Nickels & Steinhauer, 2018 ; Nickels *et al.*, 2013 ; Naito-Billen, 2012), la CPS a été systématiquement observée dans les groupes d'apprenants. Comme chez les francophones natifs, les réponses CPS chez des apprenants, quel que soit leur niveau, se répartissent largement sur les hémisphères bilatéraux, bien que des différences aient été relevées concernant la durée de leur FT. De manière générale, dans les deux groupes d'apprenants, les CPS associées aux frontières CP présentent une durée relativement courte (entre 200-300 ms), tandis que pour les frontières CT, les effets CPS sont plus prolongés, atteignant environ 900 ms.

En ce qui concerne la modulation de la CPS en fonction de niveau de français, les comparaisons effectuées sur une même condition prosodique entre les trois groupes révèlent les résultats suivants :

1) Pour la frontière CT :

Aucune différence significative n'est observée entre les trois groupes. Autrement dit, les apprenants sinophones et les francophones natifs ont une performance comparable en ce qui concerne le traitement neurocognitif de la frontière CT. Ce résultat, en accord avec les conclusions de l'étude de perception, confirme que la perception de la structure CT n'entraîne pas de difficultés particulières, y compris pour les apprenants. Comme nous l'avons déjà indiqué dans le chapitre précédent, ce résultat peut s'expliquer par les prédictions de la *Late Closure Preference* (Frazier, 1979) et de l'*Informative Boundary Hypothesis* (Clifton Jr. et al., 2002 ; Watson & Gibson, 2005 ; Carlson et al., 2009a, 2009b). La clôture tardive est considérée comme un processus de traitement syntaxique par défaut. Dans notre expérience EEG, les stimuli ont été présentés dans leur intégralité (non tronqués), fournissant ainsi tous les indices acoustiques et le contexte prosodique nécessaires, facilitant le traitement et l'intégration des informations syntaxico-prosodiques des frontières CT chez les natifs et les apprenants.

2) Pour la frontière CP :

Cette frontière active aussi la composante CPS dans les trois groupes de participants, très certainement en raison de la présence de contextes prosodiques complets dans les stimuli. Cependant, lorsque nous comparons les traitements des frontières CP entre les trois niveaux de français, l'analyse univariée de masse a identifié une différence significative entre les francophones natifs et les apprenants intermédiaires dans la FT commune [450-840] ms. D'après le cluster identifié, cette différence est observable sur l'électrode TP7, c'est-à-dire au niveau pariétal gauche. Cette différence est également très visible sur les cartes topographiques de cette FT [450-840] (Figure 4.37) :

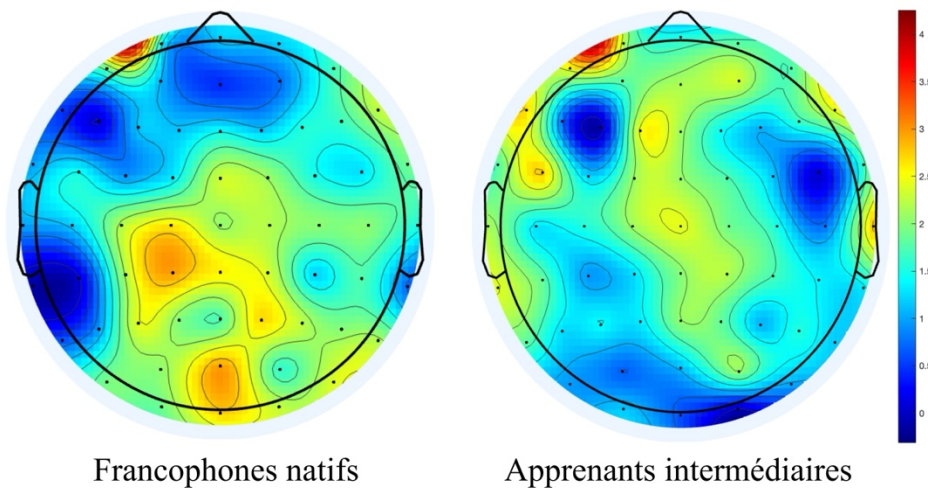


Figure 4.37 : Comparaison des cartes topographiques de la condition *FrontièreCP* dans la FT [450-840] ms entre les francophones natifs (gauche) et les apprenants intermédiaires (droite).

La Figure 4.37 montre que, dans la condition *FrontièreCP*, la polarité positive de la CPS chez les francophones natifs se situe au niveau centro-pariétal. En revanche, chez les apprenants intermédiaires, cette polarité est davantage localisée au niveau centro-frontal. Ainsi, le test comparatif du traitement neurocognitif des frontières CP entre ces deux groupes a mis en évidence un cluster de différence au niveau pariétal. Par ailleurs, les traitements neurocognitifs de la structure CP sont similaires entre les francophones natifs et les apprenants avancés, ainsi qu'entre les apprenants avancés et intermédiaires. Cette variabilité dans le traitement de la structure CP entre les trois groupes reflète l'acquisition graduelle de la prosodie en L2 : à mesure que les apprenants progressent en français, leur traitement neurocognitif des frontières prosodiques de type CP converge progressivement vers celui des locuteurs natifs. Cette observation suggère une neuroplasticité, c'est-à-dire la capacité du cerveau à changer et à s'adapter au cours de l'apprentissage d'une L2. Chez les francophones natifs, nous avons identifié une composante de négativité dans la condition CT, qui apparaît au même moment que la composante CPS de la condition « *FrontièreCP* ». Une négativité similaire a été aussi observée dans des études précédentes sur la CPS, entre autres par Steinhauer (2003), Bögels *et al.* (2013), et Nickels et Steinhauer (2018). Comme nous l'avons indiqué plus haut, cet effet a été interprété comme reflétant le prétraitement cognitif d'une future frontière prosodique ou comme une expectation/anticipation d'une frontière qui apparaîtrait éventuellement plus tard. Dans Nickels et Steinhauer (2018), qui a examiné la perception des frontières prosodiques de l'anglais chez des apprenants

germanophones et sinophones, cette négativité était la plus marquée chez les anglophones natifs, modérément présente chez les apprenants germanophones, et beaucoup moins évidente chez les sinophones. Selon eux, plus le niveau de maîtrise de la langue est avancé (voire proche de celui d'un natif), plus cette négativité devient évidente. Cela reflète un parcours d'acquisition progressif menant à un traitement descendant (*top-down*) des informations syntactico-prosodiques, similaire à celui observé chez les locuteurs natifs.

En nous inspirant de cette analyse (Nickels & Steinhauer, 2018), nous avons examiné les négativités dans les signaux ERP des deux groupes d'apprenants de notre étude. Nos observations révèlent une tendance similaire à celle mise en évidence par Nickels et Steinhauer (2018) : alors que les auditeurs francophones natifs affichent des négativités claires (Figure 4.36), les apprenants avancés présentent des négativités moins évidentes (Figure 4.38), réparties de manière diffuse dans les zones frontale et centrale.

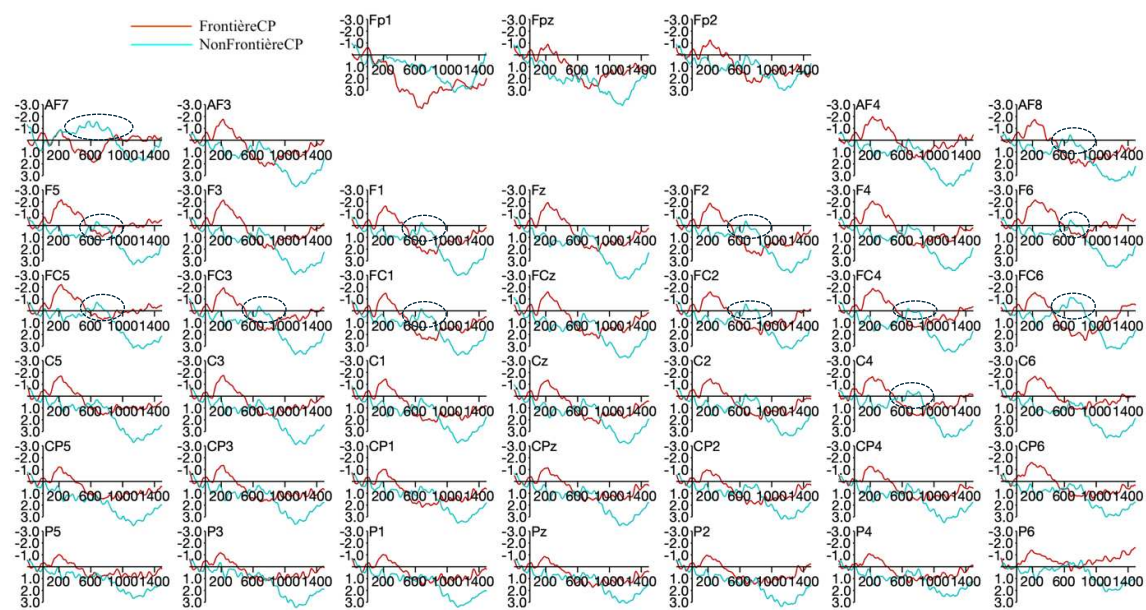


Figure 4.38 : Effets de négativité (entourés en pointillés noirs) dans la condition « *NonFrontièreCP* » sur les 42 électrodes des 3 ROIs chez les **apprenants avancés**.

Chez les apprenants intermédiaires, cette négativité d'attente est encore plus faible par rapport aux apprenants avancés (Figure 4.39) :

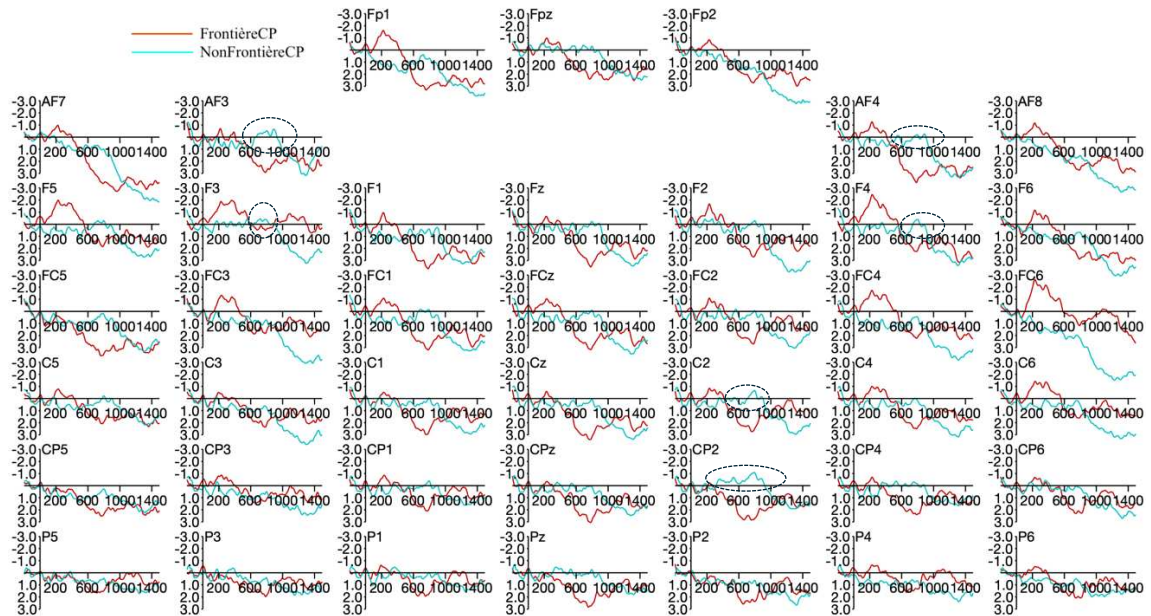


Figure 4.39 : Peu d'effets de négativité (entourés en pointillés noirs) dans la condition « *NonFrontièreCP* » sur les 42 électrodes des 3 ROIs chez les **apprenants intermédiaires**.

Bien que les CPS soient observées pour les deux groupes d'apprenants, elles sont déclenchées plutôt par les informations purement sonores présentes dans les stimuli (allongement final, pause et montée de F0). Ces CPS observées chez les apprenants peuvent être comparés aux CPS musicales qui sont activées par les indices sonores parquant les frontières entre les syntagmes musicaux (Neuhaus *et al.*, 2006)¹²³ dans lesquels les autres informations linguistiques (syntaxiques, sémantiques, etc.) sont absentes. Ainsi, comme l'ont souligné Nickels et Steinhauer (2018), dans notre étude, les CPS déclenchées chez les apprenants reflètent plutôt un traitement ascendant (*bottom-up*), contrairement au traitement descendant (*top-down*) chez les natifs.

La faiblesse de la composante « *ramp-like negativity* » chez les apprenants suggère que les indices acoustiques avant les frontières prosodiques jouent un rôle limité lors de la construction neurocognitive de la hiérarchie syntaxique et de l'établissement des bonnes relations syntaxiques entre les constituants prosodiques. Chez les natifs, le traitement des frontières prosodiques est un travail cognitif constant, qui s'étend sur l'ensemble de la phrase. Ainsi, l'absence d'une frontière à une position précoce les poussent à en anticiper une à une position tardive. Au

¹²³ Les indices sonores des frontières entre les syntagmes musicaux peuvent être, selon Neuhaus *et al.* (2006), 1) pause entre les syntagmes, 2) durée du dernier ton musical avant la pause et 3) fonction harmonique du dernier ton.

niveau syntaxique, le syntagme nominal « le rat » dans la phrase (a.) ci-dessus peut être ainsi interprété comme un complément d'objet direct. En revanche, chez les apprenants, le traitement des frontières prosodiques est davantage local et se limite à la dernière syllabe. Lorsque les indices acoustiques sont disponibles, ils parviennent à détecter la frontière, comme en témoigne la composante CPS. Toutefois, en l'absence de ces indices, aucune CPS n'est observée, et ils ne sont pas encore en mesure d'anticiper une éventuelle frontière tardive.

3) *Notre expérience perceptive a montré que les auditeurs francophones natifs et les apprenants avaient particulièrement du mal à traiter la structure CP tandis que la structure CT ne posait aucun problème. Comment cette différence comportementale se manifesterait-elle par les réponses CPS ?*

Pour répondre à cette question, nous avons comparé les traitements des structures CP et CT chez un même groupe, en calculant la différence d'amplitudes moyennes entre la présence d'une frontière et l'absence d'une frontière (i.e., « *Frontière–NonFrontière* ») pour les deux structures.

Tout d'abord, chez les francophones natifs, nous avons observé une différence significative, avec une différence de « *FrontièreCP – NonFrontièreCP* » (moyenne = 1,16 ; médiane = 1,24) légèrement plus grande que celle de « *FrontièreCT–NonFrontièreCT* » (moyenne=0,53 ; médiane=1). Cette différence se manifeste principalement au niveau frontal. Nous estimons cependant que la différence observée n'est pas liée à une plus grande facilité du traitement de la structure CP par rapport à la structure CT. Elle est plutôt liée aux différentes distributions topographiques des deux conditions : l'effet de la condition CP se manifeste sur l'ensemble des électrodes, tandis que celui de la CT se concentre au niveau centro-pariétal.

Concernant les apprenants avancés, nous avons également observé une différence significative, avec « *FrontièreCT–NonFrontièreCT* » (moyenne=2,81 ; médiane=2,40) plus grande que « *FrontièreCP–NonFrontièreCP* » (moyenne=0,81 ; médiane=2,36). Cette différence se manifeste uniquement au niveau pariétal. En effet, l'effet positif de la condition CT est largement distribué à l'ensemble des électrodes, alors que celui de la condition CP reste limité à la zone centro-frontale.

Enfin, chez les apprenants intermédiaires, aucune différence significative n'a été observée entre « *FrontièreCP–NonFrontièreCP* » et « *FrontièreCT–NonFrontièreCT* ». Cela peut s'expliquer par le fait que les traitements des deux conditions prosodiques CP et CT se concentrent dans la même zone centro-frontale. Cela suggère également, sur le plan neurocognitif, que les apprenants intermédiaires traitent les deux types de frontières prosodiques de manière similaire.

En nous fondant sur ces discussions, nous pouvons répondre à la question de recherche 3) : sur le plan neurocognitif, lorsque le contexte prosodique complet est pleinement disponible dans le signal de parole, les locuteurs natifs et les apprenants traitent de manière similaire les deux types de frontières prosodiques, CP et CT. Les différences observées se limitent à des variations mineures au niveau des distributions topographiques, en fonction de la condition CP ou CT. Ces résultats neurophysiologiques viennent compléter les résultats de notre étude perceptive et soutiennent la prédiction de l'*Informative Boundary Hypothesis* (Clifton Jr. *et al.*, 2002 ; Watson & Gibson, 2005 ; Carlson *et al.*, 2009a, 2009b). En effet, au lieu de traiter par défaut les stimuli localement ambigus et tronqués comme des structures de clôture tardive, les locuteurs natifs comme les apprenants, dans le cadre de l'expérience EEG avec un contexte prosodique complet, sont capables de distinguer les frontières prosodiques précoces et tardives dans les phrases localement ambiguës.

Notre étude électroencéphalographique avait pour objectif principal d'examiner la composante CPS chez les francophones natifs et les apprenants sinophones en réponse à la perception des frontières prosodiques du français. Les résultats de nos analyses sur la CPS reportés ci-dessus ont répondu à nos premiers questionnements. Une synthèse générale sur les résultats importants obtenus dans les trois études sera présentée dans le chapitre suivant. Outre les conclusions importantes, nous discuterons des limites et des pistes possibles pour améliorer nos trois études.

4.6 Annexes

4.6.1 Consentement de participation à l'expérience électroencéphalographique

Notice d'information

Les informations recueillies vous concernant vont faire l'objet d'un traitement dans le cadre du projet de thèse de Lei Xi (lei.xi@sorbonne-nouvelle.fr), sous la direction de M. Rachid Ridouane du Laboratoire de Phonétique et Phonologie (UMR 7018, 4 rue des Irlandais, 75005 Paris).

La Déléguée à la Protection des Données du CNRS peut être contactée pour toute question sur la protection des données personnelles. Ses coordonnées sont les suivantes : CNRS Service protection des données, 2 rue Jean Zay, 54519 Vandoeuvre-lès-Nancy ; mail : dpd.demandes@cnrs.fr

Projet de recherche et ses enjeux

Le projet de thèse porte sur : Acquisition de la prosodie du français par des apprenants sinophones : études comportementale et électrophysiologique.

Dans ce projet, nous cherchons à étudier le traitement cognitif de la perception des frontières prosodiques chez des locuteurs francophones natifs et des apprenants sinophones du français. Pour cela, nous utilisons l'électroencéphalographie (EEG), méthode expérimentale qui permet d'enregistrer l'activité électrique cérébrale au moyen d'électrodes fixées sur un bonnet que vous porterez le temps de l'expérience. Cette technique, pratiquée de façon très courante pour les recherches en neuroscience, est non-invasive, non douloureuse, et ne représente aucun danger particulier ni contre-indication connue. La séance entière, comprenant la mise en place de l'EEG et la réalisation de la tâche expérimentale, durera environ une heure trente. Il vous sera demandé d'écouter des phrases en français et de répondre par « oui » ou par « non » à des questions qui vous seront posées. Vous serez indemnisé.e à hauteur de 30 euros (en espèces ou en chèque cadeau) pour votre participation.

À ce jour, très peu d'études ont porté sur l'acquisition de la prosodie chez des bilingues tardifs (adultes). Votre participation dans cette expérience nous permettra de constituer les premières données électrophysiologiques sur l'acquisition de la prosodie du français par des sinophones.

La base légale du traitement est l'exécution d'une mission de service public.

Seules les données strictement nécessaires à la réalisation de notre recherche seront collectées et traitées.

Le doctorant et son directeur sont les seuls destinataires de tout ou partie des données à caractère personnel.

Vos données à caractère personnel seront pseudonymisées et conservées pour la durée nécessaire à l'exécution du projet.

Participation libre

Votre participation à ce projet est entièrement libre et volontaire.

Mesures de sécurité

Afin de garantir la confidentialité de vos données et éviter leur divulgation, toutes les mesures de sécurité, tant physique que logique, ont été prises.

Déroulement de l'expérience

La participation consiste en une séance d'expérimentation par EEG, d'une durée d'environ une heure trente. Vous serez installé.e dans une salle insonorisée face à un écran d'ordinateur. L'expérimentateur mesurera le périmètre de votre crâne afin de choisir le bonnet avec la taille appropriée, et commencera la mise en place des électrodes. Ensuite, il branchera l'appareil EEG avec le système Brain Vision Recorder et l'amplificateur et vérifiera que toutes les électrodes fonctionnent et sont bien branchées. Avant l'expérience, le signal EEG sera enregistré d'abord pendant 4 minutes : 2 minutes avec les yeux ouverts et 2 minutes

avec les yeux fermés, sans tâche comportementale réalisée. Ce signal pré-expérimental a pour but de vérifier le bon fonctionnement de l'appareil EEG. Après le lancement de l'expérience, le signal EEG sera enregistré pendant que vous écouterez des phrases. Ces phrases sont diffusées l'une après l'autre. Après chaque phrase, une question de compréhension s'affiche sur l'écran et vous devez y répondre en appuyant, avec la main droite, sur la touche « F » du clavier pour le choix « non » ou la touche « J » pour le choix « oui ». L'EEG enregistre le signal électrique en même temps que vous écoutez des phrases et réalisez la tâche demandée.

Retrait du consentement

Vous êtes libre de vous retirer ou de cesser votre participation à ce projet à tout moment. Ce retrait n'aura aucune conséquence.

Vous disposez des droits suivants pour l'utilisation qui est faite de vos données :

- Le droit d'opposition : vous pouvez à tout moment vous opposer au traitement de vos données et disposer du droit de retirer votre consentement ;
- Le droit d'accès et de rectification de vos données ;
- Le droit d'effacement ;
- Le droit à une utilisation restreinte lorsque vos données ne sont pas nécessaires ou ne sont plus utiles ;
- Si l'équipe constate une anomalie, elle vous orientera vers les services compétents.

Vous pouvez exercer ces droits en vous adressant à :

Lei Xi, 4 rue des Irlandais 75005 Paris, mail : lei.xi@sorbonne-nouvelle.fr

Vous pouvez également vous adresser à la Déléguée à la Protection des Données du CNRS, 2 rue Jean Zay 54519 Vandœuvre-lès-Nancy, mail : dpd.demandes@cnrs.fr

Si vous estimez, après nous avoir contactés, que vos droits Informatique et Libertés ne sont pas respectés, vous avez la possibilité d'introduire une réclamation en ligne auprès de la CNIL ou par courrier postal CNIL, 3 Place de Fontenoy, TSA 80715 – 75334 Paris Cedex 07 (<https://www.cnil.fr/>).

Consentement et Autorisation

Ce formulaire est destiné à recueillir votre consentement pour la collecte des données vous concernant, et votre autorisation sur leur exploitation, dans le cadre du projet de thèse de Lei Xi (lei.xi@sorbonne-nouvelle.fr), sous la direction de M. Rachid Ridouane du Laboratoire de Phonétique et Phonologie (UMR 7018, 4 rue des Irlandais, 75005 Paris).

En signant le formulaire de consentement, vous certifiez :

que vous avez lu et compris les renseignements communiqués dans la notice d'information,

que nous avons répondu à vos questions de façon satisfaisante,

que nous vous avons informé que de légères anomalies peuvent être repérées grâce à l'EEG enregistré,

que nous vous avons informé que vous étiez libre d'annuler votre consentement ou de vous retirer de cette recherche à tout moment, sans préjudice.

À remplir par le participant :

J'ai lu et compris les renseignements fournis dans la notice d'informations et j'accepte de plein gré de participer à cette recherche.

OUI NON

J'ai compris la procédure de l'étude et j'accepte le cadre global de l'étude.

OUI NON

J'accepte que les données sensibles (vocales et/ou électrophysiologiques) soient collectées, conservées et exploitées par l'équipe du projet.

OUI NON

J'accepte que mes données (vocales et/ou électrophysiologiques) pseudonymisées soient **diffusées** dans le cadre de publications, colloques scientifiques, séminaires ou dans toute forme de valorisation du projet.

OUI NON

J'accepte que mes données (vocales et/ou électrophysiologiques) pseudonymisées soient **utilisées** dans le cadre de publications, colloques scientifiques, séminaires ou toute forme de valorisation du projet

OUI NON

J'accepte d'être informé.e si l'éventuelle présence d'une anomalie a été constatée durant l'expérience.

OUI NON

En cas de doute sur la présence d'une anomalie, j'accepte d'être orienté.e vers les services compétents.

OUI NON

SI NON :

J'accepte d'être écarté.e de l'étude.

OUI NON

Nom, Prénom :

Date :

Signature :

Un exemplaire de ce document est remis au participant, un autre exemplaire est conservé dans le dossier du projet.

Tenir une paire de ciseaux	Cliquez ici pour choisir
Se brosser les dents	Cliquez ici pour choisir
Tenir un couteau (sans fourchette)	Cliquez ici pour choisir
Tenir une cuillère	Cliquez ici pour choisir
Tenir un balai	Cliquez ici pour choisir
Ouvrir un couvercle	Cliquez ici pour choisir
Shooter dans un ballon	Cliquez ici pour choisir

Avez-vous déjà eu des problèmes de vision , problèmes d'audition , troubles neurologiques , troubles psychiatriques , troubles du langage , ou troubles d'apprentissage ? (Cochez toutes les cases applicables).

Seriez-vous d'accord que nous vous contactions pour vous inviter à participer à nos expériences prochaines (avec une rémunération) ? Oui Non Peut-être

Merci beaucoup de votre temps !

Quel type d'épreuve	Quand ? (mois/année)	Dans quel pays ?	Niveau équivalent du CECRL
TCF/TEF			Cliquez ici
DELF			Cliquez ici
DALF			Cliquez ici
Autre (à préciser) :			Cliquez ici

Je n'ai jamais passé d'épreuve sur mon niveau de compétence du français.

Globalement, comment vous évaluez votre niveau de compétence du français actuel :

bilingue avancé intermédiaire débutant

Sur une échelle de zéro à dix, choisissez dans le menu votre niveau de compétence à l'oral et en compréhensions du français :

Oral	Cliquez ici	Compréhension de la langue orale	Cliquez ici	Lecture	Cliquez ici
------	-------------	----------------------------------	-------------	---------	-------------

Sur une échelle de zéro à dix, évaluez dans quelle mesure vous êtes actuellement exposé(e) au français dans les contextes suivants :

Interagir avec vos amis	Cliquez ici	Interagir avec vos collègues (à l'école ou au travail)	Cliquez ici
Interagir avec la famille	Cliquez ici	Regarder la télévision / Écouter la radio	Cliquez ici
Lecture (romans, journaux, etc.)	Cliquez ici	Auto-apprentissage / Auto-perfectionnement	Cliquez ici

Parlez-vous d'autre(s) langue(s) étrangère(s) excepté le français ? Oui Non

(**Attention** : « Parler » ici sous-entend que vous pouvez vous exprimer dans cette langue de façon claire et détaillée avec un degré d'aisance.)

Si oui, précisez laquelle (lesquelles) :

Âge d'apprentissage de cette(ces) langue(s) étrangère(s) : ans.

Test de latéralité manuelle (*Edinburgh Handedness Inventory*) : Merci d'indiquer votre main préférée dans les activités suivantes :

Activité	Préférence
Écrire	Cliquez ici
Dessiner	Cliquez ici
Lancer une balle	Cliquez ici
Tenir une paire de ciseaux	Cliquez ici
Se brosser les dents	Cliquez ici
Tenir un couteau (sans fourchette)	Cliquez ici
Tenir une cuillère	Cliquez ici
Tenir un balai	Cliquez ici
Ouvrir un couvercle	Cliquez ici
Shooter dans un ballon	Cliquez ici

Avez-vous déjà eu des problèmes de vision , problèmes d'audition , troubles neurologiques , troubles psychiatriques , troubles du langage ou troubles d'apprentissage ? (Cochez toutes les cases applicables).

Seriez-vous d'accord que nous vous contactions pour vous inviter à participer à nos expériences prochaines (avec une rémunération) ? Oui Non Peut-être

Merci beaucoup de votre temps !

4.6.3 Les 150 stimuli utilisés dans l'expérience électroencéphalographique

- 1) - Chaque fois que le chanteur chantait, son album devenait un succès.
- Chaque fois que le chanteur chantait son album, le produit devenait un succès.
- 2) - Chaque fois que le garçon jouait, le jeu durait des heures.
- Chaque fois que le garçon jouait le jeu, le dîner durait des heures.
- 3) - Chaque fois que l'étudiant rédigeait, son mémoire était amélioré.
- Chaque fois que l'étudiant rédigeait son mémoire, la tournure était améliorée.
- 4) - Chaque fois que le chien léchait, le bébé souriait.
- Chaque fois que le chien léchait le bébé, la nounou souriait.
- 5) - Chaque fois que le professeur enseignait, la théorie devenait claire.
- Chaque fois que le professeur enseignait la théorie, son raisonnement devenait clair.
- 6) - Chaque fois que la femme feuilletait, le journal tombait par terre.
- Chaque fois que la femme feuilletait le journal, la télécommande tombait par terre.
- 7) - Chaque fois que le serpent mangeait, le rat se cachait.
- Chaque fois que le serpent mangeait le rat, le lapin se cachait.
- 8) - Chaque fois que l'artiste dessinait, le portrait devenait plus coloré.
- Chaque fois que l'artiste dessinait le portrait, la toile devenait plus colorée.
- 9) - Chaque fois que le tigre grimpait, l'arbre tremblait.
- Chaque fois que le tigre grimpait l'arbre, les feuilles tremblaient.
- 10) - Chaque fois que l'étudiant révisait, ses notes étaient plus élevées.
- Chaque fois que l'étudiant révisait ses notes, sa moyenne était plus élevée.
- 11) - Chaque fois que l'acteur jouait, la scène amusait les spectateurs.
- Chaque fois que l'acteur jouait la scène, son costume amusait les spectateurs.
- 12) - Chaque fois que le comédien jouait, le spectacle était réussi.
- Chaque fois que le comédien jouait le spectacle, la soirée était réussie.
- 13) - Chaque fois que le patient attendait, la visite était déprogrammée.
- Chaque fois que le patient attendait la visite, l'opération était déprogrammée.
- 14) - Chaque fois que l'ennemi attaquait, nos soldats ripostaient.
- Chaque fois que l'ennemi attaquait nos soldats, la force aérienne ripostait.
- 15) - Chaque fois que le cuisinier grillait, le steak sentait très bon.
- Chaque fois que le cuisinier grillait le steak, le restaurant sentait très bon.
- 16) - Chaque fois que le serrurier vérifiait, la porte était cassée.

- Chaque fois que le serrurier vérifiait la porte, le verrou était cassé.
- 17) - Chaque fois que le technicien programmait, le site tombait en panne.
 - Chaque fois que le technicien programmait le site, le logiciel tombait en panne.
- 18) - Chaque fois que le poète contemplant, le ciel l'inspirait.
 - Chaque fois que le poète contemplant le ciel, les nuages l'inspiraient.
- 19) - Chaque fois que le voisin rénoverait, la maison était en désordre.
 - Chaque fois que le voisin rénoverait la maison, l'allée était en désordre.
- 20) - Chaque fois que le rat reniflait, le fromage tombait par terre.
 - Chaque fois que le rat reniflait le fromage, l'assiette tombait par terre.
- 21) - Chaque fois que mon père conduisait, la voiture tombait en panne.
 - Chaque fois que mon père conduisait la voiture, le phare tombait en panne.
- 22) - Chaque fois que le pilote commençait, la course captivait la foule.
 - Chaque fois que le pilote commençait la course, la vitesse captivait la foule.
- 23) - Chaque fois que l'employé appelait, le patron répondait.
 - Chaque fois que l'employé appelait le patron, la secrétaire répondait.
- 24) - Chaque fois que ma sœur dansait, le ballet semblait beau.
 - Chaque fois que ma sœur dansait le ballet, son profil semblait beau.
- 25) - Chaque fois que le vendeur fermait, le supermarché était nettoyé.
 - Chaque fois que le vendeur fermait le supermarché, les rayons étaient nettoyés.
- 26) - Chaque fois que le musicien fredonnait, le morceau était joli.
 - Chaque fois que le musicien fredonnait le morceau, la mélodie était jolie.
- 27) - Chaque fois que le garçon visitait, le château était désert.
 - Chaque fois que le garçon visitait le château, le donjon était désert.
- 28) - Chaque fois que l'entraîneur aidait, l'équipe était motivée.
 - Chaque fois que l'entraîneur aidait l'équipe, les supporters étaient motivés.
- 29) - Chaque fois que les parents racontaient, l'histoire semblait horrible.
 - Chaque fois que les parents racontaient l'histoire, l'intrigue semblait horrible.
- 30) - Chaque fois que le jardinier creusait, le puits était de plus en plus profond.
 - Chaque fois que le jardinier creusait le puits, le trou était de plus en plus profond.
- 31) - Chaque fois que le sportif nageait, la brasse semblait difficile.
 - Chaque fois que le sportif nageait la brasse, sa technique semblait difficile.
- 32) - Chaque fois que l'auteur lisait, son roman devenait vivant.

- Chaque fois que l'auteur lisait son roman, le personnage devenait vivant.
- 33) - Chaque fois que Marie poussait, le bouton était bloqué.
 - Chaque fois que Marie poussait le bouton, la télécommande était bloquée.
- 34) - Chaque fois que le chien chassait, le chat s'enfuyait.
 - Chaque fois que le chien chassait le chat, la souris s'enfuyait.
- 35) - Chaque fois que le conducteur freinait, le bus perdait le contrôle.
 - Chaque fois que le conducteur freinait le bus, le volant perdait le contrôle.
- 36) - Chaque fois que Marie soufflait, la bougie s'éteignait.
 - Chaque fois que Marie soufflait la bougie, le feu s'éteignait.
- 37) - Chaque fois que le garçon servait, le client était content.
 - Chaque fois que le garçon servait le client, le patron était content.
- 38) - Chaque fois que je pleurais, ma mère devenait triste.
 - Chaque fois que je pleurais ma mère, mon fils devenait triste.
- 39) - Chaque fois que le chien mordait, le bébé criait fort.
 - Chaque fois que le chien mordait le bébé, sa mère criait fort.
- 40) - Chaque fois que l'audience applaudissait, le directeur était satisfait.
 - Chaque fois que l'audience applaudissait le directeur, le président était satisfait.
- 41) - Chaque fois que le chef chauffait, le poulet était brûlé.
 - Chaque fois que le chef chauffait le poulet, la viande était brûlée.
- 42) - Chaque fois que le général commandait, le combat était réussi.
 - Chaque fois que le général commandait le combat, le résultat était réussi.
- 43) - Chaque fois que mon camarade copiait, mes réponses étaient bonnes.
 - Chaque fois que mon camarade copiait mes réponses, ma note était bonne.
- 44) - Chaque fois que mon père hurlait, sa colère devenait forte.
 - Chaque fois que mon père hurlait sa colère, sa voix devenait forte.
- 45) - Chaque fois que l'on partageait, la salade finissait vite.
 - Chaque fois que l'on partageait la salade, le dîner finissait vite.
- 46) - Chaque fois que la fille regardait, le film l'attirait pendant quelques heures.
 - Chaque fois que la fille regardait le film, l'histoire l'attirait pendant quelques heures.
- 47) - Chaque fois que Paul arrosait, l'arbre poussait vite.
 - Chaque fois que Paul arrosait l'arbre, les feuilles poussaient vite.
- 48) - Chaque fois que le spécialiste déchiffrait, le code devenait simple.

- Chaque fois que le spécialiste déchiffrait le code, le message devenait simple.
- 49) - Chaque fois que le comptable gérait, la dépense était dépassée.
 - Chaque fois que le comptable gérait la dépense, le budget était dépassé.
- 50) - Chaque fois que nous déménagions, la maison était désordonnée.
 - Chaque fois que nous déménagions la maison, ma chambre était désordonnée.
- 51) - Chaque fois que ma mère jouait du violon, ma sœur chantait la chanson.
- 52) - Chaque fois que j'allais à l'école, je prenais le bus 58.
- 53) - Chaque fois que mon père écoutait la radio, ma mère allumait la télévision.
- 54) - Chaque fois que les vacances s'approchaient, la famille allait à la mer.
- 55) - Chaque fois qu'il pleuvait, Paul oubliait son parapluie.
- 56) - Chaque fois que j'allais à la piscine, il y avait du monde.
- 57) - Chaque fois que mamie achetait du pain à la boulangerie, elle ne prenait que des croissants.
- 58) - Chaque fois que mon frère garait sa voiture, il heurtait d'autres véhicules.
- 59) - Chaque fois que le train entrait en gare, la voie était remplie de voyageurs.
- 60) - Chaque fois que les nuages se dispersaient, la pleine lune émergeait.
- 61) - Chaque fois qu'il neigeait pendant des jours, le trafic était perturbé.
- 62) - Chaque fois que mon copain revenait de France, il m'apportait des chocolats.
- 63) - Chaque fois que le chien aboyait, le locataire était réveillé.
- 64) - Chaque fois que le salaire augmentait, nous prenions un dîner au restaurant.
- 65) - Chaque fois que le président arrivait, l'équipe lui faisait un bon accueil.
- 66) - Chaque fois que les élèves bavardaient, le professeur arrêtait de parler.
- 67) - Chaque fois que l'eau bouillait, la femme de ménage entrait dans la cuisine.
- 68) - Chaque fois que les voisins parlaient, ma mère écoutait.
- 69) - Chaque fois que le virus circulait, on était mis en confinement.
- 70) - Chaque fois que l'on cliquait, le site ne s'ouvrait pas.

- 71) - Chaque fois que la réunion débutait, les participants applaudissaient.
- 72) - Chaque fois que les mannequins défilaient, les spectateurs prenaient des photos.
- 73) - Chaque fois que ma famille déjeunait, on prenait le dessert au jardin.
- 74) - Chaque fois que le premier ministre intervenait, le public manifestait.
- 75) - Chaque fois que mes parents se disputaient, je les conciliais.
- 76) - Chaque fois que je dormais, je me réveillais à minuit.
- 77) - Chaque fois que le concert durait des heures, je m'endormais.
- 78) - Chaque fois que la situation évoluait, on tenait une réunion.
- 79) - Chaque fois que la machine fonctionnait, elle émettait des bruits.
- 80) - Chaque fois que le sportif courait, il gagnait la course.
- 81) - Chaque fois que le patron insistait, l'employé lui obéissait.
- 82) - Chaque fois que la police intervenait, les voleurs s'enfuyaient.
- 83) - Chaque fois que ma voisine jardinait, ma mère tondait la pelouse.
- 84) - Chaque fois que le patient luttait contre la maladie, sa famille l'encourageait.
- 85) - Chaque fois que l'on marchait sur la plage, on entendait le bruit des vagues.
- 86) - Chaque fois que mon collègue partait pour Pékin, il prenait l'avion.
- 87) - Chaque fois que les clients patientaient, ils avaient l'air mécontent.
- 88) - Chaque fois que la thèse progressait, le directeur était satisfait.
- 89) - Chaque fois que je séjournais à Paris, il faisait très gris.
- 90) - Chaque fois que mes amis skiaient aux Alpes, ils se sentaient très détendus.
- 91) - Chaque fois que ma mère souriait, son visage devenait ravissant.
- 92) - Chaque fois que ma fille toussait dans la nuit, mon mari mettait le chauffage.
- 93) - Chaque fois que le maire venait au centre-ville, les habitants le conspuaient.
- 94) - Chaque fois que ses grands-parents voyageaient, ils lui apportaient des cadeaux.

- 95) - Chaque fois que je m'étais perdu dans la ville, je prenais un taxi.
- 96) - Chaque fois que le candidat baissait dans les sondages, ses adversaires étaient soulagés.
- 97) - Chaque fois que l'invité restait déjeuner, nous avions un repas copieux.
- 98) - Chaque fois que mon frère avait soif, il prenait une bière.
- 99) - Chaque fois que l'étudiant argumentait, son raisonnement était problématique.
- 100) - Chaque fois que les deux réunions coïncidaient, je n'assistais qu'à l'une d'entre elles.

4.6.4 Avis du Comité d'Éthique de la Recherche, Université Sorbonne Nouvelle (CER-USN-01-2023)



Comité d'Éthique de la Recherche Université Sorbonne Nouvelle Session 1, 2023

Lei XI
lei.xi@sorbonne-nouvelle.fr

Paris, le 30 janvier 2023

Cher collègue,

A la suite de la soumission de votre saisine au CER USN pour le projet intitulé « *Acquisition de la prosodie du français par des apprenants sinophones : étude électrophysiologique* » mené au sein de l'Université Sorbonne Nouvelle, et aux modifications que vous avez effectuées, le comité apporte

un avis FAVORABLE à la bonne conduite du projet.

Avis CER-USN-01-2023



La Présidente du CER - USN
Aliyah Morgenstern

4.6.5 Certificat du Service de Protection des Données du CNRS (2-22087)

Informations issues du registre UMR7018 / 2-22087



CNRS | Service Protection des
Données
2, rue Jean Zay | 54519 Vandoeuvre
03 83 85 64 26
dpd.demandes@cnrs.fr

La délivrance de ce **certificat** signifie que votre traitement a été porté au registre tenu par la DPD du CNRS et que votre dossier est formellement complet. Vous pouvez mettre en œuvre votre traitement.

En tout état de cause, vous êtes tenus de respecter les obligations prévues par la loi et notamment :

- 1 - la définition et le respect de la finalité du traitement.
- 2 - la pertinence des données traitées.
- 3 - la conservation pendant une durée limitée des données.
- 4 - la sécurité et la confidentialité des données.
- 5 - le respect des droits des intéressés : information sur leur droit d'accès, de modification et d'opposition.

Pour davantage d'informations, Veuillez consulter le site du [Service de Protection des Données](#).

Gaëlle BUJAN
Déléguée à la Protection des Données du CNRS

Numéro de traitement	2-22087
Enregistrement	31/03/2022
Responsable de traitement	<ul style="list-style-type: none">• UMR7018 - Laboratoire de Phonétique et Phonologie Cédric GENDROT - cedric.gendrot@sorbonne-nouvelle.fr 19 Rue des Bernardins ILPGA 75005 PARIS 05
Mise en oeuvre	<ul style="list-style-type: none">• UMR7018 - Laboratoire de Phonétique et Phonologie Lei Xi - 0659220397 - lei.xi@sorbonne-nouvelle.fr 19 Rue des Bernardins ILPGA 75005 PARIS 05
Finalité et objectif	Acquisition de la prosodie du français par des apprenants sinophones : études comportementale et électrophysiologique

CNRS | SPD | 31/03/2022

4.6.6 Signaux ERP sur les 42 électrodes dans les trois ROIs pour la comparaison entre *FrontièreCP* vs *NonFrontièreCP* et la comparaison entre *FrontièreCT* vs *NonFrontièreCT* dans chacun des trois groupes de participants

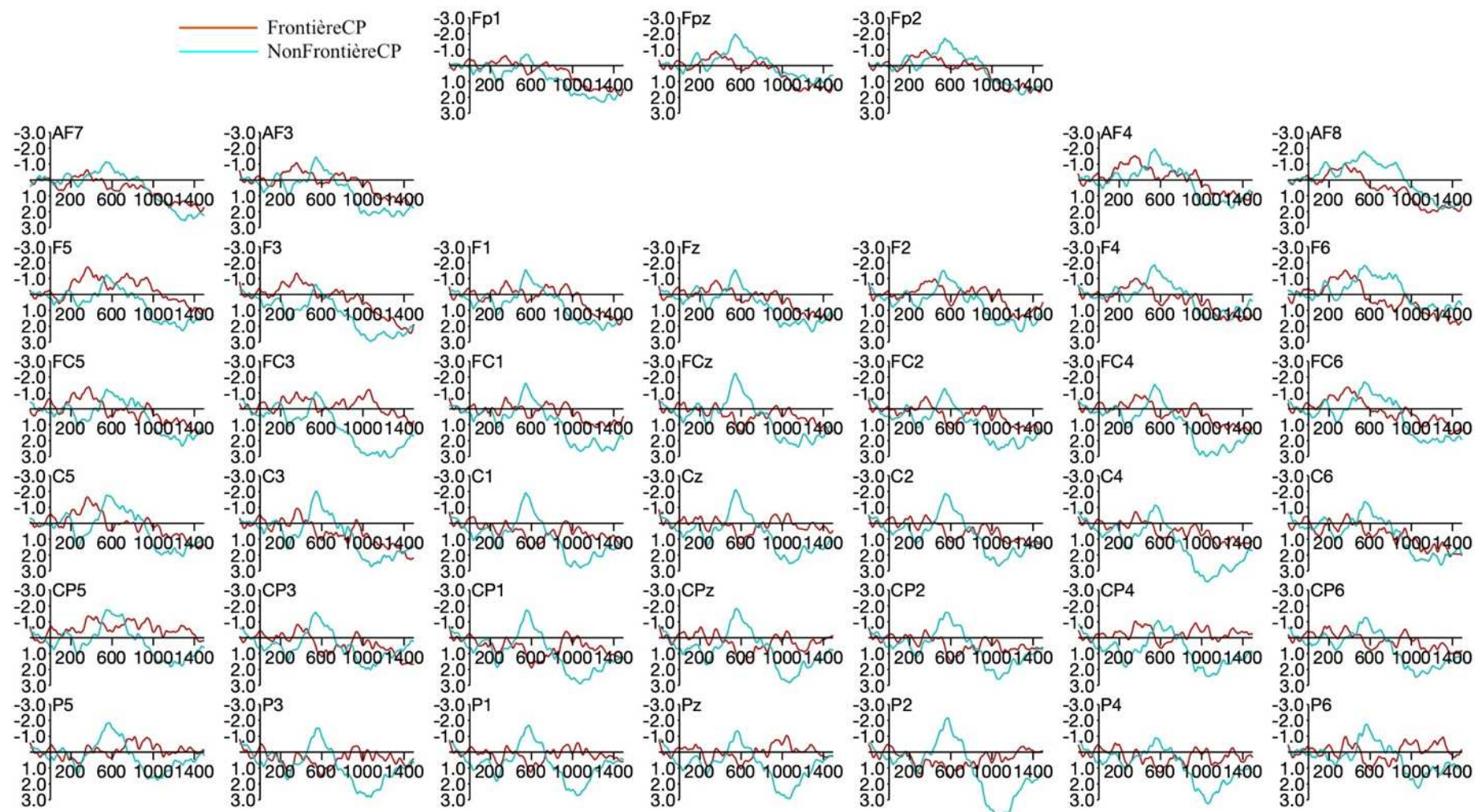


Figure 4.40 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les participants **francophones natifs**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

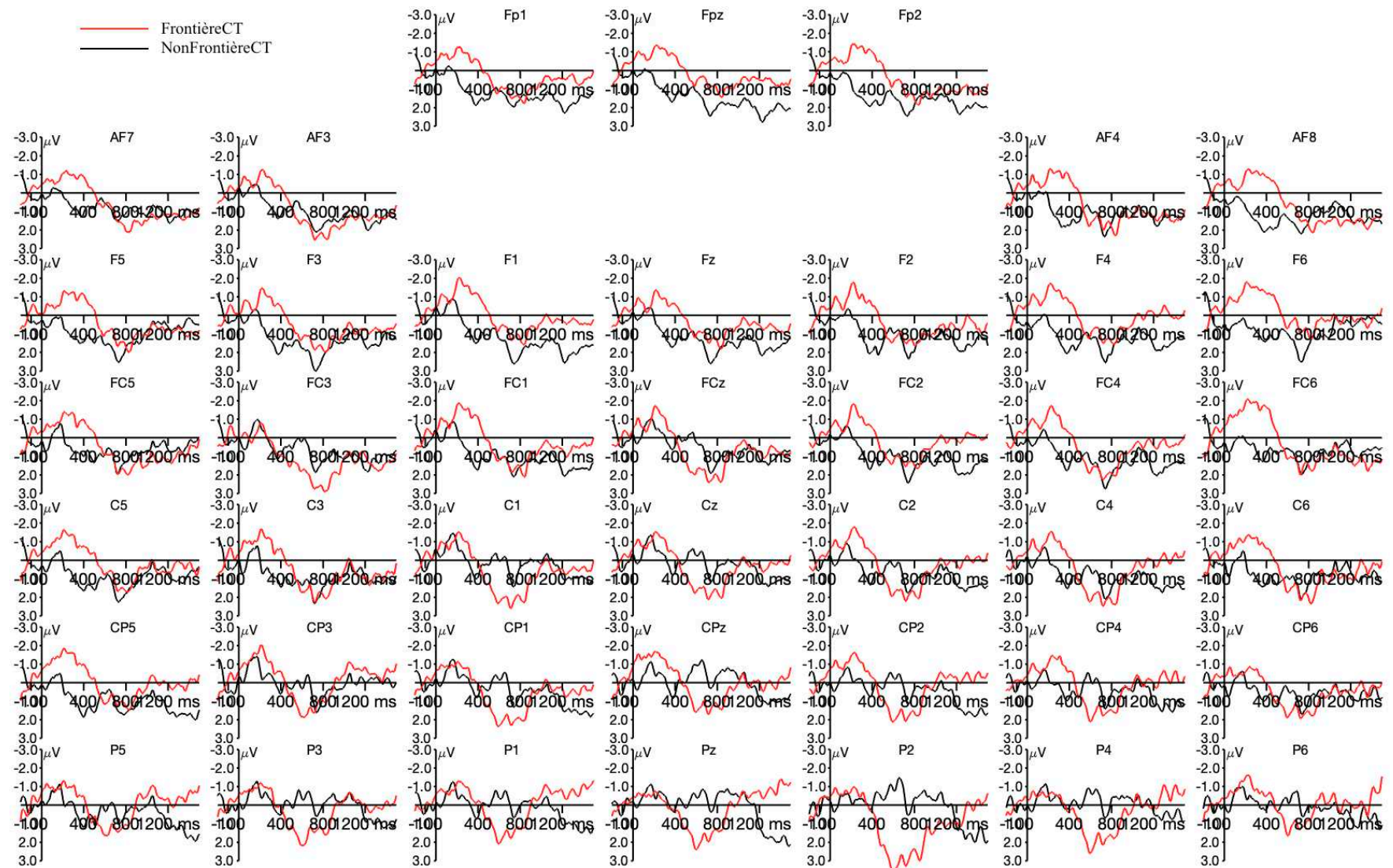


Figure 4.41 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques **FrontièreCT** vs **NonFrontièreCT** chez les participants **francophones natifs** (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

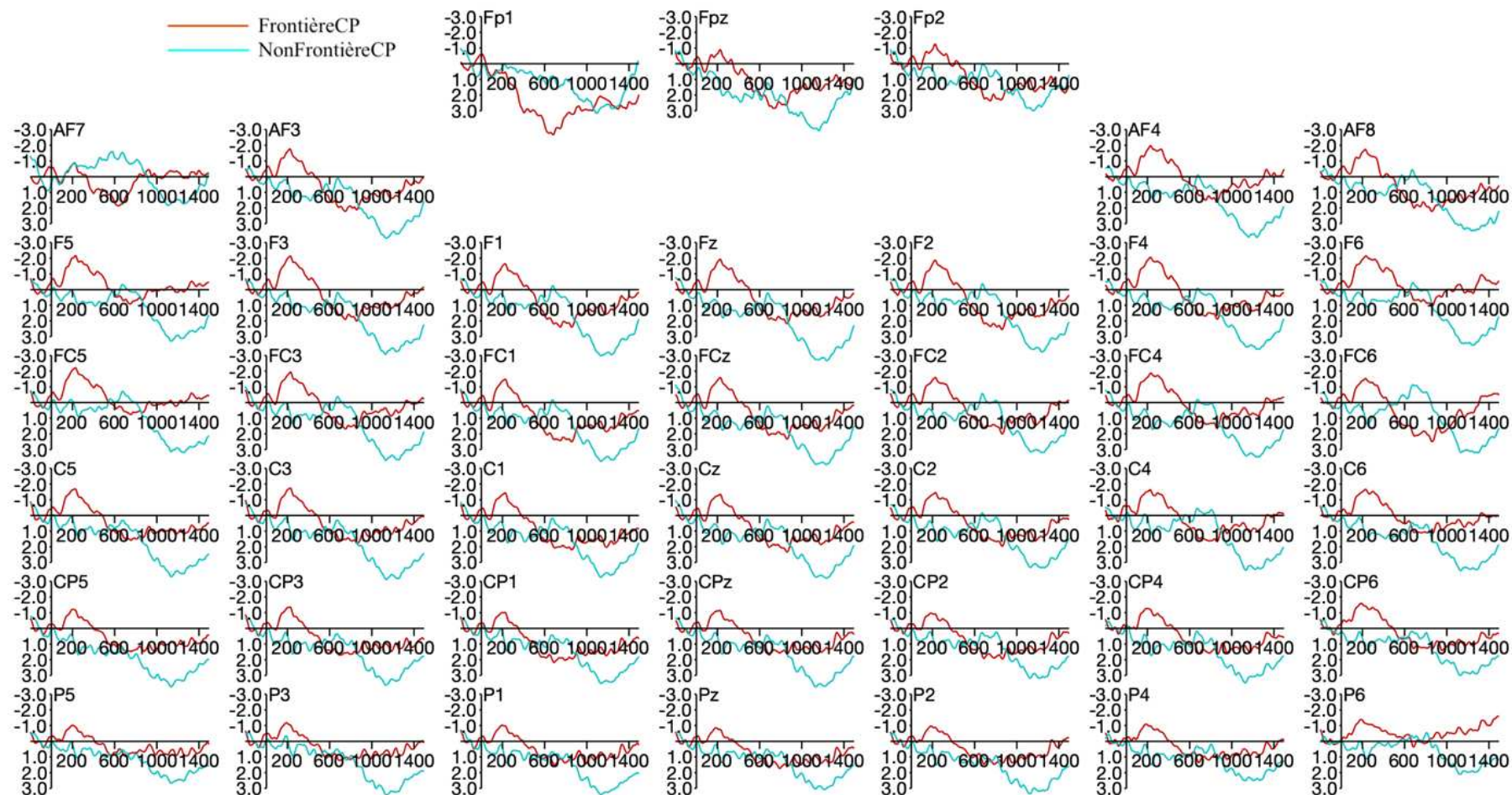


Figure 4.42 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants avancés**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

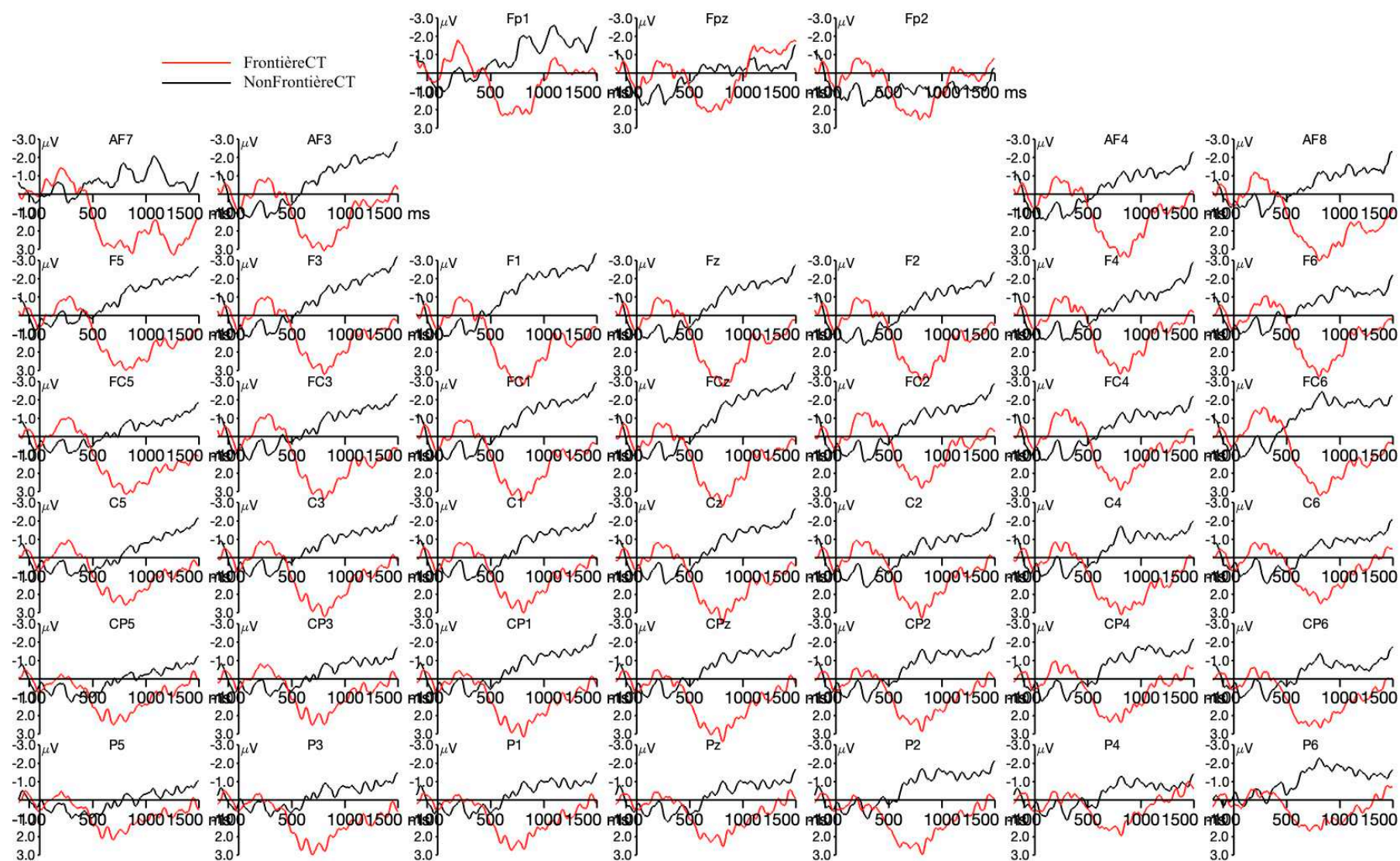


Figure 4.43 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques *FrontièreCT* vs *NonFrontièreCT* chez les **apprenants avancés**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

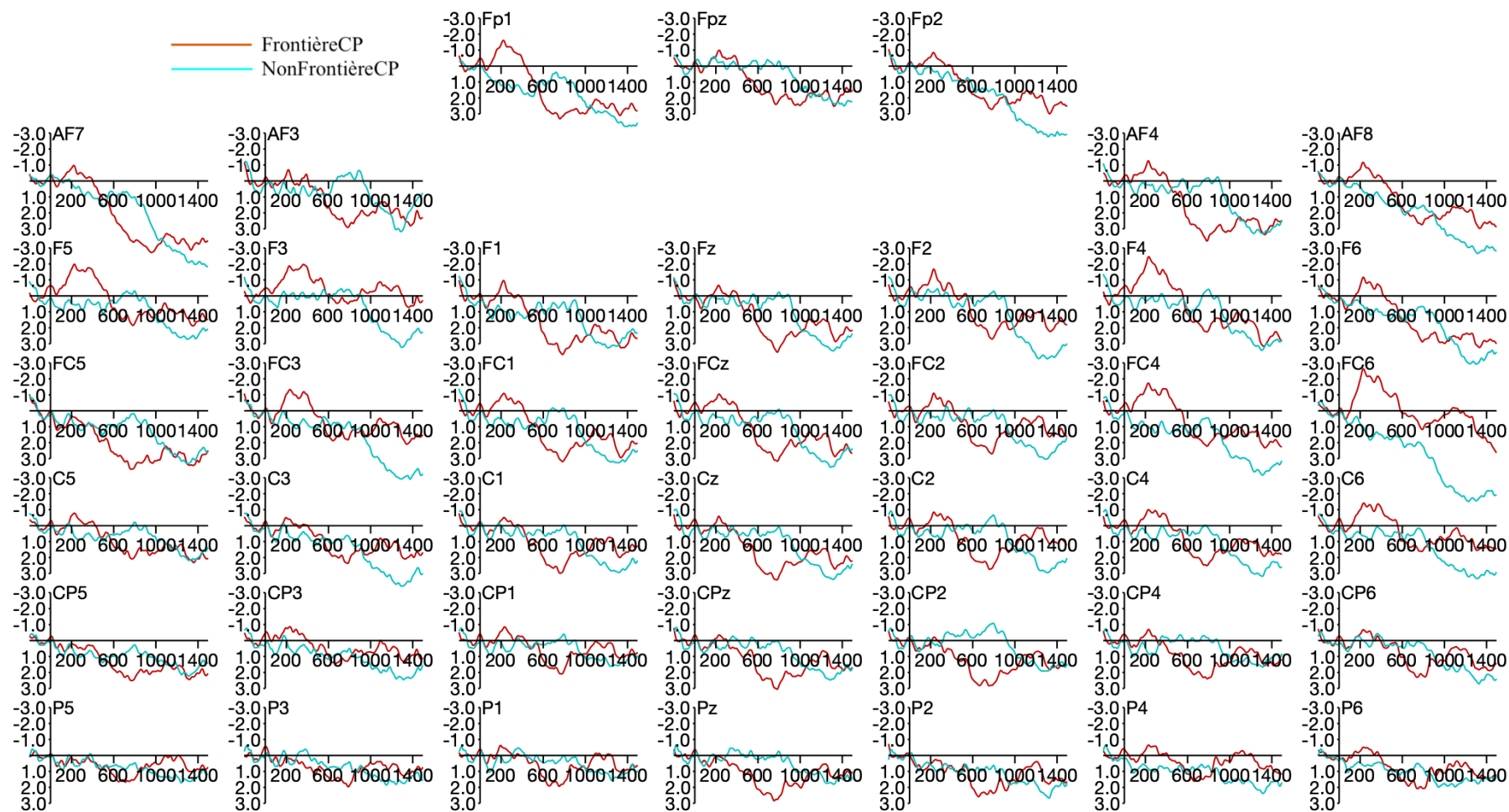


Figure 4.44 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques *FrontièreCP* vs *NonFrontièreCP* chez les **apprenants intermédiaires**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

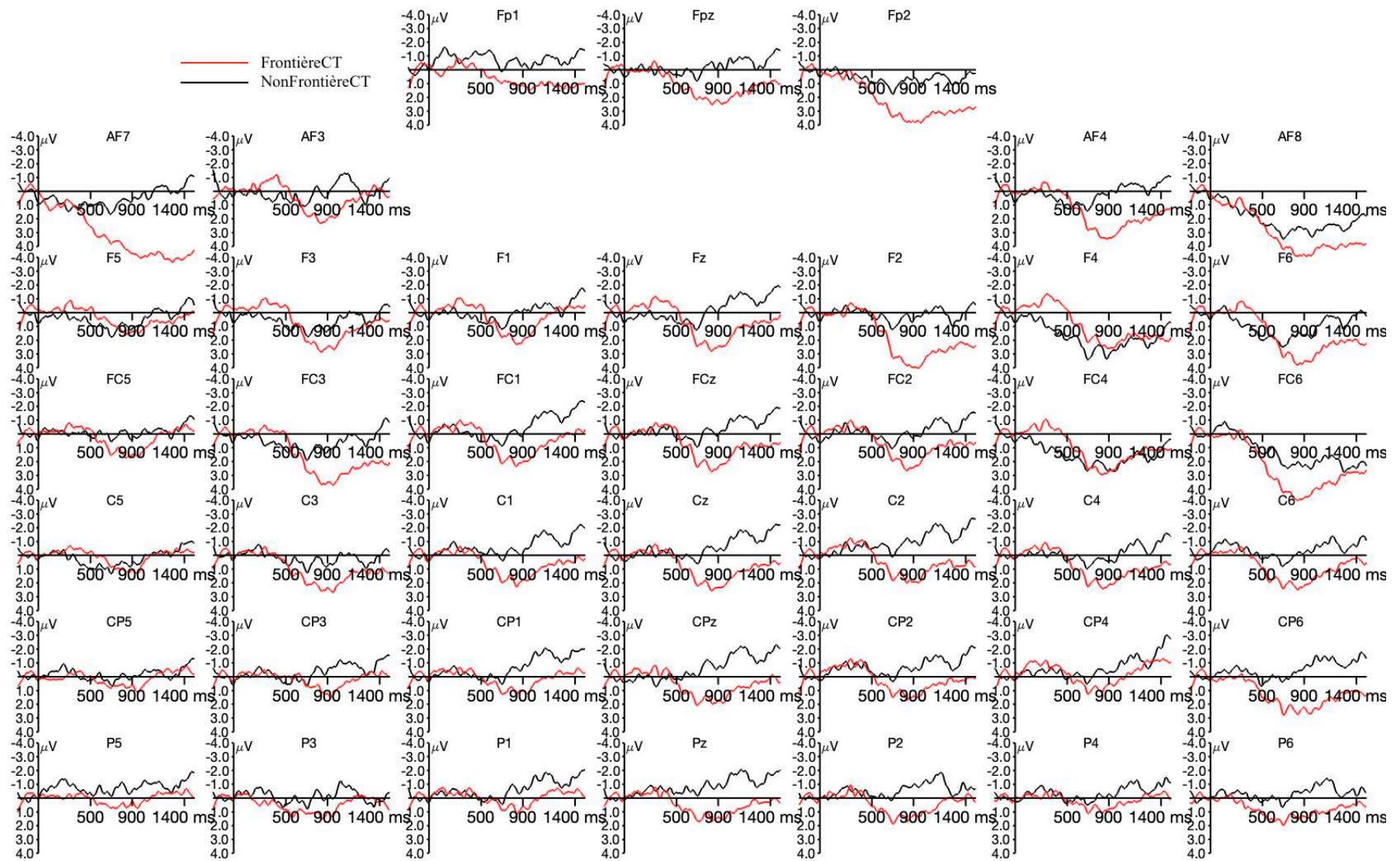


Figure 4.45 : Signaux CPS dans les 42 électrodes des trois ROIs pour les conditions prosodiques **FrontièreCT** vs **NonFrontièreCT** chez les **apprenants intermédiaires**. (Signaux filtrés avec un filtre bas de 10Hz pour une meilleure visualisation).

Chapitre 5 Synthèse et perspectives

Cette thèse s'intéresse à l'organisation et à la segmentation prosodiques du français chez deux groupes distincts : les locuteurs francophones natifs et les apprenants sinophones adultes répartis en trois niveaux (débutant, intermédiaire et avancé). Elle repose sur trois approches expérimentales complémentaires : acoustique, perceptive et électroencéphalographique. À travers une série de paires de phrases présentant deux structures prosodiques ambiguës – la clôture précoce (CP) (e.g., : « Chaque fois que le serpent mangeait, le rat se cachait. ») et la clôture tardive (CT) (e.g., : « Chaque fois que le serpent mangeait le rat, le lapin se cachait. ») –, cette recherche a exploré les questions suivantes :

- Sur le plan de la segmentation, quels indices acoustiques les locuteurs natifs et les apprenants mobilisent-ils pour signaler les frontières prosodiques ?
- Concernant l'organisation prosodique, comment les patrons rythmiques se traduisent-ils à travers des paramètres quantitatifs liés au rythme ?
- En perception, à partir des indices prosodiques disponibles, comment les auditeurs natifs et les apprenants traitent-ils, de manière comportementale, les phrases localement ambiguës ?
- Enfin, sur le plan neurocognitif, comment les locuteurs natifs et les apprenants sinophones traitent-ils ces phrases ambiguës ?

Dans ce dernier chapitre, nous ne reviendrons pas en détail sur la discussion des résultats spécifiques de chacune des trois expériences, ces discussions ayant été effectuées dans les chapitres précédents. À la place, nous proposons de synthétiser les résultats de chaque expérience, en mettant en avant les contributions et les complémentarités de ces trois approches expérimentales. Cette perspective nous permettra de mieux cerner les caractéristiques liées à l'organisation et à la segmentation prosodiques du français, tout en analysant comment ces caractéristiques évoluent chez les apprenants en fonction de leur niveau de compétence linguistique.

Ce chapitre est organisé en deux sections. La section 5.1 propose une synthèse des principaux résultats et des contributions de cette thèse, en mettant particulièrement l'accent sur la partie EEG, que nous considérons comme la plus novatrice. La section

5.2 aborde une discussion des limites méthodologiques identifiées et ouvre sur de nouvelles perspectives pour les recherches futures.

5.1 Synthèse des résultats et contributions principaux

Un des premiers objectifs de cette thèse a été de combler une lacune importante dans la recherche sur l'acquisition de la prosodie d'une langue seconde (L2), un sujet encore largement négligé dans les études sur l'apprentissage des systèmes sonores en L2, et plus particulièrement dans le cas du français par des sinophones. Pour répondre à cet objectif, trois études complémentaires ont été menées : une étude acoustique, une étude perceptive et une étude électroencéphalographique.

Dans l'étude acoustique, nous avons examiné l'organisation et la segmentation prosodiques du français en nous concentrant sur les indices mélodiques et temporels utilisés par les locuteurs natifs et les apprenants. Concernant la segmentation prosodique, les analyses acoustiques montrent que les natifs, ainsi que les apprenants intermédiaires et avancés, sont capables d'utiliser simultanément la montée de F0 et l'allongement final pour signaler les frontières prosodiques ambiguës. En revanche, les apprenants débutants ne parviennent pas encore à mobiliser ces indices, ce qui suggère que la segmentation prosodique représente une difficulté particulière à ce niveau d'apprentissage. Sur le plan de l'organisation prosodique, les paramètres quantitatifs rythmiques se sont révélés être des indicateurs pertinents pour étudier l'acquisition rythmique en L2. Nos analyses ont montré que, à mesure que le niveau de maîtrise augmente, la production rythmique des apprenants devient plus stable et tend vers une isochronie syllabique comparable à celle des natifs, particulièrement après la normalisation du débit. Ces résultats mettent en évidence l'importance du débit de parole dans l'acquisition de la prosodie du français par des sinophones. D'abord, un débit variable dans la production prosodique entraîne un patron rythmique instable, caractérisé par les valeurs élevées pour les paramètres non normalisés. Ensuite, un débit lent sur l'ensemble de la production d'un énoncé chez les apprenants débutants réduit la proéminence temporelle, notamment l'allongement final.

Dans l'étude de perception, nous avons utilisé aussi des stimuli localement ambigus (CP et CT), mais avec des stimuli tronqués. L'objectif est d'évaluer si les participants pouvaient désambigüiser les structures syntaxiques en s'appuyant uniquement sur

des indices prosodiques. Les résultats montrent que, si les conditions CT sont traitées sans difficulté, la perception des structures CP reste problématique, même pour les locuteurs natifs. Ces résultats montrent que, bien que les trois indices acoustiques (allongement final, montée de F0 et pause) soient présents, ils ne suffisent pas à permettre la perception d'une frontière CP dans des stimuli tronqués. Nous avons interprété ce phénomène, dans notre discussion de ces résultats, à travers l'hypothèse de l'*Informative Boundary Hypothesis*, selon laquelle les indices acoustiques atteignent leur efficacité maximale uniquement dans un contexte complet. En l'absence de ce contexte, les auditeurs privilégient une stratégie par défaut, attachant les nouveaux éléments à la clause en cours tant que celle-ci reste grammaticalement valide (*Late Closure Preference*).

Dans l'étude EEG, nous avons exploré le traitement neurocognitif des structures prosodiques ambiguës dans des contextes prosodiques complets. Pour la première fois, nous avons répliqué la composante CPS en français métropolitain. Les analyses montrent que les CPS sont déclenchées dans les deux conditions prosodiques (CP et CT) chez les locuteurs natifs, mais avec des différences de distribution topographique plus localisés, se concentrant essentiellement au niveau centro-frontal, chez les apprenants avancés et intermédiaires (Tableau 5.1).

Tableau 5.1 : Comparaison des distributions topographiques des composantes CPS en fonction de conditions prosodiques et de niveaux de français.

Niveau de français	Condition prosodique	Distribution topographique
Francophones natifs	CP	Sur l'ensemble des électrodes, avec une concentration au niveau centro-pariétal.
	CT	Au niveau centro-pariétal.
Apprenants avancés	CP	Au niveau centro-frontal.
	CT	Sur l'ensemble des électrodes.
Apprenants intermédiaires	CP	Sur l'ensemble des électrodes, avec une concentration au niveau centro-frontal.
	CT	Au niveau centro-frontal.

Ces résultats sont conformes à nos hypothèses. Ils confirment l'importance d'un contexte prosodique complet pour le traitement des frontières prosodiques et montrent que les indices acoustiques des structures CP et CT atteignent leur plein effet dans de tels contextes (*Informative Boundary Hypothesis*).

L'importance d'un contexte prosodique complet (*Informative Boundary Hypothesis*) et la préférence pour une clôture tardive (*Late Closure Preference*) dans la perception

de la parole ont aussi été mises en évidence sur le plan neurophysiologique par l'étude de Holzgrefe *et al.* (2013). Dans cette étude, les auteurs ont utilisé des stimuli très courts en allemand, comprenant deux conditions prosodiques, CP et CT, pour examiner le traitement des frontières prosodiques à travers la composante CPS. Les deux conditions étaient les suivantes :

(a) Condition CP : [Mona]IP # [oder Lena und Lola]IP (« Mona # ou Lena et Lola »)

(b) Condition CT : [Mona oder Lena]IP # [und Lola]IP (« Mona ou Lena # et Lola »)

Les résultats ont montré que la CPS était déclenchée uniquement dans la condition CT (b), tandis qu'aucune CPS n'était observée dans la condition CP (a). Selon nous, ces stimuli courts présentent des similitudes avec les stimuli tronqués utilisés dans notre étude perceptive. En effet, l'absence d'un contexte prosodique complet avant et après la deuxième IP a empêché un traitement neurocognitif efficace des informations acoustiques, bien qu'elles soient présentes. Dans leur analyse, Holzgrefe *et al.* (2013) ont également souligné l'importance d'un contexte prosodique suffisant et complet pour le traitement des frontières prosodiques. Ces conclusions renforcent l'idée que, sans un cadre prosodique global, les indices acoustiques ne peuvent pas être pleinement exploités par le système cognitif pour percevoir et traiter les frontières prosodiques de manière optimale.

Dans notre étude électroencéphalographique, nous avons exploité les avantages offerts par la technique EEG dans le domaine de la recherche sur l'acquisition langagière (cf. section 1.5.1). Il est important de souligner à nouveau l'un des principaux atouts de cette méthode : l'EEG permet de révéler des aspects essentiels que les méthodes comportementales ne peuvent pas détecter. Comme l'affirme Mueller (2005 : 153), « ...*differences in behavioural measures, such as reaction times, are not necessarily the result of the involvement of different neuronal structures, even if they show qualitatively different patterns. Neurophysiological measures can add valuable information about timing and degree of activation of neural networks and therefore provide physiological correlates of behavioural differences.* ». C'est précisément ce qui s'est vérifié dans nos études perceptive et EEG. Les analyses des données CPS complètent les résultats obtenus dans l'étude perceptive et apportent de nouvelles interprétations concernant le traitement des phrases ambiguës CP et CT, comme l'illustrent les deux points suivants.

Premièrement, les comparaisons des composantes CPS dans une même condition prosodique entre les trois groupes (natifs, apprenants avancés et intermédiaires) concordent avec les résultats de l'étude perceptive. Les locuteurs natifs et les apprenants ont montré un traitement neurocognitif comparable pour la condition CT. En revanche, pour la condition CP, les amplitudes moyennes des CPS étaient similaires d'une part entre les groupes natifs et avancés, et d'autre part entre les groupes avancés et intermédiaires. Ces résultats, en alignement avec ceux de l'étude perceptive, mettent en évidence l'acquisition progressive d'une structure prosodique complexe comme CP, mais cette fois d'un point de vue neurocognitif, avec des mesures quantitatives précises (notamment les amplitudes moyennes des CPS).

Deuxièmement, dans l'étude perceptive utilisant une tâche de complétion, les francophones natifs et les apprenants ont semblé traiter la condition CT de manière comparable. Cependant, les données ERP (potentiels évoqués) révèlent que ce traitement est plus complexe et interactif (de haut en bas) qu'un traitement purement prosodique basé uniquement sur des indices acoustiques (de bas en haut). Inspirés par les études de Steinhauer (2003), Pauker *et al.* (2011), Bögels *et al.* (2013), ainsi que Nickels et Steinhauer (2018), nous avons identifié un effet négatif, dit « *ramp-like negativity* », accompagné des CPS dans la condition CT. Cette négativité reflète une anticipation cognitive des auditeurs pour une frontière tardive lorsque la frontière précoce est absente (voir section 4.5 pour plus de détails). Cet effet est le plus manifeste chez les participants natifs, moins prononcé chez les apprenants avancés, et pratiquement absent chez les intermédiaires. En d'autres termes, le traitement de la condition CT chez les natifs s'appuie sur un travail cognitif continu, interactif et global. Lors de la perception de la parole, les informations prosodiques et syntaxiques, qui sont intrinsèquement liées, interagissent et se renforcent mutuellement pour intégrer les frontières prosodiques et les relations syntaxiques. En revanche, les apprenants avancés montrent une compétence plus limitée dans ce type de traitement descendant (de haut en bas), et les intermédiaires le sont encore moins. Pour ces derniers, la CPS est déclenchée par un traitement essentiellement local et restreint, reposant uniquement sur les indices acoustiques aux frontières prosodiques CT. Ils parviennent à reconnaître la condition CT dans

l'étude perceptive en raison de la nature favorable de cette structure, qui correspond au traitement initial et par défaut.

En synthèse, cette thèse apporte plusieurs contributions originales à la recherche sur la phonétique, la prosodie et l'acquisition d'une L2. Premièrement, elle fournit une description détaillée de l'acquisition de la prosodie du français chez des apprenants sinophones à trois niveaux distincts, couvrant ainsi presque toutes les étapes du processus d'apprentissage. Elle met en évidence des difficultés spécifiques à chaque niveau. Nous avons ainsi démontré qu'à partir du niveau intermédiaire, les apprenants étaient capables de signaler acoustiquement les frontières prosodiques en utilisant des indices mélodiques et temporels. Cependant, même au niveau avancé, ils continuaient à rencontrer des difficultés sur le plan perceptif pour percevoir et distinguer les frontières prosodiques ambiguës. Aussi, des différences significatives ont été mises en évidence sur le plan neurocognitif : les francophones natifs traitent et intègrent plus promptement les indices acoustiques de la condition CP (à partir de 455 ms après la dernière syllabe) par rapport aux apprenants avancés, qui les traitent à partir de 610 ms après la dernière syllabe. À notre connaissance, il s'agit de la première étude à examiner l'acquisition de la prosodie en français par des sinophones couvrant les trois niveaux d'apprentissage, contrairement aux travaux antérieurs qui se limitaient à un ou deux niveaux.

Deuxièmement, notre étude EEG réplique pour la première fois en français métropolitain la composante CPS chez les natifs et exploré son utilisation dans le cadre de l'acquisition de la prosodie en L2. Cette composante, rarement exploitée dans les recherches sur la prosodie, a permis de révéler des mécanismes neurocognitifs que les seules méthodes comportementales ne pouvaient détecter. Les résultats montrent l'utilité de l'EEG pour approfondir la compréhension des processus prosodiques dans l'acquisition langagière.

Enfin, contrairement aux travaux antérieurs souvent fragmentés et limités à un seul aspect prosodique (mélodique ou temporel), nous avons adopté une approche multimodale, intégrant des données acoustiques, perceptives et électroencéphalographiques. Cette approche globale, combinée à l'analyse des aspects mélodiques et temporels et à un échantillon conséquent de participants, nous a permis d'obtenir une vision approfondie et complète des processus d'acquisition de la prosodie du français par des apprenants sinophones.

5.2 Limites des méthodologies et perspectives

Cette section aborde les limites des méthodologies adoptées dans les expériences de cette thèse et des pistes pour y remédier dans des recherches futures. Nous commencerons par discuter les limites liées à l'étude électroencéphalographique, suivies de celles concernant l'étude acoustique, avant de conclure par des pistes d'amélioration concernant les groupes de participants.

Dans l'étude électroencéphalographique (EEG), la principale limite réside dans l'absence d'analyse approfondie sur l'interaction entre la prosodie (CPS) et la syntaxe (P600). En particulier, nous n'avons pas pu examiner comment la transitivité des verbes dans les stimuli pouvait influencer l'intégration des informations syntaxico-prosodiques et l'interaction éventuelle entre ces deux aspects. Plusieurs études antérieures ont exploré l'effet de la prosodie sur le traitement syntaxique, ou inversement, l'effet de la syntaxe sur le traitement prosodique, ainsi que leur interaction dans la parole (par exemple, Cole *et al.*, 2010 ; Eckstein et Friederici, 2006 ; Kerkhofs *et al.*, 2007). Concernant spécifiquement la composante CPS, des travaux comme ceux d'Itzhak *et al.* (2010) et de Nickels et Steinhauer (2018) ont analysé non seulement la CPS aux frontières prosodiques, mais également la P600 dans des phrases de type « *garden path* », en étudiant comment la transitivité des verbes et la structure syntaxique modulaient ces deux composantes.

Dans le traitement de la parole, la prosodie, y compris les frontières prosodiques, n'agit jamais de manière isolée, mais interagit constamment avec d'autres éléments linguistiques. Dans notre étude EEG (ainsi que dans l'étude perceptive), nous avons envisagé de mener des analyses en nous inspirant des travaux d'Itzhak *et al.* (2010) et de Nickels et Steinhauer (2018), afin de comparer les groupes de locuteurs natifs et d'apprenants. Une telle comparaison aurait permis de tirer des conclusions plus approfondies, non seulement sur l'acquisition de la prosodie en elle-même, mais aussi sur la capacité des apprenants à intégrer plusieurs éléments linguistiques dans le traitement de la parole. Cependant, comme mentionné dans la discussion du Chapitre 3 (cf. sections 3.4), l'absence d'indicateurs appropriés pour la transitivité des verbes et d'une répartition équilibrée des catégories verbales a empêché ces analyses d'aboutir. De plus, certaines limites spécifiques au protocole expérimental de l'EEG ont entravé cette analyse.

Premièrement, nous n'avons pu examiner cette question qu'à partir des stimuli enregistrés et des déclencheurs qui avaient été placés au préalable. Dans nos stimuli, aucun déclencheur n'avait été inséré à une position optimale permettant d'extraire efficacement la composante P600. Les déclencheurs actuels étaient positionnés soit sur l'onset de la dernière syllabe du syntagme verbal, soit sur celui du syntagme nominal, tous deux situés avant une frontière prosodique. Or, la CPS et la P600 sont des composantes ERP positives, caractérisées par une durée relativement longue. Cela rend difficile leur distinction à partir des déclencheurs actuellement utilisés dans les signaux. En revisitant des études ayant analysé la CPS et la P600, celle de Nickels et Steinhauer (2018) constitue un exemple pertinent. Dans leur recherche, qui portait également sur des stimuli avec des conditions prosodiques CP et CT, les auteurs avaient positionné les déclencheurs sur l'onset du dernier syntagme verbal (ligne rouge dans le Tableau 5.2), c'est-à-dire après toute frontière intonative. Cette position stratégique leur a permis d'étudier la composante P600 tout en évitant son chevauchement avec la CPS.

Tableau 5.2 : Position du déclencheur (ligne rouge) pour extraire la composante P600 dans Nickels et Steinhauer (2018).

Condition	Part 1	Splicing point	Part 2
A (LC)	<i>When a bear is approaching</i>	<i>the people # the dogs</i>	<i>come running</i>
B (EC)	<i>When a bear is approaching</i>	<i># the people</i>	<i>come running</i>

Ainsi, pour mener à bien cette analyse que nous n'avons pas pu réaliser, il aurait été nécessaire, lors de la conception de notre étude, d'ajouter des déclencheurs à des positions spécifiques, comme l'onset du dernier syntagme verbal après la frontière prosodique (ligne rouge, Tableau 5.3). Cela aurait permis d'examiner la P600 tout en évitant toute confusion avec la CPS, conformément à l'approche adoptée par Nickels et Steinhauer (2018).

Tableau 5.3 : Positions des déclencheurs à insérer sur l'onset du dernier syntagme verbal (positions indiquées par la ligne pointillée rouge) dans nos stimuli pour étudier le traitement entre la prosodie (CPS) et la syntaxe (P600) (déclencheurs actuels : lignes pointillées noires).

CP :	Chaque fois que le serpent mang	geait # le	rat	se cachait.	
CT :	Chaque fois que le serpent mang	geait	le	rat # le lapin	se cachait.

Deuxièmement, nos stimuli ne comprenaient que des phrases où les informations syntaxiques et prosodiques étaient congruentes. Pour analyser l'interaction syntaxico-prosodique, il aurait été pertinent d'ajouter des stimuli incongruents, comme des phrases de type « *garden path* », afin d'examiner comment les participants traitent des conflits entre la prosodie et la syntaxe. Malgré ces limites, nos résultats EEG ont répondu à nos objectifs initiaux concernant les effets CPS chez les francophones natifs et les apprenants sinophones. Ils ouvrent également de nouvelles perspectives pour explorer plus en détail l'interaction entre la prosodie et la syntaxe dans de futures recherches.

Concernant l'étude acoustique, une limite importante réside dans la taille réduite du corpus. Dans l'expérience de production de la parole, seuls cinq paires de phrases localement ambiguës ont été produites par les locuteurs. Un corpus plus étendu, comprenant idéalement 10 à 20 paires de phrases ambiguës, aurait permis de recueillir davantage de données acoustiques pour des analyses plus robustes et des comparaisons plus approfondies entre les groupes de participants.

Enfin, une limite notable concerne les groupes de participants dans les trois expériences. Dans l'étude acoustique, nous avons inclus des francophones natifs et des apprenants de trois niveaux (débutant, intermédiaire et avancé). Cependant, dans les études perceptive et EEG, seuls les natifs et les apprenants intermédiaires et avancés ont été inclus. Cette restriction ne s'explique pas uniquement par la difficulté de recruter un nombre suffisant d'apprenants débutants correspondant à

la tranche d'âge ciblée dans le milieu universitaire en France. Cette restriction s'explique aussi par la faible compétence en compréhension orale des apprenants débutants, notamment face à des phrases tronquées. Même si cela aurait pu limiter la pertinence des réponses dans une tâche de complétion, il aurait été intéressant d'examiner la composante CPS chez ce groupe. En effet, l'un des avantages majeurs de la technique EEG est qu'elle permet de tester des participants qui ne possèdent pas encore les compétences comportementales nécessaires, comme les apprenants débutants. Recueillir des données CPS chez ces apprenants offrirait une meilleure compréhension du parcours d'acquisition de la prosodie en français et de son rôle dans la désambiguïsation syntaxique chez les sinophones.

Pour conclure, mieux comprendre comment la prosodie contribue à l'encodage et au décodage des informations linguistiques globales implique d'examiner son interaction avec plusieurs dimensions langagières fondamentales. La dimension syntaxique joue clairement un rôle primordial, étant donnée l'interaction étroite entre prosodie et structure syntaxique pour faciliter l'analyse et la résolution des ambiguïtés, comme en témoignent les études neurophysiologiques rapportées dans cette thèse. La dimension sémantique est également cruciale car la prosodie module la perception des incongruités sémantiques et guide la compréhension en soulignant les éléments linguistiques pertinents. Enfin, la dimension interlinguistique révèle que ces interactions varient considérablement selon les langues, notamment dans les langues tonales comme le mandarin, où les variations de hauteur transmettent simultanément des informations pertinentes sur les plans lexicale et prosodique. Cette approche multidimensionnelle est donc indispensable pour développer une compréhension approfondie des mécanismes par lesquels la prosodie facilite le traitement global du langage, permettant ainsi d'améliorer tant les modèles théoriques que les applications pratiques dans l'enseignement des langues et les technologies vocales.

Bibliographie

- Abboub, N., Nazzi, T., & Gervain, J. (2016). Prosodic grouping at birth. *Brain and Language*, 162, 46-59.
- Abercrombie, D. (1965). *Studies in phonetics and linguistics*. Oxford University Press.
- Abercrombie, D. (1967). *Elements of general phonetics*. Edinburgh University Press.
- Aguilera, M., El Yagoubi, R., Espesser, R., & Astésano, C. (2014). Event-Related Potential Investigation of Initial Accent Processing in French. *Proc. Speech Prosody 2014*, 383-387.
- Allbritton, D., McKoon, G., & Ratcliff, R. (1996). Reliability of prosodic cues for resolving syntactic ambiguity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22(3), 714-735.
- Allen, G. (1975). Speech rhythm : Its relation to performance universals and articulatory timing. *Journal of Phonetics*, 3(2), 75-86.
- Altmann, G., van Nice, K., Garnham, A., & Henstra, J. (1998). Late Closure in Context. *Journal of Memory and Language*, 38(4), 459-484.
- Angoujard, J. (1997). *Théorie de la syllabe : Rythme et qualité*. CNRS ÉDITIONS.
- Ardal, S., Donald, M., Meuter, R., Muldrew, S., & Luce, M. (1990). Brain responses to semantic incongruity in bilinguals. *Brain and Language*, 39(2), 187-205.
- Arvaniti, A. (2009). Rhythm, Timing and the Timing of Rhythm. *Phonetica*, 66(1-2), 46-63.
- Arvaniti, A. (2012). The usefulness of metrics in the quantification of speech rhythm. *Journal of Phonetics*, 40(3), 351-373.
- Arvaniti, A., & Rodriquez, T. (2013). The role of rhythm class, speaking rate, and F0 in language discrimination. *Laboratory Phonology*, 4(1), 7-38.
- Arvaniti, A., Ross, T., & Ferjan, N. (2008). On the reliability of rhythm metrics. *Journal of the Acoustical Society of America*, 124(4_Supplement), 2495.
- Astésano, C. (2016). 5. Prosodic characteristics of Reference French. In S. Detey, J. Durand, B. Laks, & C. Lyche (Éds.), *Varieties of Spoken French* (p. 68-85). Oxford University Press.
- Astésano, C., Besson, M., & Alter, K. (2004). Brain potentials during semantic and prosodic processing in French. *Cognitive Brain Research*, 18(2), 172-184.
- Asu, E., & Nolan, F. (2005). Estonian rhythm and the Pairwise Variability Index. *Proceedings of Fonetik 2005*, 29-32.
- Barry, W., Andreeva, B., Russo, M., Dimitrova, S., & Kostadinova, T. (2003). Do Rhythm Measures Tell us Anything about Language Type? In M. Solé, D. Recasens, & J. Romero (Éds.), *Proceedings of the 15th International Congress of Phonetic Sciences* (p. 2693-2696). Causal Productions Pty Ltd.
- Barry, W., Andreeva, B., & Koreman, J. (2009). Do Rhythm Measures Reflect Perceived Rhythm? *Phonetica*, 66(1-2), 78-94.
- Barry, W., & Russo, M. (2003). Measuring rhythm. Is it separable from speech rate? In A. Mettouchi & G. Ferré (Éds.), *Actes du congrès international AAL: Prosodic Interfaces* (p. 15-20).
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- Batista, A., Catronas, D., Folia, V., & Silva, S. (2023). Increased Pre-Boundary Lengthening Does Not Enhance Implicit Intonational Phrase Perception in European Portuguese : An EEG Study. *Brain Sciences*, 13(3), 441.

- Beach, C. (1991). The interpretation of prosodic patterns at points of syntactic structure ambiguity : Evidence for cue trading relations. *Journal of Memory and Language*, 30(6), 644-663.
- Beckman, M. (1996). The Parsing of Prosody. *Language and Cognitive Processes*, 11(1-2), 17-68.
- Beckman, M., & Pierrehumbert, J. (1986). Intonational structure in Japanese and English. *Phonology Yearbook*, 3, 255-309.
- Bergmann, A., Armstrong, M., & Maday, K. (2008). Relative clause attachment in English and Spanish : A production study. *Proc. Speech Prosody 2008*, 505-508.
- Bertinetto, P. (1989). Reflections on the dichotomy « stress » vs. « syllable-timing ». *Revue de phonétique appliquée*, 91(93), 99-130.
- Bertinetto, P., & Bertini, C. (2008). On modeling the rhythm of natural languages. *Proc. Speech Prosody 2008*, 427-430.
- Bertinetto, P., & Bertini, C. (2010). Towards a unified predictive model of natural language rhythm. In *Prosodic Universals. Comparative studies in rhythmic modeling and rhythm typology* (p. 43-78). Aracne.
- Bertinetto, P., Bertini, C., & Zhi, N. (2012). Rhythm in Mandarin Chinese and Italian : The role of sentence accents. *Proc. Speech Prosody 2012*, 520-523.
- Bigi, B. (2015). SPPAS - Multi-lingual Approaches to the Automatic Annotation of Speech. *The Phonetician. Journal of the International Society of Phonetic Sciences*, 111-112, 54-69.
- Bigi, B. (2016). A Phonetization Approach for the Forced-Alignment Task in SPPAS. In Z. Vetulani, H. Uszkoreit, & M. Kubis (Éds.), *Human Language Technology. Challenges for Computer Science and Linguistics* (Vol. 9561, p. 397-410). Springer International Publishing.
- Bigi, B., & Meunier, C. (2018). Automatic Segmentation of Spontaneous Speech / Segmentação automática da fala espontânea. *Revista de Estudos da Linguagem*, 26(4), 1489-1530.
- Billières, M. (2002). Chapitre 1. Le corps en phonétique corrective. In R. Renard (Ed.), *Apprentissage d'une langue étrangère/seconde 2. La phonétique verbo-tonale* (p. 35-70). De Boeck Supérieur.
- Bloch, B. (1950). Studies in Colloquial Japanese IV Phonemics. *Language*, 26(1), 86-125.
- Böcker, K., Bastiaansen, M., Vroomen, J., Brunia, C., & De Gelder, B. (1999). An ERP correlate of metrical stress in spoken word recognition. *Psychophysiology*, 36(6), 706-720.
- Boersma, P., & Weenink, D. (2024). *Praat* (Version 6.4.06) [Logiciel].
- Bögels, S., Schriefers, H., Vonk, W., Chwilla, D., & Kerkhofs, R. (2010). The Interplay between Prosody and Syntax in Sentence Processing : The Case of Subject- and Object-control Verbs. *Journal of Cognitive Neuroscience*, 22(5), 1036-1053.
- Bögels, S., Schriefers, H., Vonk, W., & Chwilla, D. (2011). Prosodic Breaks in Sentence Processing Investigated by Event-Related Potentials. *Language and Linguistics Compass*, 5(7), 424-440.
- Bögels, S., Schriefers, H., Vonk, W., Chwilla, D., & Kerkhofs, R. (2013). Processing consequences of superfluous and missing prosodic breaks in auditory sentence comprehension. *Neuropsychologia*, 51(13), 2715-2728.
- Borrell, A., & Salsignac, J. (2002). Chapitre 4. Importance de la prosodie en didactique des langues (application au FLE). In R. Renard (Ed.), *Apprentissage d'une langue étrangère/seconde 2. La phonétique verbo-tonale* (p. 163-182). De Boeck Université.

- Boucher, V. (2006). On the Function of Stress Rhythms in Speech : Evidence of a Link with Grouping Effects on Serial Memory. *Language and Speech*, 49(4), 495-519.
- Bowden, H., Steinhauer, K., Sanz, C., & Ullman, M. (2013). Native-like brain processing of syntax can be attained by university foreign language learners. *Neuropsychologia*, 51(13), 2492-2511.
- Breen, M. (2014). Empirical Investigations of the Role of Implicit Prosody in Sentence Processing. *Language and Linguistics Compass*, 8(2), 37-50.
- Brooks, J., Zoumpoulaki, A., & Bowman, H. (2017). Data-driven region-of-interest selection without inflating Type I error rate. *Psychophysiology*, 54(1), 100-113.
- Cao, J. (2000a). Intonation structure of spoken Chinese : Universality and Characteristics. *From Traditional Phonology to Modern Speech Processing*.
- Cao, J. (2000b). Rhythm of spoken Chinese—Linguistic and paralinguistic evidences. *Proc. 6th International Conference on Spoken Language Processing (ICSLP 2000)*, 2, 357-360.
- Cao, J. (2001). THE RHYTHM OF MANDARIN CHINESE. *Journal of Chinese Linguistics*, 17, 53-80.
- Cao, J. (2004a). Intonation Structure of Spoken Chinese : Universality and Specificity. *From Traditional Phonology to Modern Speech Processing*, 31-38.
- Cao, J. (2004b). Tonal Aspects in Spoken Chinese : Global and Local Perspectives. *Proc. International Symposium on Tonal Aspects of Languages: With Emphasis on Tone Languages*, 17-20.
- Carlson, K., Clifton Jr, C., & Frazier, L. (2001). Prosodic Boundaries in Adjunct Attachment. *Journal of Memory and Language*, 45(1), 58-81.
- Carlson, K., Clifton Jr, C., & Frazier, L. (2009a). Nonlocal effects of prosodic boundaries. *Memory & Cognition*, 37(7), 1014-1025.
- Carlson, K., Frazier, L., & Clifton Jr, C. (2009b). How prosody constrains comprehension : A limited effect of prosodic packaging. *Lingua*, 119(7), 1066-1082.
- Chao, Y. (2011). *A grammar of spoken Chinese*. Commercial Press.
- Chen, A., & Tseng, S. (2012). Conceptual planning in conversational Mandarin : Pitch variation in prosodic phrasing. *Proceedings of the 6th International Conference on Speech Prosody*, 547-550.
- Chen, L., Shu, H., Liu, Y., Zhao, J., & Li, P. (2007). ERP signatures of subject-verb agreement in L2 learning. *Bilingualism: Language and Cognition*, 10(2), 161-174.
- Chen, L., & Zechner, K. (2011). Applying rhythm features to automatically assess non-native speech. *Proc. Interspeech 2011*, 1861-1864.
- Chow, I. (2005). Resolving temporary syntactic ambiguity by prosodic devices in Mandarin : An acoustic study. *Journal of Chinese Linguistics*, 33(2), 181-232.
- Christodoulides, G. (2018). Acoustic Correlates of Prosodic Boundaries in French A Review of Corpus Data. *Revista de Estudos da Linguagem*, 26(4), 1531-1549.
- Chun, D., & Levis, J. (2020). Chapter 45 Prosody in Second Language Teaching : Methodologies and effectiveness. In C. Gussenhoven & A. Chen (Éds.), *The Oxford Handbook of Language Prosody* (p. 619-630). Oxford University Press.
- Clahsen, H., & Felser, C. (2006). How native-like is non-native language processing? *Trends in Cognitive Sciences*, 10(12), 564-570.
- Clifton Jr, C., Carlson, K., & Frazier, L. (2002). Informative Prosodic Boundaries. *Language and Speech*, 45(2), 87-114.
- Cohen, M. (2011). It's about Time. *Frontiers in Human Neuroscience*, 5, 1-15.

- Cole, J., Mo, Y., & Baek, S. (2010). The role of syntactic structure in guiding prosody perception with ordinary listeners and everyday speech. *Language and Cognitive Processes*, 25(7-9), 1141-1177.
- Crystal, D. (2008). *A dictionary of linguistics and phonetics* (6th ed). Blackwell Pub.
- Cumming, R. (2008). Should rhythm metrics take account of fundamental frequency? *Cambridge Occasional Papers Linguistics*, 4, 1-16.
- Cumming, R. (2010). *Speech rhythm The language-specific integration of pitch and duration* [Thèse de Doctorat]. University of Cambridge.
- Cushing, I., & Dellwo, V. (2010). The role of speech rhythm in attending to one of two simultaneous speakers. *Proc. Speech Prosody 2010*.
- Cutler, A. (1991). Linguistic rhythm and speech segmentation. In J. Angoujard, L. Nord, & R. Carlson (Éds.), *Music, Language, Speech and Brain* (p. 157-166). Macmillan Education UK.
- Cutler, A. (1994a). Segmentation problems, rhythmic solutions. *Lingua*, 92, 81-104.
- Cutler, A. (1994b). The perception of rhythm in language. *Cognition*, 50, 79-81.
- Cutler, A. (1998). Prosodic Structure and Word Recognition. In A. Friederici, *Language Comprehension : A Biological Perspective* (p. 41-70). Springer Berlin Heidelberg.
- Cutler, A., & Butterfield, S. (1992). Rhythmic cues to speech segmentation : Evidence from juncture misperception. *Journal of Memory and Language*, 31(2), 218-236.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1983). A language-specific comprehension strategy. *Nature*, 304(5922), 159-160.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1986). The syllable's differing role in the segmentation of French and English. *Journal of Memory and Language*, 25(4), 385-400.
- Cutler, A., Mehler, J., Norris, D., & Segui, J. (1992). The monolingual nature of speech segmentation by bilinguals. *Cognitive Psychology*, 24(3), 381-410.
- Cutler, A., Oahan, D., & van Donselaar, W. (1997). Prosody in the Comprehension of Spoken Language : A Literature Review. *Language and Speech*, 40(2), 141-201.
- Dahan, D. (2015). Prosody and language comprehension. *WIREs Cognitive Science*, 6(5), 441-452.
- Dasher, R., & Bolinger, D. (1982). On pre-accentual lengthening. *Journal of the International Phonetic Association*, 12(2), 58-71.
- Dauer, R. (1983). Stress-timing and syllable-timing reanalyzed. *Journal of Phonetics*, 11(1), 51-62.
- De Jong, N. (2016). Predicting pauses in L1 and L2 speech : The effects of utterance boundaries and word frequency. *International Review of Applied Linguistics in Language Teaching*, 54(2), 113-132.
- Delais-Roussarie, E. (1993). Structure rythmique et prosodique du français : Pour une approche psycho-cognitive. *LINX*, 29, 45-71.
- Delais-Roussarie, E., Post, B., Avanzi, M., Buthke, C., Di Cristo, A., Feldhausen, I., Jun, S., Martin, P., Meisenburg, T., Rialland, A., Sichel-Bazin, R., & Yoo, H. (2015). Intonational phonology of French : Developing a ToBI system for French. In S. Frota & P. Prieto (Éds.), *Intonation in Romance* (p. 63-100). Oxford University Press Oxford.
- Delais-Roussarie, E., Yoo, H., & Post, B. (2011). Quand frontières prosodiques et frontières syntaxiques se rencontrent: *Langue française*, n°170(2), 29-44.
- Delattre, P. (1939). Accent de mot et Accent de Groupe. *The French Review*, 13(2), 141-146.

- Delattre, P. (1966a). A COMPARISON OF SYLLABLE LENGTH CONDITIONING AMONG LANGUAGES. *International Review of Applied Linguistics in Language Teaching*, 4(1-4), 183-198.
- Delattre, P. (1966b). Les Dix Intonations de Base Du Français. *The French Review*, 40(1), 1-14.
- Delattre, P. (1966c). *Studies in French and Comparative Phonetics*. De Gruyter.
- Dellwo, V. (2006). Rhythm and Speech Rate : A variation coefficient for ΔC . In P. Karnowski & I. Szigeti (Éds.), *Language and language-processing* (p. 231-241). Peter Lang.
- Delorme, A., & Makeig, S. (2004). EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis. *Journal of Neuroscience Methods*, 134(1), 9-21.
- Deterding, D. (2001). The measurement of rhythm : A comparison of Singapore and British English. *Journal of Phonetics*, 29(2), 217-230.
- Di Cristo, A. (1976). Indices prosodiques et structure constituante. *Cahiers de linguistique, d'orientalisme et de slavistique*, 7, 27-40.
- Di Cristo, A. (1978). *De la microprosodie à l'intonosyntaxe* [Thèse de Doctorat]. Université de Provence.
- Di Cristo, A. (1998). 11 Intonation in French. In D. Hirst & A. Di Cristo (Éds.), *Intonation Systems A Survey of Twenty Languages* (p. 195-218). Cambridge University Press.
- Di Cristo, A. (1999). Vers une modélisation de l'accentuation du français : Première partie. *Journal of French Language Studies*, 9(2), 143-179.
- Di Cristo, A. (2004). La prosodie au carrefour de la phonétique, de la phonologie et de l'articulation formes-fonctions. *Travaux Interdisciplinaires du Laboratoire Parole et Langage d'Aix-en-Provence (TIPA)*, 23, 67-211.
- Di Cristo, A. (2013). *La prosodie de la parole*. De Boeck Supérieur.
- Di Cristo, A. (2016). *Les musiques du français parlé*. DE GRUYTER.
- Di Cristo, A., & Hirst, D. (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 Liberto, G., Nie, J., Yeaton, J., Khalighinejad, B., Shamma, S., & Mesgarani, N. (2021). Neural representation of linguistic feature hierarchy reflects second-language proficiency. *NeuroImage*, 227, 117586.
- Díaz, B., Erdocia, K., De Menezes, R., Mueller, J., Sebastián-Gallés, N., & Laka, I. (2016). Electrophysiological Correlates of Second-Language Syntactic Processes Are Related to Native and Second Language Distance Regardless of Age of Acquisition. *Frontiers in Psychology*, 7, 133.
- D'Imperio, M., Fivela, B., Baltazani, M., Post, B., & Vella, A. (2020). Chapter 16 Southern Europe. In C. Gussenhoven & A. Chen (Éds.), *The Oxford Handbook of Language Prosody* (p. 236-250). Oxford University Press.
- D'Imperio, M., & Michelas, A. (2014). Pitch scaling and the internal structuring of the Intonation Phrase in French. *Phonology*, 31(1), 95-122.
- Dittmann, A., & Llewellyn, L. (1967). The phonemic clause as a unit of speech decoding. *Journal of Personality and Social Psychology*, 6(3), 341-349.
- Dowens, M., Guo, T., Guo, J., Barber, H., & Carreiras, M. (2011). Gender and number processing in Chinese learners of Spanish – Evidence from Event Related Potentials. *Neuropsychologia*, 49(7), 1651-1659.

- Dowens, M., Vergara, M., Barber, H., & Carreiras, M. (2010). Morphosyntactic Processing in Late Second-Language Learners. *Journal of Cognitive Neuroscience*, 22(8), 1870-1887.
- Drury, J., Baum, S., Valeriotte, H., & Steinhauer, K. (2016). Punctuation and Implicit Prosody in Silent Reading: An ERP Study Investigating English Garden-Path Sentences. *Frontiers in Psychology*, 7, 1375.
- Duanmu, S. (Éd.). (2007). *The phonology of standard Chinese*. Oxford University Press.
- Duanmu, S. (2008). *Syllable structure : The limits of variation*. Oxford University Press.
- Duanmu, S. (2014). Syllable Structure and Stress. In C. Huang, Y. Li, & A. Simpson (Éds.), *The Handbook of Chinese Linguistics* (p. 422-442). Wiley.
- Dubois, J. (1991). Dictionnaire électronique des verbes français. *LINX, hors-série n°3*, 213-230.
- Dubois, J., Giacomo, M., Guespin, L., Marcellesi, C., Marcellesi, J., & Mével, J. (2001). *Dictionnaire de linguistique*. Larousse.
- Eckstein, K., & Friederici, A. (2006). It's Early : Event-related Potential Evidence for Initial Interaction of Syntax and Prosody in Speech Comprehension. *Journal of Cognitive Neuroscience*, 18(10), 1696-1711.
- Fant, G., Kruckenberg, A., & Nord, L. (1991). Durational correlates of stress in Swedish, French and English. *Journal of Phonetics*, 19(3-4), 351-365.
- Faraco, M., & Cavé, C. (2008). Déclaration ou question ? Étude expérimentale de la perception de l'intonation par des étrangers apprenant le français. *TIPA. Travaux interdisciplinaires sur la parole et le langage*, 27, 63-73.
- Ferjan, N., Ross, T., & Arvaniti, A. (2008). Second language rhythm and rhythm metrics. *Journal of the Acoustical Society of America*, 123(5_Supplement), 3427.
- Ferragne, E., & Pellegrino, F. (2004). *A comparative account of the suprasegmental and rhythmic features of British English dialects*. Proc. MIDL (*Modélisations pour l'Identification des Langues*), Paris.
- Ferreira, F. (1993). Creation of prosody during sentence production. *Psychological Review*, 100(2), 233-253.
- Ferreira, V., & Dell, G. (2000). Effect of Ambiguity and Lexical Availability on Syntactic and Lexical Production. *Cognitive Psychology*, 40(4), 296-340.
- Fields, E. (2017). *Factorial Mass Univariate ERP Toolbox* [Logiciel].
- Fields, E., & Kuperberg, G. (2020). Having your cake and eating it too : Flexibility and power with mass univariate statistics for ERP data. *Psychophysiology*, 57(2), e13468.
- Fletcher, J. (1991). Rhythm and final lengthening in French. *Journal of Phonetics*, 19(2), 193-212.
- Fletcher, J. (2010). The Prosody of Speech : Timing and Rhythm. In W. Hardcastle, J. Laver, & F. Gibbon (Éds.), *The Handbook of Phonetic Sciences* (p. 521-602). Wiley.
- Fodor, J. (2002). Psycholinguistics cannot escape prosody. *Proc. Speech Prosody 2002*, 83-90.
- Fon, J., Johnson, K., & Chen, S. (2011). Durational Patterning at Syntactic and Discourse Boundaries in Mandarin Spontaneous Speech. *Language and Speech*, 54(1), 5-32.
- Fónagy, I. (1971). *Double Coding in Speech*. 3(3), 189-222.
- Fónagy, I., Bérard, E., & Fónagy, J. (1983). CLICHÉS MÉLODIQUES. *Folia Linguistica*, 17(1-4), 153-186.
- Foucart, A., & Frenck-Mestre, C. (2011). Grammatical gender processing in L2 : Electrophysiological evidence of the effect of L1-L2 syntactic similarity. *Bilingualism: Language and Cognition*, 14(3), 379-399.

- Fougeron, C., & Keating, P. (1997). Articulatory strengthening at edges of prosodic domains. *The Journal of the Acoustical Society of America*, 101(6), 3728-3740.
- Fourcin, A., & Dellwo, V. (2009). Rhythmic classification of languages based on voice timing. *UCL Discovery online*.
- Fox, A. (Éd.). (2000). *Prosodic features and prosodic structure: The phonology of suprasegmentals*. Oxford University Press.
- Fraisse, P. (1974). *Psychologie du rythme*. Presses universitaires de France.
- François, J., Le Pesant, D., & Leeman, D. (2007). Présentation de la classification des Verbes Français de Jean Dubois et Françoise Dubois-Charlier: *Langue française*, n° 153(1), 3-19.
- Frankish, C. (1995). Intonation and auditory grouping in immediate serial recall. *Applied Cognitive Psychology*, 9(7), S5-S22.
- Frazier, L. (1979). *On comprehending sentences: Syntactic parsing strategies* [Thèse de Doctorat]. University of Connecticut.
- Frazier, L. (1987). Chapter 25 Sentence Processing: A Tutorial Review. In M. Coltheart (Éd.), *Attention and Performance XII* (p. 559-586). Routledge.
- Frazier, L., Carlson, K., & Clifton Jr, C. (2006). Prosodic phrasing is central to language comprehension. *Trends in Cognitive Sciences*, 10(6), 244-249.
- Frazier, L., & Fodor, J. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4), 291-325.
- Friederici, A. (2004). Event-related brain potential studies in language. *Current Neurology and Neuroscience Reports*, 4(6), 466-470.
- Friederici, A. (2005). Neurophysiological markers of early language acquisition: From syllables to sentences. *Trends in Cognitive Sciences*, 9(10), 481-488.
- Friederici, A., & Thierry, G. (Éds.). (2008). *Early language development: Bridging brain and behaviour*. John Benjamins Publishing Company.
- Fromont, L., Royle, P., & Steinhauer, K. (2020). Growing Random Forests reveals that exposure and proficiency best account for individual variability in L2 (and L1) brain potentials for syntax and semantics. *Brain and Language*, 204, 104770.
- Fry, D. (1964). The Function of the Syllable. *STUF - Language Typology and Universals*, 17(1-6), 215-222.
- Gabriel, C., Stahnke, J., & Thulke, J. (2014). On the acquisition of French speech rhythm in a multilingual classroom: Evidence from linguistic and extra-linguistic data. *SHS Web of Conferences*, 8, 1267-1283.
- Ganga, R., Geutjes, J., Van Niekerk, E., Reshetnikova, V., & Chen, A. (2024). Processing prosodic boundaries in Dutch coordinated constructions. *Porc. Speech Prosody 2024*, 985-989.
- Geyer, A., Holcomb, P., Midgley, K., & Grainger, J. (2011). Processing words in two languages: An event-related brain potential study of proficient bilinguals. *Journal of Neurolinguistics*, 24(3), 338-351.
- Gibbon, D., & Gut, U. (2001). Measuring speech rhythm. *Proc. Eurospeech 2001*, 95-98.
- Gilbert, A., Boucher, V., & Jemel, B. (2011). The Role of Rhythmic Chunking in Speech: Synthesis of Findings and Evidence from Statistical Learning. *Proc. ICPhS XVII*, 747-750.
- Gilbert, A., Lee, J., Wolpert, M., & Baum, S. (2023). Phrase parsing in a second language as indexed by the closure positive shift: The impact of language experience and acoustic cue salience. *European Journal of Neuroscience*, 58(8), 3838-3858.

- Glushko, A., Steinhauer, K., DePriest, J., & Koelsch, S. (2016). Neurophysiological Correlates of Musical and Prosodic Phrasing : Shared Processing Mechanisms and Effects of Musical Expertise. *PLOS ONE*, *11*(5), e0155300.
- Goldman, J., Frauenfelder, U., & Content, A. (1996). Comparaison des structures syllabiques en français et en anglais. *Actes des XXIèmes Journées d'Étude de la Parole*, 119-122.
- Grabe, E., & Low, E. (2002). Durational variability in speech and the Rhythm Class Hypothesis. In C. Gussenhoven & N. Warner (Éds.), *Laboratory Phonology 7* (p. 515-546). Mouton de Gruyter.
- Grenon, I., & White, L. (2008). Acquiring rhythm : A comparison of L1 and L2 speakers of Canadian English and Japanese. In H. Chan, H. Jacob, & E. Kipia (Éds.), *Proceedings of the 32nd Annual Boston University Conference on Language Development (BUCLD 32)* (p. 155-166). Cascadilla Press.
- Grompton, A. (1980). Timing Patterns in French. *Phonetica*, *37*(4), 205-234.
- Groppe, D., Urbach, T., & Kutas, M. (2011a). Mass univariate analysis of event-related brain potentials/fields I: A critical tutorial review. *Psychophysiology*, *48*(12), 1711-1725.
- Groppe, D., Urbach, T., & Kutas, M. (2011b). Mass univariate analysis of event-related brain potentials/fields II: Simulation studies. *Psychophysiology*, *48*(12), 1726-1737.
- Gussenhoven, C., & Chen, A. (Éds.). (2020). *The Oxford Handbook of Language Prosody*. Oxford University Press.
- Gut, U. (2012). Rhythm in L2 speech. *Speech and Language Technology*, *14/15*, 83-94.
- Gutiérrez Díez, F., Dellwo, V., Gavaldà, N., & Rosen, S. (2008). The development of measurable speech rhythm during second language acquisition. *Journal of the Acoustical Society of America*, *123*(5_Supplement), 3886.
- Han, M. (1962). The feature of duration in Japanese. *音声の研究 (日本音声学會誌) The Study of Sounds (The Phonetic Society of Japan)*, *10*, 65-80.
- Harriman, J., & Buxton, H. (1979). The influence of prosody on the recall of monaurally presented sentences. *Brain and Language*, *8*(1), 62-68.
- Hayes, B. (1985). Iambic and Trochaic Rhythm in Stress Rules. *Proceedings of the Eleventh Annual Meeting of the Berkeley Linguistics Society*, 429-446.
- Hilton, M., Råling, R., Wartenburger, I., & Elsner, B. (2019). Parallels in Processing Boundary Cues in Speech and Action. *Frontiers in Psychology*, *10*, 1566.
- Hilton, M., Wartenburger, I., & Elsner, B. (2021). Kinematic boundary cues modulate 12-month-old infants' segmentation of action sequences: An ERP study. *Neuropsychologia*, *159*, 107916.
- Hirst, D. (2024). *Speech Prosody: From Acoustics to Interpretation*. Springer Berlin Heidelberg.
- Hirst, D., & Di Cristo, A. (1984). French intonation : A parametric approach. *Die Neueren Sprachen*, *83*(5), 554-569.
- Holzgrefe, J., Wellmann, C., Petrone, C., Truckenbrodt, H., Höhle, B., & Wartenburger, I. (2013). Brain response to prosodic boundary cues depends on boundary position. *Frontiers in Psychology*, *4*(421).
- Honbolygó, F., & Csépe, V. (2013). Saliency or template ? ERP evidence for long-term representation of word stress. *International Journal of Psychophysiology*, *87*(2), 165-172.
- Honbolygó, F., Kóbor, A., German, B., & Csépe, V. (2020). Word stress representations are language-specific : Evidence from event-related brain potentials. *Psychophysiology*, *57*(5), e13541.

- Honbolygó, F., Török, Á., Bánréti, Z., Hunyadi, L., & Csépe, V. (2016). ERP correlates of prosody and syntax interaction in case of embedded sentences. *Journal of Neurolinguistics*, 37, 22-33.
- Hulst, H., & Ritter, N. (Éds.). (1999). 15 The syllable in Chinese. In *The Syllable Views and Facts* (p. 477-500). De Gruyter Mouton.
- Hwang, H., & Steinhauer, K. (2011). Phrase Length Matters : The Interplay between Implicit Prosody and Syntax in Korean “Garden Path” Sentences. *Journal of Cognitive Neuroscience*, 23(11), 3555-3575.
- Igoa, J., Carreiras, M., & Meseguer, E. (1998). A Study on Late Closure in Spanish : Principle-grounded vs. Frequency-based Accounts of Attachment Preferences. *The Quarterly Journal of Experimental Psychology Section A*, 51(3), 561-592.
- Ip, M., & Cutler, A. (2018). Asymmetric Efficiency of Juncture Perception in L1 and L2. *Proc. Speech Prosody 2018*, 289-293.
- Ip, M., & Cutler, A. (2022). Juncture prosody across languages : Similar production but dissimilar perception. *Laboratory Phonology*, 13(1), 1-49.
- Isel, F., Alter, K., & Friederici, A. (2005). Influence of Prosodic Information on the Processing of Split Particles : ERP Evidence from Spoken German. *Journal of Cognitive Neuroscience*, 17(1), 154-167.
- Itô, J. (1988). *Syllable Theory in Prosodic Phonology*. Routledge.
- Itzhak, I., Pauker, E., Drury, J., Baum, S., & Steinhauer, K. (2010). Event-related potentials show online influence of lexical biases on prosodic processing. *NeuroReport*, 21(1), 8-13.
- JASP Team. (2025). *JASP* (Version 0.19.3) [Logiciel]. University of Amsterdam.
- Jongman, A., & Tremblay, A. (2020). Chapter 43 Word Prosody in Second Language Acquisition. In C. Gussenhoven & A. Chen (Éds.), *The Oxford Handbook of Language Prosody* (p. 594-604). Oxford University Press.
- Jun, S. (Éd.). (2005). Prosodic Typology. In *Prosodic Typology: The Phonology of Intonation and Phrasing* (p. 430-458). Oxford University Press.
- Jun, S. (2012). Prosodic typology revisited : Adding macro-rhythm. *Proc. Speech Prosody 2012*, 535-538.
- Jun, S. (Éd.). (2014). Prosodic typology : By prominence type, word prosody, and macro-rhythm. In *Prosodic Typology II: The Phonology of Intonation and Phrasing* (p. 520-539). Oxford University Press.
- Jun, S., & Fougeron, C. (1995). The accentual phrase and the prosodic structure of French. *Proceedings of the 13th International Congress of Phonetic Sciences*, 2, 722-725.
- Jun, S., & Fougeron, C. (2000). A Phonological Model of French Intonation. In A. Botinis (Éd.), *Intonation Analysis, Modelling and Technology* (Vol. 15, p. 209-242). Springer Netherlands.
- Jun, S., & Fougeron, C. (2002). Realizations of accentual phrase in French intonation. *Probus*, 14(1), 147-172.
- Kaan, E. (2007). Event-Related Potentials and Language Processing : A Brief Overview. *Language and Linguistics Compass*, 1(6), 571-591.
- Kaan, E., Barkley, C., Bao, M., & Wayland, R. (2008). Thai lexical tone perception in native speakers of Thai, English and Mandarin Chinese : An event-related potentials training study. *BMC Neuroscience*, 9(1), 53.
- Kahng, J. (2018). The effect of pause location on perceived fluency. *Applied Psycholinguistics*, 39(3), 569-591.

- Kasparian, K., & Steinhauer, K. (2016). Confusing similar words : ERP correlates of lexical-semantic processing in first language attrition and late second language acquisition. *Neuropsychologia*, *93*, 200-217.
- Kaushanskaya, M., Blumenfeld, H., & Marian, V. (2020). The Language Experience and Proficiency Questionnaire (LEAP-Q) : Ten years later. *Bilingualism: Language and Cognition*, *23*(5), 945-950.
- Keil, A., Debener, S., Gratton, G., Junghöfer, M., Kappenman, E., Luck, S., Luu, P., Miller, G., & Yee, C. (2014). Committee report : Publication guidelines and recommendations for studies using electroencephalography and magnetoencephalography. *Psychophysiology*, *51*(1), 1-21.
- Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. (2007). Discourse, Syntax, and Prosody: The Brain Reveals an Immediate Interaction. *Journal of Cognitive Neuroscience*, *19*(9), 1421-1434.
- Kerkhofs, R., Vonk, W., Schriefers, H., & Chwilla, D. (2008). Sentence processing in the visual and auditory modality : Do comma and prosodic break have parallel functions? *Brain Research*, *1224*, 102-118.
- Kilner, J. (2013). Bias in a common EEG and MEG statistical analysis and how to avoid it. *Clinical Neurophysiology*, *124*(10), 2062-2063.
- Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*(1), 15-47.
- Kjelgaard, M., & Speer, S. (1999). Prosodic Facilitation and Interference in the Resolution of Temporary Syntactic Closure Ambiguity. *Journal of Memory and Language*, *40*(2), 153-194.
- Knight, R. (2011). Assessing the temporal reliability of rhythm metrics. *Journal of the International Phonetic Association*, *41*(3), 271-281.
- Knösche, T., Neuhaus, C., Haueisen, J., Alter, K., Maess, B., Witte, O., & Friederici, A. (2005). Perception of phrase structure in music. *Human Brain Mapping*, *24*(4), 259-273.
- Kockaert, H., & Li, J. (2008). *Guide pratique de prononciation française pour des apprenants chinois*. Harmattan.
- Kohler, K. (2009). Rhythm in Speech and Language A New Research Paradigm. *Phonetica*, *66*(1-2), 29-45.
- Kohno, M., & Tanioka, T. (1990). The Nature of Timing Control in Language. *Proc. First International Conference on Spoken Language Processing (ICSLP 1990)*, 1-4.
- Kohno, M., & Tsushima, T. (1989). Rhythmic Phenomena in a Child's Babbling and One-Word Sentences. *The Bulletin The Phonetic Society of Japan*, *191*, 6-13.
- Kotz, S. (2009). A critical review of ERP and fMRI evidence on L2 syntactic processing. *Brain and Language*, *109*(2-3), 68-74.
- Kotz, S., & Paulmann, S. (2007). When emotional prosody and semantics dance cheek to cheek : ERP evidence. *Brain Research*, *1151*, 107-118.
- Krivokapić, J. (2007). Prosodic planning : Effects of phrasal length and complexity on pause duration. *Journal of Phonetics*, *35*(2), 162-179.
- Krivokapić, J., Styler, W., & Parrell, B. (2020). Pause postures : The relationship between articulation and cognitive processes during pauses. *Journal of Phonetics*, *79*, 100953.
- Kutas, M., & Hillyard, S. (1980). Reading Senseless Sentences : Brain Potentials Reflect Semantic Incongruity. *Science*, *207*(4427), 203-205.
- Kutas, M., & Van Petten, C. (1994). Chapter 4 Psycholinguistics Electrified : Event-Related Brain Potential Investigations. In M. Gernsbacher (Éd.), *Handbook of Psycholinguistics* (p. 83-143). Academic Press.

- Labrune, L. (2005). Autour de la syllabe: les constituants prosodiques mineurs en phonologie. In N. Nguyen, S. Wauquiers, & J. Durand (Éds.), *Phonétique et phonologie—Forme et substance* (p. 95-116). Lavoisier.
- Lacheret, A. (2011). La prosodie au cœur du verbal. *Rééducation orthophonique*, 246, 87-104.
- Lacheret-Dujour, A., & Beaugendre, F. (1999). *La prosodie du français*. CNRS ÉDITIONS.
- Ladd, D. (1983). Peak Features and Overall Slope. In A. Cutler & D. Ladd (Éds.), *Prosody : Models and Measurements* (Vol. 14, p. 39-52). Springer Berlin Heidelberg.
- Ladefoged, P. (1975). *A course in phonetics*. Harcourt Brace Jovanovich.
- Lai, C., Evanini, K., & Zechner, K. (2013). Applying rhythm metrics to non-native spontaneous speech. *Proc. Speech and Language Technology in Education (SLaTE 2013)*, 159-163.
- Landron, S. (2007). *Enseigner la prosodie du français à des apprenants taiwanais* [Mémoire de Master non publié]. Université de Nantes.
- Landron, S., Gao, J., Chang, Y., & Tian, Y. (2016). 30 Les sinophones. In S. Detey, I. Racine, Y. Kawaguchi, & J. Eychenne (Éds.), *La prononciation du français dans le monde : Du natif à l'apprenant* (p. 193-199). CLE International.
- Langus, A., Marchetto, E., Bion, R., & Nespors, M. (2012). Can prosody be used to discover hierarchical structure in continuous speech? *Journal of Memory and Language*, 66(1), 285-306.
- Langus, A., Mehler, J., & Nespors, M. (2017). Rhythm in language acquisition. *Neuroscience & Biobehavioral Reviews*, 81, 158-166.
- Lauret, B. (2007). *Enseigner la prononciation du français, questions et outils*. Hachette FLE.
- Leckey, M., & Federmeier, K. (2019). Chapter 3 Electrophysiological Methods in the Study of Language Processing. In G. De Zubicaray & N. Schiller (Éds.), *The Oxford Handbook of Neurolinguistics* (p. 42-71). Oxford University Press.
- Lee, E., & Garnsey, S. (2012). Do contrastive accents modulate the effect of intonational phrase boundaries in parsing? *Lingua*, 122(14), 1763-1775.
- Lee, P. (1987). *Contribution à l'étude de l'intonation française prononcée par les Chinois en vue de l'apprentissage* [Thèse de Doctorat]. Université de Franche-Comté.
- Lehiste, I. (1972). The Timing of Utterances and Linguistic Boundaries. *Journal of the Acoustical Society of America*, 51(6B), 2018-2024.
- Lenneberg, E. (1967). *Biological Foundations of Language*. John Wiley & Sons, Inc.
- Léon, P. (2012). *Phonétisme et prononciation du français*. Armand Colin.
- Leonard, L. (1973). The Role of Intonation in the Recall of Various Linguistic Stimuli. *Language and Speech*, 16(4), 327-335.
- Li, W., Wang, L., Li, X., & Yang, Y. (2008). Chapter 88 Closure Positive Shifts Evoked by Different Prosodic Boundaries in Chinese Sentences. In R. Wang, E. Shen, & F. Gu (Éds.), *Advances in Cognitive Neurodynamics ICCN 2007* (p. 505-509). Springer Netherlands.
- Li, W., & Yang, Y. (2009). Perception of prosodic hierarchical boundaries in Mandarin Chinese sentences. *Neuroscience*, 158(4), 1416-1425.
- Li, W., Zhang, H., Zheng, Z., & Li, X. (2019). Prosodic phrase priming during listening to Chinese ambiguous phrases in different experimental tasks. *Journal of Neurolinguistics*, 51, 135-150.
- Lin, H., & Wang, Q. (2007). Mandarin Rhythm : An Acoustic Study. *Journal of Chinese Language and Computing*, 17(3), 127-140.
- Lin, Y. (2007). *The sounds of Chinese*. Cambridge Univ. Press.

- Liss, J., White, L., Mattys, S., Lansford, K., Lotto, A., Spitzer, S., & Caviness, J. (2009). Quantifying Speech Rhythm Abnormalities in the Dysarthrias. *Journal of Speech, Language, and Hearing Research*, 52(5), 1334-1352.
- Liu, B., Wang, Z., & Jin, Z. (2010). The effects of punctuations in Chinese sentence comprehension : An ERP study. *Journal of Neurolinguistics*, 23(1), 66-80.
- Llanos, F., German, J., Gnanateja, G., & Chandrasekaran, B. (2021). The neural processing of pitch accents in continuous speech. *Neuropsychologia*, 158, 107883.
- Lopez-Calderon, J., & Luck, S. (2014). ERPLAB: an open-source toolbox for the analysis of event-related potentials. *Frontiers in Human Neuroscience*, 8, 213.
- Loukina, A., Kochanski, G., Shih, C., Keane, E., & Watson, I. (2009). Rhythm measures with language-independent segmentation. *Proc. Interspeech 2009*, 1531-1534.
- Loukina, A., Kochanski, G., Rosner, B., Keane, E., & Shih, C. (2011). Rhythm measures and dimensions of durational variation in speech. *The Journal of the Acoustical Society of America*, 129(5), 3258-3270.
- Low, E., Grabe, E., & Nolan, F. (2000). Quantitative Characterizations of Speech Rhythm : Syllable-Timing in Singapore English. *Language and Speech*, 43(4), 377-401.
- Luck, S. (2005). 2 Ten Simple Rules for Designing and Interpreting ERP Experiments. In T. Handy (Éd.), *Event-Related Potentials : A Methods Handbook* (p. 17-32). The MIT Press.
- Luck, S. (2014). *An Introduction to the Event-Related Potential Technique* (2nd ed). The MIT Press.
- Luck, S., & Gaspelin, N. (2017). How to get statistically significant effects in any ERP experiment (and why you shouldn't). *Psychophysiology*, 54(1), 146-157.
- Magne, C., Astesano, C., Aramaki, M., Ystad, S., Kronland-Martinet, R., & Besson, M. (2007). Influence of Syllabic Lengthening on Semantic Processing in Spoken French : Behavioral and Electrophysiological Evidence. *Cerebral Cortex*, 17(11), 2659-2668.
- Maia, M., Fernández, E., Costa, A., & Lourenço-Gomes, M. (2007). Early and late preferences in relative clause attachment in Portuguese and Spanish. *Journal of Portuguese Linguistics*, 6(1), 227-250.
- Mairano, P., & Romano, A. (2010). UN CONFRONTO TRA DIVERSE METRICHE RITMICHE USANDO CORRELATORE. In S. Schmid, M. Schwarzenbach, & D. Studer (Éds.), *La dimensione temporale del parlato* (Vol. 5, p. 79-100). EDK.
- Mairano, P., & Romano, A. (2011). Rhythm metrics for 21 languages. *Proc. of the 17th International Congress of Phonetic Sciences*, 1318-1321.
- Männel, C. (2008). 1. The method of event-related brain potentials in the study of cognitive processes : A tutorial. In A. Friederici & G. Thierry (Éds.), *Early Language Development Bridging brain and behaviour* (p. 1-22). John Benjamins Publishing Company.
- Männel, C., & Friederici, A. (2008). Event-related brain potentials as a window to children's language processing : From syllables to sentences. In I. Sekerina, E. Fernández, & H. Clahsen (Éds.), *Developmental psycholinguistics : On-line methods in children's language processing* (p. 29-72). John Benjamins Publishing Company.
- Männel, C., & Friederici, A. (2009). Pauses and Intonational Phrasing : ERP Studies in 5-month-old German Infants and Adults. *Journal of Cognitive Neuroscience*, 21(10), 1988-2006.
- Männel, C., & Friederici, A. (2011). Intonational phrase structure processing at different stages of syntax acquisition : ERP studies in 2-, 3-, and 6-year-old children. *Developmental Science*, 14(4), 786-798.

- Männel, C., Schipke, C., & Friederici, A. (2013). The role of pause as a prosodic boundary marker : Language ERP studies in German 3- and 6-year-olds. *Developmental Cognitive Neuroscience*, 5, 86-94.
- Marian, V., Blumenfeld, H., & Kaushanskaya, M. (2007). The Language Experience and Proficiency Questionnaire (LEAP-Q) : Assessing Language Profiles in Bilinguals and Multilinguals. *Journal of Speech, Language, and Hearing Research*, 50(4), 940-967.
- Martin, P. (2012). Neurophysiological research explains prosodic structures constraints. *Revista de Estudos da Linguagem*, 20(2), 13-22.
- Martin, P. (2013). La structure prosodique opère-t-elle avant ou après la syntaxe ? *TIPA. Travaux interdisciplinaires sur la parole et le langage*, 29, 1-12.
- Martin, P. (2014). Ondes cérébrales et contraintes de la structure prosodique. *Synergies Europe*, 9, 161-176.
- Martin, P. (2015). *The Structure of Spoken Language Intonation in Romance* (Cambridge University Press).
- Martin, P. (2018). *Intonation, structure prosodique et ondes cérébrales-introduction à l'analyse prosodique*. ISTE Editions.
- Martin, P. (2020). Enseignement de l'intonation et neurolinguistique. *Corela*, HS-30.
- Mary, L., & Yegnanarayana, B. (2008). Extraction and representation of prosodic features for language and speaker recognition. *Speech Communication*, 50(10), 782-796.
- McLaughlin, J., Tanner, D., Pitkänen, I., Frenck-Mestre, C., Inoue, K., Valentine, G., & Osterhout, L. (2010). Brain Potentials Reveal Discrete Stages of L2 Grammatical Learning. *Language Learning*, 60(s2), 123-150.
- Mehler, J., Dommergues, J., Frauenfelder, U., & Segui, J. (1981). The syllable's role in speech segmentation. *Journal of Verbal Learning and Verbal Behavior*, 20(3), 298-305.
- Meulman, N., Seton, B., Stowe, L., & Schmid, M. (2016). Chapter 6 EEG and Event-Related Brain Potentials. In M. Schmid, S. Berends, C. Bergmann, S. Brouwer, N. Meulman, B. Seton, S. Sprenger, & L. Stowe (Éds.), *Designing Research on Bilingual Development* (p. 81-107). Springer International Publishing.
- Meulman, N., Stowe, L., Sprenger, S., Bresser, M., & Schmid, M. (2014). An ERP study on L2 syntax processing : When do learners fail? *Frontiers in Psychology*, 5, 1-17.
- Meynadier, Y. (2001). La syllabe phonétique et phonologique : Une introduction. *Travaux interdisciplinaires du Laboratoire Parole et Langage*, 20, 91-148.
- Michelas, A., & D'Imperio, M. (2009). Is there an Intermediate Phrase in French? *Phonetics and Phonologie in Iberia*.
- Michelas, A., & D'Imperio, M. (2010). Indices phonétiques et contraintes phonologiques : Caractérisation du syntagme intermédiaire en français. *Actes des XXVIIIème Journées d'Étude sur la Parole*, 77-80.
- Michelas, A., & D'Imperio, M. (2011). Uncovering the Role of the Intermediate Phrase in the Syntactic Parsing of French. *Proc. ICPhS XVII*, 1374-1377.
- Mickan, A., & Lemhöfer, K. (2020). Tracking Syntactic Conflict between Languages over the Course of L2 Acquisition : A Cross-sectional Event-related Potential Study. *Journal of Cognitive Neuroscience*, 32(5), 822-846.
- Miller, G. (1956). The magical number seven, plus or minus two : Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81-97.
- Millotte, S., René, A., Wales, R., & Christophe, A. (2008). Phonological phrase boundaries constrain the online syntactic analysis of spoken sentences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34(4), 874-885.

- Millotte, S., Wales, R., & Christophe, A. (2007). Phrasal prosody disambiguates syntax. *Language and Cognitive Processes*, 22(6), 898-909.
- Mok, P. (2009). On the syllable-timing of Cantonese and Beijing Mandarin. *Chinese Journal of Phonetics*, 2, 148-154.
- Mok, P., & Dellwo, V. (2008). Comparing native and non-native speech rhythm using acoustic rhythmic measures : Cantonese, Beijing Mandarin and English. *Proc. Speech Prosody 2008*, 423-426.
- Moreno, E., Rodríguez-Fornells, A., & Laine, M. (2008). Event-related potentials (ERPs) in the study of bilingual language processing. *Journal of Neurolinguistics*, 21(6), 477-508.
- Morgan-Short, K. (2014). Electrophysiological Approaches to Understanding Second Language Acquisition : A Field Reaching its Potential. *Annual Review of Applied Linguistics*, 34, 15-36.
- Morgan-Short, K., Faretta-Stutenberg, M., & Bartlett-Hsu, L. (2015). Contributions of event-related potential research to issues in explicit and implicit second language acquisition. In P. Rebuschat (Éd.), *Implicit and Explicit Learning of Languages* (p. 349-383). John Benjamins Publishing Company.
- Morgan-Short, K., Finger, I., Grey, S., & Ullman, M. (2012a). Second Language Processing Shows Increased Native-Like Neural Responses after Months of No Exposure. *PLoS ONE*, 7(3), e32974.
- Morgan-Short, K., Sanz, C., Steinhauer, K., & Ullman, M. (2010). Second Language Acquisition of Gender Agreement in Explicit and Implicit Training Conditions : An Event-Related Potential Study. *Language Learning*, 60(1), 154-193.
- Morgan-Short, K., Steinhauer, K., Sanz, C., & Ullman, M. (2012b). Explicit and Implicit Second Language Training Differentially Affect the Achievement of Native-like Brain Activation Patterns. *Journal of Cognitive Neuroscience*, 24(4), 933-947.
- Morgan-Short, K., & Tanner, D. (2014). 6 Event-related potentials (ERPs). In *Research methods in second language psycholinguistics* (p. 127-152). Routledge.
- Mueller, J. (2005). Electrophysiological correlates of second language processing. *Second Language Research*, 21(2), 152-174.
- Näätänen, R., & Kreegipuu, K. (2012). Chapter 6 The Mismatch Negativity (MMN). In S. Luck & E. Kappenman (Éds.), *The Oxford Handbook of Event-Related Potential Components* (p. 143-157). Oxford University Press.
- Näätänen, R., Paavilainen, P., Rinne, T., & Alho, K. (2007). The mismatch negativity (MMN) in basic research of central auditory processing : A review. *Clinical Neurophysiology*, 118(12), 2544-2590.
- Naito-Billen, Y. (2012). *Effects of Prosody while Disambiguating Ambiguous Japanese Sentences in the Brain of Native Speakers and Learners of Japanese : A Proposition for Pronunciation and Prosody Training* [Thèse de Doctorat]. University of Kansas.
- Nazzi, T., Bertoni, J., & Mehler, J. (1998). Language discrimination by newborns : Toward an understanding of the role of rhythm. *Journal of Experimental Psychology: Human Perception and Performance*, 24(3), 756-766.
- Neuhaus, C., Knösche, T., & Friederici, A. (2006). Effects of Musical Expertise and Boundary Markers on Phrase Perception in Music. *Journal of Cognitive Neuroscience*, 18(3), 472-493.
- Neurobehavioral Systems, Inc. (s. d.). *Presentation* (Version Version 24.0) [Logiciel].
- Nickels, S., Opitz, B., & Steinhauer, K. (2013). ERPs show that classroom-instructed late second language learners rely on the same prosodic cues in syntactic parsing as native speakers. *Neuroscience Letters*, 557, 107-111.

- Nickels, S., & Steinhauer, K. (2018). Prosody–syntax integration in a second language : Contrasting event-related potentials from German and Chinese learners of English using linear mixed effect models. *Second Language Research*, 34(1), 9-37.
- Nishinuma, Y., & Duez, D. (1989). Perceptual optimization of syllable duration in short French sentences. *Proc. First European Conference on Speech Communication and Technology (Eurospeech 1989)*, 2694-2697.
- Nooteboom, S. (1997). The Prosody of Speech : Melody and Rhythm. In W. Hardcastle & J. Laver (Éds.), *The Handbook of Phonetic Sciences* (p. 640-673). Wiley-Blackwell.
- O'Connell, D., Turner, E., & Onuska, L. (1968). Intonation, Grammatical Structure, and Contextual Association in Immediate Recall. *Journal of Verbal Learning and Verbal Behavior*, 7(1), 110-116.
- O'Dell, M., & Nieminen, T. (1999). Coupled oscillator model of speech rhythm. *Proceedings of the XIVth international congress of phonetic sciences*, 2, 1075-1078.
- Ohala, J., & Gilbert, J. (1979). *Listeners' ability to identify languages by their prosody*. 2, 123-131.
- Ojima, S., Nakata, H., & Kakigi, R. (2005). An ERP Study of Second Language Learning after Childhood: Effects of Proficiency. *Journal of Cognitive Neuroscience*, 17(8), 1212-1228.
- Oldfield, R. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113.
- Ordin, M., & Polyanskaya, L. (2015a). Acquisition of speech rhythm in a second language by learners with rhythmically different native languages. *Journal of the Acoustical Society of America*, 138(2), 533-544.
- Ordin, M., & Polyanskaya, L. (2015b). Perception of speech rhythm in second language : The case of rhythmically similar L1 and L2. *Frontiers in Psychology*, 6, 316.
- O'Shea, M. (Éd.). (2005). *The brain : A very short introduction*. Oxford University Press.
- Osterhout, L., & Holcomb, P. (1995). 6 Event-related potentials and language comprehension. In M. Rugg & M. Coles (Éds.), *Electrophysiology of Mind : Event-related Brain Potentials and Cognition* (p. 171-216). Oxford University Press.
- Osterhout, L., McLaughlin, J., Kim, A., Greenwald, R., & Inoue, K. (2004). Chapter 14 Sentences in the Brain : Event-Related Potentials as Real-Time Reflections of Sentence Comprehension and Language Learning. In M. Carreiras & C. Clifton Jr. (Éds.), *The On-Line Study of Sentence Comprehension : Eyetracking, ERP, and Beyond* (p. 271-308). Psychology Press.
- Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., Frenck-Mestre, C., & Hirschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of Neurolinguistics*, 21(6), 509-521.
- Otake, T. (2015). 12 Mora and mora-timing. In H. Kubozono (Éd.), *Handbook of Japanese Phonetics and Phonology* (p. 493-524). DE GRUYTER.
- Paape, D., & Vasishth, S. (2021). Is reanalysis selective when regressions are consciously controlled? *Glossa Psycholinguistics*, 1(1)(2), 1-34.
- Pagel, D., Madeleni, É., & Wioland, F. (2012). *Le rythme du français parlé*. Hachette FLE.
- Pamies, A. (2010). Quelques malentendus à propos du concept de rythme en linguistique. In M. Russo (Éd.), *Prosodic universals : Comparative studies in rhythmic modeling and rhythm typology* (p. 227-263). Aracne.
- Pannekamp, A., Toepel, U., Alter, K., Hahne, A., & Friederici, A. (2005). Prosody-driven Sentence Processing : An Event-related Brain Potential Study. *Journal of Cognitive Neuroscience*, 17(3), 407-421.

- Paris, C., Thomas, M., Gilson, R., & Kincaid, J. (2000). Linguistic Cues and Memory for Synthetic and Natural Speech. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 42(3), 421-431.
- Pascual-Marqui, R. (1999). Review of Methods for Solving the EEG Inverse Problem. *INTERNATIONAL JOURNAL OF BIOELECTROMAGNETISM*, 1(1), 75-86.
- Pascual-Marqui, R., Esslen, M., Kochi, K., & Lehmann, D. (2002). Functional imaging with low resolution brain electromagnetic tomography (LORETA): Review, new comparisons, and new validation. *Japanese Journal of Clinical Neurophysiology*, 30, 81-94.
- Pascual-Marqui, R., Michel, C., & Lehmann, D. (1994). Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain. *International Journal of Psychophysiology*, 18(1), 49-65.
- Patel, A. (2008). *Music, language, and the brain*. Oxford University Press.
- Pauker, E. (2013). *How multiple prosodic boundaries of varying sizes influence syntactic parsing : Behavioral and ERP evidence* [Thèse de Doctorat]. McGill University.
- Pauker, E., Itzhak, I., Baum, S., & Steinhauer, K. (2011). Effects of Cooperating and Conflicting Prosody in Spoken English Garden Path Sentences : ERP Evidence for the Boundary Deletion Hypothesis. *Journal of Cognitive Neuroscience*, 23(10), 2731-2751.
- Paulmann, S. (2023). Chapter 21 Neurophysiological Underpinnings of Prosody. In M. Grimaldi, E. Brattico, & Y. Shtyrov (Éds.), *Language Electrified : Principles, Methods, and Future Perspectives of Investigation* (p. 669-685). Springer US.
- Payne, B., Ng, S., Shantz, K., & Federmeier, K. (2020). Chapter Four—Event-related brain potentials in multilingual language processing : The N's and P's. In K. Federmeier & H. Huang (Éds.), *Psychology of Learning and Motivation* (Vol. 72, p. 75-118). Elsevier.
- Pegolo, C. (1985). The role of rhythm and intonation in the silent reading of French as a foreign language. *Reading in a Foreign Language*, 3(1), 313-327.
- Peng, G. (2006). Temporal and tonal aspects of Chinese syllables : A corpus-based comparative study of Mandarin and Cantonese. *Journal of Chinese Linguistics*, 34(1), 134-154.
- Perani, D., & Abutalebi, J. (2005). The neural basis of first and second language processing. *Current Opinion in Neurobiology*, 15(2), 202-206.
- Peter, V., McArthur, G., & Crain, S. (2014). Using event-related potentials to measure phrase boundary perception in English. *BMC Neuroscience*, 15(1), 129.
- Pettorino, M., Maffia, M., Pellegrino, E., Vitale, M., & De Meo, A. (2013). VtoV: a perceptual cue for rhythm identification. In M. Piet & S. Catherine (Éds.), *Proceedings of the Prosody-Discourse Interface Conference* (p. 101-106).
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation* [Thèse de Doctorat]. Massachusetts Institute of Technology.
- Pike, K. (1945). *The intonation of American English*. University of Michigan Press.
- Pillot-Loiseau, C., & Xie, H. (2018). Transfert rythmique du chinois mandarin au français dans l'apprentissage du Français Langue Étrangère : Acoustique et perception. *SHS Web of Conferences*, 46, 09001.
- Poeppl, D., & Assaneo, M. (2020). Speech rhythms and their neural foundations. *Nature Reviews Neuroscience*, 21(6), 322-334.
- Pollack, I., & Pickett, J. (1957). Cocktail Party Effect. *The Journal of the Acoustical Society of America*, 29(11_Supplement), 1262.

- Polyanskaya, L., Ordin, M., & Busa, M. (2017). Relative Saliency of Speech Rhythm and Speech Rate on Perceived Foreign Accent in a Second Language. *Language and Speech*, 60(3), 333-355.
- Polyanskaya, L., & Ordin, M. (2019). The effect of speech rhythm and speaking rate on assessment of pronunciation in a second language. *Applied Psycholinguistics*, 40(03), 795-819.
- Post, B. (2000). *Tonal and Phrasal Structures in French intonation*. Thesus.
- Pratt, E. (2018). Prosody in Sentence Processing. In E. Fernández & H. Cairns (Éds.), *The Handbook of Psycholinguistics* (p. 365-391). Wiley.
- Purves, D., Augustine, G., Fitzpatrick, D., Hall, W., LaMantia, A., Mooney, R., Platt, M., & White, L. (Éds.). (2018). *Neuroscience* (Sixth edition). Oxford University Press.
- Ramus, F. (1999). *Rythme des langues et acquisition du langage* [Thèse de Doctorat]. Ecole des Hautes Etudes en Sciences Sociales (EHESS).
- Ramus, F., & Mehler, J. (1999). Language identification with suprasegmental cues : A study based on speech resynthesis. *Journal of the Acoustical Society of America*, 105(1), 512-521.
- Ramus, F., Nespors, M., & Mehler, J. (1999). Correlates of linguistic rhythm in the speech signal. *Cognition*, 73(3), 265-292.
- Rathcke, T., & Smith, R. (2015). Speech timing and linguistic rhythm : On the acoustic bases of rhythm typologies. *Journal of the Acoustical Society of America*, 137(5), 2834-2845.
- Reiterer, S. (2010). Chapter 17. The cognitive neuroscience of second language acquisition and bilingualism : Factors that matter in L2 acquisition – A neuro-cognitive perspective. In M. Kail & M. Hickmann (Éds.), *Language Acquisition across Linguistic and Cognitive Systems* (p. 307-321). John Benjamins Publishing Company.
- Riazantseva, A. (2001). SECOND LANGUAGE PROFICIENCY AND PAUSING A Study of Russian Speakers of English. *Studies in Second Language Acquisition*, 23(4), 497-526.
- Ridouane, R., Meynadier, Y., & Fougeron, C. (2011). La syllabe : Objet théorique et réalité physique. *Faits de Langues*, 37(1), 213-234.
- Roach, P. (1982). On the distinction between « stress-timed » and « syllable-timed » languages. *Linguistic controversies*, 73, 79.
- Roll, M., & Horne, M. (2011). Interaction of right- and left-edge prosodic boundaries in syntactic parsing. *Brain Research*, 1402, 93-100.
- Romano, A., & Mairano, P. (2010). Speech rhythm measuring and modelling : Pointing out multi-layer and multi-parameter assessments. In M. Russo (Éd.), *Prosodic Universals Comparative Studies in Rhythmic Modeling and Rhythm Typology* (p. 79-116). ARACNE editrice.
- Romano, A., Mairano, P., & Calabrò, L. (2011). Measures of speech rhythm in East-Asian tonal languages. *Proceedings of the 17th International Congress of Phonetic Sciences*, 1714-1717.
- Rommers, J., & Federmeier, K. (2018). 13 Electrophysiological Methods. In A. de Groot & P. Hagoort (Éds.), *Research Methods in Psycholinguistics and the Neurobiology of Language A Practical Guide* (p. 247-265). Wiley Blackwell.
- Rosner, B., Grabe, E., Nicholson, H., Owen, K., & Keane, E. (2004). Prosody, Memory Load, and Memory for Speech. In E. Grabe & D. Wright (Éds.), *Oxford University Working Papers in Linguistics* (p. 85-102).
- Royle, P., & Steinhauer, K. (2014). Component changes in ERP profiles during language acquisition. *International Journal of Psychophysiology*, 94(2), 158.

- RStudio Team. (2025). *RStudio : Integrated Development Environment for R*. (Version 2024.12.1+563) [Logiciel]. Posit Software.
- Russo, M., & Barry, W. (2008a). Isochrony reconsidered. Objectifying relations between rhythm measures and speech tempo. *Proc. Speech Prosody 2008*, 419-422.
- Russo, M., & Barry, W. (2008b). Calcul rythmique et quantification de la vitesse d'élocution. *Proc. XXVIIèmes Journées d'Études sur la Parole*, cdrom n°1562.
- Sabourin, L., Brien, C., & Tremblay, M. (2013). Chapter 11. Electrophysiology of second language processing : The past, present and future. In M. Mayo, M. Mangado, & M. Adrián (Éds.), *Contemporary Approaches to Second Language Acquisition* (p. 221-242). John Benjamins Publishing Company.
- Sabourin, L., & Stowe, L. (2008). Second language processing : When are first and second languages processed similarly? *Second Language Research*, 24(3), 397-430.
- Sanders, L., & Neville, H. (2003a). An ERP study of continuous speech processing : I. Segmentation, semantics, and syntax in native speakers. *Cognitive Brain Research*, 15(3), 228-240.
- Sauzedde, B. (2014). Découpage syllabique du français et difficulté de lecture. *立命館言語文化研究*, 26(2), 91-106.
- Schafer, A. (1997). *PROSODIC PARSING: THE ROLE OF PROSODY IN SENTENCE COMPREHENSION* [Thèse de Doctorat]. University of Massachusetts Amherst.
- Selkirk, E. (1984). *Phonology and syntax : The relation between sound and structure*. MIT press.
- Shattuck-Hufnagel, S., & Turk, A. (1996). A prosody tutorial for investigators of auditory sentence processing. *Journal of Psycholinguistic Research*, 25(2), 193-247.
- Shen, X. (1989). *The prosody of Mandarin Chinese*. University of California Press.
- Shen, X. (1990). ABILITY OF LEARNING THE PROSODY OF AN INTONATIONAL LANGUAGE BY SPEAKERS OF A TONAL LANGUAGE: CHINESE SPEAKERS LEARNING FRENCH PROSODY. *IRAL - International Review of Applied Linguistics in Language Teaching*, 28(2).
- Shen, X. (1992). A Pilot Study on the Relation between the Temporal and Syntactic Structures in Mandarin. *Journal of the International Phonetic Association*, 22, 35-43.
- Shen, X. (1993). The Use of Prosody in Disambiguation in Mandarin. *Phonetica*, 50, 261-271.
- Shih, C. (2000). 11 A Declination Model of Mandarin Chinese. In A. Botinis (Éd.), *Intonation Analysis, Modelling and Technology* (p. 243-268). Springer Netherlands.
- Simon, A., & Christodoulides, G. (2016). Frontières prosodiques perçues : Corrélats acoustiques et indices syntaxiques: *Langue française*, N° 191(3), 83-106.
- Speer, S., & Blodgett, A. (2006). Prosody. In *Handbook of Psycholinguistics* (p. 505-537). Elsevier.
- Stack, C., & Watson, D. (2023). Pauses and Parsing : Testing the Role of Prosodic Chunking in Sentence Processing. *Languages*, 8(3), 157.
- Steinhauer, K. (2003). Electrophysiological correlates of prosody and punctuation. *Brain and Language*, 86(1), 142-164.
- Steinhauer, K. (2014a). Event-related Potentials (ERPs) in Second Language Research : A Brief Introduction to the Technique, a Selected Review, and an Invitation to Reconsider Critical Periods in L2. *Applied Linguistics*, 35(4), 393-417.
- Steinhauer, K. (2014b). Learning and forgetting languages—An introduction to ERP approaches. *International Journal of Psychophysiology*, 94(2), 157.

- Steinhauer, K., Abada, S., Pauker, E., Itzhak, I., & Baum, S. (2010). Prosody-syntax interactions in aging : Event-related potentials reveal dissociations between on-line and off-line measures. *Neuroscience Letters*, 472(2), 133-138.
- Steinhauer, K., Alter, K., & Friederici, A. (1999). Brain potentials indicate immediate use of prosodic cues in natural speech processing. *Nature Neuroscience*, 2(2), 191-196.
- Steinhauer, K., & Connolly, J. (2008). Chapter 9 Event-Related Potentials in the Study of Language. In B. Stemmer & H. Whitaker (Éds.), *Handbook of the Neuroscience of Language* (p. 91-104). Elsevier.
- Steinhauer, K., & Friederici, A. (2001). Prosodic Boundaries, Comma Rules, and Brain Responses : The Closure Positive Shift in ERPs as a Universal Marker for Prosodic Phrasing in Listeners and Readers. *Journal of Psycholinguistic Research*, 30(3), 267-295.
- Steinhauer, K., White, E., & Drury, J. (2009). Temporal dynamics of late second language acquisition : Evidence from event-related brain potentials. *Second Language Research*, 25(1), 13-41.
- Stockmal, V., Markus, D., & Bond, D. (2005). Measures of Native and Non-Native Rhythm in a Quantity Language. *Language and Speech*, 48(1), 55-63.
- Swaab, T., Ledoux, K., Camblin, C., Boudewyn, M., Luck, S., & Kappenman, E. (2012). Chapter 15 Language-Related ERP Components. In *The Oxford Handbook of Event-Related Potential Components* (p. 397-439). Oxford University Press.
- Tanner, D., Mclaughlin, J., Herschensohn, J., & Osterhout, L. (2013). Individual differences reveal stages of L2 grammatical acquisition : ERP evidence. *Bilingualism: Language and Cognition*, 16(2), 367-382.
- Tanner, D., & Van Hell, J. (2014). ERPs reveal individual differences in morphosyntactic processing. *Neuropsychologia*, 56, 289-301.
- Te Rietmolen, N., El Yagoubi, R., Espesser, R., Magnen, C., & Astésano, C. (2016). Investigating the phonological status of the initial accent in French : An Event-Related Potentials study. *Proc. Speech Prosody 2016*, 1181-1185.
- Te Rietmolen, N., El Yagoubi, R., Ghio, A., & Astésano, C. (2017). The Phonological Status of the French Initial Accent and its Role in Semantic Processing : An Event-Related Potentials Study. *Proc. Interspeech 2017*, 2436-2440.
- Te Rietmolen, N., El Yagoubi, R., & Astésano, C. (2019a). *MisMatch Negativity-study showing pre-lexical sensitivity to both primary Final Accent and secondary Initial Accent in French*. Neuroscience.
- Te Rietmolen, N., El Yagoubi, R., & Astésano, C. (2019b). *Interplay between metrical and semantic processing in French : An N400 study*. Neuroscience.
- The MathWorks Inc. (2024). *MATLAB* (Version R2024a) [Logiciel].
- Tokowicz, N., & MacWhinney, B. (2005). IMPLICIT AND EXPLICIT MEASURES OF SENSITIVITY TO VIOLATIONS IN SECOND LANGUAGE GRAMMAR: An Event-Related Potential Investigation. *Studies in Second Language Acquisition*, 27(02), 173-204.
- Tolentino, L., & Tokowicz, N. (2011). ACROSS LANGUAGES, SPACE, AND TIME: A Review of the Role of Cross-Language Similarity in L2 (Morpho)Syntactic Processing as Revealed by fMRI and ERP Methods. *Studies in Second Language Acquisition*, 33(1), 91-125.
- Tortel, A. (2009). *EVALUATION QUALITATIVE DE LA PROSODIE D'APPRENANTS FRANCAIS : APPORT DE PARAMETRISATIONS PROSODIQUES* [Thèse de Doctorat]. Université de Provence - Aix-Marseille I.

- Trouvain, J., & Braun, B. (2020). Sentence Prosody in a Second Language. In C. Gussenhoven & A. Chen (Éds.), *The Oxford Handbook of Language Prosody* (p. 605-618). Oxford University Press.
- Turk, A., & Shattuck-Hufnagel, S. (2013). What is speech rhythm? A commentary on Arvaniti and Rodriquez, Krivokapić, and Goswami and Leong. *Laboratory Phonology*, 4(1).
- Turk, A., & Shattuck-Hufnagel, S. (2014). Timing in talking : What is it used for, and how is it controlled? *Philosophical Transactions of the Royal Society B: Biological Sciences*, 369(1658), 20130395.
- Urbach, T., & Kutas, M. (2018). Chapter 40 Cognitive Electrophysiology of Language. In S. Rueschemeyer & M. Gaskell (Éds.), *The Oxford Handbook of Psycholinguistics* (p. 930-954). Oxford University Press.
- Vaissière, J. (1974). On French prosody. *Quarterly Progress Report, Research Laboratory of Electronics*, 114, 212-223.
- Vaissière, J. (1975). Further note on French prosody. *Quarterly Progress Report, Research Laboratory of Electronics*, 115, 251-262.
- Vaissière, J. (1983). 5. Language-Independent Prosodic Features. In A. Cutler & D. Ladd (Éds.), *Prosody : Models and Measurements* (Vol. 14, p. 53-66). Springer-Verlag.
- Vaissière, J. (1991). Rhythm, accentuation and final lengthening in French. In J. Sundberg, L. Nord, & R. Carlson (Éds.), *Music, Language, Speech and Brain* (p. 108-120). Macmillan Education UK.
- Vaissière, J. (2002). Cross-linguistic prosodic transcription : French vs. English. In N. Volskaya, N. Svetozarova, & P. Skrelin (Éds.), *Problems and methods of experimental phonetics. In honour of the 70th anniversary of Pr. L.V. Bondarko* (p. 147-164). St Petersburg State University Press.
- Vaissière, J. (2010). Le français, langue à frontières par excellence. In D. Delomier & M. Morel (Éds.), *Frontière, du linguistique au sémiotique* (p. 10-20). Lambert Lucas.
- Vaissière, J. (2020). *La phonétique* (4e éd.). Presses Universitaires de France.
- Vaissière, J., & Michaud, A. (2006). Prosodic Constituents in French : A Data-Driven Approach. In Y. Kawaguchi, I. Fónagy, & T. Moriguchi (Éds.), *Prosody and Syntax* (p. 47-64). John Benjamins Publishing Company.
- van Hell, J., & Tokowicz, N. (2010). Event-related brain potentials and second language learning : Syntactic processing in late L2 learners at different L2 proficiency levels. *Second Language Research*, 26(1), 43-74.
- van Heuven, V. (1994). What is the smallest prosodic domain. *Papers in laboratory phonology*, 3, 76-98.
- Wagner, P. (2010). A time-delay approach to speech rhythm visualization, modeling and measurement. In R. Michela (Éd.), *Prosodic Universals. Comparative studies in rhythmic modeling and rhythm typology* (p. 117-146). Aracne.
- Wagner, P., & Dellwo, V. (2004). Introducing YARD (yet another rhythm determination) and re-introducing isochrony to rhythm research. *Proc. Speech Prosody 2004*, 227-230.
- Wang, B., Kügler, F., & Genzel, S. (2022). The interaction of focus and phrasing with downstep and post-low-bouncing in Mandarin Chinese. *Frontiers in Psychology*, 13, 884102.
- Wang, B., Xu, Y., & Ding, Q. (2018). Interactive Prosodic Marking of Focus, Boundary and Newness in Mandarin. *Phonetica*, 75(1), 24-56.
- Wang, B., Yang, Y., & Lu, S. (2005). The acoustic characteristics of large information units' boundaries in monologue discourse. *ACTA ACUSTICA*, 30(2), 177-183.

- Wang, C., Xu, Y., & Zhang, J. (2019). MANDARIN AND ENGLISH USE DIFFERENT TEMPORAL MEANS TO MARK MAJOR PROSODIC BOUNDARIES. *Proceedings of the 19th International Congress of Phonetic Sciences*, 2906-2910.
- Wang, H. (2004). 试论汉语的节奏类型——松紧型. *语言科学*, 3(3), 21-28.
- Wang, M. (2016). A study on the realization of pitch in Mandarin Chinese intonation. *The International Journal of Engineering And Science (IJES)*, 5(4), 8-14.
- Warren, P., Grabe, E., & Nolan, F. (1995). Prosody, phonology and parsing in closure ambiguities. *Language and Cognitive Processes*, 10(5), 457-486.
- Watson, D., & Gibson, E. (2005). Intonational phrasing and constituency in language production and comprehension*. *Studia Linguistica*, 59(2-3), 279-300.
- Weber-Fox, C., & Neville, H. (1996). Maturational Constraints on Functional Specializations for Language Processing : ERP and Behavioral Evidence in Bilingual Speakers. *Journal of Cognitive Neuroscience*, 8(3), 231-256.
- Wenk, B., & Wioland, F. (1982). Is French really syllable-timed? *Journal of Phonetics*, 10(2), 193-216.
- Wennerstrom, A. (Éd.). (2001). *The music of everyday speech : Prosody and discourse analysis*. Oxford University Press.
- Wheeldon, L., & Lahiri, A. (1997). Prosodic Units in Speech Production. *Journal of Memory and Language*, 37(3), 356-381.
- White, E., Genesee, F., & Steinhauer, K. (2012). Brain Responses before and after Intensive Second Language Learning: Proficiency Based Changes and First Language Background Effects in Adult Learners. *PLoS ONE*, 7(12), e52318.
- White, L., & Malisz, Z. (2020). Chapter 11 Speech Rhythm and Timing. In C. Gussenhoven & A. Chen (Éds.), *The Oxford Handbook of Language Prosody* (p. 165-180). Oxford University Press.
- White, L., & Mattys, S. (2007a). Rhythmic typology and variation in first and second languages. In P. Prieto, J. Mascaró, & M. Solé (Éds.), *Current Issues in Linguistic Theory* (p. 237-257). John Benjamins Publishing Company.
- White, L., & Mattys, S. (2007b). Calibrating rhythm : First language and second language studies. *Journal of Phonetics*, 35(4), 501-522.
- Wiget, L., White, L., Schuppler, B., Grenon, I., Rauch, O., & Mattys, S. (2010). How stable are acoustic metrics of contrastive speech rhythm? *The Journal of the Acoustical Society of America*, 127(3), 1559-1569.
- Winkler, I., Kujala, T., Tiitinen, H., Sivonen, P., Alku, P., Lehtokoski, A., Czigler, I., Csépe, V., Ilmoniemi, R., & Näätänen, R. (1999). Brain responses reveal the learning of foreign language phonemes. *Psychophysiology*, 36(5), 638-642.
- Wioland, F. (1991). *Prononcer les mots du français. Des sons et des rythmes*. Hachette FLE.
- Wolff, S., Schlesewsky, M., Hirotsu, M., & Bornkessel-Schlesewsky, I. (2008). The neural mechanisms of word order processing revisited : Electrophysiological evidence from Japanese. *Brain and Language*, 107(2), 133-157.
- Xi, L., & Pillot-Loiseau, C. (2019). Acoustic and perceptive studies of Chinese learners producing three basic French intonation patterns. In S. Calhoun, P. Escudero, M. Tabain, & P. Warren (Éds.), *Proceedings of the 19th International Congress of Phonetic Sciences* (p. 2263-2267).
- Xi, L., Wachs, S., & Ridouane, R. (2020). Production of French final stressed syllables in Accentual Phrase by Chinese learners : A pilot study. *Proc. 10th International Conference on Speech Prosody 2020*, 895-899.

- Xie, H. (2014). *Études acoustique et perceptive du transfert rythmique du chinois mandarin au français dans l'apprentissage phonétique du Français Langue Étrangère* [Mémoire de Master non publié]. Université Sorbonne Nouvelle.
- Xu, Y. (2015). Chapter 36 Intonation in Chinese. In W. Wang & C. Sun (Éds.), *The Oxford Handbook of Chinese Linguistics* (p. 490-502). Oxford University Press.
- Xu, Y., & Wang, M. (2009). Organizing syllables into groups—Evidence from F0 and duration patterns in Mandarin. *Journal of Phonetics*, 37(4), 502-520.
- Xue, J., Yang, J., Zhang, J., Qi, Z., Bai, C., & Qiu, Y. (2013). An ERP study on Chinese natives' second language syntactic grammaticalization. *Neuroscience Letters*, 534, 258-263.
- Yang, X., Shen, X., Li, W., & Yang, Y. (2014). How Listeners Weight Acoustic Cues to Intonational Phrase Boundaries. *PLoS ONE*, 9(7), e102166.
- Yang, Y., & Wang, B. (2002). Acoustic correlates of hierarchical prosodic boundary in Mandarin. *Proc. Speech Prosody 2002*, 707-710.
- Yip, M. (2002). *Tone*. Cambridge University Press.
- Ylinen, S., Strelnikov, K., Huotilainen, M., & Näätänen, R. (2009). Effects of prosodic familiarity on the automatic processing of words in the human brain. *International Journal of Psychophysiology*, 73(3), 362-368.
- Zagar, D., Pynte, J., & Rativeau, S. (1997). Evidence for Early Closure Attachment on First Pass Reading Times in French. *The Quarterly Journal of Experimental Psychology Section A*, 50(2), 421-438.
- Zawiszewski, A., Gutiérrez, E., Fernández, B., & Laka, I. (2011). Language distance and non-native syntactic processing: Evidence from event-related potentials. *Bilingualism: Language and Cognition*, 14(3), 400-411.
- Zhang, J., Duanmu, S., & Chen, Y. (2020). Chapter 22 China and Siberia. In C. Gussenhoven & A. Chen (Éds.), *The Oxford Handbook of Language Prosody* (p. 332-343). Oxford University Press.
- Zhang, N., & Zhang, Q. (2019). Rhythmic pattern facilitates speech production : An ERP study. *Scientific Reports*, 9(1), 12974.
- Zhang, W., & Kappenman, E. (2024). Maximizing signal-to-noise ratio and statistical power in ERP measurement: Single sites versus multi-site average clusters. *Psychophysiology*, 61(2), e14440.
- Zhi, N., Bertinetto, P., & Bertini, C. (2011). MODELING THE SPEECH RHYTHM OF BEIJING CHINESE IN THE CCI FRAMEWORK. *Proc. ICPHS XVII*, 2316-2319.
- Zhu, X., & Wang, C. (2015). Chapter 37 Tone. In W. Wang & C. Sun (Éds.), *The Oxford Handbook of Chinese Linguistics* (p. 503-515). Oxford University Press.

L'acquisition de la prosodie du français par des apprenants sinophones : Études acoustique, perceptive et électroencéphalographique

Bien que centrale dans la communication parlée, la prosodie reste un domaine moins exploré dans les travaux de recherche sur l'acquisition des langues secondes. Cette thèse traite de la segmentation et de l'organisation prosodiques en français chez des locuteurs natifs francophones et des apprenants sinophones de trois niveaux (débutants, intermédiaires, avancés).

La prosodie assume notamment deux rôles clés : délimiter les unités syntaxiques et structurer le flux temporel-rythmique. Le mandarin, utilisant des tons lexicaux, mobilise différemment ces éléments, comparé au français. Notre recherche cherche à montrer comment les apprenants sinophones acquièrent la prosodie française pour traiter des ambiguïtés syntaxiques locales, illustrées par des phrases à clôture précoce (CP : « *Chaque fois que le serpent mangeait, le rat se cachait.* ») ou tardive (CT : « *Chaque fois que le serpent mangeait le rat, le lapin se cachait.* »). Pour cela, des données acoustiques (31 locuteurs : 10 natifs, 21 apprenants), perceptives (60 auditeurs : 20 natifs, 40 apprenants) et électroencéphalographiques (60 participants : 20 natifs, 40 apprenants) ont été analysées, totalisant 151 individus. Les résultats acoustiques révèlent que dès le niveau intermédiaire, les apprenants utilisent la montée du F0 et l'allongement final pour signaler les frontières CP et CT. Concernant l'organisation prosodique, les valeurs de neuf paramètres quantitatifs (%V, ΔC et ΔV , VarcoV et VarcoC, nPVI-V et rPVI-C, CCI-C et CCI-V) ont révélé une même tendance dans l'évolution du patron rythmique : il devient de plus en plus stable à mesure que le niveau de français progresse. Sur le plan perceptif, les apprenants (et, dans une moindre mesure, les natifs) rencontrent des difficultés accrues avec les structures complexes comme la CP. Les données EEG montrent que les frontières CP et CT activent une *closure positive shift* (CPS) chez tous les groupes, confirmant leur traitement neurocognitif. Une neuroplasticité est suggérée par les variations topographiques des CPS entre natifs et apprenants durant le traitement de la CP. La *ramp-like negativity*, observée pour la CT, reflète un processus *top-down* chez les natifs contre un mécanisme *bottom-up* chez les apprenants lors de l'analyse des indices acoustiques.

En conclusion, cette recherche contribue à la compréhension de l'acquisition prosodique en L2 et met en lumière les différences comportementale et neurocognitive dans le traitement des frontières prosodiques localement ambiguës entre natifs et apprenants.

Mots clés : français, chinois mandarin, prosodie, frontière prosodique, acquisition, acoustique, perception, EEG, potentiel évoqué, CPS

Acquisition of French prosody by Chinese learners: acoustic, perceptive and electroencephalographic studies

Despite its central role in spoken communication, prosody remains a less explored area in second language acquisition research. This thesis investigates performance of French prosodic segmentation and organization in native speakers and Chinese L2 learners across three proficiency levels: beginner, intermediate, and advanced. Prosody serves two primary functions in language, namely, delimiting syntactic units and regulating the temporal flow of speech. Mandarin, using lexical tones, mobilizes prosodic features differently from French. This thesis aims to investigate how Chinese L2 learners acquire French prosody to deal with local syntactic ambiguities, as exemplified by sentences exhibiting early closure (EC: « *Chaque fois que le serpent mangeait, le rat se cachait.* », "Whenever the snake was eating, the rat would hide.") and late closure (LC: « *Chaque fois que le serpent mangeait le rat, le lapin se cachait.* », "Whenever the snake was eating the rat, the rabbit would hide."). The analysis involved acoustic (31 subjects: 10 natives, 21 learners), perceptual (60 participants: 20 natives, 40 learners), and electroencephalographic (60 participants: 20 natives, 40 learners) data, encompassing a total of 151 individuals.

The acoustic results reveal that learners employ F0 rising and final lengthening to signal EC and LC boundaries from the intermediate level onward. In terms of prosodic organization, nine rhythm metrics (%V, ΔC and ΔV , VarcoV and VarcoC, nPVI-V and rPVI-C, CCI-C and CCI-V) show a consistent trend in rhythmic pattern development: it becomes increasingly more stable as proficiency in French advances. Perceptually, learners (and to a lesser extent, natives) encounter challenges with complex structures like EC. The EEG data indicate that both EC and LC boundaries elicited a closure positive shift (CPS) across all participant groups, underscoring their neurocognitive processing. Neuroplasticity is implied by the distinct topography of CPS between natives and learners during EC processing. The ramp-like negativity observed in LC condition reflects a top-down process in natives versus a bottom-up mechanism in learners when deciphering prosodic cues.

In summary, this thesis enhances our comprehension of L2 prosodic acquisition and provides novel perspectives on the behavioral and neurocognitive differences in the interpretation of local ambiguous prosodic boundaries between native French speakers and Chinese L2 learners.

Keywords: French, Mandarin Chinese, prosody, prosodic boundary, acquisition, acoustics, perception, EEG, ERP, CPS