



HAL
open science

Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole

Fanny Guitard-Ivent

► **To cite this version:**

Fanny Guitard-Ivent. Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole. Linguistique. Université Sorbonne Paris Cité, 2018. Français. NNT : 2018USPCA077 . tel-02147399

HAL Id: tel-02147399

<https://theses.hal.science/tel-02147399>

Submitted on 4 Jun 2019

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.

UNIVERSITÉ SORBONNE PARIS CITÉ
UNIVERSITÉ SORBONNE NOUVELLE – PARIS 3
École Doctorale 268 : Langage et langues
Laboratoire de Phonétique et Phonologie (UMR 7018)



THÈSE

en vue de l'obtention du grade de docteur en phonétique

présentée par
Fanny GUITARD-IVENT

COARTICULATION C-à-V EN FRANÇAIS : INTERACTION AVEC LE TYPE DE VOYELLE, LA POSITION PROSODIQUE ET LE STYLE DE PAROLE

Sous la direction de
Mme Cécile FOUGERON

Soutenue à Paris le 12 septembre 2018 devant le jury composé de :

Olivier CROUZET (examinateur), Maître de Conférence à l'Université de Nantes

Cécile FOUGERON (Directrice de thèse), Directeur de Recherche CNRS, LPP, Paris

Jonathan HARRINGTON (Pré-rapporteur), Professeur à l'Université de Munich

Christine MEUNIER (Pré-rapporteur), Chargé de Recherche (HDR) CNRS, LPL, Aix-en-Provence

Daniel RECASENS (Président, examinateur), Professeur à l'Université Autonome de Barcelone

2018

COARTICULATION C-à-V EN FRANÇAIS : INTERACTION AVEC LE TYPE DE
VOYELLE, LA POSITION PROSODIQUE ET LE STYLE DE PAROLE

Fanny GUITARD-IVENT



RÉSUMÉ

Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole

Cette thèse étudie la coarticulation C-à-V en français et son interaction avec d'autres sources de variations dans le but de mieux comprendre ce qui la module et ce qui gouverne la variation dans la parole. Pour cela, à partir de grands corpus de parole, nous avons testé comment la coarticulation C-à-V était fonction : 1) des caractéristiques articulatoires des consonnes et voyelles impliquées à partir de 18.5k voyelles /i, e, ε, a, x, u, o, ə/ (/x/= /ø, œ, ə/) en contexte ALVéolaire, UVulaire et VÉLaire ; 2) de la position prosodique occupée par les voyelles, en comparant le degré de coarticulation de 17k séquences CV et VC, V=/i, e, a, ə/ et C=ALV|UV, en position initiale de groupe intonatif, avec celui de séquences semblables en position interne de mot ; et 3) du style de parole, en analysant le degré de coarticulation dans 22k séquences CV et VC, V= /i, E, a, u, ə/ (/E/= /e, ε/) et C=ALV|UV, issues de parole journalistique et conversationnelle. Cette thèse montre qu'en plus de dépendre des caractéristiques articulatoires des segments, la coarticulation est aussi modulée par des facteurs linguistiques, liés à l'organisation prosodique du message, et des facteurs communicationnels dépendant de la situation de communication. Cependant, certains résultats suggèrent que la modulation de la coarticulation par la position prosodique et le style de parole, ont des fonctions linguistiques différentes dont les implications sur la variation dans la parole seront discutées. Enfin, une réflexion sur les changements de sons en lien avec la préférence universelle pour l'antériorisation des voyelles postérieures fermées sera proposée à partir des différences observées entre les voyelles.

Mots clés : grand corpus, acoustique, coarticulation C-à-V, position prosodique, style de parole, français, voyelles

ABSTRACT

C-to-V coarticulation in French: interaction with vowel type, prosodic position and speech style

This dissertation examines C-to-V coarticulation in French and its interaction with others sources of variation in order to better understand what modulates and governs variation in speech. Based on data from large speech corpora, we tested how C-to-V coarticulation is a function of: 1) the articulatory properties of the tested segments, i.e. 18.5k vowels /i, e, ε, a, x, u, o, ə/ (/x/=/ø, œ, ə/) in ALveolar, UVular et VELar contexts; 2) the prosodic position occupied by the vowels, comparing the degree of coarticulation of 17k CV and VC sequences V=/i, e, a, ə/ and C=ALV|UV in initial position of intonational phrases, to similar sequences in internal word position; 3) the speech style, by analyzing the degree of coarticulation in 22k CV and VC sequences, V = /i, E, a, u, ə/ (/E/ = /e, ε/) and C = ALV|UV, in journalistic and conversational speech. The thesis demonstrates that coarticulation, in addition to being dependent on the articulatory characteristics of segments, is also modulated by other linguistic factors, related to the prosodic organization of the message, and communicational factors depending on the communication situation. Indeed, vowels are more resistant to coarticulation in strong prosodic positions as in formal speech. However, some results suggest that the modulation of coarticulation by prosodic position and speech style have different linguistic functions whose implications for speech variation will be discussed. Finally, a reflection on sound changes related to the universal preference for the anteriorization of back closed vowels will be proposed from the observed differences between the vowels.

Key words: large corpora, acoustics, C-to-V coarticulation, prosodic position, speech style, French, vowels

À ma mère, dont l'énergie et la douceur ne cesseront de résonner en moi...

À Théa

REMERCIEMENTS

Je tiens tout d'abord à remercier infiniment Cécile Fougeron pour avoir encadré ce travail et m'avoir permis de suivre ce cheminement scientifique. Tes idées, conseils, critiques et nombreux retours m'ont tellement appris que je ne saurais comment t'en remercier. Merci de ta bienveillance et de ta disponibilité que je trouve admirables.

J'exprime toute ma gratitude à l'ensemble des membres de mon jury de thèse Christine Meunier, Jonathan Harrington, Daniel Recasens et Olivier Crouzet devant qui j'ai l'honneur de soutenir cette thèse.

Je remercie particulièrement Giuseppina Turco avec qui j'ai collaboré à plusieurs reprises dans le cadre de ce travail. Giusy, merci pour ces échanges enrichissants qui m'ont tant appris. Ce fut un plaisir de travailler avec toi. J'espère que cela continuera...

Je suis également très reconnaissante envers Martine Adda-Decker pour ses conseils précieux dont j'ai pu bénéficier au début de ma thèse et pour m'avoir initiée à l'analyse des grands corpus oraux.

Un grand merci à l'ensemble de l'équipe du LPP pour l'accueil toujours très agréable et les retours dont j'ai pu bénéficier durant l'élaboration de ce travail. Je remercie particulièrement mes voisins de bureau avec qui j'ai toujours eu plaisir à discuter : Jane que je félicite pour sa belle réussite, Yaru (ma jumelle de thèse), Djedjiga, Bowei, Emre et Rosario.

Merci à ma copine Nora pour ces bons moments passés ensemble et ses coups de pouce en anglais.

Je remercie particulièrement mes relecteurs qui ont révisé ce document en un temps record : Sofie, Jean et Valérie. Merci aussi à vous trois de vos encouragements et de l'intérêt que vous avez toujours porté pour mon travail !

Sans oublier les amis d'ici et d'ailleurs pour être là, pour leurs encouragements et leur confiance. À mes amies de toujours Pauline et Maëli, aux tautavellois Céline et Nico, mais aussi aux parisiens-landais Alex, Christelle et Rémy, à la funky family, sans oublier les toulousaines Laura et Vanessa.

Puis, je remercie bien sûr :

- Mon petit papa pour son amour et sa confiance qui m'ont toujours portée. Je ne te remercierai jamais assez de tout ce que tu as fait pour moi.
- ma sœur Emeline, mon beau frère Alex et mes neveux Célia, Mathis et Mahé. Tout en suivant de loin mon parcours, votre présence à mes côtés et moments partagés ensemble, ont été précieux pour me faire avancer.
- Flo qui se demande comment on peut avoir autant de choses à dire sur un son et surtout qui me soutient depuis toujours.

Enfin et surtout, merci à mon formidable mari, aussi super-papa, super-cuistot, super prof... bref, l'Homme de la situation ! Ces derniers mois n'ont pas été de tout repos pour toi non plus et cette thèse, je te la dois. Tu m'as permis de me consacrer pleinement à ma rédaction sans cesser de m'encourager tout en gérant mon stress. Quelle chance j'ai de t'avoir, je t'aime !

Ah et toi ma Théa, tu n'as pas vraiment compris tout ce qu'il se passait ces derniers temps mais tu comprenais l'essentiel, comme tu me le montrais en me demandant nuit et jour : « tu [pavajø] maman ? ». Quel plaisir j'aurais maintenant de te répondre : « NON ! ». Juste merci de me faire oublier tout ce que j'ai pu faire pendant des heures, dès l'instant où je te vois...

Fanny

TABLE DES MATIÈRES

CHAPITRE 1. INTRODUCTION GÉNÉRALE	1
1.1. INTRODUCTION À LA COARTICULATION	2
1.1.1. Terminologie	2
1.1.2. Les différents types de coarticulation	3
1.1.3. La coarticulation : un phénomène variable	4
1.2. LA COARTICULATION VARIE : POURQUOI ?	7
1.2.1. Contraintes perceptives	7
1.2.2. Contraintes biomécaniques : la durée	10
1.3. LA SENSIBILITÉ À LA COARTICULATION EST-ELLE LIÉE AUX CARACTÉRISTIQUES ARTICULATOIRES DES CONSONNES ET VOYELLES IMPLIQUÉES ?	13
1.3.1. Des sons plus ou moins agressifs et résistants à la coarticulation	13
1.3.2. La coarticulation : fonction de la stabilité intrinsèque des voyelles	18
1.4. LA SENSIBILITÉ À LA COARTICULATION EST-ELLE LIÉE À POSITION PROSODIQUE OCCUPÉE PAR LES SEGMENTS ?	25
1.4.1. La prosodie et les variations segmentales	25
1.5. LA SENSIBILITÉ À LA COARTICULATION EST-ELLE LIÉE À DES FACTEURS COMMUNICATIONNELS COMME LE STYLE DE PAROLE ?	34
1.5.1. Qu'est-ce que le style de parole ?	34
1.5.2. Style de parole et coarticulation	35
1.6. LES GRANDS CORPUS ORAUX	41
1.6.1. Quelques notes sur les grands corpus oraux	41
1.6.2. Les corpus sélectionnés pour nos études	43
1.7. RÉSUMÉ DES OBJECTIFS DE LA THÈSE	46
CHAPITRE 2. ÉTUDE DE L'EFFET DU TYPE DE VOYELLE SUR LA COARTICULATION C-À-V	48
2.1. QUESTIONS DE RECHERCHE & HYPOTHÈSES	49
2.1.1. Prédications	50
2.1.2. Intéractions CV	54
2.2. METHODE	57
2.2.1. Matériel linguistique et prétraitement	57

2.2.2.	Métriques acoustiques	63
2.2.3.	Analyses Statistiques	67
2.3.	RESULTATS	69
2.3.1.	Quels déplacements formantiques selon les propriétés de C & V ?	69
2.3.2.	La coarticulation joue-t-elle sur la variabilité des voyelles ?	75
CHAPITRE 3. ÉTUDE DES INFLUENCES PROSODIQUES SUR LA COARTICULATION C-À-V		95
3.1.	QUESTIONS ET OBJECTIFS DE L'ÉTUDE	96
3.2.	MÉTHODE	98
3.2.1.	Matériel linguistique et prétraitements	98
3.2.2.	Mesure de coarticulation	102
3.2.3.	Analyses statistiques	103
3.3.	RÉSULTATS	105
3.3.1.	Étude préliminaire : contrôle de la durée	105
3.3.2.	Modulation de la coarticulation par la PP en séquence vc et cv	108
3.3.3.	Modulation de la coarticulation par la PP : effet dépendant de la voyelle ?	112
CHAPITRE 4. ÉTUDE DES INFLUENCES STYLISTIQUES SUR LA COARTICULATION C-À-V		118
4.1.	QUESTIONS ET OBJECTIFS DE L'ÉTUDE	119
4.2.	METHODE	121
4.2.1.	Matériel linguistique et pretraitements	121
4.2.2.	Mesure de coarticulation & normalisation des données	125
4.2.3.	Analyses statistiques	125
4.3.	RESULTATS	127
4.3.1.	La voyelle /i/	127
4.3.2.	La voyelle /E/	128
4.3.3.	La voyelle /a/	129
4.3.4.	La voyelle /ɔ/	130
4.3.5.	La voyelle /u/	132
4.3.6.	les effets observés dépendent-ils de la voyelle testée ?	133
CHAPITRE 5. DISCUSSION		137
5.1.	QUESTIONS RELATIVES À LA TEMPORALITÉ DE LA PAROLE	139

5.1.1.	Une modulation de la coarticulation par le style de parole malgré le contrôle de la durée des voyelles	139
5.1.2.	Une modulation de la coarticulation par la position prosodique malgré un contrôle de la durée sur un sous-échantillon de données	140
5.2.	LA COARTICULATION : UN PHÉNOMÈNE VARIABLE DÉPENDANT DE LA VOYELLE ET DE LA DIRECTION DES EFFETS COARTICULATOIRES	142
5.2.1.	Des effets contextuels dépendants des voyelles	142
5.2.2.	Une modulation de la coarticulation dépendante de la directionnalité de la coarticulation	154
5.3.	UNE PLUS GRANDE RÉSISTANCE DES VOYELLES À LA COARTICULATION EN PAROLE FORMELLE ET EN IPi : QUELLE(S) FONCTION(S) LINGUISTIQUE(S) ?	160
5.4.	RÉFLEXION SUR LES POSSIBLES IMPLICATIONS DE CERTAINS RÉSULTATS SUR LES CHANGEMENTS DE SONS	164
	CHAPITRE 6. CONCLUSION	171
	BIBLIOGRAPHIE	175
	ANNEXES	188
	ANNEXE I. Détail du matériel linguistique utilisé dans l'étude I.	189
	ANNEXE II. Contenu de l'échantillon de référence : corpus BREF de parole lue.	203
	ANNEXE III. Analyse Linéaire Discriminante : détail des valeurs du classifieur automatique construit à partir de l'échantillon de référence sur le corpus de parole lue.	212

TABLE DES FIGURES

FIGURE 1. Représentations schématiques, tirées de Meunier (2005) de deux types de variations étudiées : a) DISPERSION intra-catégorie et b) DÉPLACEMENTS.	20
FIGURE 2. Illustration de la relation non linéaire, existant entre paramètre articuloire et paramètre acoustique, sur laquelle repose la Théorie Quantique. L'image est extraite de Stevens (1972).	21
FIGURE 3. Nomogramme de Fant pour la voyelle /a/ illustrant la correspondance entre le rapprochement des deux premiers formants et la zone de stabilité (plateau) acoustique malgré les variations d'un paramètre articuloire (le lieu d'articulation).	23
FIGURE 4. Illustration des séquences sélectionnées après un premier filtrage des données.	58
FIGURE 5. Illustration des contextes consonantiques testés dans cette étude.	61
FIGURE 6. Illustration de la mesure des effets contextuels sur les déplacements formantiques des voyelles.	63
FIGURE 7. Représentation schématique, sur F1/F2, de la mesure de variabilité intra-catégorie utilisée dans cette thèse.	64
FIGURE 8. F1 (z-score) des voyelles en fonction du contexte consonantique (ALV, VEL, UV).	70
FIGURE 9. F2 (z-score) des voyelles en fonction du contexte consonantique.	72
FIGURE 10. F3 (à gauche) et F4 (à droite) en z-score de la voyelle /i/ en fonction du contexte consonantique.	75
FIGURE 11. Racine carrée des distances euclidiennes F1-F2-F3-F4 (en z-score) par voyelle.	77
FIGURE 12. Distances aux centroïdes des voyelles par formant a) F1, b) F2, c) F3, d) F4.	81
FIGURE 13. Confusions entre voyelles issues de la classification. Les chiffres rapportés sont en pourcentages et indiquent que le type de confusion représente x% des erreurs pour la voyelle en question. Les voyelles sont regroupées en trois catégories selon un code couleur (jaune pour les postérieures, rouge pour les centrales et gris pour les antérieures).	84
FIGURE 14. Racine carrée des distances euclidiennes F1-F2-F3-F4 des voyelles en fonction du contexte consonantique (alv, vel, uv).	88
FIGURE 15. Caractéristiques de nos deux positions prosodiques IPI et Wm en fonction du type de séquence (VC et CV).	100
FIGURE 16. Spectrogrammes illustrant les deux positions prosodiques testées avec la voyelle /a/ en contexte alvéolaire pour les séquences VC. La position IPI est illustrée avec le mot « <u>attends</u> » et la position Wm avec le mot « <u>privatisa</u> tion » où la consonne à gauche de V est une labiale.	101

FIGURE 17. Spectrogrammes illustrant les variations formantiques de la voyelle /a/ en fonction du contexte consonantique: a) alvéolaire (suite /dap/ extraite du mot « adapter ») et b) uvulaire (suite /ɣav/ extraite du mot « grave ») en séquence CV ₁ . Le contexte droit étant une labiale.....	103
Figure 18. Distribution de la durée de la voyelle /a/ en séquence VC sur la tranche 60-80 ms par position prosodique (à gauche) et résumé du matériel disponible (à droite).....	106
FIGURE 19. F1 (a) et F2 (b) en Bark de la voyelle /a/ (70-80 ms) en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence V ₁ C.....	107
FIGURE 20. F1 (a) et F2 (b) en Bark de la voyelle /a/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence V ₁ C.....	110
FIGURE 21. F1 (a) et F2 (b) en Bark de la voyelle /a/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence CV ₁	111
FIGURE 22. F1 (a) et F2 (b) en Bark de la voyelle /i/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence V ₁ C.....	113
FIGURE 23. F1 (a) et F2 (b) en Bark de la voyelle /e/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence V ₁ C.....	114
FIGURE 24. F1 (a) et F2 (b) en Bark de la voyelle /ɔ/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence V ₁ C.....	116
FIGURE 25. Spectrogrammes illustrant les séquences CV et VC pour la voyelle /a/ par contexte consonantique (alvéolaire vs. uvulaire) extraites du corpus de parole conversationnelle : le contexte alvéolaire en VC est illustré avec le mot « <u>formation</u> » et en séquence CV « <u>tables</u> » ; le contexte uvulaire en contexte VC est illustré avec le mot « <u>pareil</u> » et en séquence CV « <u>rapporter</u> ».	123
FIGURE 26. Distribution de la durée des voyelles en fonction du style de parole (J ou C) par voyelle (/i, E, a, ɔ, u/) et par type de coarticulation (VC ou CV).....	125
FIGURE 27. F2 (en z-score) de la voyelle /i/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel).....	127

FIGURE 28. F2 (en z-score) de la voyelle /E/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel.....	129
FIGURE 29. F2 (en z-score) de la voyelle /a/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel.....	130
FIGURE 30. F2 (en z-score) de la voyelle /ɔ/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel.....	131
FIGURE 31. F2 (en z-score) de la voyelle /u/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel.....	132
FIGURE 32. Écart-types (ET) des distances entre F2-F1 (ligne pleine) et F3-F4 (ligne en pointillé) pour les différentes voyelles étudiées.....	145
FIGURE 33. Triangles vocaliques sur les dimensions : a) F1/F2, b) F2/F3, c) F3/F4.....	147
FIGURE 34. F2 (en z-score) des voyelles : a) /i/, b) /a/, c) /u/, d) /E/ (/E/ = /e, ε/), e) /ɔ/, provenant des corpus ESTER et NCCFr, en fonction du contexte consonantique (alvéolaire vs. uvulaire) et de la durée des voyelles (courte = 50 ms, moyenne = 60-80 ms, longue = 90 – 150 ms) selon d'autres études sur le français menées sur des corpus semblables (Gendrot & Adda-Decker, 2005) ou identiques (Audibert et al., 2015).	153
FIGURE 35. Fréquences d'occurrence des voyelles en contextes gauche et droit de labiales, d'alvéolaires, de postalvéolaires, de /R/, de la liquide /l/ et de palato-vélaires mesurées sur le corpus journalistique. a) : le contexte gauche correspond à la consonne /R/ ; b): le contexte gauche correspond aux alvéolaires. Pour chaque contexte gauche les histogrammes indiquent la distribution des voyelles sur un des trois corpus étudiés en fonction des 6 classes de consonnes en contexte droit. Graphiques tirés de Adda-Decker (2006).....	168

LISTE DES TABLES

TABLE 1. Résumé des facteurs liés à la situation de communication cités dans la littérature selon le principe d'Hypo- et Hyper-articulation.....	8
TABLE 2. Couples CV compatibles (dont les trajectoires articulatoires ne sont pas en conflit) sur les dimensions hauteurs de la langue et position de la langue.	17
TABLE 3. Résumé des prédictions des différentes théories abordées : la théorie quantique (QT à gauche), la Théorie de la Dispersion-Focalisation (TDF, au milieu) et la théorie du Degré de Contrainte Articulatoire (DAC, à droite), en ce qui concerne la stabilité des voyelles.	23
TABLE 4. Résumé des résultats des études portant sur l'interaction PP et coarticulation. Quelques précisions sur les sigles utilisés : P=phrase, GI=Groupe Intonatif, GA=Groupe Accentuel, M=Mot ; Fmaj=Frontière majeure ; Fmin = Frontière mineure. PP=Position prosodique.	29
TABLE 5. Résumé des travaux traitant de l'interaction entre style de parole et coarticulation. COA DIR. indique la direction de la coarticulation pers = persévératrice et ant = anticipatrice. Mes. = mesure effectuée.	37
TABLE 6. Résumé des particularités de chacun des deux corpus étudiés.....	43
TABLE 7. Résumé des prédictions des différentes théories abordées : la théorie quantique (QT à gauche), la Théorie de la Dispersion-Focalisation (TDF, au milieu) et la théorie du Degré de Contrainte Articulatoire (DAC, à droite), en ce qui concerne la stabilité des voyelles.	50
TABLE 8. Prédiction des effets des consonnes sur les formants des voyelles.....	54
TABLE 9. Couples CV compatibles sur les dimensions hauteurs de la langue et position de la langue.....	54
TABLE 10. Résumé des prédictions des effets contextuels sur la stabilité des voyelles.....	56
TABLE 11. Valeurs limites des 3 premiers formants de chaque voyelle testée pour les hommes et les femmes pour le filtrage des valeurs aberrantes selon Gendrot & Adda-Decker (2005).	58
TABLE 12. Moyenne (M), écart-type (ET) et seuils minimum (min) et maximum (max) correspondant à $M \pm 2ET$ pour chaque voyelle et nombre d'exemplaires rejetés et pourcentage correspondant entre parenthèses après application de ce filtre.....	60
TABLE 13. Détail de la procédure de sélection des données étape par étape et effectifs.	62
TABLE 14. Résumé du matériel sélectionné.	62
TABLE 15. Moyennes et écarts-types des formants de chaque voyelle de l'échantillon de référence construit à partir de parole lue.....	67

TABLE 16. Résultats de l'analyse des contrastes du facteur CONTEXTE.....	70
TABLE 17. Résultats de l'analyse des contrastes du facteur contexte par voyelle sur F1.....	71
TABLE 18. Résultats de l'analyse des contrastes du contexte sur le F2 des voyelles.....	72
TABLE 19. Résultats de l'analyse des contrastes du facteur contexte par voyelle sur F2.....	73
TABLE 20. Résultats de l'analyse des contrastes du facteur contexte sur le F3 et le F4 de la voyelle /i/......	75
TABLE 21. Échelle de stabilité établie à partir des résultats obtenus de l'analyse des distances euclidiennes F1-F2-F3-F4, et report des prédictions initiales selon les différentes théories (QT, TDF, DAC).....	78
TABLE 22. Échelle de stabilité établie à partir des résultats obtenus de l'analyse de la dispersion des voyelles par formant F1, F2, F3 et F4, et report des prédictions initiales selon les différentes théories (QT, TDF, DAC) et des résultats de l'analyse multidimensionnelle (dispersion sur F1,2,3,4).....	79
TABLE 23. Scores de sensibilité et de précision par catégorie de voyelles, issus de l'analyse linéaire discriminante.....	83
TABLE 24. Échelle de stabilité des voyelles selon leur distinctivité dérivée du score <i>F1</i> (mêlant sensibilité et précision du modèle) de la classification et report des prédictions initiales selon les différentes théories (QT, TDF, DAC) et des résultats de l'analyse multidimensionnelle (dispersion sur F1,2,3,4).....	86
TABLE 25. Résultats de l'analyse des contrastes du facteur CONTEXTE.....	88
TABLE 26. Résultats de l'analyse des contrastes du facteur CONTEXTE par voyelle.....	89
TABLE 27. Matrices de confusions (%) de notre classification automatique par contexte consonantique (ALV en haut, UV au milieu et VEL en bas).....	93
TABLE 28. Résumé du matériel linguistique sélectionné et distribution des 16617 voyelles /a, i, e, ə/ étudiées par conditions de structure syllabique (CV ou VC) et lieu d'articulation de la consonne C (ALV=alvéolaire, UV=uvulaire).....	102
TABLE 29. Résumé du matériel linguistique utilisé dans cette étude et distribution des voyelles dans chaque contexte consonantique (ALV = alvéolaire et UV = uvulaire) par type de coarticulation (CV et VC) issues du corpus de parole conversationnelle (C) à gauche et du corpus de parole journalistique (J) à droite.....	124
TABLE 30. Test de vraisemblance des modèles mixtes linéaires testant les effets principaux des prédicteurs CONTEXTE CONSONANTIQUE*SEQUENCE*STYLE pour les cinq voyelles testées: / i, E, a, ə, u /. Les valeurs χ^2 (avec un degré de liberté) sont rapportées en tant qu'estimations de la taille de l'effet. « ALV » et « UV » pour alvéolaire et uvulaire, "J" et "C" pour le style journalistique et conversationnel. ALV>UV, par exemple, signifie que F2 de V est plus élevé lorsque V est en contexte alvéolaire qu'en contexte uvulaire. Les valeurs p significatives sont indiquées par le symbole astérisque * (p <.05 = *)......	133

TABLE 31. Test de vraisemblance des modèles mixtes linéaires testant l'interaction entre CONTEXTE CONSONANTIQUE*SEQUENCE*STYLE pour les 5 voyelles: / i, E, a, ə, u / . Les valeurs de χ^2 (avec un degré de liberté) sont rapportées en tant qu'estimations de la taille d'effet; Les coefficients β , les erreurs-types (ET) et les valeurs t du modèle sont également rapportés. Les valeurs p significatives sont indiquées par le symbole * ($p < .05 = *$).....	134
TABLE 32. Test de vraisemblance des modèles mixtes linéaires testant l'interaction entre CONTEXTE CONSONANTIQUE*TYPE DE COARTICULATION*STYLE pour les 5 voyelles: / i, E, a, ə, u / . Les valeurs de χ^2 (avec un degré de liberté) sont rapportées en tant qu'estimations de la taille d'effet; Les coefficients β , les erreurs-types (ET) et les valeurs t du modèle sont également rapportés. Les valeurs p significatives sont indiquées par le symbole * ($p < .05 = *$).....	135
TABLE 33. Résumé de l'ensemble des résultats obtenus dans cette thèse. COA = coarticulation, PP = position prosodique, IPi = Position initiale de groupe intonatif, Wm = position interne de mot, J = parole journalistique, C = parole conversationnelle, C _{ALV} = contexte alvéolaire, C _{VEL} = contexte vélaire, C _{UV} = contexte uvulaire, $\text{su} = \text{su}$	138
TABLE 34. Rappel des prédictions des différentes théories de la parole (qt = théorie quantique, tdf = théorie de la dispersion-focalisation, dac = degré de contrainte articulatoire) sur la stabilité des voyelles.	143
TABLE 35. Résultats de notre étude préliminaire visant à déterminer si certaines voyelles sont plus stables que d'autres (DF1234 = Dispersion multidimensionnelle, DF1 = Dispersion sur F1, DF2 = Dispersion sur F2, DF3 = Dispersion sur F3, DF4 = Dispersion sur F4, Distinctivité à partir du score $F1$).	143
TABLE 36. Résultats des déplacements formantiques causés par les différents contextes consonantiques.....	148
TABLE 37. Résumé des effets du contexte consonantique (ALV = alvéolaire, VEL = vélaire et UV = uvulaire) sur la dispersion des voyelles.	151
TABLE 38. Taux de classification correcte (ou score de sensibilité) en pourcentage des voyelles tous contextes confondus (TCC) et par contexte : ALV = alvéolaire, VEL = vélaire et UV = uvulaire.....	152
TABLE 39. Résumé des résultats de l'effet de la position prosodique (PP) et du style de parole (STYLE) sur l'acoustique des voyelles testées /i, e, a, ə/ pour PP (IPi = position initiale de groupe intonatif, Wm = position interne de mot) et /i, e, a, u, ə/ (/ɛ/ = /e, E/) pour le STYLE (J = parole journalistique, C = parole conversationnelle, J > C indique que le F2 de la voyelle en question est plus élevé en parole journalistique).	160

INTRODUCTION GÉNÉRALE

Dans ce chapitre, nous allons tout d'abord définir la coarticulation dans sa pluralité selon sa directionnalité et sa portée, puis nous montrerons que la coarticulation est un phénomène variable (1.1). Cela nous amènera à nous questionner sur les raisons possibles de ces variations. On s'intéressera aux contraintes perceptives et biomécaniques relatives à la durée des segments motivant le choix des facteurs étudiés dans cette thèse (1.2). Ainsi on cherchera à savoir si la sensibilité à la coarticulation est liée : a) aux caractéristiques articulatoires des segments impliqués (1.3) ; b) à la position prosodique occupée par les segments (1.4) ; et c) à des facteurs communicationnels comme le style de parole (1.4). Ensuite, on montrera pourquoi les grands corpus oraux offrent des perspectives intéressantes dans l'étude de la variation à partir d'une brève introduction à la linguistique de corpus (1.5). Après avoir présenté les corpus sélectionnés pour nos études, nous terminerons ce chapitre d'introduction par un résumé des objectifs de la thèse (1.6).

I.1. INTRODUCTION À LA COARTICULATION

I.1.1. TERMINOLOGIE

La coarticulation concerne les variations articulatoires et acoustiques subies par un son au contact d'un son voisin. Ces variations contextuelles sont dues au fait que des sons successifs sont articulés ensemble (co-produits / co-articulés). Cela est la conséquence de la continuité et de la non segmentalité de la parole :

« [...] *the movements of different articulators for the production of successive phonetic segments overlap in time and interact with one another: as a consequence, the vocal tract configuration at any point in time is influenced by more than one segment.* » (Farnetani & Recasens, 2010 : 316).

La coarticulation est un phénomène universel dans le sens où elle répond à des contraintes physiologiques ne permettant pas l'articulation séquentielle des segments de manière indépendante. Les variations contextuelles sont de ce fait graduelles et peuvent dépendre de l'action, voire de l'interaction de différents facteurs.

Dans le cas où la variation est catégorielle, on parle alors d'assimilation. Cela renvoie à l'adoption d'une ou plusieurs caractéristiques d'un son (cible) par un son voisin (déclencheur). De ce fait, l'assimilation serait la contrepartie phonologique de la coarticulation. Cependant, la question relative à 'est-ce que coarticulation et assimilation renvoient à des processus différents' fait encore débat. Les théories de la coarticulation ont toutes un point de vue différent sur cette question selon leur conception de la coarticulation et notamment en ce qui concerne le domaine d'application, la fonction de celle-ci et son contrôle (Farnetani & Recasens, 2010). Cette thèse n'adopte pas un cadre théorique particulier, mais nous traitons du phénomène qu'est la coarticulation et des questions qu'il soulève en ce qui concerne les théories phonétiques. Pour cela, nous ne ferons pas une revue de la littérature des différents modèles existants, selon les points abordés dans cette introduction, des aspects de certains modèles seront développés.

1.1.2. LES DIFFÉRENTS TYPES DE COARTICULATION

Si la tendance est de parler de « la coarticulation » au singulier, pour traiter de l'ensemble des phénomènes de variations contextuelles, il est évident qu'il n'existe pas une coarticulation mais des coarticulations selon la portée de celle-ci et direction des effets coarticulatoires.

En effet, il convient de distinguer les phénomènes de coarticulation locale comme la coarticulation consonne à voyelle C-à-V étudiée dans cette thèse, des phénomènes de coarticulation à distance du type V-à-V. Un exemple de coarticulation locale serait l'antériorisation de la voyelle /u/ au contact d'une consonne alvéolaire comme dans le mot *toute* [tʊt]. Les travaux de Öhman (1966) ont permis d'aller plus loin en ce qui concerne la portée de la coarticulation. À partir des transitions de F2 en séquence V_1CV_2 , l'auteur a montré que les transitions de V_2 influençaient V_1C et inversement. Ainsi, les voyelles seraient produites en continu avec superposition des gestes consonantiques.

On distingue deux types de coarticulation en fonction de la directionnalité de celles-ci. La coarticulation anticipatrice concerne les cas où un son est affecté par le son qui le suit. Si on a deux sons A et B, la coarticulation anticipatrice concerne les variations acoustiques de A sous l'influence du son suivant B que l'on peut écrire $A < B$ (< indique la directionnalité). Ce type de coarticulation est aussi appelé coarticulation régressive. En anglais, on trouve aussi les termes « left-to-right » et « forward ». Un exemple de coarticulation anticipatrice est l'anticipation de l'articulation labiale des voyelles durant la production de la fricative précédente en séquences [si] et [su]. Dans le premier cas, la fricative est produite avec les lèvres étirées par anticipation de la voyelle /i/ alors que dans le second cas, on aura un /s/ labialisé du type [s^w]. La coarticulation persévératrice ou progressive (aussi appelée « left-to-right », « backward » ou « carry-over »), traite du prolongement d'un son durant la production du son suivant du type $A > B$. Le dévoisement du /ʁ/ précédé d'une occlusive sourde en français est un exemple de coarticulation persévératrice comme dans *truc* où le /ʁ/ se dévoise au contact de l'occlusive sourde [tʁyk].

La directionnalité de la coarticulation est importante car elle renvoie à des processus différents. La coarticulation persévératrice serait la conséquence de l'inertie des articulateurs et donc purement mécanique, alors que la coarticulation anticipatrice serait planifiée puisque le locuteur anticipe l'articulation d'un son pendant la production d'un son le précédant. La direction de la coarticulation est aussi importante pour décrire les effets qu'un son particulier aura sur les sons environnants. Le modèle DAC, pour Degré de Contrainte Articulaire, de coarticulation (Recasens, Pallarès & Fontdevila, 1997) que nous décrirons plus en détail (section 1.3.1.1) porte une attention particulière à la directionnalité des patterns coarticulatoires. D'après les auteurs, selon leurs caractéristiques articulatoires en lien avec des contraintes articulatoires pour leur production, certains sons causeront plus de coarticulation anticipatrice que persévératrice ou inversement. Par exemple, une consonne apicale causera plus d'effets coarticulatoires sur le qui la précède (coarticulation anticipatrice) que sur celui qui la suit (coarticulation persévératrice). Alors que les consonnes dorsales auront un effet semblable sur les sons voisins, quelle que soit leur position vis-à-vis de la consonne, qu'ils la précèdent ou la suivent (Recasens, 2014).

1.1.3. LA COARTICULATION : UN PHÉNOMÈNE VARIABLE

Bien qu'universelle la coarticulation varie en fonction des langues (Öhman 1966 ; Clumeck, 1976 ; Manuel & Krakow, 1984 ; Cohn, 1990 ; Manuel 1990 ; Beddor, Harnsberger & Lindemann, 2002 ; Ma, Perrier & Dang, 2009). Par exemple, Manuel et Krakow (1984) ont montré que le degré de coarticulation V-à-V était inversement corrélé à la taille de l'inventaire phonémique de la langue en comparant une langue ayant un système vocalique dense, l'anglais (13 à 15 voyelles selon le dialecte) avec deux langues Bantoue à 5 voyelles (le shona et swahili). Ainsi, la coarticulation V-à-V était plus importante pour les langues Bantoues ayant un inventaire vocalique réduit, et donc peu de compétiteurs acoustiques entre les voyelles, qu'en anglais où la marge de manœuvre entre les voyelles est faible du fait de la forte densité du système vocalique.

La taille du système phonémique ne semble pas être la seule raison des variations observées entre les langues. En effet, il semblerait qu'il existe des variations de coarticulation entre différents dialectes d'une même langue (Recasens, 2014) associées à des stratégies articulatoires différentes. Les consonnes /s/ et /ʃ/ sont plus proches l'une de l'autre en catalan valencien qu'en catalan de l'est, et cette différence articulatoire a un impact sur le degré de résistance à la coarticulation des voyelles à l'une ou aux deux fricatives. Plus les consonnes sont palatales, plus leur effet coarticulatoire sur les voyelles environnantes est fort et donc, moins les voyelles sont résistantes aux effets coarticulatoires de ces fricatives.

Au sein d'un même système linguistique et d'une même variété dialectale, des différences entre locuteurs peuvent être aussi observées. La revue de la littérature faite par Kühnert et Nolan (1997) montre que ces patterns coarticulatoires particuliers aux locuteurs vont au-delà de simples différences anatomiques et seraient plutôt des caractéristiques idiosyncratiques (Nolan, 1983 ; 1985). Ceci étant, comme le soulignent les auteurs, peu d'études portent sur le sujet et celles qui existent, traitent souvent d'un nombre restreint de locuteurs. De plus, la pertinence des corrélats acoustiques de la coarticulation pour la reconnaissance du locuteur fait débat : certaines études affirment que c'est un indice utile (Su, Li & Fu, 1974) alors que d'autres indiquent le contraire (Kahn, 2011).

Au-delà des propriétés idiosyncratiques, le comportement coarticulatoire des locuteurs varie en fonction de leur âge. Les études sur le sujet se centrent principalement sur l'acquisition des comportements coarticulatoires des enfants en comparant la coarticulation des enfants avec celle des adultes. Dans leur revue de la littérature, Kühnert et Nolan (1997) montrent qu'il existe des positions opposées vis-à-vis de l'acquisition des patterns coarticulatoires des enfants. La première indique que les enfants produisent une parole plus séquentielle, et donc moins coarticulée, que les adultes (Kent, 1983 ; Katz, Kripke & Tallal, 1991) alors que d'autres trouvent au contraire que les enfants coarticulent plus que les adultes (Nittrouer, Studdert-Kennedy & McGowan, 1989 ; Nittrouer & Walhen, 1989). Tout cela a bien sûr des implications théoriques sur l'acquisition du langage mais nous ne rentrerons pas dans ce débat (voir la revue de la littérature citée ci-dessus pour plus de détails à ce sujet). Ceci étant, comme le rapporte Recasens (2018) dans une revue

de la littérature sur la coarticulation, les résultats de certaines études articulatoires (ultrason) récentes confirment l'hypothèse d'un plus fort degré de coarticulation chez l'enfant (Zharkova, Hewlett & Hardcastle, 2011 ; 2012) alors que d'autres ne trouvent pas de différences entre enfants et adultes (Noiray, 2013 ; Goffman, Smith, Heisler & Ho 2008).

La coarticulation est aussi influencée par des facteurs lexicaux comme la fréquence lexicale du mot contenant les segments coarticulés (Bybee, 2001 ; Shi, Gick, Kanwischer & Wilson, 2005 ; Scarborough & Zellou, 2013 ; Renwick & Cassidy, 2015). Par exemple, dans une étude acoustique menée sur de la parole spontanée Renwick et Cassidy (2015) ont montré que le degré de palatalisation de la fricative /s/ finale de mot au contact d'un mot commençant par la semi-consonne /j/ était corrélé à la fréquence lexicale du mot. Plus un mot était fréquent, plus le degré de palatalisation du /s/ augmentait. Scarborough (2004), a aussi montré que la coarticulation était fonction de la densité lexicale. Plus la densité lexicale est forte, plus le mot le moins fréquent présentera un taux de coarticulation important. Ce résultat est indépendant du type de coarticulation étudié (coarticulation nasale ou voyelle à voyelle), de la direction (anticipatrice ou persévératrice), de la prééminence syllabique (syllabe accentuée vs. non accentuée) et de la langue (français et anglais).

Dans cette thèse, nous nous intéressons à trois facteurs en particulier pouvant affecter la coarticulation C-à-V que nous présenterons plus en détail sections 1.3, 1.4 et 1.5 : 1) le type de voyelle et les contraintes de productions spécifiques aux segments (C et V) impliqués (variation phonético-phonologique) ; 2) la position prosodique occupée par les segments (variation liée à la structuration du message linguistique) ; et 3) le style de parole (variation liée à la situation de communication).

1.2. LA COARTICULATION VARIE : POURQUOI ?

Dans cette section, nous nous intéresserons aux raisons pour lesquelles la coarticulation peut varier sous l'effet de différentes sources de variation, justifiant le choix des facteurs d'intérêt de cette thèse que nous décrirons par la suite. Nous nous focaliserons sur deux points en particulier : les contraintes perceptives et les contraintes biomécaniques relatives à la durée des segments.

1.2.1. CONTRAINTES PERCEPTIVES

Le modèle *Hyper-& Hypo-speech* ou H&H (Lindblom, 1990) met la situation de communication au centre des problématiques liées à la variation et donc à la variation contextuelle (coarticulation). Ainsi, la variation serait le résultat d'une adaptation continue de la production de la parole selon les besoins / contraintes de la situation de communication. Cette théorie repose sur deux principes : l'économie de l'effort (orienté locuteur) et la clarté (orienté auditeur). Le locuteur va essayer de minimiser le coût de sa production, ce qui aura pour conséquence de créer de la réduction dans la parole. Mais la réduction du coût de production ne doit pas se faire au détriment de l'intelligibilité du message, qui reste l'objectif premier de toute communication. Le locuteur s'adapte donc à la situation de communication et au besoin de l'auditeur selon le principe de clarté. Selon ces deux principes, toute production linguistique se place sur un continuum allant de l'hypo-articulation, traduisant une économie maximale, à l'hyper-articulation, où la clarté est maximale, en fonction des besoins imposés par la situation de communication comme on peut le voir Table 1.

FACTEUR	HYP0- ARTICULATION	HYPER- ARTICULATION	RÉFÉRENCES
STYLE DE PAROLE	relâché	soutenu	Smiljanić, R., & Bradlow, A.R (2005); Johnson, et al., 1993 ; Picheny et al., 1986 ; Audibert et al., 2015 ; Gendrot & Adda-Decker, 2005.
BRUIT ENVIRONNEMENTAL	absence	présence	Lane & Tranel, 1971 ; Lane et al., 1970 ; Lombard, 1911
INFORMATIONS VISUELLES	présence	absence	Anderson et al., 1997
PRÉDICTIBILITÉ À PARTIR DU CONTEXTE	forte	faible	Lieberman, 1963 ; Bard & Anderson, 1983 ; Fowler & Housum, 1987 ; Hunnicutt, 1985
POIDS INFORMATIONNEL	faible	fort	Chafe, 1974; Hawkins & Warren, 1994

TABLE 1. Résumé des facteurs liés à la situation de communication cités dans la littérature selon le principe d'Hypo- et Hyper-articulation

Selon cette théorie, **la fonction de la coarticulation serait l'économie de l'effort** en adoptant une stratégie de production orientée locuteur comme l'expliquent très clairement Farnetani & Recasens :

« Coarticulation, manifested as a reduced displacement and a shift of articulatory movements towards the contextual phonetic segments, is a low-cost motor behavior, an economical way of speaking. Its pervasiveness indicates that the speech motor system, like other kinds of motor behavior, is governed by the principle of economy. » (2010 : 329)

Ainsi, la coarticulation peut varier en fonction de facteurs mettant en jeu la distinctivité / l'intelligibilité des segments. C'est pour cette raison que dans cette thèse, nous chercherons à mieux comprendre comment le style de parole et la position prosodique (voir section 1.4 et 1.5 pour les questions relatives à ces facteurs de variation) affectent le degré de coarticulation C-à-V.

En effet, la question relative au style de parole prend tout son sens dans la perspective de cette théorie. S'il convient de se demander pourquoi le style de parole devrait affecter la

coarticulation, la réponse apparaît clairement ici. En style de parole relâchée où la stratégie de production est plus orientée locuteur, l'économie de l'effort est privilégiée sur la clarté (voir Table 1) et la coarticulation augmentée par rapport à une parole plus formelle.

Cependant, les études traitant du rôle de la coarticulation dans la parole semblent indiquer que la coarticulation soit une **stratégie de production orientée auditeur**. En effet, **la pertinence de la coarticulation pour le processus de perception de la parole est désormais bien établie** (Ali, Gallagher, Goldstein & Daniloff, 1971 ; Beddor, 2009 ; Whalen, 1984 ; Fowler, 1984 ; Assman, 1995 ; Meunier & Floccia, 1997, 1999). Pour une revue des travaux de psycholinguistique montrant le rôle des effets coarticulatoires dans la reconnaissance des mots, voir Nguyen (2001).

L'implication de ces résultats n'est pas sans conséquences sur les prédictions que l'on peut faire en relation du style et de la coarticulation. Si la coarticulation facilite le processus de perception de la parole, alors elle devrait se trouver augmentée en parole plus formelle dans laquelle le système de production est plus orienté auditeur. Ceci dit, les études traitant de l'effet du style de parole sur l'acoustique des voyelles indiquent une hyper-articulation des segments en parole formelle. Par 'hyper-articulation', comprendre un renforcement des caractéristiques phonétiques des sons, comme la réalisation de voyelles plus périphériques (Gendrot & Adda-Decker, 2005 ; Audibert, Fougeron, Gendrot & Adda-Decker, 2015 parmi d'autres), le fait d'atteindre pleinement ses cibles articulatoires *overshot*. Ainsi, il paraît difficile qu'un locuteur puisse à la fois coarticuler et hyper-articuler pour répondre aux caractéristiques de la parole formelle. Sauf si l'articulateur est libre (hyper-articulation labiale accompagnée d'une coarticulation nasale ou labiale par exemple). Les résultats de Scarborough et Zellou (2013) illustrent très bien ce fait. Les auteurs trouvaient qu'en parole claire, issue d'une situation de communication authentique, les locuteurs produisaient des voyelles plus périphériques et augmentaient en même temps leur degré de coarticulation nasale C-à-V. Comme l'anticipation du geste nasal n'entre pas en compétition avec la réalisation du geste lingual pour l'articulation de la voyelle, le locuteur peut renforcer les caractéristiques orales des voyelles tout en anticipant le geste nasal de la consonne surtout qu'en anglais, il n'existe pas de voyelle nasale. En

revanche, si les auteurs s'étaient intéressés à la coarticulation linguale de la consonne sur la voyelle, ils auraient probablement trouvé des résultats opposés. L'anticipation du geste lingual de la consonne entrant en compétition avec l'articulation de la voyelle, celle-ci n'aurait pu se faire qu'au détriment de la préservation de la voyelle. On peut donc imaginer que, même si la coarticulation sert à la perception, le locuteur réduise le degré de coarticulation en parole claire afin de préserver et de renforcer les caractéristiques acoustiques des sons lorsque les segments concernés sollicitent l'action d'un même articulateur. D'un autre côté, le locuteur augmentera la coarticulation d'articulateurs n'entrant pas en conflit (nasalité de la consonne, labialité de la voyelle) afin de faciliter l'identification des mots.

La fonction perceptive du renforcement prosodique initial, autre facteur étudié dans cette thèse, n'a jamais été établie même si les résultats des études sur le sujet suggèrent que ce phénomène contribue à maximiser les contrastes syntagmatiques et paradigmatiques de la langue (voir section 1.4). On peut donc imaginer que le renforcement des contrastes ait une fonction perceptive. Le renforcement des segments initiaux contribuerait à marquer la frontière et ainsi à aider les auditeurs à analyser le signal de parole en facilitant la segmentation lexicale. Cette hypothèse n'est pas nouvelle puisqu'elle a été émise dès les premiers travaux portant sur le sujet (Fougeron & Keating, 1997). D'ailleurs, des résultats donnent du poids à la fonction perceptive du renforcement initial (Cho, 2006). Selon ces faits, on pourrait s'attendre à une réduction de la coarticulation en position initiale de constituant prosodique.

1.2.2. CONTRAINTES BIOMÉCANIQUES : LA DURÉE

La durée des segments est la cause de nombreuses variations dans la parole. On sait notamment que les voyelles courtes sont plus fortement influencées par le contexte consonantique (Lindblom, 1963 ; Vaissière, 1985 ; Gendrot & Adda-Decker, 2005). Selon Lindblom, l'amplitude des déplacements de formants dépend de la durée de la voyelle et de la distance entre la cible de la consonne et celle de la voyelle. Lorsqu'un articulateur est sollicité par des commandes motrices successives qu'il doit réaliser dans un intervalle

temporel court, par exemple en débit de parole rapide, il n'a pas le temps de réaliser pleinement une première commande, qu'il doit déjà produire la seconde. Ainsi, les différentes commandes sont produites simultanément conduisant à des raccourcissements des voyelles et des déplacements formantiques. C'est aussi cette vue qu'adopte la Phonologie Articulatoire (Browman & Goldstein, 1990b ; 1992) pour rendre compte de la coarticulation :

« In casual rapid speech, subsequent consonantal gestures can so far overlap as to hide each other when they occur on different tiers, or to completely blend their characteristics when they occur on the same tier. » (Farnetani et Recasens, 2010, page 343)

Ainsi, les variations de la coarticulation dues à la durée des segments sont la conséquence de contraintes biomécaniques. Or, on sait que différents facteurs affectent la temporalité de la parole. C'est le cas par exemple du débit de parole, du style de parole ou de la position prosodique. La coarticulation pourra donc varier sous l'effet de ces facteurs agissant sur la temporalité de la parole.

Cependant, la coarticulation ne peut être que la conséquence de contraintes temporelles. En effet, si plusieurs études confirment bien le lien existant entre coarticulation et débit de parole (Hardcastle, 1985 ; Matthies et al., 2001 ; Agwuele et al., 2008), celui-ci n'a pas toujours pu être établi (Kuehn & Moll, 1976 ; Gay, 1978). Dans une étude récente sur le catalan, Recasens (2015) montre aussi que le débit de parole joue sur la coarticulation mais il n'affecte pas de la même manière tous les segments. Certains sons, connus pour résister plus fortement aux influences contextuelles, restent plus résistants à la coarticulation malgré l'augmentation du débit de parole. Il semblerait donc que la durée des segments, fortement dépendante du débit de parole, ne soit pas la seule cause de la coarticulation (en ce qui concerne la sensibilité spécifique des segments à la coarticulation, nous en reparlerons en 1.3).

Toutes ces observations montrent que la durée des segments détermine bon nombre de variations dans la parole mais elle n'explique pas tous les phénomènes observés. La durée interagit avec d'autres sources de variation et contribue à dessiner les contours du signal produit, mais on ne peut pas résumer le signal sortant au simple produit de l'inertie des

articulateurs. Dans cette thèse, nous contrôlerons autant que possible la durée des segments et nos résultats seront discutés en tenant compte de ce paramètre.

1.3. LA SENSIBILITÉ À LA COARTICULATION EST-ELLE LIÉE AUX CARACTÉRISTIQUES ARTICULATOIRES DES CONSONNES ET VOYELLES IMPLIQUÉES ?

Dans cette section, nous traiterons de la coarticulation en tant que phénomène de variation structurelle. Ce terme a été utilisé par Meunier (2005, page 365) pour décrire « les formes de variation issues de la structure même de la langue. », soit les variations résultantes de « l'organisation structurelle des sons dans une langue ». Ce qui nous intéresse particulièrement ici, est de voir comment selon les caractéristiques articulatoires des sons et du degré de conflit existant entre eux (sont en conflit des sons dont la trajectoire articulatoire s'oppose, terme que nous définirons plus en détail 1.3.1.1), la coarticulation varie.

1.3.1. DES SONS PLUS OU MOINS AGRESSIFS ET RÉSISTANTS À LA COARTICULATION

Le terme de « résistance à la coarticulation » a été développé par Bladon et Al-bamerni (1976) lors d'une étude acoustique des allophones du phonème /l/ anglais ([l] clair vs. [ɫ] sombre), rapportant une plus grande résistance à la coarticulation du [ɫ] comparé au [l]. Ce qui veut dire que le [ɫ] est moins sensible aux effets du contexte segmental que le [l].

1.3.1.1. Une sensibilité à la coarticulation dépendante du degré de contrainte articulatoire

Depuis, cette notion de « résistance à la coarticulation » a été très employée dans l'étude de la coarticulation, et notamment par Recasens et ses collaborateurs dans le cadre du modèle DAC, pour degré de contrainte articulatoire, de coarticulation (Recasens & al., 1997). Selon ces auteurs, le degré de résistance à la coarticulation d'un segment serait inversement corrélé à son niveau de contrainte articulatoire. Ceci veut dire que plus un segment est articulatoirement contraint, et particulièrement au niveau du dos de la langue, plus celui-ci est résistant à la coarticulation. Au même titre qu'il est résistant à la coarticulation, un segment tendra plus fortement à imposer ses caractéristiques phonétiques aux segments

environnants. Il sera ainsi à la fois résistant à la coarticulation et agressif envers les segments voisins (Farnetani, 1990 ; Récasens & Espinosa, 2009).

Afin de classer les sons selon leur sensibilité à la coarticulation (résistance / agressivité) Recasens et al. (1997) proposent une classification en termes de Degré de Contrainte Articulaire. Un DAC allant de 1 à 3 est attribué à chaque son (voyelles comme consonnes), plus le DAC est élevé, plus le son est articulairement contraint et plus il sera résistant aux variations contextuelles. Les consonnes alvéo-palatales du type /ʃ/ seront maximalement contraintes (DAC=3) du fait que leur articulation nécessite une élévation du dos de la langue importante et une large zone de contact linguo-palatale. De même, une consonne avec une double articulation, comme le /w/ ayant une constriction vélaire et une constriction labiale, sera très contrainte. En revanche, une consonne bilabiale, sera très faiblement contrainte puisque son articulation n'implique pas de déplacement du dos de la langue (/p/=1). Le mode d'articulation peut aussi influencer sur le degré de contrainte articulaire. Par exemple, les consonnes alvéolaires comme /n/ ou /s/ sont moyennement contraintes (DAC=2) du fait que le dos de la langue ne soit pas directement impliqué dans leur articulation, mais qu'un couplage avec l'articulateur principal (l'apex de la langue) opère. Cependant, du fait de la précision articulaire requise pour l'articulation des fricatives, un DAC de 3 sera assigné à la consonne /s/ alors que d'autres consonnes alvéolaires comme /n/ ou /t/ présenteront un DAC intermédiaire de 2 (voir Récasens et al., 1997).

Le modèle DAC prévoit aussi une classification des voyelles sur la même base que celle des consonnes. Pour cela, Recasens et ses collaborateurs, adoptent la classification de Wood (1979) spécifiant 4 lieux d'articulations pour les voyelles : le palais dur /i, e, ε/ ; le palais mou /u/ ; le haut pharynx /o, ɔ/ ; et le bas pharynx /a/. Ainsi les voyelles palatales, au même titre que les consonnes palatales, sont très contraintes articulairement au niveau du dos de la langue, ont un DAC élevé et sont plus résistantes à la coarticulation que les autres voyelles. La voyelle /i/ est encore plus contrainte étant donné le rétrécissement important de la constriction. On notera cependant qu'en français, la constriction du /i/ est plutôt pré-palatale, zone articulaire où le F3 est maximalement haut et donc plus proche

de F4 alors que pour un /i/ palatal c'est F2 qui est maximalelement élevé et se rapproche de F3 (Vaissière, 2009). On peut donc se demander si le degré d'antériorisation du /i/ a un impact sur le degré de contrainte articulatoire de la voyelle et donc de sa résistance à la coarticulation. Pour revenir à la classification du DAC, le schwa articulatoirement neutre (aucune constriction dans le conduit vocal) est minimalelement contraint et sera plus sensible aux effets contextuels que les autres voyelles. Les voyelles postérieures sont contraintes au niveau de l'articulation labiale mais très peu contraintes au niveau du corps de la langue et auront un DAC intermédiaire de 2, le même que celui de la voyelle /a/ (Recasens et al., 1997) même si plusieurs études placent cette voyelle au dessus des voyelles postérieures en termes de résistance à la coarticulation (Recasens & Espinosa, 2009).

Au-delà du DAC respectif des segments, un élément fondamental déterminant les effets contextuels concerne la compatibilité mutuelle des gestes articulatoires dans la séquence (Recasens, 1984b). Lorsqu'un même articulateur est impliqué dans la réalisation de deux segments adjacents et que le mouvement qu'il doit produire pour atteindre chacune des deux cibles est contraire, l'articulateur n'aura pas d'autre choix que d'effectuer une accommodation pour rendre possible la réalisation des deux gestes dits *incompatibles*. Cette notion de compatibilité mutuelle, fait écho à la notion de « conflit » introduite par Bell-Berti & Harris (1981) reprise par Recasens (1985) selon lequel ce terme manque d'une définition précise. Ce dernier la considère de la manière suivante :

« [...] articulatory gesture for given phonemes differ as to the degree of compatibility with respect to adjacent gestures. Thus, it is claimed that the degree of constraint determines the extent to which an articulatory gesture conflict with an adjacent gesture. » (Recasens, 1985 : page 99).

Les effets contextuels dépendront donc du DAC des segments impliqués et de leur trajectoire articulatoire respective. Pour les cas de coarticulation C-à-V, dont il est question dans cette thèse, si C et V ont des trajectoires comparables, à DAC égal ou si le DAC de la voyelle est supérieur à celui de la consonne comme dans /ʃi/ (DAC C = DAC V = 3) ou /ni/ (DAC C = 2 et DAC V = 3) respectivement, la voyelle sera peu affectée par la consonne. En revanche, si la consonne a un DAC supérieur à celui de la voyelle comme dans la séquence /la/ avec un /l/ sombre [ɫ] (DAC C = 3 et DAC V = 2) alors la

voyelle subira des effets coarticulaires considérables de la consonne. Si les trajectoires de C et V s'opposent et que la consonne a un DAC supérieur à celui de la voyelle /ɲa/ (DAC C = 3 et DAC V = 2), les effets coarticulaires de C sur V augmentent avec le degré de conflit entre les segments. Pour les cas où C et V ont un DAC comparable comme dans /li/ lorsque /l/ est sombre [ɫ] (DAC C = DAC V = 3), les effets coarticulaires sont aussi prééminents et cela est fonction du degré de conflit existant entre C et V en lien avec leur trajectoire articulatoire. Enfin, lorsque la voyelle a un DAC supérieur à celui de la consonne, les effets coarticulaires C-à-V sont insignifiants.

Une autre acceptation du terme « conflit » concerne la distance existant entre la consonne et la voyelle. En effet, il a souvent été rapporté que l'amplitude des déplacements de formants dépendait de la distance entre le locus des consonnes et la cible des voyelles (Lindblom, 1963 ; Stevens & House, 1963). Comme deux segments, impliquant un même articulateur mais dans des directions opposées, sont en conflit, on peut voir deux segments articulatoirement très éloignés, concernant notamment l'articulation linguale, comme étant aussi en conflit. Ainsi une consonne articulée à l'avant du conduit vocal serait en conflit avec une voyelle postérieure et inversement.

Les deux conceptions du « conflit » présentées ici (compatibilité mutuelle et distance CV) sont très proches l'une de l'autre mais pas équivalentes. Par exemple la distance séparant les consonnes alvéolaires de la voyelle /ɛ/ est courte mais ces consonnes sont peu compatibles avec cette voyelle du fait que la voyelle /ɛ/ requiert un abaissement important de la partie antérieure de la langue et une rétraction de celle-ci.

Dans cette thèse nous ne distinguerons pas ces deux aspects et nous parlerons de conflits (et par opposition de compatibilité) entre C et V lorsque les distances entre C et V sont importantes ou que le degré de compatibilité entre les deux segments est faible selon leur articulation linguale. On distinguera les dimensions de hauteur de la langue et de position afin justement de pouvoir graduer ces effets. La Table 2 présente les couples CV, selon les consonnes et voyelles testées dans cette thèse, n'étant pas en conflit et donc compatibles selon les dimensions de hauteur et de position de la langue.

COMPATIBILITÉ ENTRE C ET V:

	Hauteur de la langue -y	Position de la langue - x
ALVÉOLAIRE	/i, u, e, o/	/i, e, ε/
VÉLAIRE	/i, u, e, o/	/u/ + neutre autres V
UVULAIRE	/ε, a, ɔ/	/u, o, ɔ/

TABLE 2. Couples CV compatibles (dont les trajectoires articulatoires ne sont pas en conflit) sur les dimensions hauteurs de la langue et position de la langue.

Ce tableau sera commenté plus en détail au début de notre première étude section 2.1.

1.3.1.2. Une sensibilité à la coarticulation selon la notion de chemin du modèle Window et de l'économie de l'effort

Un autre modèle de coarticulation qui traite notamment des principes de résistance et d'agression est le « Window model » développé par Keating (1985, 1988, 1990a). Dans ce modèle la notion de cible en tant que point fixe à atteindre n'existe pas. Elle est remplacée par la notion de fenêtre « window » qui correspond à une plage de variation plus ou moins grande, définie par un seuil minimum et un seuil maximum. Chaque segment possède une « window » sur chacune des dimensions articuloire comme acoustique. La largeur de la fenêtre dépendra principalement de la spécification du trait concerné pour chaque segment. Si un segment est spécifié pour le trait X, alors la fenêtre sera étroite et le segment tolérera peu de variation contextuelle sur cette dimension. Le deuxième aspect important de ce modèle est le « path ». Cet aspect concerne le rapport entre les segments. Pour une dimension donnée, chacun des trois segments d'une suite CVC par exemple aura une fenêtre bien définie et le but de l'articulation sera de connecter ses « window » ensemble en respectant le principe d'économie. Le chemin choisi pour connecter les différentes fenêtres devra être celui qui minimise l'effort articuloire selon le principe développé par Lindblom (1990). Selon ce modèle, un segment avec une fenêtre étroite sera résistant aux effets contextuels et sera plus agressif envers les segments environnants.

1.3.2. LA COARTICULATION : FONCTION DE LA STABILITÉ INTRINSÈQUE DES VOYELLES

L'idée que certaines voyelles sont intrinsèquement plus stables que d'autres est au cœur de différentes théories de la production de la parole. L'objectif de cette partie est de montrer en quoi ces théories suggèrent des différences de stabilité entre les voyelles et de dresser, à partir de ces postulats, des échelles de variabilité pour les voyelles du français que nous testerons dans cette étude. Le premier point sera bien sûr de définir ce que l'on entend par « stable », « stabilité » pour pouvoir ensuite en discuter.

1.3.2.1. Questions relatives à la notion de « stabilité »

i. Qu'entendons-nous par « plus stable » ?

Le mot « stable » renvoie dans son acceptation la plus générale à un état, une situation qui se maintient durablement selon les définitions qu'en donne le Centre National de Ressources Textuelles et Lexicales (CNRTL) ou le Larousse en ligne. Une des définitions qu'en donne le CNRTL est relative à « une idée d'absence de variation » à la fois dans le temps « conservant ses caractéristiques initiales, peu susceptibles de transformations » comme dans l'espace, avec une idée d'immobilité ou de sédentarité, pour un être inanimé « qui ne bouge pas, ne se déplace pas ». Selon cette définition générale du mot « stable », des voyelles plus stables seraient des voyelles qui ne varieraient pas, à la fois dans le temps, on pourrait y voir une certaine immuabilité sur la dimension diachronique, mais aussi dans l'espace, renvoyant à une certaine constance dans les réalisations acoustiques de certaines voyelles dont la cible serait toujours atteinte (pour une discussion sur la notion de « cible » voir plus bas).

ii. Stabilité et parole deux notions aux antipodes ?

Si la recherche d'invariants dans la parole a été pendant très longtemps au cœur des recherches dans la parole, il est communément accepté aujourd'hui que la parole est par essence variable faisant de la phrase « la parole est variable », je cite : « un véritable pléonasme » (Meunier, 2005, page 350). Ne serait-il donc pas vain de chercher à savoir si certaines voyelles sont stables (et donc ne varient pas) ? Et puis, stables par rapport à quoi ?

Comment définir une référence robuste lorsqu'on sait que deux réalisations d'une même voyelle d'un seul locuteur ne seront jamais identiques ?

La vision de la variation que l'on adopte dans cette étude est justement celle d'une parole par essence variable. Et nous pensons donc qu'il n'est pas nécessaire de définir une référence à partir de laquelle il y aurait de la variation. Cependant, on peut très bien imaginer que certains segments soient moins variables que d'autres et donc plus stables que d'autres. C'est d'ailleurs ce que semblent indiquer différentes théories de production de la parole comme on le verra dans la prochaine sous-section. C'est donc cette idée de stabilité que nous voulons étudier. On ne parlera pas de stabilité absolue, chose qui serait vaine, mais d'une stabilité relative (comprendre par rapport à d'autres éléments plus variables) et nous utiliserons les termes « stable » ou « stabilité », et « variable » ou « variabilité » comme des antonymes. Tester des mesures de variation sur les voyelles, nous permettra de discuter de la stabilité de celles-ci.

iii. *Comment décrire la variabilité, et donc la stabilité ?*

Lorsqu'on étudie la variation, il convient de définir quels sont les aspects de la parole qui varient, comment ils varient et pourquoi ils varient. Dans cette étude, on s'intéressera aux variations acoustiques des voyelles orales du français, selon les propriétés intrinsèques des voyelles. Aussi, on regardera si le contexte consonantique affecte la stabilité des voyelles. Par *plus stable* on entendra donc *plus stable acoustiquement, d'une répétition à une autre, et/ou, d'un contexte à un autre*. La variation sera étudiée sous différents angles renvoyant à différentes formes de variation (trois au total).

La première forme de variation d'intérêt est la DISPERSION acoustique des exemplaires des voyelles par rapport au centre de leur catégorie vocalique (voir Figure 1a). Cette forme de variation rend compte de l'importance de la variation au sein d'une catégorie (Meunier, 2005) et d'une certaine manière de la consistance dans leur réalisation (une voyelle ayant des exemplaires peu dispersés autour du centre de la catégorie, sera consistante dans sa réalisation acoustique. Plus une voyelle occupera un espace de variation important, plus ses exemplaires seront susceptibles d'être éloignés du centre de la catégorie et de s'approcher, voire d'empiéter sur des zones acoustiques des catégories vocaliques voisines. En ce sens on

considèrera une catégorie vocalique avec des exemplaires très dispersés comme peu stable. À l'inverse, une voyelle occupant un espace de variation restreint indiquera une faible dispersion des exemplaires au centre de la catégorie reflétant une certaine constance, ou stabilité, dans ses réalisations.

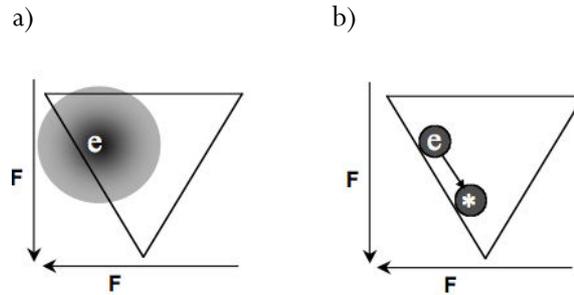


FIGURE 1. Représentations schématiques, tirées de Meunier (2005) de deux types de variations étudiées : a) DISPERSION intra-catégorie et b) DÉPLACEMENTS.

Le deuxième aspect de variation analysé sera la **DISTINCTIVITÉ** acoustique des voyelles. Est-ce que certaines voyelles se distinguent plus des autres à partir de leurs propriétés acoustiques ? Est-ce que certaines voyelles de par la zone acoustique qu'elles occupent sont plus distinctives que d'autres ? En effet, on peut imaginer qu'une catégorie de voyelle ayant des exemplaires très dispersés sera malgré tout distinctive si la distance acoustique qui sépare cette catégorie de voyelle d'une autre catégorie est importante. La distinctivité dépendrait donc plus du nombre de compétiteurs acoustiques que de la dispersion des exemplaires de la catégorie vocalique. De ce fait, une voyelle ayant de nombreux compétiteurs acoustiques aura moins de marge de manœuvre dans sa dispersion sans quoi son manque de constance se ferait au détriment de sa distinctivité.

Cette forme de variation étudiée permettra de rendre compte de la stabilité des différentes voyelles vis-à-vis du maintien de leur pouvoir distinctif malgré l'impact de différentes sources de variations interagissant dans le signal de parole et notamment les effets contextuels.

1.3.2.2. Les particularités des voyelles quantiques : une stabilité dans les rapports acoustico-articulatoires

Selon la **théorie quantique** de la parole (Stevens, 1972 ; 1989), les voyelles produites dans les régions naturelles du conduit vocal où la variation articulaire n'entraîne pas de changements importants dans la sortie acoustique, c'est-à-dire les régions quantiques, sont acoustiquement plus stables que les autres.

Cette théorie repose sur l'hypothèse que les relations entre les mouvements articulaires et leur sortie acoustique ne sont pas linéaires. Comme on peut le voir Figure 2, des mouvements articulaires importants peuvent ne pas engendrer de variation dans la sortie acoustique lorsqu'ils se font dans les zones I et III, alors que dans la zone II, un mouvement articulaire, même léger, provoquera une grande modification de la sortie acoustique. Ainsi, des voyelles produites dans des zones de type I et III, sont acoustiquement plus stables que les autres. C'est le cas des voyelles ayant une articulation « consonantique », comprendre produite avec une constriction importante en un point du conduit vocal, soit : les voyelles /i/ et /y/ présentant une constriction palatale ; la voyelle /u/ produite avec un resserrement du conduit vocal important au niveau du voile du palais ; et la voyelle /a/ produite avec une constriction au niveau du pharynx. Les formants de ces voyelles sont plus sensibles aux variations de degré de constriction que de lieu d'articulation, ce qui explique pourquoi ces voyelles seraient plus stables sur la dimension F2.

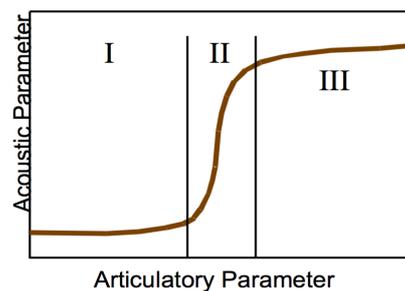


FIGURE 2. Illustration de la relation non linéaire, existant entre paramètre articulaire et paramètre acoustique, sur laquelle repose la Théorie Quantique. L'image est extraite de Stevens (1972).

1.3.2.3. Les particularités des voyelles focales : une stabilité dans les rapports acoustico-perceptifs

Si, selon les prévisions de la théorie quantique, la stabilité des voyelles dépend des relations entre articulatoire et acoustique, la **Théorie de la Dispersion-Focalisation** (Schwartz, Boë, Vallée & Abry, 1997) offre une réponse à partir du lien existant entre l'acoustique et la perception. En effet, selon cette dernière théorie, les voyelles focales, produites avec un rapprochement de deux formants, sont des voyelles plus saillantes perceptivement. Ceci s'explique du fait que les deux formants rapprochés voient leur amplitude renforcée et soient perçus comme un seul et même pic par l'oreille humaine.

C'est ce renforcement de l'amplitude d'une certaine zone fréquentielle qui est pertinent pour la perception de ces voyelles focales. Donc, même s'il y a de la variation dans le signal, du moment que ce principe de « focalité » est respecté, les voyelles seront bien perçues. Ainsi les voyelles focales sont, d'une certaine manière, perceptivement plus stables que les autres. Or, cette caractéristique focale n'est pas totalement éloignée de la théorie quantique, puisque la plupart des zones de stabilité acoustique correspondent à des zones de rapprochement de deux formants comme les illustrent très bien les nomogrammes de Fant (1960) comme illustré Figure 3 avec la voyelle /a/.

Les voyelles focales du français sont : la voyelle /i/ (rapprochement des formants F3 et F4), la voyelle /y/ sur les dimensions F2-F3, et enfin les voyelles postérieures /u, o, ɔ/ et la voyelle /a/ F1-F2 dont les deux premiers formants sont très proches (Vaissière, 1987).

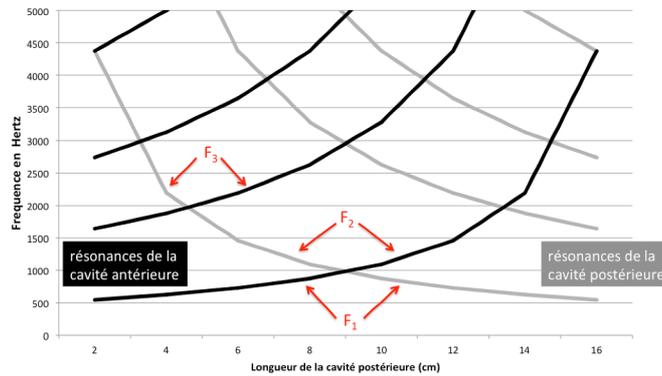


FIGURE 3. Nomogramme de Fant pour la voyelle /a/ illustrant la correspondance entre le rapprochement des deux premiers formants et la zone de stabilité (plateau) acoustique malgré les variations d'un paramètre articulaire (le lieu d'articulation).

I.3.2.4. Récapitulatif des échelles de stabilité prédites par les différents modèles

Que ce soit en raison d'une relation articulaire-acoustique, ou acoustico-perceptive, particulière ou en raison de contraintes articulaires inhérentes, le degré de stabilité dépendant de la voyelle peut être prédit comme le montre la Table 3. Ces prédictions ne concernent que les voyelles testées dans cette étude, à savoir les 8 catégories vocaliques orales suivantes /i, e, ε, a, x, u, o, ɔ/ (avec x=∅, œ, ə/).

	Prédictions QT	Prédictions TDF	Prédictions DAC
	/i, u, a/	/i, u, o, ɔ, a/	/i/
			/e/
+			/ε/
			/a/
-			/u, o, ɔ/
	/e, ε, x, o, ɔ/	/e, ε, x/	/x/

TABLE 3. Résumé des prédictions des différentes théories abordées : la théorie quantique (QT à gauche), la Théorie de la Dispersion-Focalisation (TDF, au milieu) et la théorie du Degré de Contrainte Articulaire (DAC, à droite), en ce qui concerne la stabilité des voyelles.

Pour résumer, dans cette section nous avons vu que la coarticulation dépend fortement des propriétés structurelles de la langue. La coarticulation n'affecte pas tous les sons de la même manière, certains sont moins sensibles aux influences contextuelles que d'autres et

imposeront plus facilement leurs caractéristiques phonétiques aux autres segments. Cela dépendra de leurs contraintes articulatoires, mais aussi et surtout du degré de conflit existant entre les segments en contexte. Enfin, nous avons vu que selon leurs particularités articulatoires, acoustico-articulatoires ou acoustico-perceptives certaines voyelles seraient plus stables que d'autres. Est-ce que ces voyelles « particulières » sont moins affectées que les autres par les variations contextuelles ? La coarticulation au-delà des déplacements formantiques affecte-t-elle la variabilité des voyelles (dans leurs réalisations acoustiques et leur distinctivité) ? Certains contextes amènent-ils plus de variabilité que d'autres ? C'est à ces questions que nous tenterons de répondre dans notre première étude, CHAPITRE 2.

1.4. LA SENSIBILITÉ À LA COARTICULATION EST-ELLE LIÉE À POSITION PROSODIQUE OCCUPÉE PAR LES SEGMENTS ?

Les études traitant du renforcement prosodique initial montrent que la position prosodique joue sur les contrastes syntagmatiques CV et les contrastes paradigmatiques en lien avec la distinctivité des sons. En position initiale prosodique de haut rang, le contraste entre C et V est maximisé et les particularités relatives aux différents sons se trouvent renforcées. Or la coarticulation a aussi des conséquences sur ces contrastes : les variations contextuelles entre C et V diminuent le contraste syntagmatique et on peut imaginer que selon les caractéristiques articulatoires de la consonne et de la voyelle, la distinctivité de la voyelle s'en trouverait augmentée (lorsque C et V sont compatibles) ou diminuée (si C et V sont en conflit). En étudiant comment la position prosodique affecte le degré de coarticulation C-à-V, nous chercherons à comprendre comment les effets de la position prosodique et ceux de la coarticulation sur les contrastes linguistiques sont compatibles.

Dans cette section, nous commencerons par un bref exposé des principaux effets des variations phonétiques subies par les segments selon leur position dans la structure prosodique allant dans le sens d'un renforcement des contrastes syntagmatiques et paradigmatiques. Puis, nous rapporterons les résultats en rapport avec la coarticulation. Ce dernier point nous permettra de soulever les incohérences observées dans la littérature et de définir le contour des questions de recherche posées dans ce travail.

Remarque : tout au long de cette thèse, nous utiliserons le symbole « # » pour représenter une frontière prosodique.

1.4.1. LA PROSODIE ET LES VARIATIONS SEGMENTALES

Dans certaines positions prosodiques, les segments subissent des variations dans leurs propriétés acoustiques ou articulatoires. C'est le cas des segments : 1) sous proéminences prosodiques (syllabes : sous focus, accent de phrase, ou accent tonique) ou 2) proches des limites de domaines/constituants prosodiques (position initiale et finale de constituants). Dans cette thèse, nous nous concentrerons sur le second type de position prosodique

(marqueur de frontière), et notamment sur les phénomènes de variations segmentales se produisant en position initiale de domaine/constituant prosodique, connus sous le nom de « renforcement prosodique initial » ou « renforcement initial ».

1.4.1.1. Consonnes et voyelles en position initiale de constituant prosodique

Une conséquence du renforcement prosodique initial concerne l'expansion spatiale et/ou temporelle des segments initiaux (ou segments post-frontaliers) dont l'étendue dépend du niveau hiérarchique de la frontière prosodique : plus la frontière est forte, plus le renforcement est important (Fougeron & Keating, 1997 ; Cho & Keating, 2001 ; Cole, Kim, Choi & Hasegawa-Johnson, 2007 ; Cho, Lee & Kim, 2011 ; 2014).

Les effets du renforcement initial sur les propriétés phonétiques des consonnes sont très bien documentés. Les résultats d'études menées sur différentes langues (comme le français, l'anglais, le coréen et le japonais, par exemple) et impliquant différents types de structures articulatoires (articulation linguale, nasale ou encore glottale) se rejoignent pour dire qu'en position initiale de domaines prosodiques élevés, les consonnes sont produites avec plus de force articulatoire (Fougeron et Keating, 1997 ; Fougeron, 1998, 2001 ; Cho, 1998 ; 2001 ; Keating, Cho, Fougeron & Hsu, 2003 ; Onaka, 2003 ; Cho et Keating, 2009) ce qui se traduit par une augmentation de la consonantalité des consonnes (orales comme nasales) en position initiale de constituant prosodique.

Les voyelles aussi subissent les effets du renforcement prosodique initial (Fougeron, 1998, 2001 ; Lehnert-LeHouillier McDonough, & McAleavey, 2010 ; Gendrot, Gerdes et Adda-Decker, 2011 ; Kim & Cho, 2012 ; Georgeton, 2014 ; Georgeton & Fougeron, 2014). Par exemple, dans ses travaux de thèse, Georgeton (2014) s'est intéressée à l'articulation labiale et aux propriétés acoustiques des voyelles orales du français dans deux positions prosodiques (position initiale de Groupe Intonatif « IP » ou position médiane de Groupe Intonatif « W »). Les analyses articulatoires ont montré que les voyelles orales du français étaient toutes produites avec une ouverture aux lèvres plus élevée et un plus grand étirement des lèvres en position initiale d'IP. D'un point de vue acoustique, le renforcement initial des voyelles se fait en fonction des spécificités des voyelles. En effet, l'auteur a trouvé

qu'en position initiale d'IP, 1) les voyelles ouvertes et mi-ouvertes /a, ε, œ, ɔ/ étaient produites avec un F1 plus élevé ; 2) les voyelles antérieures /e, ø, ε, œ/ étaient produites avec un F2 plus élevé alors que le contraire était observé pour la voyelle postérieure /o/ ; 3) aussi le F3 des voyelles non-arrondi /i, e, ε/ se trouvait plus élevé dans cette position alors qu'il se retrouvait abaissé pour les voyelles arrondies /y, ø/. Enfin, l'auteur rapportait un rapprochement de F1 et F2 pour les voyelles postérieures /u, o, ɔ/ comme de F2 et F3 pour la voyelle /y/ indiquant un rehaussement du caractère focal de ces voyelles en position initiale d'IP.

Les résultats de cette étude pris conjointement avec la revue de Cho (2011) sur le sujet, indiquent que le renforcement prosodique initial aurait une fonction linguistique en contribuant à maximiser les contrastes syntagmatiques et paradigmatiques. De plus, dans plusieurs études on s'est posé la question de savoir si les phénomènes observés en position initiale de constituants prosodiques, n'étaient pas liés à des modifications de la coarticulation en position initiale (allant dans le sens d'une réduction des variations contextuelles).

1.4.1.2. La coarticulation en position initiale de constituant prosodique

Les effets du renforcement prosodique initial ont aussi été observés sur le timing inter-gestuel des segments, soit sur le degré de chevauchement entre segments (c.-à.-d. leur coarticulation). Nous distinguerons deux types de travaux : 1) ceux qui traitent des segments transfrontaliers « *_#_* » dont les segments en question ne sont pas directement adjacents ; et 2) ceux qui portent sur les séquences post-frontalières « *#__* » (comprendre que les deux segments testés sont du même côté de la frontière, tous deux au début du constituant prosodique).

La Table 4 page suivante résume l'ensemble des résultats observés dans la littérature que nous développerons plus en détail dans les sections suivantes.

REF	LANG.	TYPE COA.	PP #	METH.	EFFETS
Cho (2004)	anglais	V#V [Ci#Ca] [Ca#Ci]	Gl iGl Mot	EMA	Moins de coarticulation linguale entre /i/ et /a/ lorsque # = Gl.
Tabain (2003)	français	V#C [a#b,d,g]	P, Gl, GA, Mot	transition de F2	Plus la frontière est faible, plus F2 de /a/ tend vers le locus des C.
Byrd & al. (2000)	tamil	C ₁ #C ₂ [m#n] [n#m]	Fmaj Fmin Mot,	EMA	Augmentation du temps de latence entre les pics des deux gestes sur les trois niveaux de frontière prosodique testés. Pic de C ₁ produit proportionnellement plus tôt durant la phase de fermeture de C ₂ en frontière forte.
Cho & al. (2014)	anglais	C#V [m#a] #CV [#ma]	Gl Mot	EMA	C#V : Déplacement du geste de la consonne pré-frontalière vers la gauche quand # = Gl #CV : pas de modification du timing intergestuel selon PP
Meynadier, (2003)	français	VC ₁ #C ₂ V [at#ka] [al#ka] [ak#ta] [ak#la] #C ₂ V	Gl GA abs F	EPG	C ₁ #C ₂ : retard du geste de C ₂ par rapport à C ₁ avec frontière intonative, proportionnel à la force de la frontière pour les /tk/ et /lk/. #C ₂ V : pas de modification du timing intergestuel selon PP
Byrd & Choi (2010)	anglais	C#C [s#p] [s#k] [k#l] #CC	Mot, GA, Gl, U	EMA	C#C : augmentation graduelle du temps de latence entre les pics des deux consonnes en montant dans la hiérarchie prosodique sur les 4 niveaux prosodiques testés. #CC : pas de modification du timing inter-gestuel selon PP
Bombien et al. (2011)	allemand	#CC [#k] [#kn]	Fmin Fmaj Mot	EPG	Réduction du chevauchement entre C ₁ et C ₂ lorsque C ₁ est précédée d'une frontière prosodique forte.

Byrd (2000)	anglais	#CV [mi]	Fmaj Fmin Mot	EMA	Retard du geste /i/ par rapport à C en FM traduisant moins de chevauchement CV observé que pour 1 locuteur sur 3 testés.
Cho et al. (2017)	anglais	#CV [NV]	GI Mot	A1-P0	GI = augmentation de la différence A1-P0 indiquant une augmentation de la sonorité de V soit moins de coarticulation NV.

TABLE 4. Résumé des résultats des études portant sur l'interaction PP et coarticulation. Quelques précisions sur les sigles utilisés : P=phrase, GI=Groupe Intonatif, GA=Groupe Accentuel, M=Mot ; Fmaj=Frontière majeure ; Fmin = Frontière mineure. PP=Position prosodique.

i. La coarticulation entre segments transfrontaliers « _#_ »

Toutes les études traitant de la coarticulation entre segments transfrontaliers s'accordent pour dire que ces segments voient leur degré de chevauchement diminuer proportionnellement au niveau de la frontière prosodique qui les sépare. C'est le cas de séquences de consonnes se trouvant séparées par une frontière prosodique (C#C) en français (Meynadier, 2003) en anglais (Byrd & Choi, 2010) et en Tamil (Byrd, Kaun, Narayanan & Saltzman, 2000). Des résultats semblables ont été rapportés pour des séquences impliquant des voyelles. En effet, Cho (2004) a trouvé une réduction de la coarticulation V-à-V en position prosodique de haut rang en anglais, en analysant les mouvements linguaux de séquences CV₁#CV₂ (où '#' était soit une frontière de mot, soit une frontière de Groupe Intonatif ou une Groupe Intonatif intermédiaire, V les voyelles /i, a/ et C était toujours un /b/) à partir de données EMA. Les résultats dépendaient de la directionnalité de la coarticulation, de la voyelle testée et de la dimension observée (position de la langue *x* ou hauteur de la langue *y*) :

- pour les cas de coarticulation anticipatrice V₂-à-V₁, une réduction de la coarticulation de V₁ en position prosodique forte était observée pour les deux voyelles /i/ et /a/, mais la réduction de la coarticulation de la voyelle /a/ se faisait à la fois sur les dimensions *x* et *y*, alors que pour /i/ seule la dimension *y* était concernée.

- pour les cas de coarticulation persévératrice V_1 -à- V_2 , une réduction de la coarticulation en position prosodique forte était observée uniquement lorsque V_2 était un /a/ sur les deux dimensions testées.
- l'analyse des influences réciproques entre V_1 et V_2 indiquait que la réduction de la coarticulation réciproque en position prosodique forte apparaissait sur la dimension de hauteur de la langue (y) uniquement.

Aussi, en coréen, Cho, Youn & Kim (2014) ont trouvé, à partir de l'étude de la séquence /m#a/ (où '#' était soit une frontière de mot, soit une frontière de Groupe Intonatif), un déplacement du geste de la consonne pré-frontalière vers la gauche quand la frontière séparant C et V était une frontière de Groupe Intonatif. Ce résultat était interprété comme une réduction du chevauchement inter-gestuel des séquences C#V lorsque # était une frontière de haut rang. Une réduction de la coarticulation en frontière prosodique de haut niveau a aussi été rapportée pour les séquences V#C : Tabain (2003) a étudié les transitions formantiques de la voyelle finale /a/ en fonction des consonnes suivantes /b, d, g/ selon que la frontière séparant les segments était une frontière de Mot, de Groupe Accentuel, de Groupe Intonatif ou de Phrase. Les données de trois locuteurs ont montré que plus la frontière, séparant la voyelle /a/ et les consonnes /g/ et /d/ était faible, plus le F2 de la voyelle tendait vers le locus de ces consonnes. Autrement dit, lorsque la frontière prosodique était de bas niveau, le degré de coarticulation anticipatoire de C sur V augmentait.

Si la réduction de la coarticulation de segments transfrontaliers ne semble pas dépendre de la langue étudiée, ou du type de séquence étudié, on notera cependant qu'elle peut dépendre des segments étudiés. Par exemple, dans son étude électro-palatographique sur la coarticulation des consonnes transfrontalières /tk//lk/, /kt/et /kl/ en séquence VC#CV où '#' soit représentait une absence de frontière, une frontière de Groupe Accentuel, ou une frontière de Groupe Intonatif, Meynadier (2003) a montré que la présence d'une frontière intonative retardait le geste de la consonne post-frontalière C2 par rapport à la consonne pré-frontalière C1 et que ce retard était proportionnel à la force de la frontière. Cependant, ce résultat ressortait uniquement pour les clusters /tk/ et /lk/, soit dans les

cas où C1 était une consonne antérieure et C2 une consonne postérieure. De plus, l'effet était plus robuste pour le cluster /tk/.

Il convient aussi de nuancer la gradation des effets selon la force de la frontière. Très souvent, tous les niveaux prosodiques testés dans les études ne se différencient pas en termes de degré de coarticulation. Par exemple, dans leur étude sur le tamil, Byrd et al. (2000) sur les trois niveaux testés, seul le plus haut niveau de frontière (Frontière intonative majeure) se distinguait des deux autres (Frontière intonative mineure et mot).

Enfin, on notera que des variations entre locuteurs existent. Par exemple Tabain (2003) rapportait un effet de la frontière prosodique sur le degré de coarticulation C-à-V de la consonne post-frontalière sur la voyelle pré-frontalière V#C pour tous les locuteurs mais la gradation de l'effet ne permet pas de distinguer tous les niveaux de frontière et cela est fonction du locuteur (P <GI <GA =M ; P<GI =GA<M ; P=GI=GA<M).

Ces derniers constats nous conduisent à dire que la réduction de la coarticulation entre des segments séparés par une frontière prosodique de haut niveau, est un résultat robuste, mais il n'est pas pour autant systématique. Dans cette thèse, on se demandera plus particulièrement si les effets du renforcement initial sont liés à une diminution de la coarticulation en position initiale de constituant prosodique. Pour cela, il apparaît fondamental que les deux segments en question se trouvent du même côté de la frontière prosodique pour s'assurer que l'effet observé soit bien dû aux effets du renforcement initial et non à ceux de la frontière prosodique.

ii. *La coarticulation entre segments post-frontaliers « #__ »*

Il existe très peu de littérature sur les effets du renforcement initial sur des séquences post-frontalières et les résultats des différentes études s'opposent. Certaines études rapportent une réduction du chevauchement entre segments initiaux lorsque la frontière qui les précède est de rang hiérarchique élevé alors que d'autres ne rapportent aucune modification.

C'est le cas notamment des séquences consonantiques #CC. Chez Bombien, Mooshammer, Hoole, Rathcke, et Kühnert (2011) ont étudié, à partir de données électro-

palatographiques, les séquences initiales /#kl/ et /#kn/ produites par sept locuteurs de l'allemand où « # » pouvait être une frontière de mot, une frontière dite de bas niveau (ton de frontière sans pause) et une frontière de haut niveau (ton de frontière avec pause). Les auteurs rapportaient une réduction du chevauchement inter-gestuel entre les consonnes /kn/ et /kl/ positionnées après une frontière de rang hiérarchique élevé. Cependant, il convient de noter que ce résultat n'était avéré que pour certains locuteurs. En effet, quatre des sept locuteurs testés réduisaient le chevauchement des séquences /kn/ et trois locuteurs, pas forcément les mêmes, le faisaient pour les séquences /kl/. Mais ce résultat est en opposition avec ceux rapportés par Byrd et Choi (2010) dont nous avons détaillé la méthode précédemment pour les séquences C#C (voir Table 4). Dans cette étude, en plus des séquences C#C, les auteurs avaient aussi regardé les variations du timing inter-gestuel des séquences consonantiques post-frontalière #CC (#sk, #sp, #kl) selon la force de la frontière (mot, groupe accentuel, groupe intonatif). Ils n'avaient trouvé aucune modification du timing inter-gestuel dans ces séquences post-frontalières.

Ce genre de contradiction apparaît aussi pour les séquences CV post-frontalières. Par exemple, parmi les études décrites dans la section précédente page 26 traitant des séquences C#C, deux d'entre elles s'intéressaient aussi à la syllabe post-frontalière #CV (Cho et al. 2014 pour le coréen avec la séquence #ma et Meynadier, 2003 pour le français avec les séquences #ka, #la et #ta). Les analyses ne révélaient alors aucune variation du timing inter-gestuel de la syllabe post-frontalière selon la force de la frontière. La même chose a été rapportée pour l'anglais. À partir de données EMA, Byrd (2000) a analysé le timing inter-gestuel de la syllabe #mi post-frontalières à partir des séquences /mamə(#)mi(mi)/ où '#' était soit une frontière de mot, une frontière prosodique mineure ou une frontière prosodique majeure. Le retard du geste de la voyelle /i/ par rapport à celui de /m/ en frontière prosodique majeure, traduisant une réduction du chevauchement inter-gestuel, n'apparaissait que pour un seul des trois locuteurs enregistrés. L'effet restait donc très limité et difficilement interprétable. En revanche, dans une étude acoustique récente portant sur l'effet de la frontière prosodique sur le degré de coarticulation nasale de la syllabe post-frontalière #NV en anglais, Cho, Kim & Kim

(2017) montraient que les voyelles en position IPi avaient un A1-P0 (correspondant à la différence entre l'amplitude du premier formant A1 et celle du premier pic nasal P0) plus faible. Cela indique que les voyelles en position IPi sont moins nasalisées par l'effet de la consonne nasale précédente (#NV). Les auteurs interprétaient ce résultat comme une augmentation de la sonorité de la voyelle dans cette position prosodique

En résumé, on peut dire qu'il existe très peu de travaux portant sur les influences de la position prosodique sur la coarticulation de séquences post-frontalières et que les résultats ne font pas consensus. Ceci étant, il semblerait que de plus amples investigations soient nécessaires pour mieux comprendre ce qui se joue au niveau des segments post-frontaliers et pouvoir affirmer que la modification de la coarticulation des segments est une caractéristique du renforcement initial. De plus, aucune étude ne porte sur la coarticulation C-à-V lorsque la voyelle se trouve directement à la frontière de constituant prosodique. Notre étude s'intéressera donc particulièrement aux effets du renforcement prosodique initial sur la coarticulation C-à-V persévératrice (#CV) et anticipatrice (#VC) impliquant un changement de position de la voyelle vis-à-vis de la frontière prosodique.

1.5. LA SENSIBILITÉ À LA COARTICULATION EST-ELLE LIÉE À DES FACTEURS COMMUNICATIONNELS COMME LE STYLE DE PAROLE ?

L'objet de cette section est de faire un état des lieux de la question concernant les relations existant entre style de parole et coarticulation. Dans un premier temps, nous questionnerons le terme même de *style de parole* afin de pouvoir discuter ensuite des résultats observés dans la littérature.

1.5.1. QU'EST-CE QUE LE STYLE DE PAROLE ?

Le style désigne « une façon particulière dont chacun exprime sa pensée, ses émotions, ses sentiments » (Larousse). D'un point de vue linguistique, ou plutôt phonostylistique, cela renvoie à étudier la valeur expressive des sons et des éléments verbaux du langage. La dimension individuelle est donc fondamentale dans l'étude du style puisque le lien avec la construction de la personnalité est désormais envisagé (Eckert, 2003, 2005).

Ceci étant, une conception du style, au-delà de l'individu, est aussi développée et c'est à celle-ci que l'on va s'intéresser dans cette thèse, le style :

« ...comme un ensemble d'ajustements à une situation, dépendant entre autres de la familiarité que l'on a avec son interlocuteur et le sujet abordé, qui font que l'on ne parle pas de la même façon dans toutes les circonstances de la vie (Bolinger, 1989 ; Lindblom, 1990 ; Eskénazi, 1993). » (Boula de Mareüil, 2014 : 9)

L'étude du style, tel qu'il est défini ci-dessus, est donc étroitement liée à la situation de communication, situation dans laquelle un locuteur est impliqué, avec éventuellement un ou plusieurs interlocuteurs.

Dans sa revue de la littérature, Boula de Mareüil (2014) reprend la typologie de Labov (1976) distinguant trois principaux *styles contextuels* (parole familière ou *casual speech*, « en contexte ordinaire, quotidien, de communication familiale », parole lue ou *reading speech*, « lecture de textes et de listes de mots isolés », et parole surveillée ou *careful speech*, « correspondant à un registre soutenu, en situation d'interview ») à laquelle il propose d'ajouter une gradation supplémentaire selon le degré de spontanéité « plus ou moins libres

ou contraints, plus ou moins soignés ou relâchés ». Ainsi, la situation de production renvoyant aux conditions d'enregistrement, le type de parole, directement lié à la tâche de parole, ainsi que le registre de langue sont englobés dans la notion de style.

La description du style qui est faite ci-dessus montre à quel point traiter de ce sujet n'est pas chose aisée tant la notion est complexe. Cependant, c'est ce qui la rend d'autant plus intéressante car elle met la parole au centre de la situation de communication. Or le but premier de tout acte de parole est la communication.

1.5.2. STYLE DE PAROLE ET COARTICULATION

La prononciation est très sensible aux variations de style de parole. Dans la littérature, les changements acoustiques rapportés concernent les propriétés spectrales et temporelles des sons. En particulier, des études antérieures ont montré un allongement des segments (Picheny, Durlach & Braida, 1986 ; Moon et Lindblom 1994 ; Ferguson et Kewley-Port 2002, 2007) et une hyperticulation, souvent mesurée par une expansion de l'espace acoustique vocalique (Picheny et al., 1986 ; Johnson, Flemming et Wright 1993 ; Smiljanic & Bradlow, 2005 ; Gendrot et Adda-Decker, 2005 ; Audibert, Fougeron, Gendrot & Adda-Decker, 2015), en tant que signaux robustes de la parole « claire » (hyper-articulée) par rapport à d'autres formes de parole élicitée (forme de citation) aussi bien que la parole relâchée (souvent appelée parole spontanée).

Les travaux traitant des effets stylistiques sur la coarticulation sont eux moins bien nombreux. Comme on peut le voir sur la Table 5, la plupart d'entre eux impliquent différents types de coarticulation et différentes langues, révèlent un effet du style de parole sur la coarticulation (Krull, 1989 ; Duez, 1992 ; Moon & Lindblom, 1994 ; Dicanio, Nam, Amith, García & Whallen, 2015 ; Scarborough & Zellou, 2013). Les résultats de ces études indiquent une réduction de la coarticulation en parole plus formelle. Cependant, d'autres travaux affichent un résultat partiel (Matthies, Perrier, Perkell & Zandipour, 2001), une réduction de la coarticulation est observée à partir des données acoustiques seulement, voire aucune modification de la coarticulation en fonction du style (Bradlow, 2002).

RÉF	LANG.	COA. DIR	STYLES	EFFET DU STYLE SUR LA COA ?
Krull (1989)	suédois	V-à-C ant CV	Parole spontanée : -raconter une histoire après lecture -interview -dialogue Parole lue (dit claire) : -lecture de liste de mots produits dans la partie spontanée.	Mes. : équation du locus <u>OUI</u> Moins de coarticulation en parole claire
Moon & Lindblom (1994)	anglais	C-à-V pers CV	Parole lue variant en clarté : -lecture liste de mot normalement (forme de citation) -lecture liste de mot en hyperarticulé (parole claire)	Mes. : Distance formants locus C et max V <u>OUI</u> Moins de coarticulation en parole claire
Duez (1992)	français	V-à-C ant CV	Parole spontanée : -interview informelle par l'expérimentateur Parole lue (dit claire) : - lecture de liste de mots produits dans la partie spontanée à débit lent.	Mes. : équation du locus <u>OUI</u> Moins de coarticulation en parole claire
Matthies et al. (2001)	anglais	V-à-V [iC _n u] ant VC _n V	Parole lue variant en clarté : - débit normal - hyperarticulée (parole claire simulée : les locuteurs croyaient être écoutés par une personne évaluant leur prononciation)	Mes. : Art. EMA - mouvement langue & lèvres sur la durée de [iC _n u] Ac. F2 au milieu de V Art. <u>NON</u> vs. Ac. <u>OUI</u> F2 de /i/ plus élevé en parole claire, moins de coarticulation.
Bradlow (2002)	anglais / espagnol	C-à-V pers CV	Parole lue variant en clarté : - débit normal - hyperarticulée (parole claire simulée : les locuteurs imaginaient qu'ils parlaient à un malentendant ou une personne parlant mal la langue.)	F2' à 20-50-80% V <u>NON</u>
Diciano et al. (2015)	mixtèque	C-à-V pers CV	Parole spontanée : - Narration d'histoires personnelles Parole lue (dit claire) : -lecture de listes de mots	Mes. : variabilité des V distances euclidiennes F1/F2 <u>OUI</u> Moins d'effets contextuels sur la variabilité de V en parole claire

Scarborough & Zellou (2013)	anglais	C-à-V ant VN	2 situations de com. :	Mes. : différence d'amplitude du pic harmonique de F1 (A1) et celle du plus bas pic nasal (P0) : A1-P0
			<p>Fictive : parole lue</p> <ul style="list-style-type: none"> - débit normal CITATION - le plus clairement possible CLAIR - en imaginant parler à un ami A - en imaginant parler à un malentendant HOH <p>Réelle : parole spontanée</p> <ul style="list-style-type: none"> - jeu en interaction avec un pair REELLE 	<p style="text-align: center;"><u>OUI</u></p> <p>Production : coarticulation maximale en situation REELLE et minimale en HOH +</p> <p>Perception (décision lexicale) : TR plus court REELLE que HOH. = la coarticulation sert à la perception</p>

TABLE 5. Résumé des travaux traitant de l'interaction entre style de parole et coarticulation. COA DIR. indique la direction de la coarticulation pers = persévératrice et ant = anticipatrice. Mes. = mesure effectuée.

Si l'effet du style sur la coarticulation semble donc avéré, c'est la considération même du style de parole et la façon dont il est élicité qui interpellent dans ces études. C'est sur ce point, que nous allons nous attarder dans les prochains paragraphes en discutant plus en détail les travaux présentés Table 5. En effet, la plupart des études traitant des effets du style de parole sur la coarticulation, visent à comparer deux conditions de parole différentes, la parole normale, ou le style par défaut des locuteurs et la parole claire traduisant un style plus soigné dans lequel le locuteur fait attention à sa prononciation. Ces études se sont donc focalisées sur un style de parole bien particulier qui est la parole claire (pour *clear speech*). En revanche, celles-ci diffèrent dans la méthode utilisée pour éliciter ce type de parole. Or comme nous allons le voir, les différentes méthodes vont impliquer des dimensions du style différentes et pouvant avoir des répercussions sur la parole obtenue. Il n'est donc pas sûr que ce qui est appelé parole claire dans les différentes études, renvoie au même style de parole.

Moon et Lindblom (1994) comparent la lecture d'une liste de mots dans deux conditions : à débit et intensité normaux *forme de citation* ou en hyper-articulant et en parlant le plus clairement possible *parole claire*. Ce qui est manipulé ici c'est l'attention qui est portée à ce qui est lu. On peut se poser la question de la place du style de parole dans une tâche aussi

artificielle que la lecture de mot. Aussi, il convient de se demander s'il est correct de faire le rapprochement entre ces mots lus avec une attention particulière et des productions issues d'une parole soignée dans une situation de communication réelle. Dans quelle mesure peut-on donc faire des inférences sur des effets du style au sens large tel qu'il est décrit dans la section précédente 1.5.1 ?

D'autres travaux comparent de la parole spontanée avec la lecture d'une liste de mots (Krull, 1989 ; Duez, 1992 ; Dicanio et al., 2015). Si les auteurs ont pris soin de comparer des mots identiques dans les deux tâches de lecture, on peut se poser la question de l'effet de la tâche de production. En effet, les deux paroles n'ont pas grand-chose à voir l'une avec l'autre. Les auteurs comparent des mots lus, présentés dans une liste *forme de citation*, correspondant à une vision très minimaliste de l'acte de parole, avec des mots extraits d'un contexte discursif élaboré (narrations, conversations, interview etc.), *parole spontanée*. Cette comparaison permet d'éclairer les particularités de la parole spontanée et les résultats qui en ressortent sont tout à fait intéressants dans cette perspective, mais on n'est pas vraiment dans une comparaison de style de parole. De plus, peut-on réellement parler de spontanéité quand la parole enregistrée est le fruit d'instructions données à des sujets, lesquels doivent parler face à des inconnus, dans un endroit peu familier ?

Un troisième type d'étude concerne les travaux ayant recours à de la parole simulée pour éliciter ce que les auteurs appellent de la parole claire. C'est le cas par exemple de Bradlow (2002) qui, dans une étude comparative anglais-espagnol, comparait de la parole hyper-articulée avec de la parole hypo-articulée. Pour cela, les sujets devaient lire des phrases dans deux conditions différentes : 1) normalement, sans faire attention à la clarté comme s'ils parlaient à un ami qu'ils connaissaient bien (condition hypo-articulé) ; 2) en imaginant qu'ils parlaient à une personne sourde ou parlant mal la langue (condition hyper-articulée). À l'intérieur de ces phrases étaient placées des séquences CV, où C pouvait être un /b/ ou un /d/ et V une des voyelles /i, u/. À partir des valeurs de F2' en mel (valeurs de formants extraites à 20- 50- et 80 % des voyelles) et de la durée des voyelles, en condition hyper-articulée, l'auteur notait une expansion de l'espace vocalique, un allongement de la durée vocalique ainsi qu'un maintien de la coarticulation CV. L'auteur concluait que

l'hyper-articulation et la coarticulation étaient contrôlées par le locuteur dans un but perceptif.

C'est cette hypothèse de la fonction perceptive que Scarborough et Zellou ont testée dans leur étude de 2013. À l'issue d'une analyse acoustique du degré de coarticulation nasale anticipatrice sur les voyelles (VN) de mots extraits à la fois de différentes conditions de parole claire simulées et d'une situation de communication réelle (voir Table 4), les auteurs avaient identifié un degré maximal de coarticulation en situation de communication réelle et un degré minimal en condition simulée HOH (les participants devaient faire comme s'ils parlaient à un malentendant). Cela les avait conduits à mettre en place une expérience perceptive, de décision lexicale à partir de mots et non-mots, produits par une locutrice de la 1^{ère} expérience, extraits des situations REELLE et HOH. Les résultats de cette deuxième analyse montraient un clair effet du type de parole : les auditeurs répondaient plus rapidement lorsque le mot était produit en condition de communication réelle (REELLE) qu'en situation simulée (où le participant devait faire comme s'il parlait avec un malentendant – HOH). Les auteurs validaient ainsi la fonction perceptive de la coarticulation indiquant que c'était l'authenticité de la situation de communication qui induisait de la variation dans le degré de coarticulation.

Matthies et al. (2001) ont aussi eu recours à de la parole simulée pour éliciter un style plus « clair ». Ces derniers faisaient croire aux sujets que quelqu'un les écoutait, afin de contrôler la qualité de leur prononciation. Ainsi, ils comparaient de la lecture de phrase à débit normal avec de la parole lue en se croyant écouté et évalué. On peut imaginer que le fait de se sentir écouté et contrôlé peut gêner le sujet et modifier son comportement. Il risque de manquer de naturel et cela peut causer des variations particulières dans ses productions. L'analyse des données articulatoires ne révélait aucun effet du style de parole sur la coarticulation anticipatrice sur /i/ en séquence [iC_nu]. En revanche, les données acoustiques, elles, montraient une réduction de la coarticulation en parole claire (simulée). En effet, les valeurs de F2 prises au milieu de la voyelle /i/ indiquaient que la voyelle /i/, caractérisée par un F2 élevé, présentait un F2 plus haut en parole claire qu'en parole normale. La non linéarité des relations acoustique/articulatoire (Stevens, 1972) peut

expliquer ces différents résultats. De plus, la sortie acoustique est le résultat de la configuration globale du conduit vocal et du mouvement des différents articulateurs impliqués, alors que l'analyse articulatoire se focalise sur un articulateur bien particulier et un type de mouvement.

Pour terminer ce passage sur le recours à la parole simulée, on posera quelques questions relatives à ce genre de méthode : en effet, il n'est pas forcément aisé pour tout le monde de réussir à imaginer une scène et jouer en quelque sorte un rôle. On demande finalement aux locuteurs, d'être des acteurs. Peut-on être sûr qu'ils agiraient de la sorte si la situation se produisait dans la réalité ? Ceci étant, le but est bien sûr d'obtenir de la parole hyper-articulée, et c'est ce qui est observé. La consigne fonctionne. Mais peut-on vraiment faire le lien avec ce que ces locuteurs produiraient dans des situations réelles ?

Pour résumer, on peut dire qu'il existe finalement très peu de travaux traitant de l'interaction entre style de parole et coarticulation. Cependant, les quelques travaux existants relatent une réduction de la coarticulation en parole claire (sauf une étude). Ce qui pose question dans ces études, c'est plutôt la place qui est donnée au style de parole, la façon dont il est parfois élicité et le degré de « naturalité » de la parole testée. Si ces travaux ont le mérite d'ouvrir la question sur les relations entre style et coarticulation, il semble qu'il reste à mieux comprendre ce qui se joue, en élargissant les investigations à d'autres styles de parole, puisque la parole « claire », si c'en est un, n'est pas le seul style existant.

1.6. LES GRANDS CORPUS ORAUX

L'ensemble des études menées dans cette thèse repose sur l'analyse de grands corpus oraux. L'objet de cette dernière section est de présenter dans les grandes lignes, la linguistique de corpus en s'attardant sur les implications du recours à ce genre de corpus pour l'étude de la variation, les avantages (étude à grande échelle, test de différents facteurs simultanément) et les limites (impossibilité de contrôler toutes les sources de variation et forts déséquilibres dans les données) que l'on peut rencontrer. Il ne s'agit donc pas de faire une revue détaillée des études portant sur les grands corpus ou les problématiques que cela soulève. Nous terminerons par présenter les corpus utilisés dans cette thèse.

1.6.1. QUELQUES NOTES SUR LES GRANDS CORPUS ORAUX

Le recours aux grands corpus oraux est de plus en plus fréquent dans le domaine de la phonétique et de la phonologie. La particularité de la parole produite en contexte de parole plus écologique ne fait plus de doute. Par exemple, il est désormais bien établi que la parole spontanée renferme des patterns prosodiques particuliers qui n'avaient jusque là pas été observés sur des corpus contrôlés enregistrés en condition expérimentale de laboratoire (Schafer, Speer & Warren, 2005). En mettant la situation de communication au cœur de l'étude de la parole, on redonne à cette dernière sa fonction première : communiquer, transmettre un message. Ainsi, on peut tester différentes théories établies à partir de données très contrôlées à grande échelle et voir si leurs prédictions se reflètent dans la parole de tous les jours, celle qui au final nous importe de comprendre.

Le recours à de grandes bases de données (*Big data*) est en plein essor. L'on doit principalement cela aux nouvelles technologies rendant possible le traitement automatique ou semi-automatique de très grandes quantités d'informations qu'il serait impossible d'analyser manuellement. La phonétique et la phonologie bénéficient notamment des connaissances et outils développés en Traitement Automatique de la parole et en particulier

la reconnaissance et l'alignement automatique (pour plus de détails, voir Nguyen & Adda-Decker, 2013).

Comme le soulignent ces derniers auteurs, l'émergence des grands corpus a permis de développer de nouvelles méthodes d'analyse dont les bénéfices pour l'étude de la variation ne font aucun doute. Le plus important sans doute concerne le test des facteurs *a posteriori*. En effet, si jusque-là les chercheurs créaient en amont les corpus afin de tester un facteur en particulier tout en prenant soin de contrôler d'autres possibles sources de variation, c'est la démarche inverse que ce genre de corpus nous invite à adopter. On dispose de ressources riches et complexes sur lesquelles on va tester l'effet d'un, voire de plusieurs facteurs, ainsi que leur interaction. C'est très important dans l'étude de la variation puisque l'on sait que le signal de parole résulte de l'interaction de différents facteurs provenant d'origines diverses. Or, la façon dont ils interagissent, reste encore très mal comprise. Sans doute parce que l'on ne disposait pas jusqu'à il y a peu de temps, des moyens nécessaires pour analyser ces interactions complexes. Ici se situe donc un enjeu important dans l'étude des grands corpus : ouvrir les champs du possible dans l'étude de la variation afin de dresser un tableau plus précis de la façon dont les sources de variations interagissent entre elles. À la fois pour comprendre ce qui gouverne la variation dans la parole, identifier les fonctions de ces variations et pour déterminer leurs implications sur la perception de la parole et l'évolution de la langue.

Concernant les limites de ce genre de corpus, on en notera deux majeures. Paradoxalement, la première touche à l'interaction des facteurs. En effet, le revers de la médaille d'avoir des corpus très riches, c'est que l'on ne peut ni contrôler, ni tester toutes les sources de variation simultanément. Ainsi, on ne peut pas prétendre que nos résultats soient 100% attribuables au(x) facteur(s) testé(s). Ceci étant, cela est compensé par la grande quantité de données étudiées permettant d'asseoir les résultats. La seconde est relative au déséquilibre dans les données. Les grands corpus oraux reflètent très souvent les tendances naturelles de la langue. Ainsi, certains sons sont très fréquents (les consonnes alvéolaires, la voyelle /a/ par exemple en français), et donc surreprésentés dans ce genre de corpus, alors que d'autres sont plutôt rares (les voyelles postérieures, la consonne

uvulaire) et donc sous-représentés. Si on souhaite tester les influences contextuelles par exemple, on peut facilement imaginer que l'on se trouvera face à des problèmes de déséquilibre dans les données. L'analyse d'interaction des variations contextuelles avec d'autres sources de variation, augmentera très rapidement ces déséquilibres rendant certaines analyses impossibles. Le recours à des méthodes statistiques robustes vis-à-vis des échantillons déséquilibrés sera donc fondamental dans ce genre de cas de figure. Les méthodes de *bootstrapping* offrent des solutions très intéressantes pour ce type de données. Dans nos études, nous aurons recours à des modèles linéaires mixtes permettant de traiter des échantillons de données déséquilibrés. Dans certains cas de figure, des méthodes de rééchantillonnages seront utilisées afin de rééquilibrer nos différentes conditions testées.

1.6.2. LES CORPUS SÉLECTIONNÉS POUR NOS ÉTUDES

L'objectif ici est de présenter les corpus d'intérêt dans cette thèse. Seules les informations relatives aux conditions de productions et à la parole seront rapportées. Pour tout ce qui concerne la partie technique d'enregistrement, de transcription et d'alignement, se référer aux références citées. La Table 6 offre un résumé des particularités relatives à chaque corpus.

CRITÈRES DE PAROLE	CORPUS NCCFR	CORPUS ESTER
TYPOLOGIE LABOVIENNE	spontanée / familière	surveillée / soutenue
CONDITIONS D'ENREGISTREMENT	chambre sourde	parole radiophonique
SITUATION DE COMMUNICATION	dialogue en face à face entre amis	salle de radio échange face à un micro
LOCUTEURS	étudiants / locuteurs lambda	journalistes & politique/ locuteurs professionnels
PROVENANCE	Ile de France,	d'origines diverses
NOMBRE	46	> 14000
SEXE	24 F / 22 H	H > F
TYPE DE PAROLE	conversationnelle	majoritairement scriptée avec des échanges spontanés

TABLE 6. Résumé des particularités de chacun des deux corpus étudiés.

1.6.2.1. NCCFr (Torreira, Adda-Decker & Ernestus, 2010): parole spontanée

Le corpus NCCFr pour *Njimegen Corpus of Casual French* a été créé dans le but de pallier le manque de données de parole spontanée en grande quantité en français et de fournir des enregistrements de très bonne qualité permettant des analyses phonétiques détaillées. Au total 46 locuteurs (24 femmes et 22 hommes) la plupart provenant d'Île de France (12 étaient issus d'autres régions du centre et du nord de la France), étudiant à l'Université Sorbonne Nouvelle – Paris 3, ont été enregistrés au sein même de l'université dans la chambre sourde du Laboratoire de Phonétique et Phonologie. Tous les locuteurs avaient entre 18 et 27 ans sauf deux locutrices qui avaient environ 50 ans.

Trois personnes prenaient part à chaque session d'enregistrement de 90 minutes : il s'agissait toujours de trois personnes de même sexe, toutes amies. Un des locuteurs était un « confédéré » au courant des finalités de l'enregistrement et il n'était pas enregistré pour des raisons techniques. Les deux autres locuteurs étaient totalement naïfs à ce sujet. Chaque session était composée de trois parties. Durant la première, le confédéré simulant un problème de micro sortait de la pièce et les deux locuteurs se trouvant seuls et ne se pensant pas être enregistrés, discutaient entre eux. Lorsque le confédéré revenait, les trois locuteurs s'engageaient dans une conversation libre sachant cette fois qu'ils étaient écoutés. Pour la dernière partie de l'enregistrement, il était demandé aux participants d'échanger sur des sujets sociétaux particuliers durant lesquels ils devaient exprimer et défendre leur point de vue. Cette dernière partie, contrairement aux autres, avait donc un but communicatif précis.

1.6.2.2. ESTER (Galliano et al., 2006) : parole journalistique

Le corpus ESTER pour *Evaluation de Systemes de Transcription enrichie d'Emissions Radiophoniques* est un corpus audio créé afin d'évaluer les systèmes de transcription d'émissions d'informations en français. Le corpus contient des émissions radiophoniques diffusées sur les stations : France Inter, France Info, Radio France International (RFI), Radio Télévision Marocaine (RTM), France Culture et Radio Classique. Une partie du corpus contient 100h d'émissions d'informations scriptées, enregistrées en 1998, 2000,

2003 et 2004 et transcrites manuellement. Dans cette partie, 2172 locuteurs ont été recensés (744 femmes et 1398 hommes). Une seconde partie renferme 1 677 heures d'émissions non scriptées, enregistrées entre 2003 et 2004 dans lesquelles 14862 locuteurs ont été recensés.

I.7. RÉSUMÉ DES OBJECTIFS DE LA THÈSE

Cette thèse propose d'étudier la coarticulation C-à-V et son interaction avec différentes sources de variation, à la lumière de grands corpus oraux du français. Si la coarticulation est largement décrite dans la littérature, la façon dont les variations contextuelles interagissent avec d'autres sources de variation dans la parole reste mal comprise comme on a pu le voir dans les sections précédentes. Pourtant, le signal de parole résulte de l'interaction de multiples facteurs de variation provenant d'origines diverses. La variation qui résulte de l'action de ces facteurs peut donc être motivée linguistiquement, stylistiquement ou bien n'être que le résultat de contraintes biomécaniques. Toutes ces variations n'ont certes pas la même fonction mais elles affectent toutes le signal de parole qui peut s'en trouver enrichi comme appauvri. Cela ne se trouve pas sans conséquence pour la transmission du message, qui peut être facilitée par la variation lorsqu'elle enrichit le signal de parole, ou au contraire altérée, lorsqu'elle l'appauvrit.

Les principaux objectifs de ce travail sont donc d'offrir une description plus précise de la coarticulation C-à-V en français et de mieux comprendre ce qui gouverne la variation dans la parole. À partir d'une série de trois études acoustiques, nous allons tester si et comment la coarticulation C-à-V est fonction : 1) du type de voyelle et des contraintes de productions spécifiques aux segments (C et V) impliqués (variation phonético-phonologique) ; 2) de la position prosodique occupée par les segments (variation liée à la structuration du message linguistique) ; et 3) du style de parole (variation liée à la situation de communication).

Chaque étude est organisée en chapitre. Chacun d'entre eux commence par une section de présentation des objectifs et questions de recherche. Pour le reste, on suit une organisation classique : méthode, résultats, discussion. Enfin une discussion générale permettra de discuter de l'ensemble des résultats obtenus dans ces trois études et des implications qu'ils peuvent avoir en perception de la parole et ce que cela dit plus généralement sur la variation.

ÉTUDE DE L'EFFET DU TYPE DE VOYELLE SUR LA COARTICULATION C-à-V

Le but de ce chapitre est de mieux identifier les effets coarticulaires des consonnes sur les voyelles en français à partir d'une étude sur grand corpus. Au-delà des effets de durée, on sait que la coarticulation est fonction des caractéristiques articulatoires des segments impliqués : de la distance entre C et V (Lindblom, 1963) ou du degré de contrainte articulatoire des segments et de la compatibilité mutuelle des gestes les constituant (Recasens et al., 1997). De plus selon différentes théories de la parole, certaines voyelles sont décrites comme étant plus stables que d'autres. Ainsi, au vu de ces éléments nous avons voulu tester à grande échelle la direction des influences contextuelles sur les voyelles. L'originalité de cette étude est, qu'au-delà des déplacements formantiques, nous regardons si la coarticulation affecte la variabilité des voyelles selon leur stabilité intrinsèque, au niveau de la dispersion des exemplaires à une valeur de référence de leur catégorie vocalique, et de la distinctivité des voyelles, déterminée par les scores d'une classification acoustique automatique. Concernant les déplacements formantiques, nos résultats confirment la littérature (CUVULAIRE : élève F1, abaisse F2 et F3, F4 de /i/ ; CALVÉOLAIRE : abaisse F1 et élève F2 ; le contexte CVÉLAIRE abaisse F1 mais n'affecte pas le F2 des voyelles). Nos analyses de la variabilité des voyelles indiquent que la voyelle /u/ est la plus dispersée, toutes dimensions acoustiques confondues, et l'ensemble des postérieures le sont sur F2. Ces voyelles sont aussi les moins distinctives et leur distinctivité dépend plus fortement du contexte dans lequel elles ont été produites. Aussi, la voyelle /i/ n'est pas moins dispersée que les autres. En revanche, elle est de loin la plus distinctive, ce qui montre que la particularité de cette voyelle tient plus de la zone acoustique qu'elle occupe que de sa constance dans ses réalisations. Enfin, nous avons vu que les voyelles étaient moins dispersées en contexte alvéolaire et que les voyelles étaient plus distinctives lorsqu'elles apparaissaient dans un contexte renforçant leurs propriétés intrinsèques : les voyelles postérieures et les voyelles plus ouvertes sont plus distinctives en contexte uvulaire alors que les voyelles antérieures et les plus fermées le sont en contexte alvéolaire et vélaire.

2.1. QUESTIONS DE RECHERCHE & HYPOTHÈSES

Dans cette première étude nous regarderons si la coarticulation C-à-V varie en fonction de contraintes phonético-phonologiques des sons impliqués : selon les caractéristiques articulatoires des consonnes et des voyelles et du conflit existant entre les deux dépendant de leur trajectoire articulatoire respective (on rappelle que le terme de ‘compatibilité’ entre C et V sera utilisé dans les cas où les trajectoires de C et V sont comparables). Pourquoi ne pas s’en tenir à étudier la coarticulation selon les « contraintes phonétiques des segments » en ce cas-là ? Tout simplement parce que les sons étudiés évoluent au sein d’un système phonologique qui déterminera en partie leur variation. Par exemple, la densité du système vocalique peut influencer sur le degré de variabilité des voyelles (Manuel, 1990). Ainsi, le nombre de voisins acoustico-perceptifs définirait la marge de manœuvre qu’un locuteur a vis-à-vis de sa précision articulatoire. Les conséquences sur l’intelligibilité de la parole ne seront pas les mêmes si le son peut être confondu avec un autre son existant dans la langue ou si la variation n’engendre pas de telles confusions. C’est ce que l’on entend par « contraintes phonético-phonologiques ».

Dans la section suivante une série de prédictions est donnée à partir de la revue de la littérature présentée section 1.3. Cela concernera, les voyelles, selon leur stabilité intrinsèque, les consonnes, selon leur degré de contrainte articulatoire, et l’interaction CV selon que consonne et voyelle sont ou non en conflit sur le plan articulatoire. L’objectif premier de cette étude est donc de tester l’ensemble de ces prédictions à la lumière des grands corpus de parole. L’originalité de cette étude est, qu’au-delà des déplacements formantiques, nous regardons si la coarticulation joue sur la variabilité des voyelles, au niveau de la dispersion des exemplaires au centroïde de leur catégorie, et de leur distinctivité, déterminée par les scores d’une classification acoustique automatique. À partir de l’examen de 18.5k exemplaires des voyelles /i, e, ε, a, x, u, o, ɔ/ (avec x=/ø, œ, ə/) précédées ou suivies par une consonne alvéolaire /t, d, n, s, z, l/, vélaire /k,g/ ou uvulaire /ʁ/, le contexte opposé étant toujours une consonne labiale parmi /p, b, m, f, v/, nous tenterons de répondre aux questions suivantes :

- Est-ce que la coarticulation C-à-V dépend de la voyelle impliquée, de la consonne et/ou de la compatibilité de C et V ?
- Est-ce que certaines consonnes amènent plus de variabilité dans la réalisation des voyelles que d'autres ? Au niveau de la dispersion des voyelles à une valeur de référence de leur catégorie vocalique ? Au niveau de leur distinctivité ?

2.1.1. PRÉDICTIONS

2.1.1.1. Les voyelles

Que ce soit en raison d'une relation articuloire-acoustique (théorie quantique page 21), ou acoustico-perceptive particulière (théorie de la dispersion-focalisation page 22), ou en raison de contraintes articuloires inhérentes (voir le modèle DAC page 13), le degré de stabilité dépendant de la voyelle peut être prédit comme le montre la Table 7 ci-dessous (tableau déjà présenté page 23). Ces prédictions ne concernent que les voyelles testées dans cette étude à savoir les 8 catégories vocaliques orales suivantes /i, e, ε, a, x, u, o, ɔ/ (avec x=/ø, œ, ə/).

	Prédictions QT	Prédictions TDF	Prédictions DAC
	/i, u, a/	/i, u, o, ɔ, a/	/i/
			/e/
+			/ε/
			/a/
-			/u, o, ɔ/
	/e, ε, x, o, ɔ/	/e, ε, x/	/x/

TABLE 7. Résumé des prédictions des différentes théories abordées : la théorie quantique (QT à gauche), la Théorie de la Dispersion-Focalisation (TDF, au milieu) et la théorie du Degré de Contrainte Articulaire (DAC, à droite), en ce qui concerne la stabilité des voyelles.

2.1.1.2. Les consonnes

iii. Contexte alvéolaire

Ce premier groupe de consonnes que nous appelons alvéolaire comprend les consonnes dento-alvéolaires du français /t, d, z, s, n, l/. Ces consonnes sont produites avec une élévation de la partie antérieure de la langue pour former une constriction alvéo-dentale avec l'apex ou la lame de la langue. C'est le groupe de consonnes articulées le plus à l'avant du conduit vocal que nous étudions dans cette thèse. Les effets coarticulaires des consonnes alvéolaires sont connus (Öhmann, 1966 pour le suédois ; Carney & Moll, 1971 pour l'anglais ; Butcher & Weiher, 1976 pour l'allemand et Recasens, 1983, pour le catalan) selon (Recasens, 1985), notamment pour ce qui est d'attirer le F2 des voyelles autour de leur locus à 1800 Hz. De plus, les contraintes articulatoires relatives à la réalisation de la constriction buccale vont causer un abaissement du F1 des voyelles. L'antériorisation de la langue pour l'articulation des consonnes alvéolaires conduit à une réduction de la cavité antérieure et au contraire, à une augmentation de la taille de la cavité pharyngale, habituellement très resserrée pour l'articulation du /a/. Le F1 étant principalement associé aux résonances de la cavité postérieure sera donc abaissé (la cavité pharyngale étant plus grande) et le F2 affilié à la cavité antérieure sera, lui, élevé du fait de la diminution de la cavité antérieure. La voyelle /x/ étant produite sans aucune constriction dans le conduit vocal, elle devrait être très sensible aux effets des différents contextes (voir Recasens et al., 1997).

Le modèle DAC attribue un score intermédiaire de 2 à ces consonnes dont l'articulation n'implique pas directement le corps de la langue mais pour lesquelles un couplage entre l'articulateur principal et le corps de la langue est souvent observé (Recasens et al., 1997). De plus, certaines de ces consonnes se voient attribuer un DAC de 3 du fait de leur plus grande contrainte articulatoire due par exemple au mode d'articulation de ces dernières comme les consonnes fricatives. Ainsi, on s'attend à observer des effets coarticulaires considérables de la part de ces consonnes alvéolaires, notamment dû au fait que l'on regroupe toutes les consonnes alvéolaires sans différencier celles ayant un DAC maximal de 3 de celles ayant un DAC intermédiaire de 2.

iv. Contexte vélaire

Les travaux sur la coarticulation des consonnes vélares nous laissent penser que ce type de consonnes affectera peu l'acoustique de nos voyelles sur la dimension F2. En effet, il est bien connu que les consonnes vélares s'adaptent au lieu d'articulation des voyelles environnantes. Ainsi, on leur attribue un allophone palatal lorsqu'elles sont entourées de voyelles antérieures (Öhmann, 1966 ; Fant, 1973 parmi d'autres exemples). De ce fait ces consonnes seraient peu résistantes à la coarticulation sur la dimension x et par extension, peu agressives envers les autres segments. Ce résultat a été rapporté par Fowler & Brancazio (2000) sur la dimension F2 seulement. Sur la dimension F1, ces consonnes présentaient un fort degré de résistance à la coarticulation que les auteurs interprétaient comme la conséquence des exigences du mode occlusif. La mise en œuvre de l'occlusion demande une précision articulatoire importante sur la dimension y laissant peu de place à la variation.

Le modèle DAC attribue un DAC de 3 à ces consonnes vélares du fait de l'implication du dos de la langue dans leur articulation (Recasens et al., 1997). Cependant, d'autres travaux dans le cadre de ce modèle ont confirmé les résultats de Fowler et Brancazio (2000). Ainsi, les consonnes vélares sont résistantes à la coarticulation et agressives envers les sons adjacents en ce qui concerne la hauteur de la langue, mais pas sur sa position sur l'axe antéro-postérieur (Recasens et Espinosa, 2009).

On s'attend donc à un faible effet coarticulatoire des consonnes vélares sur le F2 des voyelles. Cependant, plus les voyelles seront ouvertes, plus elles seront susceptibles des variations sur la dimension F1 dans ce contexte.

v. Contexte uvulaire

Les consonnes uvulaires sont moins répandues dans les langues et leurs effets coarticulatoires sont forcément moins bien décrits. Cependant, un ouvrage récent portant sur le lien entre coarticulation et changement de sons dans les langues romanes (Recasens, 2014) traite de l'ensemble des consonnes rhotiques et notamment de la fricative uvulaire du français /ʁ/. Ainsi, l'auteur place cette consonne au même niveau que le [ʁ] sombre ou

la semi-consonne labio-vélaire /w/ et lui attribue le plus haut indice de DAC = 3. Cette consonne est articulée avec un abaissement de la partie pré-dorsale de la langue et une rétraction de la partie post-dorsale de la langue dans la zone uvulaire pour la réalisation de la constriction (Straka, 1965). Finalement, l'articulation de cette consonne se rapproche fortement de celle du [ɰ] (sans la constriction apicale). D'ailleurs, le changement de son de [r] en [ʁ] aurait commencé par la perte du contact apical et une rétraction du dos de langue du [r] en coda de syllabe (Recasens, 2014). L'acoustique de la consonne uvulaire /ʁ/ est caractérisée par un F1 élevé de par l'abaissement important de la partie pré-dorsale de la langue, un F2 bas de par sa forte postériorisation et un F3 bas.

Pourtant, la consonne /ʁ/ est très variable en français. Elle est décrite comme étant une consonne polymorphe (Meunier, 1994) du fait de sa grande sensibilité au contexte environnant (tant vocalique que consonantique). Elle peut par exemple 1) se dévoiser au contact d'un son sourd ; 2) devenir approximante ou sonante au contact d'une voyelle, ou 3) avoir une articulation vélaire au contact d'une voyelle antérieure. Cette consonne serait donc peu résistante à la coarticulation. Ceci voudrait donc dire que cette consonne aurait aussi peu d'effet sur les voyelles environnantes, si nous restons sur la correspondance totale entre résistance et agressivité. Or, nous pensons que, de par ces particularités articulatoires décrites plus haut, et notamment de l'abaissement de la mandibule, de l'apex et du corps de la langue maintenu quel que soit l'allophone réalisé de ce phonème, ce segment montrera une certaine agressivité vis-à-vis des voyelles environnantes. Nous faisons donc le postulat que pour cette consonne, résistance à la coarticulation et agressivité ne sont pas en totale correspondance. On s'attend donc à observer un fort effet de la consonne uvulaire sur les propriétés phonétiques des voyelles dans les directions suivantes : élévation de F1, abaissement de F2 et F3.

La Table 8 présente un récapitulatif des effets contextuels attendus des différents groupes de consonnes sur l'acoustique des voyelles.

CONTEXTE	CONSÉQUENCES ACOUSTIQUES
ALVÉOLAIRE	abaissement F1 élévation F2 vers 1800Hz
VÉLAIRE	abaissement F1 peu d'effets sur F2
UVULAIRE	élévation de F1 abaissement de F2 et F3

TABLE 8. Prédiction des effets des consonnes sur les formants des voyelles.

2.1.2. INTÉRACTIONS CV

Au-delà des caractéristiques articulatoires des consonnes et des voyelles, un élément qui déterminera le degré de variations contextuelles des consonnes sur les voyelles est le degré de conflit existant entre ces segments. La Table 9, rappelle les couples compatibles (les couples CV n'étant pas en conflit) selon la dimension considérée :

COMPATIBILITÉ ENTRE C ET V:		
	Hauteur de la langue - y	Position de la langue - x
ALVÉOLAIRE	/i, u, e, o/	/i, e, ε/
VÉLAIRE	/i, u, e, o/	/u/ + neutre autres V
UVULAIRE	/ε, a, ɔ/	/u, o, ɔ/

TABLE 9. Couples CV compatibles sur les dimensions hauteurs de la langue et position de la langue.

Comme on peut le voir dans le tableau ci-dessus, sur la dimension y, les voyelles les plus fermées seront compatibles avec les consonnes alvéolaires et vélaire alors que les voyelles plus ouvertes seront en accord avec la consonne uvulaire. Sur la dimension x, les consonnes

alvéolaires sont en accord avec les voyelles antérieures alors que les voyelles postérieures le seront avec la consonne uvulaire.

Les consonnes vélaires seraient plus compatibles avec la voyelle /u/ concernant la position de la langue si l'on suit la classification de Wood (1979) définissant le lieu d'articulation de la voyelle comme vélaire. Cependant, l'existence d'un allophone palatal lorsque ces consonnes vélaires se trouvent en contexte antérieur, fait que nous nous attendons à très peu d'effets contextuels des consonnes vélaires sur le deuxième formant des voyelles. C'est pour cela que nous déterminons les consonnes vélaires comme « neutres » sur la dimension horizontale.

La consonne uvulaire étant articulée à l'arrière du conduit vocal, nous la considérons comme compatible avec la voyelle /u/ même si la constriction est plus postérieure que celle du /u/. Nous noterons que la voyelle /x/ n'apparaît pas dans la Table 9. Ceci est dû au fait que cette voyelle est déterminée comme centrale et donc finalement compatible avec aucun des contextes. On rajoutera que le /x/ ne correspond pas totalement au schwa puisque les voyelles moyennes antérieures arrondies sont aussi incluses dans cette catégorie. Il se peut donc que les patterns observés ne soient pas totalement ceux attendus pour un schwa. Nous tiendrons compte de ce fait lors de la discussion de nos résultats. Enfin, la voyelle /a/ n'apparaît pas sur la dimension *x* dans la Table 9. Ceci est dû au fait que cette voyelle est articulatoirement antérieure mais définie comme étant une voyelle acoustiquement centrale. Ainsi, on n'a pas de prédiction concernant la compatibilité de cette voyelle avec quelque groupe que ce soit concernant la position de la langue, même si elle sera sans doute plus proche de celle des consonnes vélaires.

Un résumé des prédictions des effets contextuels sur les voyelles orales du français étudiées dans cette thèse est présenté Table 10. Le symbole « + » indique que l'on s'attend à des effets contextuels sur la voyelle, et donc une diminution de sa stabilité. Le symbole « - » indique au contraire que, de par une certaine compatibilité entre C et V aucun effet particulier de la consonne sur la voyelle n'est attendu. La voyelle maintiendrait ses caractéristiques dans ce contexte.

Afin de pouvoir croiser les effets contextuels avec les prédictions effectuées sur les voyelles elles-mêmes, nous avons représenté sur notre tableau les échelles de stabilité de nos trois théories (théorie quantique, théorie dispersion-focalisation et le modèle DAC). Le dégradé de gris associé aux différentes voyelles sert à représenter l'échelle de stabilité des voyelles selon le modèle DAC (plus on est dans le foncé, plus le DAC de la voyelle est élevé et plus celle-ci est stable). Les lettres « Q » et « F » en indice à côté des voyelles, permettent d'indiquer si les voyelles en question sont des voyelles stables selon les théories quantiques « Q » et de dispersion focalisation « F ».

	<i>/i/</i> ^{QF}	<i>/e/</i>	<i>/ɛ/</i>	<i>/a/</i> ^{QF}	<i>/ɤ/</i>	<i>/u/</i> ^{QF}	<i>/o/</i> ^F	<i>/ɔ/</i> ^F
ALV	-	-	+	+	+	+	+	+
VEL	-	-	+	+	+	-	-	+
UV	+	+	+	+	+	+	-	-

TABLE 10. Résumé des prédictions des effets contextuels sur la stabilité des voyelles.

Il faut donc imaginer que les effets contextuels pourront être modulés selon les particularités de nos voyelles. Par exemple, de par le conflit existant entre */ɤ/* et */i/* on s'attend à ce que le contexte uvulaire affecte l'acoustique de la voyelle */i/* même si ces deux segments ont un DAC similaire. Cependant, toutes les théories de la parole s'accordent pour donner un statut particulier à cette voyelle, la montrant très résistante à la variation. Il se peut donc que les effets attendus du contexte uvulaire sur l'acoustique de cette voyelle, restent marginaux. On peut imaginer que cette voyelle présentera peu de variation sur F2 puisqu'elle est quantique. Selon la théorie de la focalisation, on peut imaginer que cette voyelle soit variable en contexte uvulaire mais qu'elle maintienne une forte proximité entre F3 et F4 et de ce fait garde sa saillance perceptive.

2.2. METHODE

L'ensemble des analyses menées dans cette étude a été effectué à partir du corpus de parole conversationnelle NCCFr (Torreira et al. 2010) décrit en introduction générale 1.6.2.1. Un long processus de filtrage et de nettoyage des données a été effectué en amont pour éviter les erreurs d'analyses et contrôler au maximum les données. La procédure utilisée, décrite ci-dessous, permet de comprendre pourquoi d'un corpus de 36h, nous n'analysons que 18.5k voyelles.

2.2.1. MATÉRIEL LINGUISTIQUE ET PRÉTRAITEMENT

Dans cette étude, nous analysons la coarticulation C-à-V à partir des 8 catégories vocaliques suivantes : /i, e, ε, a, x, u, o, ə/ (avec /x/= /ø, œ, ə/).

2.2.1.1. Caractéristiques formantiques : procédure d'extraction & rejet des valeurs aberrantes

Les valeurs de formants (F1, F2, F3 et F4) ont été extraites en utilisant l'algorithme Burg du logiciel Praat (Boersma & Weenink, 2014) sur un alignement automatique forcé (Gauvain et al., 2002). La détection des pics d'amplitude est déterminée dans une bande inférieure à 5 kHz pour les hommes et à 5.5kHz pour les femmes. Les valeurs formantiques analysées dans ce travail correspondent à la moyenne des valeurs extraites à 1/3, 1/2 et 2/3 de chaque voyelle. Plus précisément, les mesures de formants ont été collectées dans des trames consécutives de 5 ms, les valeurs moyennes de trois trames consécutives sont calculées autour de 1/3, 1/2 et 2/3 de chaque voyelle, puis la moyenne de ces trois valeurs est calculée pour obtenir une seule valeur par voyelle. Afin d'éliminer les valeurs de formants aberrantes, un filtre a été appliqué selon les mêmes critères que ceux utilisés dans Gendrot & Adda-Decker (2005) dont les valeurs sont reportées Table 11.

HOMMES									
	/i/	/e/	/ɛ/	/a/	/ø/	/œ/	/u/	/o/	/ɔ/
F1	<750	<800	<1000	<1000	<900	<1000	<900	<900	<900
F2	1500- 2500	1100- 2400	1200- 2300	800- 2300	700- 2000	800- 2000	400- 1500	600- 1600	600- 1500
F3	>2000	>2000	>2000	>1800	>1700	>2000	>1400	>1500	>1500

FEMMES									
	/i/	/e/	/ɛ/	/a/	/ø/	/œ/	/u/	/o/	/ɔ/
F1	<900	<900	<1100	<1100	<1000	<1100	<1000	<1000	<1000
F2	1600- 3100	1400- 3000	1700- 2700	900- 2300	700- 2300	800- 2400	400- 1500	600- 2000	600- 2000
F3	>2500	>2200	>2000	>1900	>1800	>2000	>1800	>2100	>2000

TABLE 11. Valeurs limites des 3 premiers formants de chaque voyelle testée pour les hommes et les femmes pour le filtrage des valeurs aberrantes selon Gendrot & Adda-Decker (2005).

2.2.1.2. Premier filtrage des contextes consonantiques

À l'issue de ces procédures, une base de 350 851 voyelles a été constituée. Nous avons alors procédé à une première sélection de voyelles précédées ou suivies d'une consonne alvéolaire /t, d, s, z, l, n/, vélaire /k, g/, uvulaire /ʁ/ ou labiale /p, b, f, v, m/ (195k), comme illustré Figure 4.



FIGURE 4. Illustration des séquences sélectionnées après un premier filtrage des données.

2.2.1.3. Contrôle de la durée

Afin d'éliminer les valeurs aberrantes et les possibles erreurs de segmentation automatique, nous avons sélectionné les voyelles dont la durée était comprise entre 50 et 150 ms sur le modèle de (Audibert et al. 2015) dans une étude menée sur le même corpus.

2.2.1.4. Sélection des locuteurs

Afin de limiter la variabilité due aux locuteurs, notre étude ne porte que sur les productions des locuteurs masculins ayant au moins produit 90 exemplaires de chaque catégorie de voyelles étudiées. Ces critères nous ont conduits à la sélection de 15 locuteurs sur les 22 locuteurs masculins du corpus.

2.2.1.5. Nettoyage des données

Un second filtrage des données a été conduit afin d'éliminer les valeurs extrêmes pour chaque voyelle. Nous avons écarté tous les exemplaires de chaque voyelle dont les deux premiers formants se situaient à ± 2 écarts-types (ET) de la moyenne (M) de la catégorie. La Table 12 ci-dessous donne les valeurs moyennes des deux premiers formants de chaque voyelle dans notre échantillon, les seuils à ± 2 ET ainsi que le nombre d'exemplaires rejetés et, entre parenthèses, le pourcentage que cela représente pour la catégorie de voyelle en question. Chaque valeur correspond à la valeur arrondie mais tous les calculs ont été faits à partir des données brutes (non arrondies). Ceci explique pourquoi les valeurs des colonnes seuils minimum (min) et maximum (max) ne correspondent pas exactement à $M-2ET$ et $M+2ET$ respectivement.

	F1				F2				Rejet N (%)
	M	ET	min (M-2ET)	max (M+2ET)	M	ET	min (M-2ET)	max (M+2ET)	
/i/	338	99	139	536	1971	147	1678	2265	579 (11)
/e/	375	57	260	489	1867	139	1589	2145	665 (9)
/ɛ/	451	82	286	616	1752	149	1454	2050	659 (8)
/a/	564	93	378	750	1500	173	1155	1846	1090 (8)
/ɔ/	403	77	250	556	1523	168	1186	1859	431 (10)
/u/	405	108	188	621	1127	203	720	1533	108 (6)
/o/	403	62	280	526	1123	207	709	1538	102 (6)
/ɔ̃/	490	92	306	750	1273	188	898	1648	218 (7)

TABLE 12. Moyenne (M), écart-type (ET) et seuils minimum (min) et maximum (max) correspondant à M± 2ET pour chaque voyelle et nombre d'exemplaires rejetés et pourcentage correspondant entre parenthèses après application de ce filtre.

2.2.1.6. Normalisation des données

Afin de réduire au maximum les différences entre locuteurs, nous avons normalisé les valeurs de nos quatre premiers formants, initialement en Hertz, en z-score par locuteur selon la formule de Lobanov (1971) présentée Équation 1 :

$$F_{i_{lobanov}} = ((F_i - \mu_i) / (\delta_i))$$

EQUATION 1. Transformation en z-score de Lobanov.

où μ_i est la fréquence formantique moyenne calculée pour toutes les voyelles d'un locuteur pour un formant i alors que δ_i réfère à l'écart-type, et que F_i est la valeur du formant i pour une observation en particulier. Cette procédure permet de minimiser les différences entre locuteur en ré-échantillonnant le centre de l'espace vocalique de chaque locuteur et la dispersion moyenne depuis ce centre.

2.2.1.7. Les contextes examinés dans cette étude

Dans cette étude, nous considérons les influences contextuelles des consonnes sur l'acoustique des voyelles orales du français à partir de trois groupes de consonnes (tests) définis selon leur lieu d'articulation et leurs potentiels effets (voir section 2.1.1) sur nos voyelles : 1) le groupe alvéolaire « ALV » constitué des consonnes /t, d, s, z, l, n/, dont le principal effet est d'attirer le F2 des voyelles vers leur locus à 1800 Hertz ; 2) le groupe vélaire « VEL » avec les consonnes /k, g/ dont le lieu d'articulation s'adapte à celui des voyelles environnantes : palatale contexte [+antérieur], vélaire ailleurs, et 3) le groupe uvulaire « UV » constitué d'une seule consonne /ʁ/ connue pour élever le F1 des voyelles et abaisser leur F2. Enfin, les consonnes tests précédemment citées pouvaient se trouver à gauche comme à droite de la voyelle, sachant que le contexte opposé était toujours une consonne labiale « LAB » parmi les consonnes /p, b, f, v, m/. Ainsi, le contexte opposé était toujours contrôlé. Nous avons choisi les consonnes labiales car leur production n'implique pas d'articulation linguale et de ce fait leur effet sur les formants des voyelles est moindre. Un résumé des contextes sélectionnés est présenté Figure 5.

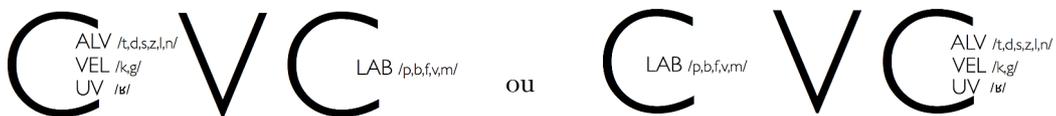


FIGURE 5. Illustration des contextes consonantiques testés dans cette étude.

Nous avons exclu les contextes symétriques du type « ALV_ALV » car cela ne faisait qu'augmenter le déséquilibre dans nos données. Le contexte alvéolaire déjà surreprésenté, du fait de sa forte fréquence en français, s'en trouvait favorisé alors que l'inclusion des contextes symétriques avait peu d'impact sur le nombre d'occurrences des deux autres contextes (il est peu fréquent de trouver une voyelle entre deux /ʁ/ par exemple).

2.2.1.8. Base finale obtenue après filtrage

À l'issue de cette procédure de sélection, nous avons obtenu une base de données de 18 578 observations des voyelles /i, e, ε, x, a, u, o, ɔ/ (avec /x/= /ø, œ, ə/). La Table 13

présente, étape par étape, la procédure de sélection des voyelles qui nous a conduits à l'obtention de cet échantillon de 18.5k voyelles.

ÉTAPES DE SÉLECTION	N
Après extraction des formants des 8 voyelles testées	350 851
Sélection des contextes d'intérêt	195 753
Filtre de durée : sélection des V de 50-150 ms	122 844
Filtre sexe locuteur : sélection des locuteurs masculins	60 897
Sélection des locuteurs avec $N \geq 90$ exemplaires des 8 voyelles	45 432
Nettoyage à +/- 2ET de F1 et F2	41 577
Base finale après le choix des contextes examinés	18 578

TABLE I3. Détail de la procédure de sélection des données étape par étape et effectifs.

Un récapitulatif du matériel disponible par voyelle est présenté Table 14. Pour une vision complète du matériel par locuteur et durée voir Annexe 1 page 189.

	/i/	/e/	/ɛ/	/a/	/ɔ/	/u/	/o/	/ɔ/
LAB_ALV	582	746	1333	2632	416	159	242	380
LAB_VEL	105	90	398	454	72	27	187	38
LAB_UV	59	67	558	775	63	412	4	313
ALV_LAB	698	1715	704	2144	861	131	55	170
VEL_LAB	130	38	27	88	107	78	10	390
UV_LAB	158	68	219	269	32	136	123	130
<i>Total</i>	<i>1732</i>	<i>2709</i>	<i>3239</i>	<i>6362</i>	<i>1551</i>	<i>943</i>	<i>621</i>	<i>1421</i>

TABLE I4. Résumé du matériel sélectionné.

2.2.2. MÉTRIQUES ACOUSTIQUES

Cette sous-section a pour but de présenter les métriques acoustiques utilisées dans cette étude pour répondre à nos objectifs. Celles-ci comprennent une métrique de coarticulation, mesurant les DÉPLACEMENTS de formants, et deux métriques de variabilité (DISPERSION et DISTINCTIVITÉ) que nous détaillerons plus bas. Si dans une étude préliminaire, dont les résultats sont présentés section 2.3.2.1, nous avons mesuré la variabilité des voyelles tous contextes confondus afin d'identifier si certaines voyelles étaient réellement moins variables que d'autres, l'intérêt principal de ces métriques de variabilité est d'identifier l'impact du contexte consonantique sur la variabilité des voyelles.

2.2.2.1. Métrique de coarticulation

i. Les déplacements formantiques comme mesure de coarticulation C-à-V

La coarticulation C-à-V est mesurée à partir des déplacements de formants en fonction du contexte consonantique (ALV, VEL, UV) sur les dimensions F1 et F2, ainsi que les dimensions F3 et F4 pour la voyelle /i/. Cela correspond aux changements de valeurs des formants de la voyelle selon le contexte consonantique dans lequel elle se trouve. La Figure 6 illustre cette mesure de coarticulation.

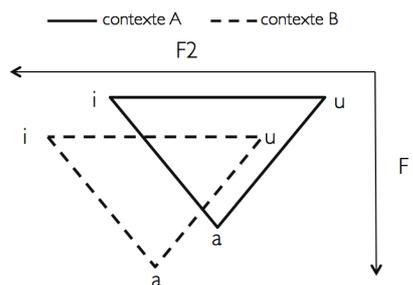


FIGURE 6. Illustration de la mesure des effets contextuels sur les déplacements formantiques des voyelles

2.2.2.2. Métriques de variabilité

i. Les distances euclidiennes sur F1-F2-F3-F4 comme mesure de dispersion

Le premier aspect de variabilité que nous testons dans cette étude concerne la dispersion des voyelles à l'intérieur de leur catégorie vocalique (i.e. la variabilité intra-catégorie), informant de la constance dans la réalisation acoustique de ces voyelles. Moins les exemplaires d'une catégorie vocalique sont dispersés, moins la zone acoustique sur laquelle elle s'étend est importante, et plus la cible acoustique de cette voyelle est atteinte avec constance. Nous avons mesuré cette dispersion à partir des distances euclidiennes multidimensionnelles des exemplaires d'une catégorie vocalique à une valeur de référence pour la catégorie en question sur les dimensions F1, F2, F3 et F4. Une représentation schématique de la mesure est fournie Figure 7. La figure ne représente que la variabilité du /a/ sur les deux dimensions F1 et F2, mais le calcul comprend aussi F3 et F4 et cela a été effectué pour chacune des 8 catégories vocaliques.

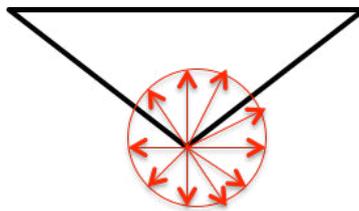


FIGURE 7. Représentation schématique, sur F1/F2, de la mesure de variabilité intra-catégorie utilisée dans cette thèse.

Nous avons calculé la racine carrée de la somme des carrés des écarts entre les exemplaires des voyelles et une valeur de référence pour la catégorie vocalique correspondante (valeur de référence que nous définirons en 2.2.2.3). Cette mesure nous informe de la dispersion des exemplaires d'une catégorie vocalique à une valeur de référence pour la catégorie vocalique, soit de la variabilité dans la réalisation des cibles acoustiques (Huet & Harmegnies, 2000 ; Audibert & Fougeron, 2012). Si on prend l'exemple de la distance F1-F2-F3-F4 pour la voyelle /a/, la formule est présentée Equation 2.

$$D_{F1234} = \sqrt{(F1ia - CF1a)^2 + (F2ia - CF2a)^2 + (F3ia - CF3a)^2 + (F4ia - CF4a)^2}$$

EQUATION 2. Distance Euclidienne F1-F4

F[1234]ia représente la valeur d'un des quatre premiers formants de l'exemplaire i de la voyelle /a/ et CF[1234]a, symbolise la valeur de référence pour chacun des quatre premiers formants de la voyelle /a/. La procédure d'obtention des valeurs de référence est décrite dans la sous-section 2.2.2.3.

ii. *Distinctivité acoustique*

L'examen de la dispersion acoustique des exemplaires sera complété par l'examen de leur distinctivité acoustique, estimée à partir des résultats d'une classification automatique des exemplaires sur la base de leurs caractéristiques acoustiques. Cette mesure nous permettra de voir : 1) si certaines voyelles sont plus « distinctives » que d'autres en regardant le taux de classification correcte et les métriques associées à la matrice de confusion (scores de sensibilité et de précision) ; 2) si cette « distinctivité » est fonction du contexte consonantique en regardant les résultats de la classification par contexte ; 3) et d'identifier les patterns de confusion en fonction des contextes. La classification acoustique est une mesure indirecte de la distinctivité acoustique des voyelles les unes par rapport aux autres dans le sens où la classification repose sur les principales propriétés acoustiques des voyelles : les quatre premiers formants (en z-score).

Pour cette analyse, nous avons eu recours à la bibliothèque 'MASS' (Venables & Ripley 2002) du logiciel R (R Development Core Team, 2008) dans le but d'effectuer une Analyse Linéaire Discriminante (LDA). Les 8 catégories de voyelles représentent nos classes d'origine et les quatre premiers formants (en z-score) nos variables discriminantes (ou prédicteurs).

Nous avons, dans un premier temps, construit notre classifieur à partir d'un échantillon d'entraînement, le même qui a servi à calculer les valeurs de référence pour la métrique de dispersion, décrit dans la sous-section 2.2.2.3. Puis, nous l'avons testé sur notre corpus de 18 578 observations de parole conversationnelle.

2.2.2.3. Valeurs de références & set d'entraînement : échantillon de parole lue

Les résultats d'une mesure de dispersion dépendent fortement des valeurs de références que l'on prend. Il est donc important de bien les choisir. Lorsque les données sont bien équilibrées, prendre la moyenne de tous les exemplaires de la catégorie convient tout à fait. Seulement, lorsque l'on travaille sur des données très déséquilibrées, comme c'est le cas dans cette étude, il est risqué de suivre la même procédure. Comment tester si le contexte consonantique X amène plus de dispersion que le contexte Y, si la mesure de dispersion a été calculée à partir d'une valeur de référence contenant majoritairement des exemplaires de la voyelle en contexte X ? Il est important de s'assurer que nos valeurs de références ne soient pas le reflet d'un déséquilibre dans les données.

Aussi, la mesure de distinctivité acoustique, consistant en une classification acoustique automatique de nos exemplaires de voyelles, implique le recours à un échantillon de référence, appelé échantillon d'entraînement, à partir duquel est construit le modèle, ou classifieur, que l'on teste dans un second temps sur un échantillon test, en l'occurrence nos 18.5k voyelles. Pour les mêmes raisons que celles précédemment citées, il est important que cet échantillon de référence soit construit sur des données équilibrées tant au niveau des contextes dans lesquels apparaissent les voyelles qu'au niveau du nombre d'exemplaires de chaque catégorie de voyelles. En effet, il est important que la probabilité a priori soit la même pour chaque voyelle pour ne pas biaiser les résultats. De plus, il est important de s'assurer qu'une partie des données de l'échantillon d'entraînement ne se retrouve pas dans l'échantillon test (une observation incluse dans l'échantillon à partir duquel a été construit le modèle aura de grandes chances d'être bien identifiée par celui-ci durant la phase test).

Pour pallier ces problèmes, nous avons décidé de créer un échantillon de référence externe au corpus étudié, en recourant au corpus BREF (Lamel, Gauvain & Eskenazi, 1991) de parole lue.

Pour obtenir notre échantillon de référence, nous avons appliqué la même procédure de prétraitement que celle décrite dans la section 2.2.1 (de 2.1.1.1 à 2.2.1.6) pour l'obtention de notre base de données. Puis nous avons procédé à un ré-échantillonnage avec remplacement, à l'aide des fonctions `group_by` et `sample_n` de la bibliothèque 'dplyr'

(Wickham et al. 2017) sur R, des données afin d'obtenir un échantillon de 50 exemplaires de chacune des 8 voyelles dans chacun des 16 contextes (c.-à.-d. toutes les combinaisons des alvéolaire, vélaire, uvulaire, labial), soit 6400 voyelles. Dix simulations de ce type ont été effectuées avant d'être regroupées au sein d'un même échantillon de 64 000 voyelles. Pour plus de détails concernant le contenu de nos 10 simulations se référer Annexe 2 page 203.

Les moyennes de chaque formant z-scoré, présentées Table 15, ont ensuite servi à la définition de nos valeurs formantiques de référence par catégorie vocalique pour les mesures de dispersion. L'ensemble des 64 000 exemplaires des 8 voyelles a servi à la création du classifieur automatique.

	F1	F2	F3	F4
/i/	-1.13 (0.46)	1.34 (0.29)	1.54 (0.82)	0.40 (0.95)
/e/	-0.41 (0.35)	0.87 (0.30)	0.11 (0.56)	-0.001 (0.90)
/ɛ/	0.26 (0.50)	0.49 (0.33)	-1.17 (0.53)	0.06 (0.95)
/a/	1.53 (0.58)	-0.46 (0.45)	-0.48 (0.59)	0.08 (1.02)
/ɔ/	-0.22 (0.45)	-0.54 (-0.53)	-0.78 (0.57)	-0.54 (0.75)
/u/	-0.62 (0.62)	-1.76 (0.54)	-0.60 (0.72)	-0.41 (0.88)
/o/	-0.29 (0.34)	-1.82 (0.43)	-0.50 (0.64)	-0.58 (0.79)
/ɔ̃/	0.22 (0.49)	-1.41 (0.45)	-0.67 (0.64)	-0.46 (0.74)

TABLE 15. Moyennes et écarts-types des formants de chaque voyelle de l'échantillon de référence construit à partir de parole lue.

2.2.3. ANALYSES STATISTIQUES

Les données ont été analysées avec une série de modèles mixtes en utilisant la bibliothèque 'lme4' (Bates, Maechler, Bolker & Walker, 2015) du logiciel R (R Development Core Team, 2008). La bibliothèque 'lmerTest' (Kuznetsova, Brockhoff & Chistensen, 2016) a

été utilisée pour obtenir des valeurs de p sur la base d'approximations Satterwaite donnant une estimation très conservatrice des valeurs de p pour les régressions linéaires. Afin de tester l'effet de chaque facteur ainsi que celui d'éventuelles interactions, nous avons utilisé le test du ratio de vraisemblance implémenté dans la fonction *anova* en comparant le modèle avec un certain facteur (ou interaction) avec un modèle sans ce facteur (ou interaction) en particulier. Enfin, la valeur R2 associée à chaque modèle a été calculée avec la fonction *r.squaredGLMM* de la bibliothèque 'MuMIn' (Kamil Barton, 2013) dans le but de déterminer le pouvoir prédictif de nos modèles. Le résultat de cette mesure est un coefficient variant entre 0 et 1. Plus on s'approche de 1, plus le pouvoir prédictif du modèle est fort. De plus, cette fonction nous permet de mesurer la pertinence de notre structure aléatoire. En effet, de cette fonction, nous obtenons deux valeurs de R : R2m ou valeur R2 marginale (indiquant la variance expliquée par les facteurs fixes uniquement) ; et R2c ou la valeur R2 conditionnelle associée à la fois aux effets fixes et aléatoires du modèle. Si R2c est plus élevé que R2m, c'est que le recours à un modèle mixte est pertinent pour l'analyse des données. Les analyses des contrastes a posteriori ont été effectuées avec la fonction *lsmeans* (bibliothèque 'emmeans', Lenth et al., 2018). Les graphiques présentés dans la section résultats, sont réalisés à partir des valeurs ajustées par les modèles mixtes obtenus à partir de la fonction *effect* de la bibliothèque 'effects' (Fox et al., 2018).

Les structures des différents modèles seront décrites une à une dans la partie résultats.

2.3. RESULTATS

2.3.1. QUELS DÉPLACEMENTS FORMANTIQUES SELON LES PROPRIÉTÉS DE C & V ?

Dans cette section, nous rapportons les résultats de l'analyse des effets coarticulaires des consonnes (ALV, VEL, UV) sur nos 8 catégories vocaliques testées /i, e, ε, a, x, u, o, ɔ/ à partir de notre métrique DEPLACEMENT (voir page 63) qui mesure les changements de formants des voyelles selon le contexte consonantique. Nos analyses reposent sur les deux premiers formants (en z-score) des voyelles, mais une analyse des effets contextuels sur les valeurs de F3 et F4 a aussi été incluse pour la voyelle /i/. L'objectif est de confirmer les effets contextuels attendus selon les propriétés articulatoires des segments impliqués.

Pour cela, nous avons construit deux modèles mixtes (un pour F1 et un pour F2) en incluant les facteurs : VOYELLE (8 niveaux) et CONTEXTE (3 niveaux) comme facteurs fixes en interaction. Pour la voyelle /i/ nous avons construit deux autres modèles pour les formants F3 et F4 avec uniquement le facteur CONTEXTE dans la structure fixe. La structure aléatoire de nos modèles est constituée : d'un intercept par locuteur et par mot ainsi que d'une pente aléatoire par locuteur pour nos deux facteurs fixes. Cela correspond à la variabilité entre locuteurs de l'effet du contexte et de la voyelle sur les formants. L'interaction n'a pas pu être intégrée à la pente aléatoire pour des problèmes de convergence. Aucune pente aléatoire n'a été intégrée au facteur aléatoire « mot » pour les mêmes raisons. Pour plus d'informations sur les fonctions utilisées se rendre section 2.2.2.3. Les résultats seront présentés formant par formant.

2.3.1.1. Déplacement sur F1 (z-score)

L'analyse menée sur le premier formant des voyelles /i, e, ε, a, x, u, o, ɔ/ indique un R^2 marginal de 0.69 et un R^2 conditionnel de 0.77 et met en évidence un effet du contexte consonantique ($\chi^2(1) = 63.39, p < .0001^{***}$). L'analyse des contrastes de ce facteur, présentée Table 16, révèle que le premier formant des voyelles est plus élevé en contexte uvulaire qu'en contexte alvéolaire ou vélaire. Comme on peut le voir Figure 8, cet effet est considérable et ressort pour toutes les voyelles testées.

contraste	β	ET	df	z.ratio	valeur de p
ALV-UV	-0.70	0.03	Inf	-24.2	<.0001
ALV-VEL	0.03	0.03	Inf	1.31	0.39
UV-VEL	0.73	0.03	Inf	20.7	<.0001

TABLE 16. Résultats de l'analyse des contrastes du facteur CONTEXTE.

Sans surprise, un effet du type de voyelle a été trouvé ($\chi^2(7) = 574.1, p < .0001^{***}$). Enfin, une interaction entre nos deux facteurs est ressortie de cette analyse ($\chi^2(14) = 149.02, p < .0001^{***}$): celle-ci indique que pour certaines voyelles, des différences entre les contextes alvéolaire et vélaire ressortent. C'est le cas de la voyelle mi-ouverte / ϵ / dont le F1 est plus élevé en contexte alvéolaire qu'en contexte vélaire ($\beta = 0.12, SE = 0.04, z.ratio = 2.81, p < .05^*$) et de la voyelle /u/ pour laquelle le F1 est plus bas en contexte alvéolaire qu'en contexte vélaire ($\beta = -0.28, SE = 0.07, z.ratio = -3.74, p = .0005^{***}$).

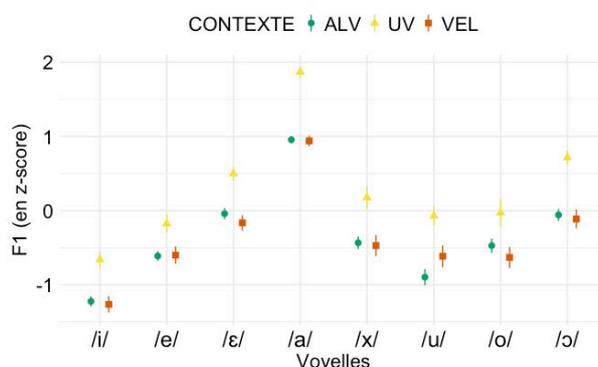


FIGURE 8. F1 (z-score) des voyelles en fonction du contexte consonantique (ALV, VEL, UV).

De plus, les observations des valeurs de z.ratio présentées Table 17, montrent que certaines voyelles sont plus affectées que d'autres par les effets contextuels sur la dimension F1. Par exemple, l'écart entre les contextes ALV-UV et VEL-UV est plus important pour les voyelles les plus ouvertes /a, ϵ , ɔ / allant dans le sens des résultats de Stevens & House

(1966) pour l'anglais américain et Recasens (1985) pour le catalan. De plus, nos résultats montrent que cela est aussi le cas pour les voyelles fermées /i, u/.

voyelles	contraste	β	SE	z.ratio	valeur de p
/i/	ALV-UV	-0.56	0.05	-10.43	<.0001
	ALV-VEL	0.04	0.05	0.78	.72
	UV-VEL	0.60	0.07	8.72	<.0001
/e/	ALV-UV	-0.44	0.06	-7.43	<.0001
	ALV-VEL	-0.01	0.05	-0.23	<.97
	UV-VEL	0.42	0.08	5.56	<.0001
/ɛ/	ALV-UV	-0.53	0.044	-12.26	<.0001
	ALV-VEL	0.12	0.043	2.81	<.014
	UV-VEL	0.66	0.057	11.61	<.0001
/a/	ALV-UV	-0.91	0.04	-24.8	<.0001
	ALV-VEL	0.01	0.03	0.47	.887
	UV-VEL	-0.93	0.05	20.21	<.0001
/ɔ/	ALV-UV	-0.61	0.08	-7.73	<.0001
	ALV-VEL	0.04	0.06	0.58	<.83
	UV-VEL	0.65	0.1	6.61	<.0001
/u/	ALV-UV	-0.83	0.07	-11.97	<.0001
	ALV-VEL	-0.28	0.07	-3.74	<.0005
	UV-VEL	0.54	0.08	6.37	<.0001
/o/	ALV-UV	-0.44	0.096	-4.60	<.0001
	ALV-VEL	0.15	0.069	2.27	<.06
	UV-VEL	0.60	0.11	5.58	<.0001
/ɔ̃/	ALV-UV	-0.77	0.053	-14.51	<.0001
	ALV-VEL	0.05	0.068	0.8	0.71
	UV-VEL	0.83	0.072	11.44	<.0001

TABLE 17. Résultats de l'analyse des contrastes du facteur contexte par voyelle sur F1.

2.3.1.2. Déplacement sur F2 (z-score)

L'analyse menée sur le second formant des voyelles /i, e, ε, a, x, u, o, ɔ/ indique un R² marginal de 0.80 et un R² conditionnel de 0.87 et met en évidence un effet du contexte consonantique ($\chi^2(1)= 59.95, p< .0001^{***}$) : Le second formant de toutes les voyelles est plus élevé en contextes alvéolaire et vélaire qu'en contexte uvulaire comme le révèle l'analyse des contrastes de ce facteur présentés Table 18. L'analyse des contrastes par voyelle révèle que cela est avéré pour toutes les voyelles. Comme attendu un effet de la voyelle ressort ($\chi^2(7)= 391.96, p< .0001^{***}$).

contraste	β	SE	df	z.ratio	valeur de p
ALV-UV	0.56	0.28	Inf	21.15	<.0001
ALV-VEL	-0.001	0.02	Inf	-0.06	0.1
UV-VEL	-0.6	0.03	Inf	-21.93	<.0001

TABLE 18. Résultats de l'analyse des contrastes du contexte sur le F2 des voyelles.

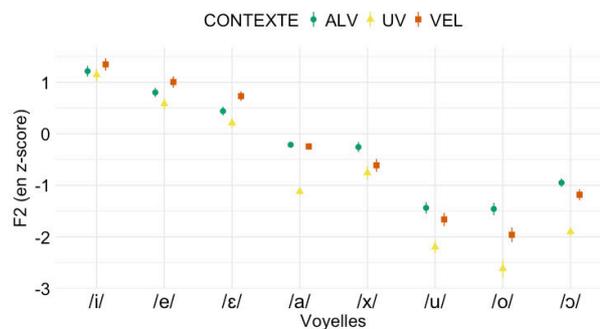


FIGURE 9. F2 (z-score) des voyelles en fonction du contexte consonantique.

Enfin, une interaction entre nos principaux facteurs est trouvée ($\chi^2(14)=994.92, p=< 2.2e-16^{***}$). Comme on peut le voir Figure 9, celle-ci indique une différence de F2 dans les trois contextes consonantiques pour les voyelles postérieures (ALV > VEL > UV) et les voyelles antérieures moyennes /e, ε/ (VEL > ALV > UV). La voyelle /i/ qui montre très peu de variation contextuelle sur F2, présente un F2 plus élevé en contexte vélaire qu'en contexte alvéolaire ou uvulaire, sans distinction entre ces deux contextes. Le F2 de /a/ est

plus élevé en contextes alvéolaire et vélaire qu'en contexte uvulaire alors que pour l'autre voyelle centrale /x/ les contextes uvulaires et vélares ne se distinguent pas.

voyelles	contraste	β	SE	z.ratio	valeur de p
/i/	ALV-UV	0.07	0.04	1.63	.23
	ALV-VEL	-0.13	0.04	-3.35	.0023
	UV-VEL	-0.20	0.05	-3.91	.0003
/e/	ALV-UV	0.22	0.05	4.66	<.0001
	ALV-VEL	-0.20	0.04	-4.94	<.0001
	UV-VEL	-0.42	0.06	-7.37	<.0001
/ɛ/	ALV-UV	0.23	0.04	6.28	<.0001
	ALV-VEL	-0.29	0.03	-8.75	<.0001
	UV-VEL	-0.52	0.04	-12.24	<.0001
/a/	ALV-UV	0.91	0.03	28.37	<.0001
	ALV-VEL	0.03	0.03	1.35	.37
	UV-VEL	-0.87	0.03	-25.70	<.0001
/ɔ/	ALV-UV	0.50	0.06	8.06	<.0001
	ALV-VEL	0.35	0.05	7.19	<.0001
	UV-VEL	-0.15	0.07	-1.97	.1
/u/	ALV-UV	0.76	0.06	13.81	<.0001
	ALV-VEL	0.22	0.06	3.9	.0003
	UV-VEL	-0.54	0.07	-8.27	<.0001
/o/	ALV-UV	1.16	0.08	15.34	<.0001
	ALV-VEL	0.5	0.05	9.41	<.0001
	UV-VEL	-0.66	0.08	-8	<.0001
/ɔ̃/	ALV-UV	0.95	0.04	21.86	<.0001
	ALV-VEL	0.23	0.05	4.45	<.0001
	UV-VEL	-0.72	0.06	-13.11	<.0001

TABLE 19. Résultats de l'analyse des contrastes du facteur contexte par voyelle sur F2.

Plus intéressant, les observations des z.ratio montrent que les effets contextuels sur F2 sont bien plus importants pour les voyelles postérieures et la voyelle /a/ que pour les autres voyelles (voir différences ALV-UV Table 19). Nos analyses révèlent très peu d'effets du contexte vélaire sur le F2 des voyelles : 1) les voyelles antérieures, caractérisées par un F2 élevé, présentent un F2 maximale haut dans ce contexte ; et 2) les voyelles postérieures, caractérisées par un F2 bas, montrent dans ce contexte des valeurs de F2 intermédiaires entre celles observées en contexte alvéolaire et celles observées en contexte uvulaire.

2.3.1.3. Déplacement sur le F3 et le F4 de la voyelle /i/

L'analyse de 1732 /i/ a révélé un effet du contexte consonantique sur les valeurs du troisième ($\chi^2(2)=34.242, p<.0001^{***}$) et quatrième formants de la voyelle ($\chi^2(2)=21.107, p<.0001^{***}$). L'analyse des contrastes, présentée Table 20, indique que le F3 et le F4 de /i/ est abaissé en contexte uvulaire comparé aux contextes alvéolaire et vélaire (voir Figure 10). Ceci s'explique du fait que les troisième et quatrième formants de la voyelle /i/ sont dépendants de la taille de la constriction palatale et de celle de la cavité antérieure (Fant, 1960). Il est intéressant de remarquer que pour cette voyelle, les effets contextuels sont plus fortement marqués sur les dimensions F3 et F4 que sur la dimension F2 (cf Figure 9).

	contraste	β	SE	df	z.ratio	valeur de p
	ALV-UV	0.91	0.1	Inf	9.42	<.0001 ^{***}
F3	ALV-VEL	-0.09	0.07	Inf	-1.13	0.5
	UV-VEL	-0.99	0.12	Inf	-8.55	<.0001 ^{***}
	ALV-UV	0.81	0.11	Inf	7.1	<.0001 ^{***}
F4	ALV-VEL	0.05	0.12	Inf	0.55	0.85
	UV-VEL	-0.76	0.12	Inf	-6.11	<.0001 ^{***}

TABLE 20. Résultats de l'analyse des contrastes du facteur contexte sur le F3 et le F4 de la voyelle /i/.

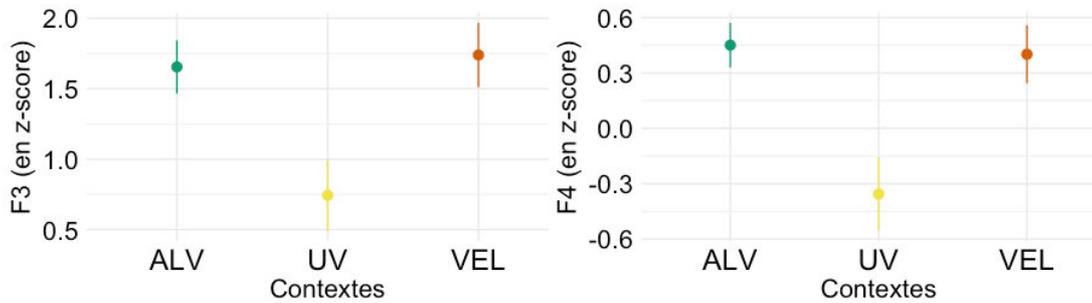


FIGURE 10. F3 (à gauche) et F4 (à droite) en z-score de la voyelle /i/ en fonction du contexte consonantique.

En résumé, l'ensemble de nos analyses sur la direction des DÉPLACEMENTS formantiques montre qu'il y a bien des effets contextuels dans notre corpus et que ces derniers vont dans les directions attendues : le contexte uvulaire élève le premier formant des voyelles alors que les contextes vélaire et alvéolaire contraints par la réalisation de l'occlusion tendent à l'abaisser. Le second formant des voyelles est lui abaissé en contexte uvulaire alors qu'il est élevé en contexte alvéolaire. Pour la voyelle /i/ un abaissement de F3 et F4, plus important que celui de F2, est aussi observé en contexte uvulaire. Très peu d'effets contextuels du contexte vélaire sont observés sur F2. Les voyelles antérieures montrent un F2 maximale élevé dans ce contexte, alors que les voyelles postérieures en contexte vélaire montrent des valeurs de F2 intermédiaires entre celles observées en contexte alvéolaire et uvulaire. Enfin, les effets contextuels sont plus forts sur certaines voyelles. C'est le cas notamment des voyelles les plus ouvertes et les plus fermées sur F1 et des voyelles postérieures et la voyelle /a/ sur F2.

2.3.2. LA COARTICULATION JOUE-T-ELLE SUR LA VARIABILITÉ DES VOYELLES ?

L'objectif de cette section est de déterminer si la coarticulation C-à-V affecte la variabilité des voyelles et d'identifier si certains contextes amènent plus de variabilité que d'autres dans les réalisations acoustiques des voyelles. À partir de nos deux métriques de variabilité,

nous testerons si, et de quelle manière, les différents contextes consonantiques affectent la variabilité des voyelles, en termes de dispersion et de distinctivité acoustiques.

Dans un premier temps, nous rapporterons les résultats d'une étude préliminaire dont le but était de déterminer si certaines voyelles étaient ou non plus stables que d'autres comme cela est suggéré par différentes théories de la parole (voir section 1.3). Ces résultats constitueront une base de référence pour la suite de nos analyses par contexte et nous permettront notamment de déterminer si les voyelles les plus variables sont aussi plus fortement affectées par le contexte consonantique et si la cause en est un contexte en particulier.

2.3.2.1. Analyse préliminaire : des voyelles intrinsèquement plus variables que d'autres ?

i. Des voyelles plus dispersées que d'autres (indépendamment du contexte consonantique) ?

Dans cette section nous rapporterons les résultats de notre mesure de dispersion multidimensionnelle consistant en des mesures de distances euclidiennes sur quatre dimensions F1-F2-F3-F4, des exemplaires des voyelles aux valeurs de référence de leur catégorie vocalique. Afin d'affiner nos observations, nous regarderons aussi les distances à la valeur de référence pour chaque formant pris individuellement. Ceci nous permettra de voir par exemple si la forte variabilité d'une voyelle est attribuable à une ou deux dimensions acoustiques en particulier ou si cela se retrouve sur toutes les dimensions.

Afin de tester si certaines voyelles sont plus dispersées que d'autres au sein de leur catégorie vocalique, nous avons analysé les relations entre nos mesures de dispersions et le facteur VOYELLE, à 8 niveaux. Pour des contraintes liées à nos données, différents types d'analyses ont été menées. Les analyses de la mesure de dispersion multidimensionnelle, ou distances euclidiennes F1-F2-F3-F4, ont été effectuées à partir d'un modèle linéaire mixte après normalisation des données en prenant la racine carrée des valeurs de distances euclidiennes (meilleure solution parmi les différentes transformations des données testées). La structure aléatoire comprenait un intercept par locuteur et par mot et une pente

aléatoire par locuteur était incluse. Cela correspond à la variabilité entre locuteurs de la voyelle sur les mesures de dispersion. Aucune pente aléatoire n'a été intégrée au facteur aléatoire « mot » pour des problèmes de convergences et du fait que les 8 voyelles ne pouvaient être extraites depuis un même mot. Concernant nos distances à la valeur de référence par formant (DF1, DF2, DF3, DF4), quatre modèles linéaires mixtes généralisés ont été construits avec la même variable VOYELLE comme facteur fixe en précisant que les données suivaient une loi gamma. Pour ces quatre derniers modèles, seul un intercept par locuteur a été inclus en raison de problèmes de convergence. Pour plus d'informations sur les fonctions utilisées se rendre section 2.2.2.3.

Dispersion multidimensionnelle F1-F2-F3-F4

L'analyse des distances euclidiennes sur quatre dimensions F1-F2-F3-F4 de 18.5k voyelles /i, e, ε, a, x, u, o, ɔ/ indique un R^2 marginal de 0.07 et un R^2 conditionnel de 0.2 et met en évidence un effet de la voyelle ($\chi^2(7)=146.33, p < .0001^{***}$) qui traduit que certaines voyelles présentent une plus grande dispersion intra-catégorie que les autres.

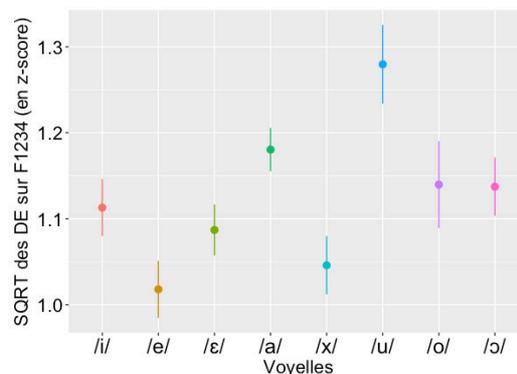


FIGURE 11. Racine carrée des distances euclidiennes F1-F2-F3-F4 (en z-score) par voyelle.

L'analyse des contrastes a posteriori et l'observation de la Figure 11 nous permettent de faire plusieurs constats : tout d'abord, on peut remarquer que la voyelle /u/ est la voyelle la plus dispersée parmi nos 8 voyelles testées ($\beta=0.099, SE=0.02, t= 4.69, p < .0001^{***}$) suivie de la voyelle /a/ ($\beta=1.18048, SE=0.013, t= 92.1, p < .0001^{***}$, valeurs Intercept). Contre toute attente, la voyelle présentant la plus faible dispersion intra-catégorie, n'est pas la voyelle /i/, mais la voyelle /e/ ($\beta = -0.16, SE=0.01, t= -11.91, p < .0001^{***}$) et la

voyelle notée /x/ qui inclut pourtant des exemplaires de /ø, œ, ə/ ($\beta=-0.13$, $SE=0.01$, $t=-9.49$, $p<.0001^{***}$). Les voyelles /i/ ($\beta=-0.07$, $SE=0.01$, $t=-4.93$, $p<.0001^{***}$), /ɛ/ ($\beta=-0.09$, $SE=0.01$, $t=-8.41$, $p<.0001^{***}$), /ɔ/ ($\beta=-0.04$, $SE=0.01$, $t=-3.065$, $p=.00446^{**}$) et /o/ ($\beta=-0.04$, $SE=0.01$, $t=-1.71$, $p=.09^{ns}$) ne se distinguent pas du point de vue de leur dispersion et se positionnent à un niveau intermédiaire entre la voyelle /u/ maximalement dispersée et les voyelles /e,x/ de dispersion minimale. La voyelle /o/ n'est pas moins dispersée que la voyelle /a/, du fait de sa forte variabilité dans sa dispersion comme on peut le voir Figure 11 avec les barres d'erreur type. Cependant, la voyelle /o/ ne se distingue pas non plus des voyelles /i, ɛ, ə/ contrairement à la voyelle /a/ que l'on place de ce fait à un degré de stabilité intermédiaire entre les voyelles /i, ɛ, ə, o/ et la voyelle /u/.

	Prédictions QT	Prédictions TDF	Prédictions DAC	Résultats DFI234
Stable + - -	/i, u, a/	/i, u, o, ə, a/	/i/ /e/ /ɛ/ /a/ /u, o, ə/ /x/	/e, x/ /i, ɛ, ə, o/ /a/ /u/
	/e, ɛ, x, o, ə/	/e, ɛ, x/		

TABLE 21. Échelle de stabilité établie à partir des résultats obtenus de l'analyse des distances euclidiennes F1-F2-F3-F4, et report des prédictions initiales selon les différentes théories (QT, TDF, DAC).

Les résultats qui ressortent de cette analyse montrent peu de consistance avec les prédictions établies à partir des différentes théories de la parole reprise dans les colonnes grisées de la Table 21). Certaines attentes du modèle DAC se retrouvent dans nos résultats comme la forte variabilité du /u/ et un /e/ contraint articulatoirement et peu variable en sortie. Cependant, un point de divergence important concerne la voyelle /x/ présentée comme étant la moins stable par ce modèle alors que nos analyses montrent qu'elle fait partie des voyelles les plus stables sur l'aspect de dispersion multidimensionnelle. Ceci

étant, il se peut que les prédictions des théories se retrouvent sur certaines dimensions. Par exemple, la théorie quantique se base principalement sur les effets de variations de lieu d'articulation principalement corrélé au F2. L'analyse des dispersions par formant sera donc l'occasion de tester plus finement ces prédictions.

Dispersion par formant : DF1, DF2, DF3, DF4

Les analyses des distances des exemplaires des 18.5k voyelles au centre de leur catégorie par formant (DF1, DF2, DF3 et DF4), ont toutes révélé un effet du type de voyelle : DF1 ($\chi^2(7)=1914.1$, $p < .0001^{***}$), DF2 ($\chi^2(7)=1474.9$, $p < .0001^{***}$), DF3 ($\chi^2(7)=581.6$, $p < .0001^{***}$) et DF4 ($\chi^2(7)=377.36$, $p < .0001^{***}$) ; indiquant que certaines voyelles sont plus dispersées que d'autres quel que soit le formant étudié. Les analyses des contrastes a posteriori montrent que le degré de dispersion des différentes voyelles est fonction du formant étudié comme on pourra le voir Figure 12, plus bas.

	Prédictions			Résultats				
	QT	TDF	DAC	DF1234	DF1	DF2	DF3	DF4
- Stable +	/i, u, a/	/i, u, o, ə, a/	/i/ /e/ /ɛ/ /a/ /u, o, ə/ /x/	/e, x/ /i, ɛ, ə, o/ /a/ /u/	/e, o/ /i, x/ /ɛ, ə/ /a/ /u/	/e/ /i, ɛ/ /x, a/ /ə/ /o/ /u/	/e,ɛ,a,x/ /ə/ /i/ /u/ /o/	/ə, x/ /e/ /o/ /a, i, ɛ/ /u/

TABLE 22. Échelle de stabilité établie à partir des résultats obtenus de l'analyse de la dispersion des voyelles par formant F1, F2, F3 et F4, et report des prédictions initiales selon les différentes théories (QT, TDF, DAC) et des résultats de l'analyse multidimensionnelle (dispersion sur F1,2,3,4).

Ces résultats des contrastes sont résumés Table 22 à partir d'une échelle de stabilité (du plus au moins stable en partant du haut). Pour rappel, les prédictions des différentes théories sont rapportées dans les colonnes grisées. Pour plus de détails sur ces analyses se

référer à l'Annexe 3 page 212. Il est intéressant de remarquer que certaines régularités dans la dispersion des voyelles apparaissent à travers les formants. En effet, on peut voir que la voyelle /e/, la voyelle la plus stable globalement (mesure multidimensionnelle de l'analyse précédente) est toujours en haut de l'échelle de stabilité quelle que soit la dimension étudiée (2^{ème} sur F4). Sa faible dispersion globale n'est donc pas attribuable à une dimension acoustique en particulier. Cette voyelle est peu variable sur l'ensemble des dimensions étudiées. Le même constat peut être fait pour la voyelle la plus dispersée /u/. La forte dispersion de cette voyelle n'est pas due à certaines dimensions acoustiques particulières mais à l'ensemble des quatre dimensions acoustiques analysées. Aussi la forte stabilité de la voyelle /x/ se retrouve principalement sur les dimensions F3 et F4. Elle sera moins stable (3^{ème} position avec /a/) sur la dimension F2.

Nous noterons aussi que les voyelles postérieures /u, o, ɔ/ présentent toutes les trois une dispersion bien plus forte que les autres voyelles sur la dimension F2 (voir Figure 12b), dimension d'intérêt pour ces voyelles focales caractérisées par un F1 et un F2 maximale ment proches.

La voyelle /a/ souvent décrite comme étant une voyelle très variable et dont la dispersion globale (comprendre incluant toutes les dimensions acoustiques) est considérable dans notre corpus, est surtout dispersée sur les dimensions F1 et F4 mais pour cette dernière dimension, elle reste équivalente à celle de la plupart des voyelles. Aussi, avec la voyelle /u/, très variable sur toutes les dimensions, ce sont les voyelles les plus ouvertes /a, ε, ɔ/ qui sont les plus variables sur la dimension F1. Ce résultat rappelle les observations effectuées sur la métrique de DÉPLACEMENTS formantiques (2.3.1.1) montrant des déplacements formantiques, dus au contexte consonantique, plus importants pour ces voyelles et la voyelle /i/. Cette dernière malgré des déplacements assez importants sur F1 montre une dispersion assez faible sur ce même formant. D'ailleurs, on peut voir que cette voyelle est plus dispersée sur les dimensions F3 et F4 (voir Figure 12c et d vs. Figure 12a et b).

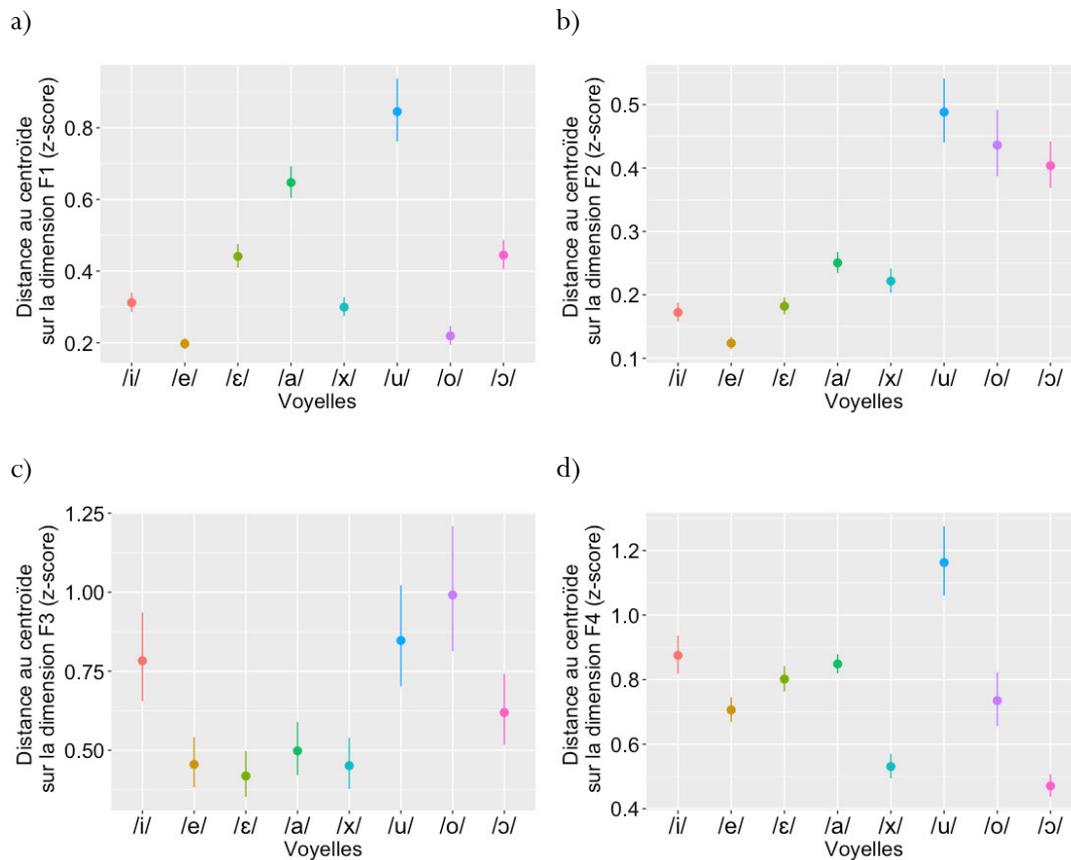


FIGURE 12. Distances aux centroides des voyelles par formant a) F1, b) F2, c) F3, d) F4.

L'ensemble de nos analyses sur la dispersion des voyelles (dispersion multidimensionnelle et par dimension acoustique) ne valide aucune des différentes prédictions établies par la théorie quantique, les modèles DAC et TDF quant à la dispersion des exemplaires des voyelles à une valeur de référence dans notre corpus de parole spontanée. Selon la théorie quantique, les voyelles /i, a, u/ devaient être plus stables que les autres sur la dimension F2 contrairement à ce que l'on pouvait s'attendre du fait qu'elles soient décrites comme moins sensibles aux variations de lieu d'articulation. Cependant, seule la voyelle /i/ présente moins de dispersion sur F2 comparé aux autres dimensions acoustiques F1, F3 et F4. Les résultats obtenus sur cette même dimension F2, se rapprochent dans les grandes lignes des prédictions faites par le modèle DAC : les voyelles antérieures présentant un DAC plus élevées sont plus stables (i.e. moins dispersées) alors que les voyelles postérieures peu contraintes articulatoirement au niveau du corps de la langue sont les plus variables. Par contre, la voyelle centrale /x/ (rassemblant les voyelles /ø, œ, ə/) ne se

retrouve pas tout en bas de l'échelle de stabilité comme le prédisait le modèle DAC. Pour terminer, on notera que les voyelles focales présentent une forte dispersion sur leurs formants focaux. La voyelle /i/ est même plus variable sur ces dimensions pertinentes F3 et F4 que sur les autres. La question que l'on peut se poser, et que nous n'avons pas examinée ici, est de savoir si l'écart entre les formants focaux de ces voyelles reste, lui, stable. Ainsi, on pourrait dire que malgré la forte dispersion des voyelles focales, leurs principales caractéristiques acoustiques restent stables (voir la discussion page 144 à ce sujet).

ii. *Des voyelles moins distinctives que d'autres (indépendamment du contexte consonantique) ?*

Dans cette section, nous présentons les résultats de l'analyse linéaire discriminante à partir desquels nous questionnerons le deuxième aspect de variabilité des voyelles : la distinctivité acoustique. Pour rappel, nous ne regardons pas encore l'influence du contexte consonantique sur cette distinctivité (ce qui sera fait en section 2.3.2.2). L'analyse précédente nous a permis de constater que certaines voyelles étaient moins dispersées que d'autres, parfois sur des dimensions acoustiques particulières. Nous allons maintenant examiner si certaines voyelles sont plus distinctives que d'autres, c'est-à-dire si elles sont mieux reconnues que d'autres à partir de leurs caractéristiques acoustiques ou si elles se confondent avec d'autres catégories vocaliques. Nous rappelons que le modèle a été construit à partir d'un échantillon d'entraînement de 64000 voyelles issues d'un autre corpus, comme décrit dans la section 2.2.2.3, avant d'être testé sur notre échantillon de parole conversationnelle de 18 568 voyelles. La partie suivante ne concerne que les résultats de la classification des voyelles de la parole conversationnelle (tous contextes consonantiques confondus). Pour toute information relative aux caractéristiques de notre modèle, consulter l'Annexe 3 page 212.

La classification acoustique automatique présente un taux de classification correcte de 61%. Ce résultat est tout à fait satisfaisant puisqu'il est bien supérieur à des résultats observés dans la littérature (Huet et Harmegnies, 2000) en parole spontanée également (52.6%) ; et que dans notre étude le seuil du hasard était fixé à 12.5% (100/8 catégories vocaliques possibles). D'ailleurs comme nous pouvons le voir Table 23 (ligne *sensitivité*), toutes nos

voyelles obtiennent des taux de classification correcte bien supérieurs à 12.5%. Dans ce tableau sont présentés les scores de *sensitivité* et de *précision* par voyelle de la classification automatique. Ces paramètres sont fondamentaux pour analyser les résultats d'une classification. La *sensitivité* nous informe de la capacité du modèle à prédire un événement, et correspond au taux de classification correcte : sur X événements Y, combien le modèle prédit-il de Y ? La *précision*, elle, nous informe de la capacité du modèle à bien isoler une catégorie : quand le modèle prédit que des événements sont des Y, combien sont réellement des Y ? La classification étant effectuée à partir des principales caractéristiques acoustiques des voyelles (les 4 formants F1, F2, F3, F4), l'analyse de ces scores nous permet de questionner la distinctivité acoustique de nos voyelles.

	/i/	/e/	/ɛ/	/a/	/ɔ/	/u/	/o/	/ɔ/
Sensitivité	85 %	73 %	46 %	62 %	78 %	41 %	31 %	42 %
Précision	73 %	56 %	49 %	90 %	42 %	48 %	29 %	47 %

TABLE 23. Scores de sensibilité et de précision par catégorie de voyelles, issus de l'analyse linéaire discriminante.

En regardant le tableau ci-dessus, on remarque qu'il existe une forte disparité entre les voyelles vis-à-vis des scores de classification. Cela semble indiquer que, oui, certaines voyelles sont plus distinctives acoustiquement que d'autres. C'est le cas principalement de la voyelle /i/ qui a été bien identifiée comme un /i/ par le modèle dans 85% des cas. Aussi, le modèle a rarement pris une autre voyelle pour un /i/ (précision = 73%), traduisant la particularité acoustique de cette voyelle peu confondue avec les 8 autres catégories vocaliques. Nous noterons que cette voyelle n'était pas pourtant des moins dispersées, notamment sur des dimensions pertinentes pour son identification, à savoir F3 et F4.

Regardons maintenant la voyelle /e/, la moins variable au vu de nos précédentes observations. Celle-ci se place au troisième rang pour la sensibilité, avec 73% (soit 12% de moins que /i/ et 5% de moins que /ɔ/), comme pour la précision (56 % contre 90% pour /a/ et 73% pour /i/). L'observation de la Figure 13, illustrant les confusions entre

voyelles, nous permet de voir que le modèle a souvent pris le /e/ pour un /i/ (16% des erreurs de classification concernant la voyelle /e/) et qu'il a souvent classé des /ε/ dans la catégorie /e/ (35% des erreurs de /ε/) d'où le plus faible score de précision par rapport au score de sensibilité.

La voyelle /ε/ est la voyelle antérieure qui a obtenu les moins bons scores de classification, avec tout de même un meilleur taux de précision de (49 %) que de sensibilité de (46 %). Comme on vient de le voir, cette voyelle a souvent été confondue avec un /e/ par notre modèle, les confusions avec la voyelle /x/ sont moins nombreuses mais restent non négligeables (10%). En revanche, on aurait pu s'attendre à plus de confusions avec la voyelle /a/ mais celles-ci ne représentent que 7% des erreurs (le contraire est plus fréquent /a/ > /ε/ = 19%).

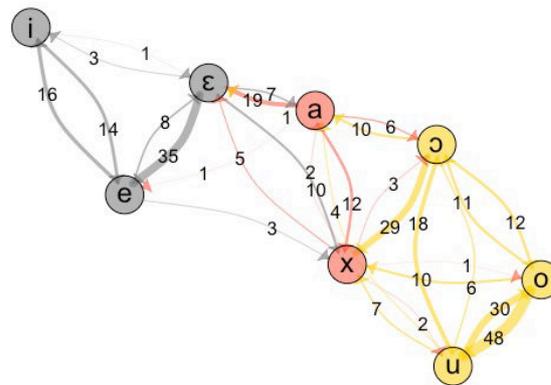


FIGURE 13. Confusions entre voyelles issues de la classification. Les chiffres rapportés sont en pourcentages et indiquent que le type de confusion représente x% des erreurs pour la voyelle en question. Les voyelles sont regroupées en trois catégories selon un code couleur (jaune pour les postérieures, rouge pour les centrales et gris pour les antérieures).

La voyelle /a/, souvent décrite comme très variable dans la littérature (Vaissière, 1985) et montrant une dispersion assez forte dans notre corpus, surtout sur F1, n'a pas été pour autant mal classifiée par le modèle. Elle l'a été correctement dans 62% des cas, ce qui en

fait la 4^{ème} voyelle la mieux classifiée par le modèle sur les 8 voyelles testées. Les confusions du /a/ avec d'autres voyelles comme les voyelles acoustiquement voisines /x, ε, ɔ/ ne sont pas pour autant négligeables. En revanche, très peu d'exemplaires d'autres voyelles ont été pris pour des /a/ par le modèle (P = 90%). Ce résultat est intéressant car il montre que le modèle a su créer une bonne catégorie de /a/ avec les propriétés acoustiques à disposition (très peu de voyelles autres que /a/ ont été prises pour des /a/). Autrement dit, même si le /a/ est très variable, c'est la seule voyelle caractérisée par un F1 maximale haut, dû à un abaissement maximal de la mandibule, ce qui en fait une catégorie bien définie et distincte des autres.

La voyelle /x/ est intéressante car elle a obtenu à la fois un très bon score de sensibilité (78 % deuxième meilleur taux) et un très mauvais score de précision (42 % deuxième plus mauvais taux). Ceci peut s'expliquer de par sa position centrale dans l'espace acoustique et de la faible dispersion dont cette voyelle fait preuve dans notre corpus (voir les précédentes analyses 2.3.1). Comme c'est la seule voyelle au centre de l'espace vocalique et qu'elle est peu dispersée, le modèle peut l'identifier facilement. En revanche, on sait qu'une forme de réduction des voyelles est la centralisation des cibles acoustiques (F1 et F2 respectivement attirés vers le centre de l'espace vocalique autour de 500 et 1500 Hz). Les voyelles tendent à se centraliser et donc à occuper l'espace vocalique du phonème /x/. La coarticulation aussi conduit parfois à la centralisation des cibles en postériorisant les voyelles antérieures et en antériorisant les voyelles postérieures. Ainsi le modèle a pu prendre de nombreuses voyelles réduites, c'est-à-dire des voyelles dont la cible n'a pas été atteinte, pour des /x/. D'ailleurs si on regarde la Figure 13, on se rend compte qu'à part la voyelle /i/, toutes les voyelles ont parfois été prises par erreur pour des /x/ par le modèle.

Plus important, les résultats des voyelles /u, o, ɔ/ montrent que ce sont les voyelles postérieures qui ont obtenu les plus mauvais scores de classification (S=38%, P =41% en moyenne). La plupart des confusions concernent les voyelles /u/ et /o/ entre elles alors que la voyelle /ɔ/ est très souvent confondue avec la voyelle centrale /x/ (Figure 13). Enfin, il est intéressant de noter que ce n'est pas la voyelle la plus dispersée, /u/, qui a

obtenu les plus mauvais scores de classification. La voyelle mi-fermée /o/ a plus souvent été confondue avec la voyelle /u/ (48% des erreurs) que l'inverse (30%).

Enfin, si on regroupe les voyelles par lieu d'articulation on observe que les voyelles antérieures sont globalement mieux identifiées (92 %) que les autres (voyelles centrales et voyelles postérieures). Une fois regroupées, les voyelles postérieures (76 %) atteignent un taux de classification tout à fait convenable et se retrouvent au même niveau que les voyelles centrales (75 %). Ceci est dû au fait que les voyelles postérieures sont principalement confondues entre elles, alors que les voyelles centrales sont souvent confondues avec des voyelles antérieures ou postérieures du fait de leur position intermédiaire.

	Prédictions			Résultats	
	QT	TDF	DAC	DISPERSION	DISTINCTIVITÉ
- Stable +	/i, u, a/	/i, u, o, ə, a/	/i/	/e, x/	/i/ (0.79)
			/e/		/a/ (0.73)
			/ɛ/	/i, ɛ, ə, o/	/e/ (0.63)
			/a/		/x/ (0.54)
			/u, o, ə/	/a/	/ɛ/ (0.47)
			/x/		/ə/ (0.45)
	/e, ɛ, x, o, ə/	/e, ɛ, x/	/x/	/u/	/u/ (0.44)
					/o/ (0.30)

TABLE 24. Échelle de stabilité des voyelles selon leur distinctivité dérivée du score *F1* (mêlant sensibilité et précision du modèle) de la classification et report des prédictions initiales selon les différentes théories (QT, TDF, DAC) et des résultats de l'analyse multidimensionnelle (dispersion sur F1,2,3,4).

Dans la Table 24, nous présentons les résultats issus de la classification sous forme d'échelle de stabilité, calculée à l'aide d'un score composite, appelé *score F1*, correspondant à une moyenne harmonique des scores de sensibilité et de précision. La distinctivité maximale est de 1 et la pire de 0. Les prédictions faites à partir des différentes théories, sont

rappelées dans les colonnes gris foncé, en gris clair un rappel des résultats de la métrique de dispersion.

Contrairement aux résultats sur la dispersion acoustique des voyelles, ceux mesurant la distinctivité des voyelles, placent la voyelle /i/ en haut de l'échelle de stabilité répondant aux prédictions des trois théories. La voyelle /a/ apparaît aussi comme très distinctive ce qui est en accord avec les théories QT et TDF mais pas avec le modèle DAC. En revanche, en ce qui concerne la voyelle /u/ ce sont les prédictions du DAC qui sont en accord avec nos résultats s'opposant aux prédictions de QT et TDF (aussi en désaccord pour /o, ə/). Les observations issues de l'analyse de nos deux aspects de la variabilité considérés ne vont pas dans le même sens indiquant que la distinctivité des voyelles ne dépend pas forcément de la dispersion dans leur réalisation acoustique mais plutôt du nombre de compétiteurs acoustiques existants.

2.3.2.2. Analyse des effets contextuels sur la variabilité des voyelles

Le profil de variabilité des voyelles de notre corpus désormais identifié, nous pouvons revenir à la question centrale de notre étude : la coarticulation influence-t-elle la variabilité des voyelles ? On cherchera particulièrement à savoir si un contexte amène plus de variabilité dans les réalisations acoustiques de nos voyelles ou si cela dépend du conflit existant entre C et V.

i. Une dispersion des voyelles dépendante du contexte consonantique ?

Dans cette section, seront présentés les résultats de l'analyse de la dispersion des voyelles en fonction du contexte consonantique. Seule la mesure multidimensionnelle (F1-F2-F3-F4) sera considérée puisqu'il nous importe de savoir si la coarticulation crée de la variabilité dans l'acoustique des voyelles en général.

Afin de tester si la dispersion des voyelles dépend du contexte consonantique dans lequel elles se trouvent, nous avons rajouté à notre modèle initial présenté dans l'étude préliminaire (page 76) le facteur CONTEXTE CONSONANTIQUE en interaction avec le facteur VOYELLE. Le facteur CONTEXTE a également été ajouté à la pente aléatoire pour le facteur LOCUTEUR mais nous n'avons pas pu inclure l'interaction entre VOYELLE et CONTEXTE à cette pente pour des problèmes de convergence.

Ce nouveau modèle ($R^2_m = 0.08$, $R^2_c = 0.21$) indique un effet de la VOYELLE ($\chi^2(7) = 146.39$, $p < .0001$ ***) et du CONTEXTE ($\chi^2(2) = 15.477$, $p < .0001$ ***) sur les valeurs distances euclidiennes F1-F2-F3-F4. Ceci indique que la dispersion des voyelles est fonction du type de voyelle : certaines voyelles sont plus dispersées que d'autres, mais cela dépend aussi du contexte dans lequel elles se trouvent. L'analyse des contrastes a posteriori, indique que les voyelles sont moins dispersées en contexte alvéolaire qu'en contexte uvulaire. Le contexte vélaire ne se détache pas de ces deux derniers comme nous pouvons le constater Table 25.

contraste	β	SE	df	z.ratio	valeur de p
ALV-UV	-0.06	0.01	Inf	-4.634	<.0001
ALV-VEL	-0.03	0.01	Inf	-2.336	0.0510
UV-VEL	0.03	0.01	Inf	1.922	0.1324

TABLE 25. Résultats de l'analyse des contrastes du facteur CONTEXTE.

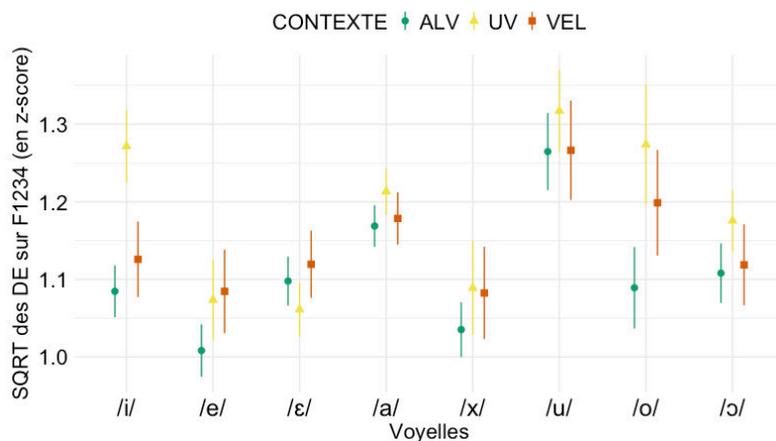


FIGURE 14. Racine carrée des distances euclidiennes F1-F2-F3-F4 des voyelles en fonction du contexte consonantique (alv, vel, uv).

Enfin, une interaction entre nos deux facteurs ressort de notre analyse ($\chi^2(14) = 107.62$, $p < .0001$ ***). Les effets du contexte consonantique sur la dispersion des voyelles dépendent de la voyelle étudiée. Par exemple, comme on peut le voir Figure 14, la dispersion des

voyelles /u/ et /x/ n'est pas fonction du contexte consonantique. Il est intéressant de noter que ce sont des voyelles qui ont des patterns de dispersion opposés : la voyelle /u/ est la voyelle la plus dispersée alors que la voyelle /x/ est une des voyelles les plus stables dans ses réalisations acoustiques.

voyelles	contraste	β	SE	z.ratio	valeur de p
/i/	ALV-UV	-0.19	0.022	-8.40	<.0001
	ALV-VEL	-0.04	0.023	-1.80	.169
	UV-VEL	0.15	0.028	5.159	<.0001
/e/	ALV-UV	-0.07	0.03	-2.61	.0245
	ALV-VEL	-0.08	0.03	-3.03	.007
	UV-VEL	-0.01	0.03	-0.34	.94
/ɛ/	ALV-UV	0.04	0.02	2.11	.09
	ALV-VEL	-0.02	0.02	-1.06	.54
	UV-VEL	-0.058	0.02	-2.53	.031
/a/	ALV-UV	-0.05	0.01	-2.96	.009
	ALV-VEL	-0.01	0.02	-0.624	.81
	UV-VEL	0.04	0.02	1.90	.14
/x/	ALV-UV	-0.053	0.03	-1.71	0.20
	ALV-VEL	-0.047	0.03	-1.635	.23
	UV-VEL	0.01	0.04	0.16	.1
/u/	ALV-UV	-0.05	0.03	-1.94	0.13
	ALV-VEL	-0.002	0.03	-0.05	.1
	UV-VEL	0.05	0.03	0.15	.3
/o/	ALV-UV	-0.19	0.04	-4.91	<.0001
	ALV-VEL	-0.11	0.03	-3.497	0.0014
	UV-VEL	0.08	0.04	1.74	0.19
/ɔ/	ALV-UV	-0.07	0.03	-3.181	0.0042
	ALV-VEL	-0.01	0.028	-0.389	.92
	UV-VEL	0.06	0.03	2.039	.1

TABLE 26. Résultats de l'analyse des contrastes du facteur CONTEXTE par voyelle.

Pour les autres voyelles, la dispersion dépend du contexte consonantique mais elle est toujours minimale en contexte alvéolaire comme on peut le voir Table 26. Il est intéressant de noter que les effets contextuels sur la dispersion des voyelles sont plus importants pour certaines voyelles comme l'indiquent les valeurs z.ratio. C'est le cas notamment de la voyelle /i/ dont la dispersion est faible en contexte alvéolaire et vélaire mais très forte en contexte uvulaire. En effet, dans ce contexte, la voyelle /i/ habituellement décrite comme peu variable atteint une dispersion équivalente à celle observée pour la voyelle /u/. Le contexte uvulaire semble donc affecter fortement les réalisations acoustiques de cette voyelle. Ceci est cohérent avec nos résultats portant sur la direction des effets contextuels sur les voyelles, section 2.3.1, indiquant des déplacements des formants F1, F2, F3 considérables en contexte uvulaire. De plus la voyelle /o/ montre une très forte dispersion en contextes uvulaire et vélaire mais aussi une forte variabilité dans sa dispersion.

ii. *Une distinctivité des voyelles dépendante du contexte consonantique ?*

Dans cette section, nous analyserons les résultats de la classification automatique en fonction du contexte consonantique dans lequel se trouvaient les voyelles. Nous examinerons les types de confusions effectuées par le modèle selon que les voyelles étaient dans un contexte ou un autre en traitant uniquement des résultats les plus importants.

Les matrices de confusions présentées Table 27 montrent les résultats de la classification automatique de nos 18.5k voyelles par contexte consonantique : le tableau du haut expose les résultats de la confusion pour le contexte alvéolaire, celui du milieu ceux observés en contexte uvulaire et enfin celui du bas rapporte les résultats pour le contexte vélaire. Il est intéressant de remarquer que la voyelle /i/ est presque toujours bien identifiée lorsqu'elle se trouve en contextes alvéolaire (90%, soit +5% par rapport aux résultats tous contextes confondus - TCC) ou vélaire (88%, soit +3% / score TCC) alors que son taux de classification correcte s'abaisse à 51% en contexte uvulaire. Dans ce dernier, elle est à la troisième place des voyelles les moins bien identifiées (-34% / score TCC). Ceci est dû au fait que dans ce contexte, 45% des exemplaires de la voyelle /i/ aient été classifiés dans la catégorie /e/ par le modèle, du fait de l'élévation du premier formant du /i/ par le /ʋ/. Un effet semblable est observé pour la voyelle /e/ dont le score de classification est très

bon en contextes alvéolaire et vélaire (74 et 70% respectivement) mais très mauvais (56%, -27% / TCC) en contexte uvulaire du fait du grand nombre de confusions pour la voyelle voisine mi-ouverte /ɛ/ dans ce contexte ouvrant.

Un autre résultat important, et peut-être moins attendu, concerne la voyelle /u/ dont le taux de classification correcte atteint les 71% en contexte alvéolaire (+30% / TTC) et 64% en contexte vélaire (+23% / TCC) (contre 20% en contexte uvulaire). Ceci s'explique par le fait que la plupart des erreurs de classification de la voyelle /u/ concernent des confusions avec la voyelle mi-ouverte /o/. Or c'est en contexte uvulaire que la voyelle /u/ est plus susceptible d'être confondue avec un /o/ à cause de l'effet élévateur de F1 du /ʁ/. Ce résultat est en accord avec le modèle DAC qui prédit de la variabilité au niveau de la hauteur des voyelles postérieures en contexte uvulaire (Recasens, 2014). En revanche, c'est l'inverse qui se produit pour la voyelle /o/ qui atteint d'excellents scores de classification correcte en contexte uvulaire 84% (soit + 52% / TCC) pour les mêmes raisons que celles évoquées pour /u/. Malgré les effets ouvrant de la consonne uvulaire, très peu de /o/ ont été pris pour un /ɔ/ dans ce contexte uvulaire (3% seulement). Sans surprise les confusions de /ɔ/ pour /x/ sont observées en contexte alvéolaire principalement (45%) dues à l'élévation de F2 par ce type de consonnes. Plus surprenant, ce même type de confusion est très fréquent lorsque la voyelle se trouve en contexte vélaire (38% des erreurs) alors que le locus des vélaires se situe au niveau de F2 pour les voyelles postérieures (Vaissière, 2006). On notera d'ailleurs que quel que soit le type de confusion, les contextes vélaire et alvéolaire présentent les mêmes patterns.

Enfin, le score de classification de la voyelle /x/ est nettement diminué en contexte uvulaire (49% soit -29% / TCC). Ceci est dû au fait que cette voyelle centrale d'une manière générale peu dispersée, voir 2.3.1, est souvent confondue avec la voyelle /a/ et la voyelle /ɔ/ dans ce contexte qui élève F1 et abaisse F2. En revanche, les scores de classifications de la voyelle /a/ (89%) et de la voyelle /ɛ/ (61%) sont nettement améliorés en contexte uvulaire (+18 et +15% / TCC respectivement) du fait que les principales confusions de voyelles concernent la catégorie voisine plus fermée : /ɛ/ > /e/ et /a/

>/ε/comme nous l'avons vu dans l'analyse tous contextes confondus. Or en contexte uvulaire on aura moins de confusions de ce type du fait que le /ʁ/ élève le F1 des voyelles.

RÉSULTATS DE LA CLASSIFICATION EN CONTEXTE ALV (N=12958)								
	/i/	/e/	/ɛ/	/a/	/ɔ/	/u/	/o/	/ɔ̃/
/i/	90	17	3	0	0	0	0	0
/e/	9	74	41	1	9	0	0	0
/ɛ/	1	7	41	24	5	0	0	1
/a/	0	0	3	57	1	4	0	11
/ɔ/	0	3	11	14	80	15	20	45
/u/	0	0	0	0	1	71	38	9
/o/	0	0	0	0	0	5	23	4
/ɔ̃/	0	0	0	5	1	5	19	29

RÉSULTATS DE LA CLASSIFICATION EN CONTEXTE UV (N=3377)								
	/i/	/e/	/ɛ/	/a/	/ɔ/	/u/	/o/	/ɔ̃/
/i/	51	2	1	0	0	0	0	0
/e/	45	56	5	0	0	0	0	0
/ɛ/	4	38	67	1	7	0	0	0
/a/	0	0	19	89	13	4	0	5
/ɔ/	0	4	8	1	49	3	0	1
/u/	0	0	0	0	2	20	13	0
/o/	0	0	0	0	6	49	84	32
/ɔ̃/	0	0	0	10	22	24	3	61

RÉSULTATS DE LA CLASSIFICATION EN CONTEXTE VEL (N=2222)

	/i/	/e/	/ɛ/	/a/	/ɤ/	/u/	/o/	/ɔ/
/i/	88	24	4	0	0	0	0	0
/e/	12	70	61	1	6	0	0	0
/ɛ/	0	4	30	16	4	0	0	1
/a/	0	0	1	59	0	4	0	13
/ɤ/	0	3	3	15	78	6	2	38
/u/	0	0	0	0	7	64	85	7
/o/	0	0	0	0	1	6	10	3
/ɔ/	0	0	0	9	4	21	3	39

TABLE 27. Matrices de confusions (%) de notre classification automatique par contexte consonantique (ALV en haut, UV au milieu et VEL en bas).

En résumé, les voyelles sont globalement moins dispersées lorsqu'elles se trouvent en contexte alvéolaire. Concernant la distinctivité acoustique des voyelles, il semblerait que les plus fermées /i,e,u/ soient plus distinctives en contextes alvéolaire et vélaire alors que les voyelles plus ouvertes /ɛ, a, ɔ/ le sont en contexte uvulaire. La voyelle /o/ diffère des autres voyelles fermées car elle est plus distinctive en contexte uvulaire. Ceci est dû à sa proximité acoustique avec le /u/ notamment sur la dimension F1. La voyelle /o/ se trouve plus distincte de cette dernière voyelle en contexte uvulaire où son F1 est élevé et son F2 abaissé. Enfin la voyelle centrale/ɤ/ est moins distinctive en contexte uvulaire qu'en contextes alvéolaire et vélaire. Ces résultats montrent que la distinctivité acoustique dépend de la compatibilité de C et V. Les voyelles sont plus distinctives, si elles apparaissent dans un contexte renforçant leurs propriétés inhérentes. Ainsi, on aura plus d'erreurs de classification si les consonnes déplacent une catégorie vocalique vers la zone acoustique d'une voyelle voisine.

ÉTUDE DES INFLUENCES PROSODIQUES SUR LA COARTICULATION C-à-V

*Le poète choisit, élit dans la masse du monde
ce qu'il lui faut préserver, à quoi s'accorde son chant.
Et le rythme est force rituelle, levier de conscience.
Il achemine vers ces puissances la richesse (la rigueur) prosodique,
garante du choix, gardienne des conquêtes la connaissance du monde
dans son épaisseur et son erre, l'envers éclairant de l'Histoire.*
Édouard Glissant
L'intention poétique

Dans ce chapitre, nous étudions comment la coarticulation C-à-V est modulée par la position prosodique (PP) à partir de données extraites de deux grands corpus de parole du français. Dans certaines positions prosodiques les segments subissent des variations dans leurs propriétés acoustiques ou articulatoires. C'est le cas des segments : 1) sous prééminences prosodiques (syllabes : sous focus, accent de phrase, ou accent tonique) ou 2) proches des limites de domaines/constituants prosodiques (position initiale et finale de constituants). Dans cette thèse, nous nous concentrerons sur le second type de position prosodique (marqueur de frontière), et notamment sur les phénomènes de variations segmentales se produisant en position initiale de domaine/constituant prosodique, connus sous le nom de « renforcement prosodique initial » ou « renforcement initial ». Les principaux travaux existant sur la coarticulation sous l'effet du renforcement initial, portent sur la coarticulation de segments transfrontaliers. L'originalité de ce travail est de traiter des segments post-frontaliers où les segments coarticulés se trouvent du même côté de la frontière. Ainsi, on dissocie effet du renforcement initial, de l'effet de frontière. Pour cela nous avons comparé la coarticulation de séquences CV et VC issues de deux positions prosodiques différentes : la position initiale de groupe intonatif (IPi) et la position de syllabe à l'intérieur d'un mot (Wm). Nos résultats montrent une réduction de la coarticulation C-à-V en position prosodique forte (IPi) pour toutes les voyelles testées /i, e, a, ɔ/ en séquences VC. Soit, pour les cas de coarticulation anticipatrice où, en condition IPi, les voyelles V se trouvent directement à la frontière du constituant prosodique. En revanche, aucune modification de la coarticulation en fonction de la position prosodique n'a été rapportée pour les cas de coarticulation persévératrice CV, où la seule voyelle testée était un /a/, lorsque c'est la consonne qui se trouve à la frontière du constituant prosodique.

3.1. QUESTIONS ET OBJECTIFS DE L'ÉTUDE

Comme on a pu le voir dans la revue de la littérature (section 1.4) les effets du renforcement prosodique initial sur la coarticulation des segments post-frontaliers (i.e. en position initiale de constituant) ne font pas consensus. Si peu de travaux existent, la majorité d'entre eux ne rapportent aucune modification du timing inter-gestuel des séquences post-frontalières quelle que soit la force de la frontière (Meynadier, 2003 ; Byrd & Choi, 2010 ; Cho et al., 2014). Cependant, quelques études indiquent une réduction du chevauchement inter-gestuel lorsque la frontière prosodique qui précède les segments est une frontière de haut rang (Byrd, 2000 ; Bombien et al., 2011 ; Cho et al., 2017). L'objectif premier de notre travail sera donc d'éclaircir ce point à la lumière de nouveaux résultats sur le français. À partir de données acoustiques de 7k /a/, nous testerons l'effet de la position prosodique sur le degré de coarticulation C-à-V, lorsque 'C' peut être une consonne ALVéolaire ou une consonne UVulaire, en comparant le degré de coarticulation de séquences en position initiale de Groupe Intonatif (IPi) avec celui de séquences en position interne de mot (Wm).

Les travaux existants ne se sont intéressés qu'aux syllabes post-frontalières de type CV. Autrement dit, lorsque c'est la consonne qui se trouve en position strictement initiale de constituant prosodique. En effet, dans cette configuration la voyelle se trouve en syllabe initiale de constituant prosodique mais elle n'occupe pas la position initiale absolue. Or, on sait que l'effet du renforcement initial peut s'étendre à toute la syllabe initiale mais il affecte plus fortement les segments directement adjacents à la frontière (Byrd & Saltzman, 2003). Ce constat nous amène à notre second objectif qui est de déterminer si l'effet de la frontière prosodique sur le degré de coarticulation C-à-V dépend du type de séquence observé : CV (coarticulation persévératrice) ou VC (coarticulation anticipatoire), impliquant une différence dans la position de la voyelle coarticulée par rapport à la frontière.

Dans notre première étude, nous avons vu que toutes les voyelles n'étaient pas affectées de la même manière par les variations contextuelles tant au niveau des déplacements

formantiques que sur leur variabilité mesurée en termes de dispersions acoustiques et de distinctivité. La question que l'on peut donc se poser est de savoir si la modulation de la coarticulation C-à-V selon la position prosodique est ou non dépendante du type de voyelle. Est-ce que la coarticulation d'une voyelle globalement plus résistante aux variations contextuelles (par exemple le /i/ sur F2), sera modulée par la position prosodique de la même manière qu'une voyelle peu résistante à la coarticulation (les voyelles postérieures sur F2 par exemple) ? Notre troisième objectif est d'identifier si les effets observés sont ou non fonction des voyelles en étendant nos analyses aux voyelles /i, e, ə/ pour lesquelles le matériel disponible était suffisant.

Les questions auxquelles nous tenterons de répondre dans cette étude sont :

- La coarticulation C-à-V est-elle réduite en position initiale d'IP ?
- La coarticulation en CV et VC est-elle modulée de façon identique par la position prosodique ?
- La modulation de la coarticulation par la position prosodique est-elle fonction des voyelles ?

3.2. MÉTHODE

3.2.1. MATÉRIEL LINGUISTIQUE ET PRÉTRAITEMENTS

Pour cette étude, nous avons regroupé les données des deux corpus étudiés dans cette thèse : ESTER (Gravier et al., 2006) et NCCFr (Torreira et al., 2010) présentés section 1.6.2. En effet, nos premières observations ont montré des résultats semblables concernant l'effet de la position prosodique sur le degré de coarticulation C-à-V. Nous avons donc choisi de privilégier le nombre d'occurrences en unifiant nos deux bases de données.

La même procédure d'extraction des formants des voyelles que celle qui est décrite en 2.2.1.1 pour le corpus NCCFr a été appliquée au corpus ESTER sur un alignement automatique forcé (Galliano et al., 2005). Pour toute information relative à cette procédure d'extraction et le rejet de valeurs aberrantes, se référer à la section précédemment citée page 57. Ainsi nous disposons d'une base de 350 851 voyelles provenant du corpus NCCFr et une autre de 445 542 voyelles issues du corpus ESTER.

3.2.1.1. Les positions prosodiques testées

Afin de tester l'effet de la position prosodique sur le degré de coarticulation de C-à-V, les séquences CV et VC, ont été observées à la fois : 1) position initiale de Groupe Intonatif que l'on appellera « position IPi » en référence au terme anglais *Intonational Phrase* et 2) en position de syllabe interne de mot ou « Wm » pour *Word medial*.

Le Groupe Intonatif (IPi) est le plus grand constituant prosodique. Il est accepté dans la plupart des travaux portant sur la prosodie du français. Il se caractérise par la présence d'un ton de frontière à son bord droit, un allongement final important et il est souvent suivi d'une pause (Delais-Roussarie et al., 2015).

À la position initiale de Groupe Intonatif, nous comparerons la position de syllabe interne de mot (Wm). La syllabe est le constituant prosodique de rang hiérarchique le plus bas. Nous avons choisi d'analyser les syllabes internes de mots se trouvant à l'intérieur d'un Groupe Intonatif afin d'avoir deux positions prosodiques extrêmes, se distinguant très

bien, pour tester facilement les effets du Renforcement Initial sur les influences contextuelles.

3.2.1.2. Critères de définition du phrasé prosodique

Afin d'étudier les constituants prosodiques, il faut au préalable définir le phrasé de la phrase produite. Selon les méthodes de phonétique expérimentale et de phonologie de laboratoire, l'obtention d'un phrasé particulier reflétant la structure prosodique voulue, passe par la conception d'un corpus très contrôlé en amont. Ainsi, les expérimentateurs créent des phrases cadres en manipulant leur structure morphosyntaxique et en jouant sur la ponctuation de celles-ci, tout en prenant soin de placer le segment test à la position prosodique souhaitée. Le phrasé produit est ensuite vérifié à partir d'un codage prosodique des productions sur la base de critères prosodiques comme la présence ou l'absence de pause, le degré d'allongement final de la voyelle précédant le segment test ou encore le type de contour intonatif. Dans ce genre d'étude le phrasé est donc élicité à partir de critères syntaxiques et vérifié sur des critères prosodiques (intonatifs et temporels). Il est facile de s'imaginer qu'une telle méthode, aussi répandue et robuste soit-elle, est difficilement applicable à de grands corpus. En effet, ces derniers n'ont pas été conçus pour tester l'influence de la structure prosodique sur les variations segmentales. Ceci renvoie aux notions de test des facteurs a priori vs. a posteriori abordées en introduction (je vous renvoie à la section 1.6 pour plus d'informations). De plus l'évaluation des contours intonatifs sur un si grand nombre de données serait très fastidieuse.

Le recours à des méthodes facilement automatisables, fondamental pour l'étude sur grands corpus, implique d'aborder la structure prosodique différemment. Ainsi, dans notre étude, la position initiale de Groupe Intonatif est déterminée à partir de la détection automatique des pauses dans le signal sur la base des méthodes utilisées dans de précédents travaux (Gendrot & Gerdes, 2011 ; Gendrot et al., 2011). En effet, si un groupe intonatif n'est pas forcément délimité par une pause, la présence d'une pause indique toujours une frontière de Groupe Intonatif. Ce critère est donc quelque peu restrictif, puisqu'il exclut tous les

Groupes Intonatifs non marqués par une pause, mais c'est aussi ce qui en fait un critère fiable, et qui plus est, très facilement automatisable.

La Figure 15 illustre les caractéristiques particulières à nos deux positions prosodiques.

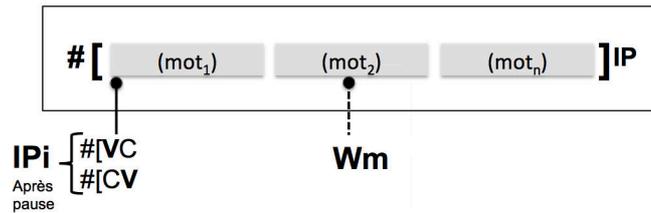


FIGURE 15. Caractéristiques de nos deux positions prosodiques IPI et Wm en fonction du type de séquence (VC et CV).

La position IPI concerne les séquences strictement initiales de Groupes Intonatifs. La voyelle se trouve donc soit en position initiale absolue de Groupe Intonatif (c'est le cas de V en séquences VC), soit en syllabe initiale de Groupe Intonatif (c'est le cas de V en séquences CV). Toutes nos séquences observées en position IPI, sont précédées d'une pause marquée par le symbole #. Aucune restriction concernant la durée de cette pause n'a été déterminée mais la durée moyenne de ces pauses est de 70ms. La position Wm regroupe les séquences se trouvant en position interne de mot. Nous avons considéré la position interne au sens strict et avons exclu les séquences en syllabe initiale et finale de mots. La Figure 16, présente deux spectrogrammes illustrant ces deux positions prosodiques. La position IPI est illustrée avec le mot « attends » et la position Wm avec le mot « privatisation » où la consonne à gauche de la voyelle /a/ est une labiale.

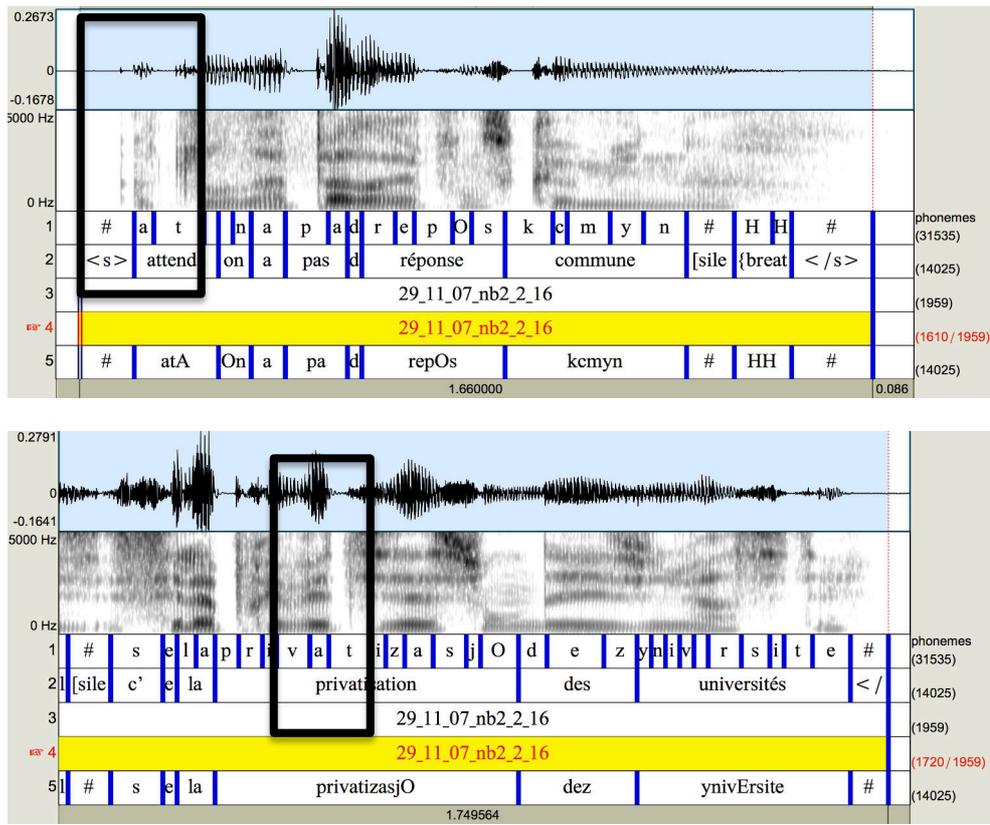


FIGURE 16. Spectrogrammes illustrant les deux positions prosodiques testées avec la voyelle /a/ en contexte alvéolaire pour les séquences VC. La position IP_i est illustrée avec le mot « attends » et la position W_m avec le mot « privatisation » où la consonne à gauche de V est une labiale.

3.2.1.3. Sélection des séquences CV_1 & V_1C

Après la sélection des voyelles répondant à nos critères prosodiques, nous avons cherché des séquences CV et VC afin d'étudier la coarticulation C-à-V, où la consonne était une consonne alvéolaire (ALV=/t,d,s,z,n,l/, ex : *dépanner* /depane/) ou une consonne uvulaire (UV=/ʁ/, ex : *appareil* /apaʁɛj/). Nous avons choisi ces consonnes car les effets coarticulatoires observés dans notre première étude (CHAPITRE 2) étaient les plus marqués et allaient dans des directions opposées (UV : élévation de F1 et abaissement de F2 vs. ALV : abaissement de F1 et élévation de F2). En position W_m , le contexte opposé (gauche en VC et droit en CV) était toujours une consonne labiale /p,b,f,v,m/. Les mots monosyllabiques étaient directement exclus. Il est important de noter que dans les

séquences CV la consonne et la voyelle faisaient toujours partie de la même syllabe alors que pour les séquences VC cela n'a pas pu être contrôlé. Pour ce dernier type de séquences la consonne et la voyelle pouvaient être tauto-syllabiques comme hétéro-syllabiques.

À partir de ces critères, et selon la quantité de matériel disponible, nous avons trouvé des séquences VC avec les voyelles /i, e, a, ɔ/ et des séquences CV avec la voyelle /a/ uniquement. Au total plus de 16.6k voyelles ont été analysées. Le résumé du matériel utilisé est disponible Table 28.

SÉQUENCE	VOYELLE	CONTEXTE	POSITION PROSODIQUE	
			IPi	Wm
CV	/a/	ALV	148	1454
		UV	56	435
VC	/a/	ALV	1638	1896
		UV	189	1197
	/i/	ALV	271	3773
		UV	44	325
	/e/	ALV	493	34
		UV	531	2014
	/ɔ/	ALV	52	141
		UV	648	1278

TABLE 28. Résumé du matériel linguistique sélectionné et distribution des 16617 voyelles /a, i, e, ɔ/ étudiées par conditions de structure syllabique (CV ou VC) et lieu d'articulation de la consonne C (ALV=alvéolaire, UV=uvulaire)

Comme on peut la voir Table 28, il existe un fort déséquilibre entre nos groupes (ALV vs. UV ; IPi vs. Wm). La position IPi et le contexte UV étant sous représentés par rapport à la position Wm et au contexte ALV. C'est d'ailleurs pour cette raison que nous n'avons pas pu tester plus de voyelles, notamment en séquences CV.

3.2.2. MESURE DE COARTICULATION

Comme pour la première étude (voir page 63), les effets contextuels ont été mesurés par les changements des deux premiers formants (F1 et F2) de la voyelle V₁ en fonction du lieu

d'articulation de la consonne adjacente C. La FIGURE 17 illustre l'effet du contexte sur les propriétés acoustiques de /a/ en séquence CV₁. Les caractéristiques articulaires et acoustiques des consonnes étudiées sont décrites 51.

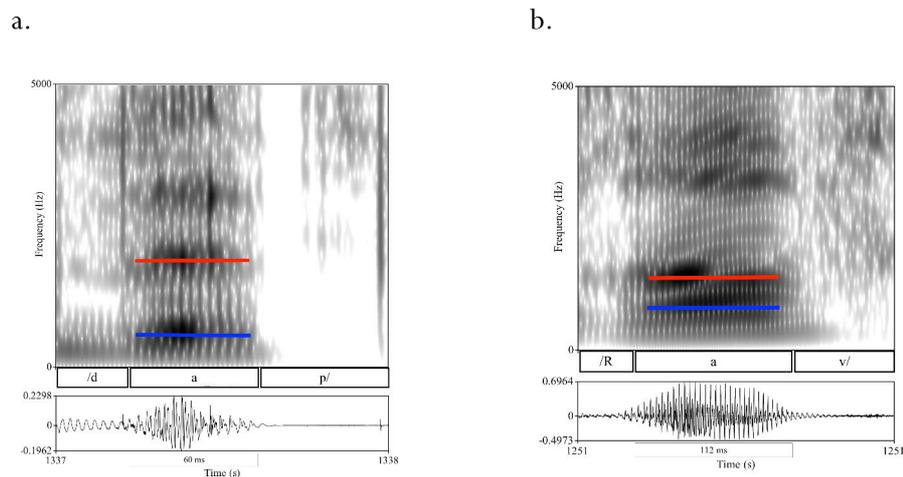


FIGURE 17. Spectrogrammes illustrant les variations formantiques de la voyelle /a/ en fonction du contexte consonantique: a) alvéolaire (suite /dap/ extraite du mot « adapter ») et b) uvulaire (suite /ʁav/ extraite du mot « grave ») en séquence CV₁. Le contexte droit étant une labiale.

Les analyses reposent sur les valeurs des deux premiers formants en Bark (Traunmüller, 1997) selon la formule présentée Équation 3 où F_i est un formant, F1 ou F2 dans notre cas, et F_{i_bark} la valeur de ce formant en Bark. Pour toute information relative aux mesures formantiques se référer à la section 2.2.1.1.

$$F_{i_bark} = 26.81 / (1 + 1960 / F_i) - 0.53$$

EQUATION 3. Formule de transformation en Bark (Traunmüller, 1997)

3.2.3. ANALYSES STATISTIQUES

Afin de tester si le degré de coarticulation C-à-V est modulé par la position prosodique, nous avons construit 10 modèles mixtes (un par voyelle, par type de séquence, CV₁ et V₁C pour la voyelle /a/, et par formant (F1 et F2)) à l'aide du logiciel R (R development Core

Team, 2008) et de la bibliothèque 'lme4' (Bates et al., 2015). Ainsi, nous avons testé les relations entre les valeurs de F1 et F2 (Bark) de /a/ et les facteurs suivants : 1) position prosodique (IPi vs. Wm) et 2) contexte (C_{ALV} vs. C_{UV}). La structure fixe contenait aussi une interaction entre les deux facteurs. Nous avons modélisé un intercept par locuteur et par mot. Afin d'éviter un taux élevé d'erreur de Type I (Cunnings, 2012) une pente aléatoire par locuteur a été incluse pour nos deux facteurs fixes. Cela correspond à la variabilité entre locuteurs de l'effet du contexte et de la position prosodique sur les formants. L'interaction n'a pas pu être intégrée à la pente aléatoire pour des problèmes de convergence. Aucune pente aléatoire n'a été intégrée au facteur aléatoire « mot » pour les mêmes raisons. Les valeurs de p ont été obtenues par approximations de type Satterthwaite à l'aide de la fonction lmerTest (Kuznetsova et al., 2013). Le seuil de référence a été fixé à $p < .05$. L'effet de chaque facteur fixe et de leur interaction a été testé par comparaison de modèles avec la fonction anova. Les valeurs de R^2 associées à chaque modèle ont été obtenues à l'aide de la fonction r.squaredGLMM intégrée dans la bibliothèque 'MuMIn'. Les analyses des contrastes a posteriori ont été effectuées avec la fonction lsmeans (bibliothèque 'emmeans', Lenth et al., 2018).

3.3. RÉSULTATS

Cette section résultat s'organise en trois parties. Dans un premier temps, nous rapporterons les résultats d'une étude préliminaire dans laquelle nous avons contrôlé la durée de la voyelle pour s'assurer que d'éventuels effets de la position prosodique sur le degré de coarticulation C-à-V ne soient pas attribuables à des variations de durée causées par les positions prosodiques testées. Dans un deuxième temps, seront présentés les résultats obtenus pour la voyelle /a/ en séquences CV et VC, étant donné que c'est la seule voyelle testée dans les deux conditions, afin de déterminer si la modulation de la coarticulation par la position prosodique est dépendante du type de coarticulation. Enfin, nous terminerons par la présentation des résultats obtenus pour les trois autres voyelles /i, e, ə/ en séquence VC qui nous permettront de déterminer si la modulation de la coarticulation par la position prosodique est fonction de la voyelle impliquée.

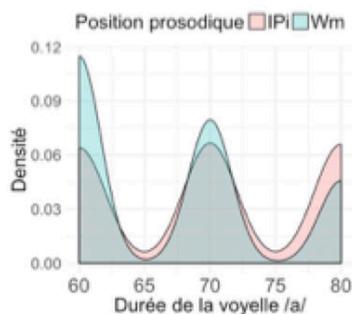
3.3.1. ÉTUDE PRÉLIMINAIRE : CONTRÔLE DE LA DURÉE

L'objectif de cette étude préliminaire est de s'assurer que la durée ne soit pas à l'origine des éventuels effets de la position prosodique sur le degré de coarticulation C-à-V. En effet, comme on a pu le voir en introduction la durée des segments explique une grande partie des phénomènes de coarticulation et une des conséquences du renforcement prosodique initial est l'allongement des segments. Même si celui-ci n'est pas systématique en français (Fougeron, 1998 ; Georgeton, 2014), il nous paraissait important de s'assurer que la modulation de la coarticulation par la position prosodique, si elle a lieu, ne soit pas que le produit de variations de durée dans les différentes positions prosodiques testées.

Deux solutions étaient envisageables : 1) intégrer la durée comme facteur fixe à nos modèles statistiques ; ou 2) contrôler la durée des segments en ne gardant que les segments d'une certaine durée. La première option a vite été écartée par manque de matériel. En effet, rajouter la durée en facteur fixe risquait d'appauvrir l'analyse en augmentant les déséquilibres (déjà importants) entre les groupes. En revanche, le contrôle de la durée s'est

révéler possible en séquence VC. Nous avons observé la distribution des /a/ de durée moyenne, de 60 à 80 ms (Gendrot & Adda-Decker, 2005 ; Audibert & al., 2015).

Comme nous pouvons le voir Figure 18 en position IPi, les exemplaires de /a/ se répartissent équitablement sur toute la tranche de durée 60-80 ms. Il y a autant de voyelles de 60 ms que voyelles de 70 ou 80 ms. En position Wm, la répartition est moins équitable, avec une surreprésentation des voyelles d'une durée de 60 ms et une sous représentation des exemplaires de 80 ms. Bien que les deux distributions se superposent et que nous n'ayons pas tous les exemplaires de /a/ de 80 ms en position IPi et ceux de 60 ms en position Wm, nous avons préféré écarter les exemplaires de 60 ms et faire une première analyse sur la tranche de durée 70-80 ms. Le but de cette analyse préliminaire est de tester l'effet de la position prosodique sur la coarticulation C-à-V sur des sous-groupes de même durée en IP et en Wm, en contexte ALV et UV.



	60ms	70ms	80ms
IPi	198	206	204
Wm	516	357	205

Figure 18. Distribution de la durée de la voyelle /a/ en séquence VC sur la tranche 60-80 ms par position prosodique (à gauche) et résumé du matériel disponible (à droite).

Pour cette analyse du /a/ en séquence V_1C avec contrôle de la durée, nous présenterons d'abord les résultats de l'analyse du premier formant (F1) puis ceux de l'analyse du second formant (F2).

3.3.1.1. F1 :

L'analyse indique un R^2 marginal de 0.26 et un R^2 conditionnel de 0.67 et révèle un effet du contexte consonantique ($\chi^2(1)=142.84$, $p < .0001^{***}$) et de la position prosodique ($\chi^2(1)=93.21$, $p < .0001^{***}$). Ainsi, le F1 de la voyelle /a/ est plus élevé lorsque la consonne adjacente est une consonne uvulaire et lorsque la voyelle se trouve en position prosodique forte (IPi). Plus important, notre analyse révèle une interaction entre nos deux facteurs fixes ($\chi^2(1) = 15.214$, $p < .0001^{***}$) : les effets du contexte consonantique sur le F1 de /a/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta=0.54$, $SE=0.14$, $t=3.99$, $p.0001^{***}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 19a.

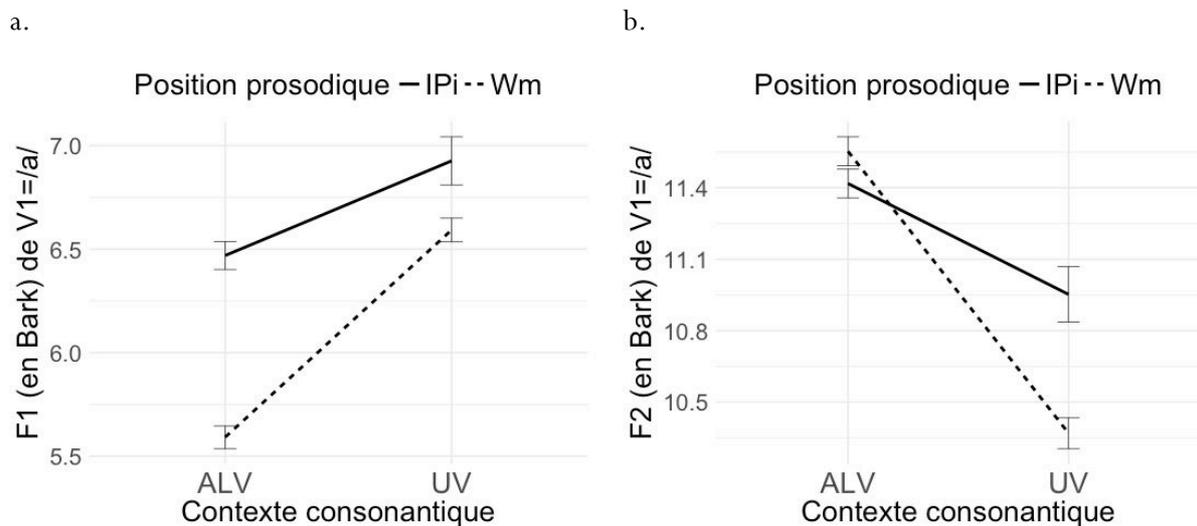


FIGURE 19. F1 (a) et F2 (b) en Bark de la voyelle /a/ (70-80 ms) en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPi » vs. position interne de mot « Wm ») en séquence V₁C.

3.3.1.2. F2 :

L'analyse menée sur le second formant de /a/ indique un R^2 marginal de 0.25 et un R^2 conditionnel de 0.80 et met en évidence un effet du contexte consonantique ($\chi^2(1)=139.49$, $p < .0001^{***}$) : Le F2 de la voyelle /a/ est plus élevé lorsqu'elle se trouve en contexte alvéolaire comparé au contexte uvulaire. En revanche, contrairement au premier formant, aucun effet de la position prosodique sur le second formant de /a/

n'apparaît ($\chi^2(1)=0.07$, $p=0.79$). Enfin, comme pour le premier formant, une interaction entre nos deux facteurs ressort de notre analyse ($\chi^2(1)=28.23$, $p<.0001^{***}$) : les effets du contexte consonantique sur le F2 de /a/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta=-0.72$, $SE=0.13$, $t=-5.50$, $p<.0001^{***}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 19b.

Notre analyse menée sur près de 1k /a/ d'une durée comprise entre 70 et 80 ms en séquence V₁C révèle une modulation de la coarticulation C-à-V par la position prosodique. En effet, les résultats obtenus, pour F1 comme pour F2, montrent moins de variations contextuelles en position IPi qu'en position Wm. Ainsi, nous pouvons affirmer que les effets de la position prosodique sur le degré de coarticulation C-à-V ne sont pas uniquement dus à des variations de durée. Nous pouvons donc étendre nos observations à toutes nos voyelles sans tenir compte de leur durée.

3.3.2. MODULATION DE LA COARTICULATION PAR LA PP EN SÉQUENCE VC ET CV

3.3.2.1. Séquences VC

L'analyse précédente a montré que la position prosodique affecte le degré de coarticulation C-à-V et que ce résultat n'est pas le simple produit de variations de durée. Nous pouvons désormais élargir nos analyses à l'ensemble de notre échantillon de 5k exemplaires de la voyelle /a/ en séquence V₁C sans contrôle de la durée (min=30ms, max=480ms M=61ms).

i. F1 :

Les résultats de cette nouvelle analyse ($R^2_m = 0.22$, $R^2_c = 0.57$) corroborent les résultats obtenus à partir du set de durée contrôlée (70-80 ms) pour F1. À savoir, un effet du contexte consonantique ($\chi^2(1) = 310.67$, $p<.0001^{***}$) et de la position prosodique ($\chi^2(1)=253.04$, $p<.0001^{***}$) sur les valeurs premier formant de /a/ indiquant un F1 plus élevé en contexte uvulaire et en position IPi. De plus, l'interaction entre nos deux facteurs

est maintenue ($\chi^2(1)=15.563, p < .0001^{***}$) : les effets du contexte consonantique sur le F1 de /a/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta=0.34, SE=0.09, t=3.99, p < .0001^{***}$). Ainsi, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 20a.

ii. F2 :

L'analyse du second formant à partir de l'échantillon complet ($R^2_m = 0.29, R^2_c = 0.77$) corrobore les résultats obtenus à partir de l'analyse du set de durée contrôlée (70-80 ms) à une exception près. En effet, comme dans l'analyse précédente nous observons un effet du contexte consonantique sur le F2 de /a/ ($\chi^2(1) = 447.9, p < .0001^{***}$), indiquant un F2 plus élevé en contexte alvéolaire, et une interaction entre la position prosodique et le contexte consonantique ($\chi^2(1) = 73.588, p < .0001^{***}$). Ce dernier résultat montre que les effets du contexte consonantique sur le F2 de /a/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta=-0.67, SE=0.08, t=-8.77, p < .0001^{***}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 20b. Enfin et contrairement à l'analyse précédente, un effet de la position prosodique sur le F2 de /a/ ressort aussi sur ce corpus plus étoffé ($\chi^2(1) = 8.728, p = 0.003^{**}$) : le F2 de /a/ est plus élevé en position IPi comparé à la position Wm. La Figure 20b montre que cet effet de la position prosodique est principalement dû au /a/ en contexte uvulaire.

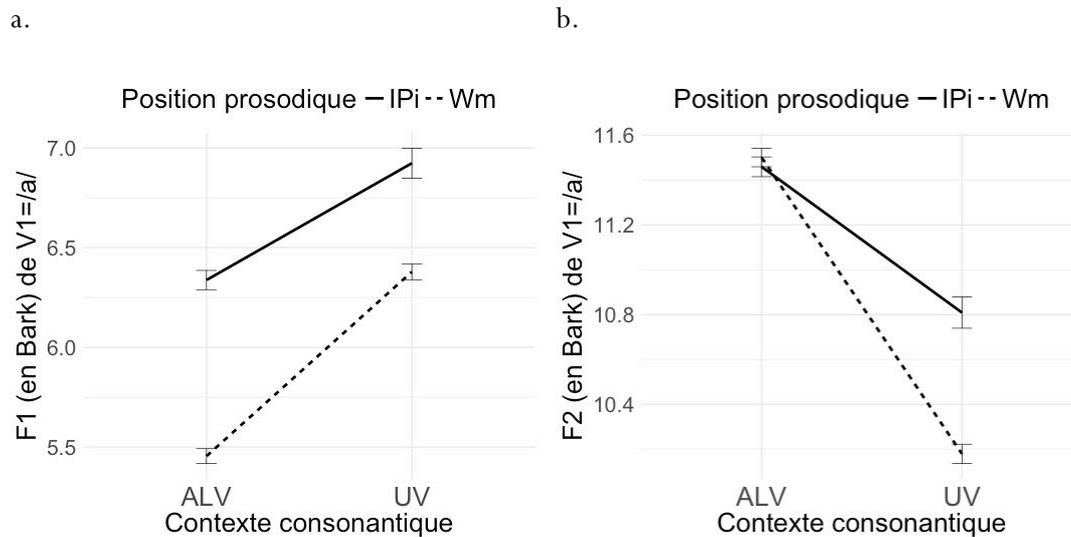


FIGURE 20. F1 (a) et F2 (b) en Bark de la voyelle /a/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence V₁C.

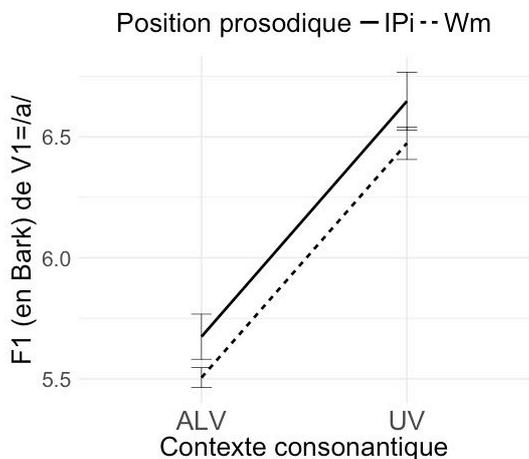
3.3.2.2. Séquences CV₁

Dans cette sous-section, nous rapporterons les résultats de l'analyse des séquences CV₁ à partir desquelles nous avons testé l'effet de la position prosodique sur le degré de coarticulation C-à-V persévératrice. La configuration des séquences, fait que la voyelle coarticulée n'est plus en position initiale absolue de phrase intonative, mais en syllabe initiale absolue. C'est la consonne (alvéolaire ou uvulaire) qui se trouve donc en position initiale absolue. La structure des modèles mixtes reste inchangée.

i. F1 :

L'analyse indique un R² marginal de 0.21 et un R² conditionnel de 0.73 et révèle un effet du contexte consonantique ($\chi^2(1) = 161.8$, $p < .0001^{***}$) et de la position prosodique ($\chi^2(1) = 4.6118$, $p < .05^*$). Ainsi, le F1 de la voyelle /a/ est plus élevé lorsque la consonne adjacente est une consonne uvulaire et lorsque la voyelle se trouve en position prosodique forte (IPI). En revanche, notre modèle ne révèle pas d'interaction entre nos deux facteurs ($\chi^2(1) = 9^{e-04}$, $p = 0.98$). En séquence CV₁, les effets contextuels sur le F1 de /a/ ne sont pas modulés par la position prosodique comme illustré Figure 21a.

a.



b.

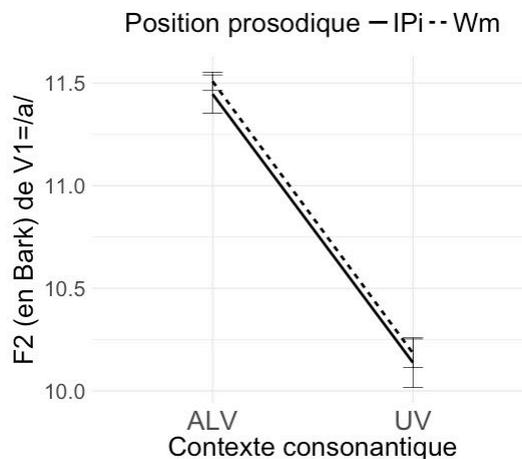


FIGURE 21. F1 (a) et F2 (b) en Bark de la voyelle /a/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPI » vs. position interne de mot « Wm ») en séquence CV₁.

ii. F2 :

L'analyse menée sur le second formant de /a/ indique un R² marginal de 0.31 et un R² conditionnel de 0.85 et met en évidence un effet du contexte consonantique ($\chi^2(1) = 227.83, p < .0001^{***}$) : Le F2 de la voyelle /a/ est plus élevé lorsqu'elle se trouve en contexte alvéolaire comparé au contexte uvulaire. En revanche, contrairement au premier formant, aucun effet de la position prosodique sur le second formant de /a/ n'apparaît ($\chi^2(1) = 0.51, p = 0.5$). Plus important, comme pour F1, notre analyse n'a pas révélé d'interaction entre nos deux facteurs ($\chi^2(1) = 0.01, p = 0.9$). En séquence CV₁, les effets contextuels sur le F2 de /a/ ne sont pas modulés par la position prosodique comme nous pouvons le voir Figure 21b.

Nos analyses menées sur la voyelle /a/ ont montré que la position prosodique affectait le degré de coarticulation C-à-V et que cet effet ne pouvait être totalement expliqué par des variations de durée. De plus, l'examen des séquences CV₁ et V₁C nous a permis de constater que cette modulation de la coarticulation par la position prosodique était fonction du type de séquence. En effet, alors qu'une réduction de la coarticulation est observée en position IPI en séquence V₁C (coarticulation anticipatoire) aucune modulation n'apparaît en séquence CV₁ (coarticulation persévératrice).

La question désormais est de savoir si l'interaction entre effets contextuels et position prosodique est fonction des voyelles, et donc spécifique, ou si au contraire ce phénomène est constant, généralisable et plus systématique. Notre première étude ayant montré que toutes les voyelles n'étaient pas affectées de la même manière par les variations contextuelles, tant au niveau des déplacements formantiques que sur leur variabilité mesurée en termes de dispersions acoustiques et de distinctivité, nous avons voulu savoir si la modulation de la coarticulation C-à-V selon la position prosodique est ou non dépendante du type de voyelle.

3.3.3. MODULATION DE LA COARTICULATION PAR LA PP : EFFET DÉPENDANT DE LA VOYELLE ?

Pour répondre à cette question, nous avons analysé les trois autres voyelles pour lesquelles nous avons suffisamment de matériel, les voyelles /i, e, ə/ en séquence V₁C. Nous présenterons les résultats par voyelle et par formant.

3.3.3.1. Voyelle /i/

i. F1 :

Le modèle mixte construit sur 4413 occurrences de la voyelle /i/ ($R^2_m = 0.05$, $R^2_c = 0.60$) révèle un effet du contexte consonantique sur les valeurs du premier formant de la voyelle ($\chi^2(1) = 98.362$, $p < .0001^{***}$) indiquant un F1 plus élevé en contexte uvulaire. En revanche aucun effet de la position prosodique n'est observé ($p=1$). Plus important, notre

modèle révèle une interaction entre nos deux facteurs ($\chi^2(1) = 7.55, p = .006^{**}$) : les effets du contexte consonantique sur le F1 de /i/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta = 0.32, SE = 0.12, t = 2.78, p = .006^{**}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 22a.

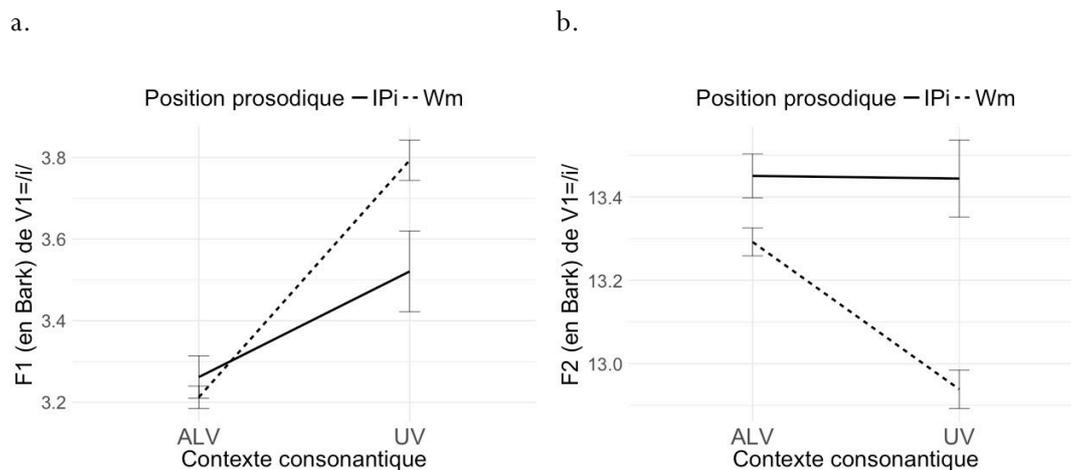


FIGURE 22. F1 (a) et F2 (b) en Bark de la voyelle /i/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPi » vs. position interne de mot « Wm ») en séquence V1C.

ii. F2 :

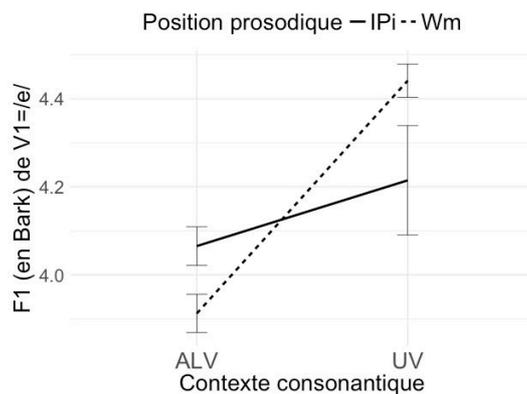
L'analyse menée sur le second formant de /i/ indique un R^2 marginal de 0.02 et un R^2 conditionnel de 0.72 et met en évidence un effet du contexte consonantique ($\chi^2(1) = 50.738, p < .0001^{***}$) et de la position prosodique ($\chi^2(1) = 23.758, p < .0001^{***}$) sur les valeurs de F2. La voyelle /i/ a un F2 plus élevé lorsqu'elle se trouve en contexte alvéolaire et lorsqu'elle apparaît en position IPi. Crucialement, notre modèle révèle une interaction entre nos deux facteurs ($\chi^2(1) = 12.209, p = .0005^{***}$) : les effets du contexte consonantique sur le F2 de /i/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta = -0.35, SE = 0.09, t = -3.57, p = .0004^{***}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 22b.

3.3.3.2. Voyelle /e/

i. F1 :

L'observation des 3k /e/ ($R^2_m = 0.01$, $R^2_c = 0.52$) met en évidence un effet du contexte consonantique sur les valeurs du premier formant ($\chi^2(1)=85.671$, $p < .0001^{***}$) indiquant un F1 plus élevé en contexte uvulaire comparé au contexte alvéolaire. Aucun effet de la position prosodique n'est en revanche révélé ($p=.08$) alors qu'une interaction entre contexte consonantique et position prosodique ressort ($\chi^2(1)=7.97$, $p=0.005^{**}$). Les effets du contexte consonantique sur le F1 de /e/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta=0.38$, $SE=0.13$, $t=2.85$, $p=.005^{**}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 23a.

a.



b.

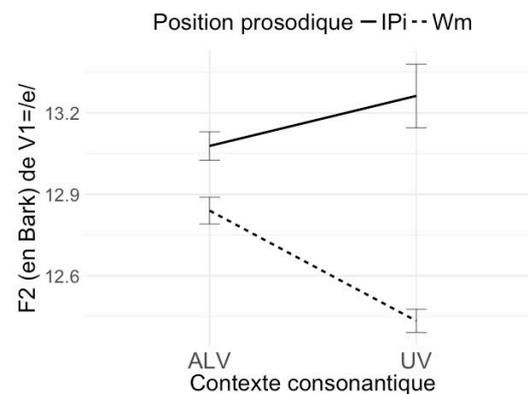


FIGURE 23. F1 (a) et F2 (b) en Bark de la voyelle /e/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPi » vs. position interne de mot « Wm ») en séquence V₁C.

ii. F2 :

L'analyse menée sur le second formant de /e/ indique un R^2 marginal de 0.1 et un R^2 conditionnel de 0.71 et met en évidence un effet du contexte consonantique ($\chi^2(1)= 39.945$, $p < .0001^{***}$) et de la position prosodique ($\chi^2(1)= 34.179$, $p < .0001^{***}$) sur les valeurs de F2. La voyelle /e/ a un F2 plus élevé lorsqu'elle se trouve en contexte

alvéolaire et lorsqu'elle apparaît en position IPi. Plus important, notre modèle révèle une interaction entre nos deux facteurs ($\chi^2(1) = 20.16, p = .0005^{***}$) : les effets du contexte consonantique sur le F2 de /e/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta = -0.59, SE = 0.13, t = -4.63, p < .0001^{***}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) ainsi que le montre la Figure 23b.

3.3.3.3. Voyelle /ɔ/

i. F1 :

Notre modèle construit à partir de 2 119 /ɔ/ indique un R^2 marginal de 0.28 et un R^2 conditionnel de 0.60 et révèle un effet du contexte consonantique ($\chi^2(1) = 276.07, p < .0001^{***}$). Ainsi, le F1 de la voyelle /ɔ/ est plus élevé lorsque la consonne adjacente est une consonne uvulaire. En revanche, notre modèle ne révèle pas d'effet de la position prosodique sur le F1 de /ɔ/ ($p = .09$) ni d'interaction entre nos deux facteurs ($p = .163$). Contrairement à toutes les autres voyelles en séquences V_1C testées, les effets contextuels sur le F1 de /ɔ/ ne sont pas modulés par la position prosodique comme illustré Figure 24a.

ii. F2 :

L'analyse du second formant ($R^2_m = 0.21, R^2_c = 0.65$) met en évidence un effet du contexte consonantique ($\chi^2(1) = 165.39, p < .0001^{***}$) sur le F2 de /ɔ/. Sans surprise, le F2 de /ɔ/ est plus élevé en contexte alvéolaire qu'en contexte uvulaire. La position prosodique n'affecte pas le second formant de la voyelle ($p = .53$). En revanche, comme pour les autres voyelles en séquence VC testées et contrairement aux résultats observés lors de l'analyse du F1 de /ɔ/, une interaction entre contexte consonantique et position prosodique apparaît ($\chi^2(1) = 26.692, p < .0001^{***}$). Ce dernier résultat montre que les effets du contexte consonantique sur le F2 de /ɔ/ sont moins forts lorsque la voyelle se trouve en position IPi comparé à la position Wm ($\beta = -0.98, SE = 0.19, t = -5.26,$

$p < .0001^{***}$). Autrement dit, on observe moins de coarticulation C-à-V en position prosodique forte (IPi) comme illustré Figure 24b.

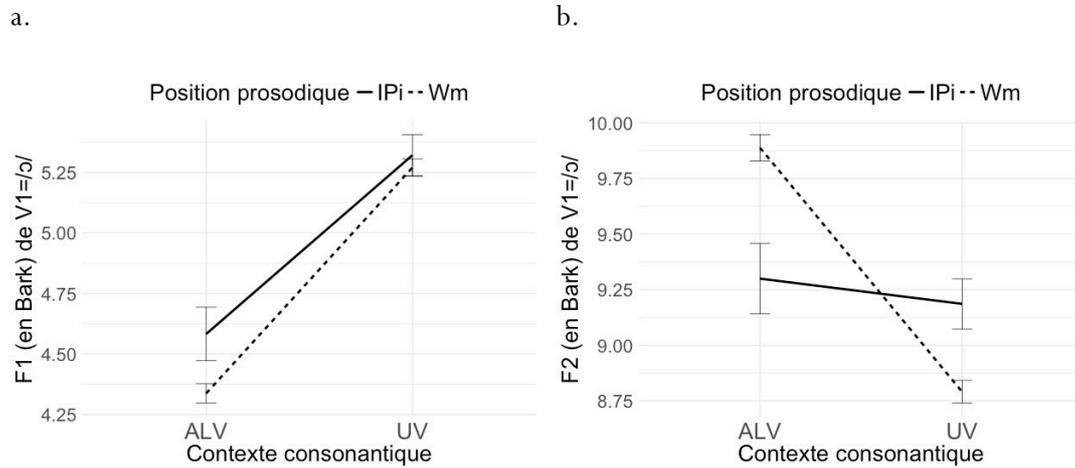


FIGURE 24. F1 (a) et F2 (b) en Bark de la voyelle /ɔ/ en fonction du contexte consonantique (alvéolaire « ALV » et uvulaire « UV ») et de la position prosodique (position initiale de groupe intonatif « IPi » vs. position interne de mot « Wm ») en séquence V1C.

On notera une différence entre les voyelles antérieures /i, e/ et la voyelle postérieure /ɔ/. En effet, en ce qui concerne F2, la réduction des effets contextuels en position IPi se fait surtout par la réduction de l'effet du contexte uvulaire pour les voyelles antérieures (voir Figure 22b et Figure 23b) alors que pour la voyelle /ɔ/ la réduction se fait autant sur le contexte alvéolaire qu'uvulaire (Figure 24b). De plus, sur F1, la réduction de la coarticulation des voyelles antérieures (mi)-fermées /i, e/ opère en contexte uvulaire (voir Figure 22a et Figure 23a), contexte élevant F1 avec lequel nos voyelles caractérisées par un F1 bas, sont en conflits.

L'analyse des voyelles /i, e, ɔ/ indique que les effets de la position prosodique sur le degré de coarticulation C-à-V anticipatoire (VC) n'est pas dépendant de la voyelle étudiée. Seuls les résultats de l'analyse du premier formant de la voyelle /ɔ/ n'indiquent pas de modulation de la coarticulation selon la position prosodique. Pour le reste, une réduction considérable de la coarticulation est observée en position IPi. D'ailleurs pour le F2 de ces trois voyelles, on remarque une inhibition totale des effets contextuels dans cette position prosodique forte. En effet, en position IPi, les valeurs de F2 restent inchangées d'un contexte consonantique à l'autre.

ÉTUDE DES INFLUENCES STYLISTIQUES SUR LA COARTICULATION C-à-V

Le style est un sentiment du monde.

André Malraux

Les voix du silence

Dans ce chapitre, il est question d'étudier comment la coarticulation C-à-V varie en fonction du style de parole à partir de deux grands corpus oraux du français variant en style. Un corpus de parole conversationnelle (C), et un corpus de parole journalistique (J) constitué d'émissions radiophoniques. Par style de parole nous incluons toutes les variations liées à ces deux types de parole bien particuliers (conversation *vs.* parole scriptée ; locuteur lambda s'adressant à un ami en face à face *vs.* locuteur professionnel face à un micro ; enregistrement en chambre sourde *vs.* parole radiophonique). L'étude des interactions entre style de parole et coarticulation s'est souvent limitée à comparer de la parole spontanée avec de la lecture de liste de mots ; ou des mots lus, insérés dans une phrase cadre ou à partir d'une liste, en hyper-articulant (parole claire parfois simulée par des consignes données aux locuteurs) avec les mêmes mots lus normalement. Dans cette étude, nous proposons de tester comment le style de parole affecte la coarticulation C-à-V en adoptant une approche plus écologique de la parole et en analysant un très grand nombre de données. Pour cela nous avons comparé le degré de coarticulation de séquences CV et VC issues du corpus J avec celles provenant du corpus C. Ainsi, nous avons montré, à partir de 22k exemplaires des voyelles /i, E, a, u, ɔ/ (/E/= /e, ε/) en contexte ALvéolaire ou uvulaire, que la coarticulation C-à-V était réduite en parole journalistique plus formelle et que cela apparaissait principalement en séquence VC (coarticulation anticipatrice). En effet, seule la voyelle /E/ présente une réduction de la coarticulation en parole formelle de même ampleur en séquence VC et CV. Pour toutes les autres voyelles, la modulation de la coarticulation selon le style de parole est plus forte en séquence VC (/a, ɔ/) voire n'apparaît qu'en séquence VC (/i, u/).

4.1. QUESTIONS ET OBJECTIFS DE L'ÉTUDE

Selon la théorie H&H (Lindblom, 1990) et les implications de la coarticulation sur la perception de la parole, discutées en introduction générale section 1.2.1, on peut s'attendre à ce que le style de parole affecte le degré de coarticulation C-à-V. Les études sur le sujet ne sont pas nombreuses et portent toutes sur un style de parole bien particulier : la parole claire. Ces travaux relatent une réduction de la coarticulation en parole claire (sauf une étude). Ce qui pose question dans ces études c'est plutôt la place qui est donnée au style de parole, la façon dont il est parfois élicité et le degré de « naturalité » de la parole testée. Si ces travaux ont le mérite d'ouvrir la question sur les relations entre style et coarticulation, il semble qu'il reste à mieux comprendre ce qui se joue, en élargissant les investigations à d'autres styles de parole, puisque la parole « claire », si s'en est un, n'est pas le seul style existant.

Sur la base de ces constats, le présent travail étudie les effets potentiels du style de parole sur la coarticulation C-à-V dans deux grands corpus de parole naturelle du français. Un corpus de parole conversationnelle, très familière, et un corpus de parole journalistique, de style plus formel. Contrairement aux précédents travaux sur le sujet, principalement effectués sur de la parole élicitée, nous étudions la relation entre le style de parole et la coarticulation dans une condition de parole plus écologique. Ceci est d'autant plus pertinent qu'il a été récemment montré que la modulation de la coarticulation en fonction du style de parole dépendait de l'authenticité de la situation de communication (Scarborough & Zellou, 2013). Le premier objectif est donc de déterminer si le degré de coarticulation varie en fonction du style de parole et si oui dans quelle direction et pour quelles voyelles.

Dans notre précédente étude, nous avons vu que l'effet de la position prosodique sur la coarticulation C-à-V en français était fonction du type de coarticulation. Une modulation de la coarticulation en position prosodique forte était observée uniquement pour les cas de coarticulation anticipatrice (VC). Cela peut être dû aux différences relatives à la position qu'occupe la voyelle vis-à-vis de la frontière de constituant, puisqu'en condition VC, la

voyelle est directement à la frontière alors qu'en condition CV c'était la consonne qui se trouvait à la frontière. On peut cependant aussi imaginer que ce résultat soit dû au fait que les deux types de coarticulations résultent de mécanismes différents : la coarticulation persévératrice CV est due à l'inertie des articulateurs mis en jeu pour l'articulation de C, qui étendent leur influence sur la voyelle V suivante, alors que la coarticulation anticipatrice est due à la planification anticipée de l'articulation de la consonne à venir pendant la voyelle. L'inertie des articulateurs est un phénomène biomécanique que l'on ne peut pas vraiment contrôler, contrairement à l'anticipation qui, elle, est contrôlée, donc modifiable en fonction de la position prosodique. Dans ce cas là, on devrait aussi observer des différences dans la relation de ces deux types de coarticulation avec le style de parole. Notre second objectif est donc de déterminer si les variations de style affectent de la même manière la coarticulation C-à-V anticipatrice (VC) et persévératrice (CV).

Enfin, les effets de style étant souvent attribués à des variations de durée, nous contrôlerons la durée des segments pour s'assurer que les éventuels effets du style de parole sur le degré de coarticulation C-à-V, ne soient pas le simple produit de variations de durée. Ainsi, nous pourrions discuter de la question d'interdépendance entre variation de style et variation de durée à la lumière de nos résultats.

Les questions auxquelles nous tenterons ici de répondre sont :

- Est-ce que le degré de coarticulation C-à-V varie en fonction du style de parole ? Si oui, dans quelle direction et pour quelles voyelles ?
- Les influences stylistiques affectent-elles de la même manière la coarticulation anticipatrice VC que persévératrice CV ?
- Observe-t-on un effet du style de parole sur la coarticulation C-à-V lorsque la durée de la voyelle est contrôlée ?

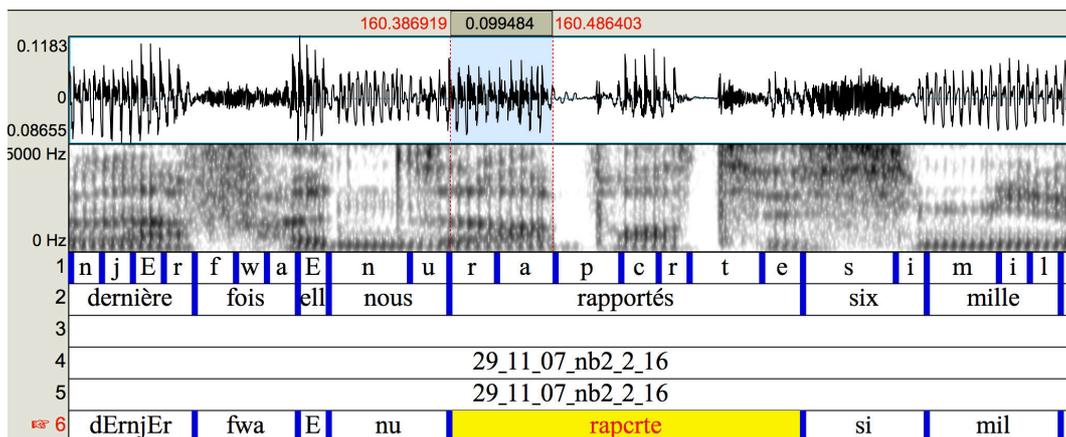
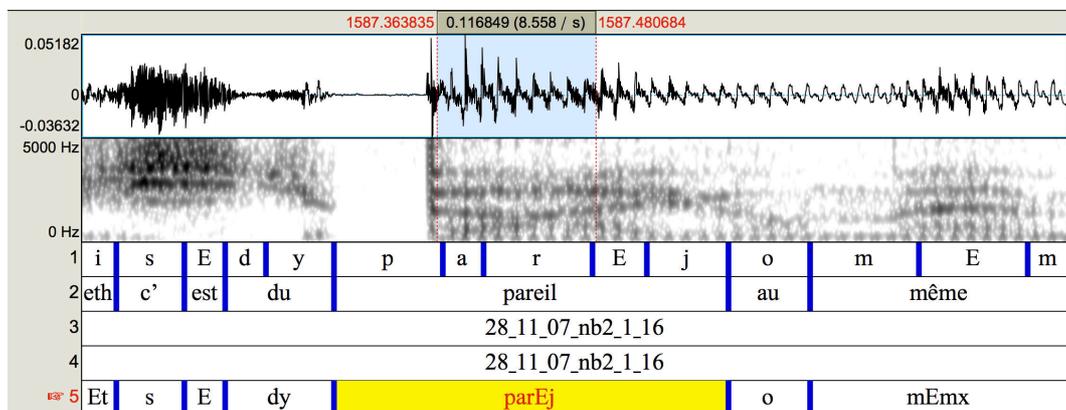
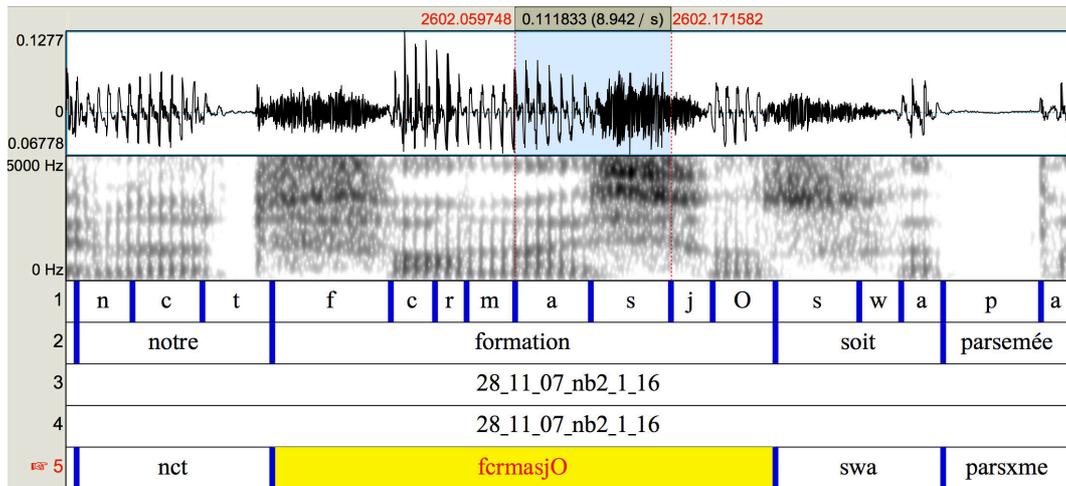
4.2. METHODE

Afin de tester si le degré de coarticulation C-à-V varie en fonction du style de parole, nous comparons des séquences CV et VC extraites du corpus NCCFr (Torreira et al., 2010), corpus de parole conversationnelle (C), avec des séquences similaires issues du corpus ESTER (Gravier et al., 2006) qui est un corpus de parole journalistique (J). Toutes les informations relatives à ces corpus et à leurs particularités sont décrites en introduction générale. Pour plus de détails se référer à la section 1.6.2.

Nous sommes partis de la base de données construite pour l'étude 1 (2.2.1) à partir du corpus NCCFr comportant les données de 15 locuteurs masculins et nous avons appliqué la même procédure au corpus ESTER. Après examen des données disponibles, nous avons sélectionné 8 locuteurs du corpus ESTER. Nous avons cherché à maximiser le nombre d'observations par locuteur en vue de faire une normalisation des données par locuteur. Ainsi, nous avons sélectionné les locuteurs disposant de plus de matériel dans ce second corpus en cherchant à obtenir un nombre total d'observations semblables dans les deux corpus.

4.2.1. MATÉRIEL LINGUISTIQUE ET PRETRAITEMENTS

À partir des données de ces 23 locuteurs (15 NCCFr, 8 ESTER), nous avons cherché des séquences CV et VC où la consonne testée C devait être, soit une consonne alvéolaire /t, d, s, z, n, l/, soit une consonne uvulaire /ʁ/. Nous avons choisi ces consonnes car les effets coarticulatoires observés dans notre première étude (voir CHAPITRE 2) étaient les plus marqués et allaient dans des directions opposées (uv : élévation de F1 et abaissement de F2 vs. ALV : abaissement de F1 et élévation de F2). Le contexte opposé, gauche en VC et droit en CV, était toujours une consonne labiale /p, b, m, f, v/. Les spectrogrammes présentés Figure 25 illustrent les types de séquences analysées dans ce travail.



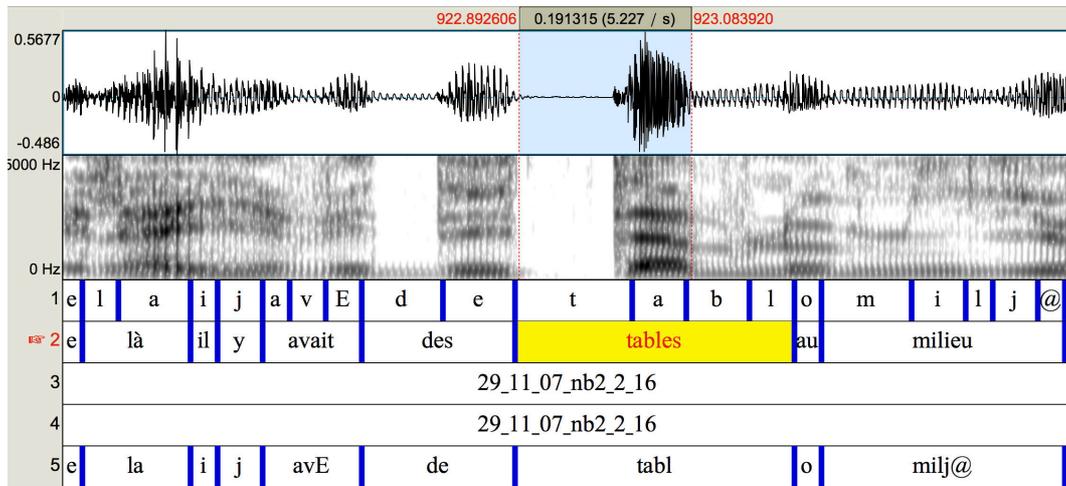


FIGURE 25. Spectrogrammes illustrant les séquences CV et VC pour la voyelle /a/ par contexte consonantique (alvéolaire vs. uvulaire) extraites du corpus de parole conversationnelle : le contexte alvéolaire en VC est illustré avec le mot « formation » et en séquence CV « tables » ; le contexte uvulaire en contexte VC est illustré avec le mot « pareil » et en séquence CV « rapporter ».

Il est important de noter que dans les séquences CV la consonne et la voyelle faisaient toujours partie de la même syllabe alors que pour les séquences VC cela n'a pas pu être contrôlé. Pour ce dernier type de séquences la consonne et la voyelle pouvaient être tautosyllabiques comme hétéro-syllabiques.

Enfin, nous avons aussi contrôlé la durée des voyelles afin de s'assurer que les éventuels effets du style de parole sur la coarticulation C-à-V ne puissent être attribués à des variations de durée dans les deux styles de parole. Pour cela, nous n'avons considéré que les voyelles d'une durée comprise entre 50 et 80 ms. La tranche de durée a été définie afin de favoriser le nombre d'observations. En effet, les voyelles courtes et moyennes sont les mieux représentées dans nos corpus. Cependant, nous avons préféré écarter les exemplaires de moins de 50 ms pour éviter les erreurs de mesures plus fréquentes sur des segments très courts.

À l'issue de cette procédure de prétraitement des données, nous avons examiné le matériel disponible pour les différentes voyelles. Nous avons pu constater que dans chaque corpus, nous avons suffisamment d'exemplaires des voyelles /i, E, a, ə, u/ (avec /E/= /e, ε/).

Cela correspond à 22k exemplaires de nos cinq voyelles analysées (11326 en parole C et 11286 en parole J) comme on peut le voir en détail Table 29.

	Corpus NCCFR				Corpus ESTER			
	<i>parole Conversationnelle C</i>				<i>parole Journalistique J</i>			
	VC		CV		VC		CV	
	ALV	UV	ALV	UV	ALV	UV	ALV	UV
/i/	439	41	520	104	1090	98	907	159
/E/	1644	472	2001	220	935	810	1493	328
/a/	1940	593	1759	222	759	638	1272	363
/ɔ/	265	238	112	206	497	296	334	342
/u/	122	293	91	104	130	589	202	84
<i>total</i>	6047		5239		5484		5239	
TOTAL	11326				11286			

TABLE 29. Résumé du matériel linguistique utilisé dans cette étude et distribution des voyelles dans chaque contexte consonantique (ALV = alvéolaire et UV = uvulaire) par type de coarticulation (CV et VC) issues du corpus de parole conversationnelle (C) à gauche et du corpus de parole journalistique (J) à droite.

La Figure 26 illustre la distribution de la durée des différentes voyelles en fonction du style de parole par type de coarticulation (ou séquences) VC et CV. Ainsi, on peut voir que la distribution de la durée des voyelles se répartit équitablement dans les deux styles de parole. La superposition est quasi totale en VC alors qu'en CV des différences sur certaines voyelles sont observables. C'est le cas pour les voyelles moyennes /E, ɔ/ et la voyelle ouverte /a/. Contre toute attente, on peut voir que les voyelles les plus courtes (50-60 ms) sont plus nombreuses en parole journalistique qu'en parole conversationnelle. Si les résultats observés pour ces voyelles en séquences CV divergent de ceux des autres voyelles, il faudra tenir compte de ces observations pour les interpréter.

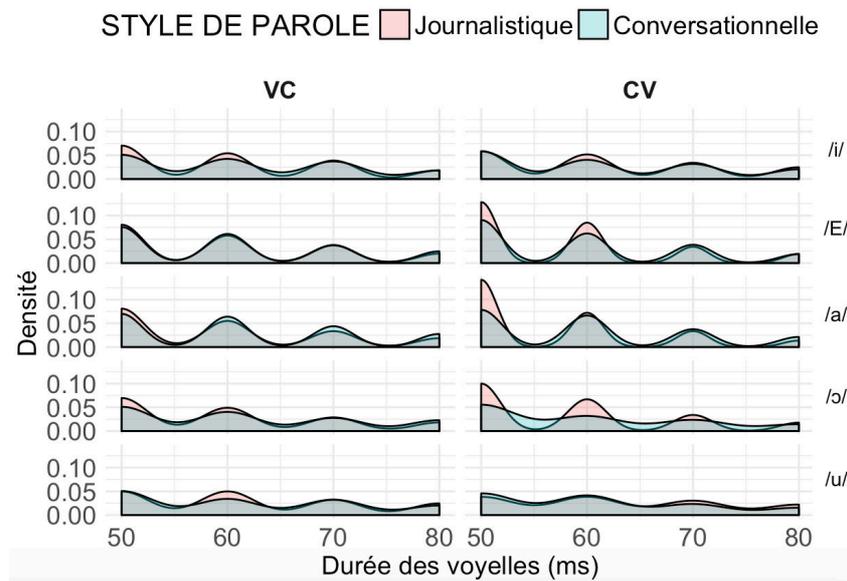


FIGURE 26. Distribution de la durée des voyelles en fonction du style de parole (J ou C) par voyelle (/i, E, a, ɔ, u/) et par type de coarticulation (VC ou CV).

4.2.2. MESURE DE COARTICULATION & NORMALISATION DES DONNÉES

La coarticulation C-à-V est mesurée par les changements du second formant (F2) en fonction du CONTEXTE CONSONANTIQUE (alvéolaire « ALV » ou uvulaire « UV ») sur la base de ce qui a été effectué pour la première étude (2.2.2.1). Pour toutes questions relatives aux valeurs formantiques voir la section 2.2.1.1. Afin de réduire la variabilité due aux différences anatomiques entre locuteurs, nous avons procédé à une normalisation des valeurs de F2 en z-score par locuteur selon la formule de Lobanov (1971) détaillée section 2.2.1.6.

4.2.3. ANALYSES STATISTIQUES

Afin de répondre à nos questions de recherches détaillées section 4.1, nous avons construit avec la bibliothèque lme4 (Bates et al., 2015) de R (R development Core Team, 2008), un modèle mixte par voyelle analysée, soit 5 modèles au total, pour tester les relations entre le F2 de chaque voyelle (en z-score), le CONTEXTE CONSONANTIQUE (C : ALV vs. UV), le TYPE DE COARTICULATION (CV ou VC) et le STYLE DE PAROLE (journalistique J ou

conversationnel C). Le locuteur a été intégré comme facteur aléatoire. Afin d'éviter un taux élevé d'erreur de Type I (Cunnings, 2012) une pente aléatoire par locuteur a été incluse pour nos facteurs fixes CONTEXTE CONSONANTIQUE et TYPE DE COARTICULATION en interaction. Cela correspond à la variabilité entre locuteurs de l'effet du contexte interagissant avec le type de coarticulation sur le second formant des voyelles. Les valeurs de p ont été obtenues par approximations de type Satterthwaite à l'aide de la fonction lmerTest (Kuznetsova et al., 2013). Le seuil de référence a été fixé à $p < .05$. L'effet de chaque facteur fixe et de leur interaction a été testé par comparaison de modèles avec la fonction anova. Les valeurs de R^2 associées à chaque modèle ont été obtenues à l'aide de la fonction r.squaredGLMM intégrée dans la bibliothèque 'MuMIn'. Les analyses des contrastes a posteriori ont été effectuées avec la fonction lsmeans (bibliothèque 'emmeans', Lenth et al., 2018).

4.3. RESULTATS

Dans cette partie, seront présentés les résultats voyelle par voyelle. Pour chaque voyelle testée, nous répondrons à nos différentes questions, puis nous reprendrons les différents résultats pour faire un résumé général des résultats. Seuls les résultats statistiques relatifs à nos questions de recherches seront présentés et discutés.

4.3.1. LA VOYELLE /i/

L'analyse de 3k exemplaires de la voyelle /i/ indique un R^2 marginal de 0.01 et un R^2 conditionnel de 0.23 et révèle un effet du CONTEXTE CONSONANTIQUE ($\chi^2(1)=26$, $p < .0001^{***}$) et du STYLE DE PAROLE ($\chi^2(1)=5$, $p = .02^*$). Ainsi, le F2 de la voyelle /i/ est plus élevé en parole journalistique et lorsque la consonne adjacente est une consonne alvéolaire. En revanche, aucun effet du TYPE DE COARTICULATION ne ressort ($\chi^2(1)=0.4$, $p = .5$). Le F2 de /i/ ne varie pas selon qu'il soit en séquence CV ou VC.

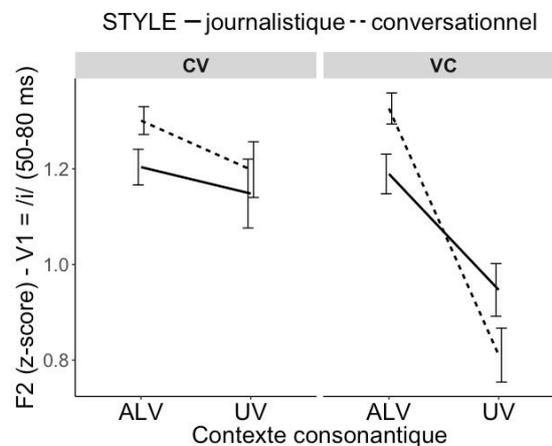


FIGURE 27. F2 (en z-score) de la voyelle /i/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel).

Plus intéressant, notre analyse révèle une interaction entre le CONTEXTE CONSONANTIQUE et le STYLE DE PAROLE ($\chi^2(1)=6$, $p < .02^*$). Celle-ci indique un degré de coarticulation C-à-V plus faible en parole journalistique comparé à la parole conversationnelle ($\beta = -0.27$, $SE = 0.09$, $t = 3.10$, $p < .003^{**}$). Enfin, le modèle révèle aussi une triple

interaction entre nos trois facteurs fixes CONTEXTE CONSONANTIQUE, STYLE et TYPE DE COARTICULATION. Comme on peut le voir Figure 27, l'interaction entre STYLE DE PAROLE et CONTEXTE CONSONANTIQUE, n'apparaît que pour les séquences VC, soit pour les cas de coarticulation anticipatrice ($\beta=0.22$, $SE=0.09$, $t=2.41$, $p=.02^*$). La coarticulation C-à-V est plus faible en parole journalistique uniquement pour les cas de coarticulation anticipatrice.

Pour cette voyelle, on a pu voir que la coarticulation C-à-V, mesurée par les changements de F2 selon le lieu d'articulation de consonne adjacente, était moins importante en parole journalistique qu'en parole conversationnelle pour les cas de coarticulation anticipatrice (séquences VC) uniquement.

4.3.2. LA VOYELLE /E/

L'analyse de 8k exemplaires de la voyelle /E/ (regroupant /e, ε/ indique un R2 marginal de 0.2 et un R² conditionnel de 0.25 et révèle un effet du CONTEXTE CONSONANTIQUE ($\chi^2(1)=6$, $p<.01^*$) et du TYPE DE COARTICULATION ($\chi^2(1)=11$, $p=.0015^{**}$). Ainsi, le F2 de la voyelle /E/ est plus élevé lorsque la consonne adjacente est une consonne alvéolaire et lorsqu'elle se trouve en séquence CV. En revanche, aucun effet du STYLE DE PAROLE ne ressort ($\chi^2(1)=0.05$, $p=.8$). Les valeurs de F2 de /E/ restent inchangées d'un style de parole à l'autre.

Plus important, notre analyse révèle une interaction entre le CONTEXTE CONSONANTIQUE et le STYLE DE PAROLE ($\chi^2(1)=20$, $p<.0001^{***}$). Celle-ci indique un degré de coarticulation C-à-V plus faible en parole journalistique comparé à la parole conversationnelle ($\beta=-0.24$, $SE=0.04$, $t=-5.75$, $p<.0001^{***}$). Contrairement à la voyelle /i/, notre modèle n'indique pas de triple interaction pour la voyelle /E/ ($\chi^2(1)=0.09$, $p=.8$). Ceci signifie que l'interaction entre STYLE DE PAROLE et CONTEXTE CONSONANTIQUE ne dépend pas du type de coarticulation. Pour cette voyelle, on observe moins de coarticulation C-à-V en parole journalistique comparé à la parole conversationnelle en séquence VC (coarticulation

anticipatrice) comme CV (coarticulation persévératrice). Ces résultats sont présentés Figure 28.

Pour cette voyelle /E/, on a pu voir que la coarticulation C-à-V, mesurée par les changements de F2 selon le lieu d'articulation de consonne adjacente, était moins importante en parole journalistique qu'en parole conversationnelle quel que soit le type de coarticulation (anticipatrice comme persévératrice).

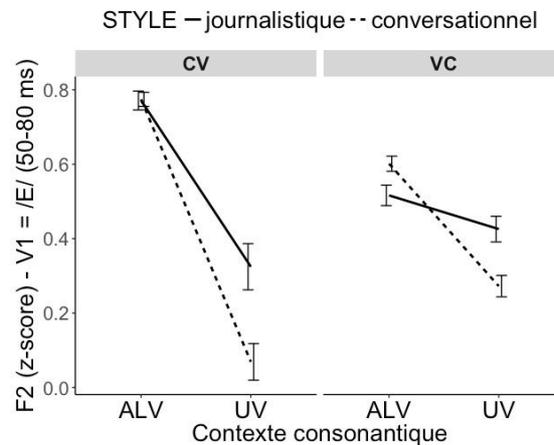


FIGURE 28. F2 (en z-score) de la voyelle /E/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel).

4.3.3. LA VOYELLE /a/

L'analyse de 7.5k exemplaires de la voyelle /a/ indique un R^2 marginal de 0.46 et un R^2 conditionnel de 0.53 et révèle un effet du CONTEXTE CONSONANTIQUE ($\chi^2(1)=65$, $p < .0001^{***}$), du STYLE DE PAROLE ($\chi^2(1)=6$, $p=.014^*$) et du TYPE DE COARTICULATION ($\chi^2(1)=11$, $p < .0001^{***}$). Ainsi, le F2 de la voyelle /a/ est plus élevé en parole journalistique, lorsque la consonne adjacente est une consonne alvéolaire et quand elle se trouve séquence CV.

Plus important, notre analyse révèle une interaction entre le CONTEXTE CONSONANTIQUE et le STYLE DE PAROLE ($\chi^2(1)=13$, $p < .001^{**}$). Celle-ci indique un degré de coarticulation C-à-V plus faible en parole journalistique comparé à la parole conversationnelle ($\beta=-0.35$, $SE=0.06$, $t=-5.79$, $p < .001^{**}$). Enfin, le modèle révèle aussi une triple

interaction entre nos trois facteurs fixes CONTEXTE CONSONANTIQUE, STYLE et TYPE DE COARTICULATION. Comme on peut le voir Figure 29, l'interaction entre STYLE DE PAROLE et CONTEXTE CONSONANTIQUE apparaît pour les deux types de coarticulation (CV et VC), mais la réduction de la coarticulation en parole journalistique est plus forte dans les cas de coarticulation anticipatrice VC.

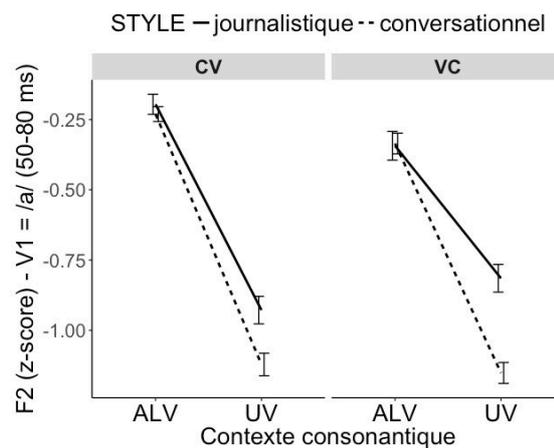


FIGURE 29. F2 (en z-score) de la voyelle /a/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel).

L'analyse de la voyelle /a/ indique un degré de coarticulation C-à-V moins important en parole journalistique comparé à de la parole conversationnelle. Si cet effet apparaît pour nos deux types de coarticulation (CV et VC), celui-ci est plus fort pour les cas de coarticulation anticipatrice VC.

4.3.4. LA VOYELLE /ɔ/

L'analyse de 2k exemplaires de la voyelle /ɔ/ indique un R^2 marginal de 0.64 et un R^2 conditionnel de 0.77 et révèle un effet du CONTEXTE CONSONANTIQUE ($\chi^2(1)=40$, $p < .0001^{***}$). Sans surprise, le F2 de la voyelle est plus élevé en contexte alvéolaire. En revanche, aucun effet du STYLE DE PAROLE ($\chi^2(1)=2$, $p=.02^{ns}$) ni du TYPE DE COARTICULATION ($\chi^2(1)=2$, $p=0.14^{ns}$) n'est ressorti de l'analyse.

Plus important, notre analyse révèle une interaction entre le CONTEXTE CONSONANTIQUE et le STYLE DE PAROLE ($\chi^2(1)=24, p < .0001^{***}$). Celle-ci indique un degré de coarticulation C-à-V plus faible en parole journalistique comparé à la parole conversationnelle ($\beta=-0.72, SE=0.09, t=-7.67, p < .0001^{***}$). Enfin, le modèle montre aussi une triple interaction entre nos trois facteurs fixes CONTEXTE CONSONANTIQUE, STYLE et TYPE DE COARTICULATION. Comme on peut le voir Figure 30 l'interaction entre STYLE DE PAROLE et CONTEXTE CONSONANTIQUE apparaît pour les deux types de coarticulation (CV et VC), mais la réduction de la coarticulation en parole journalistique est plus forte dans les cas de coarticulation anticipatrice VC.

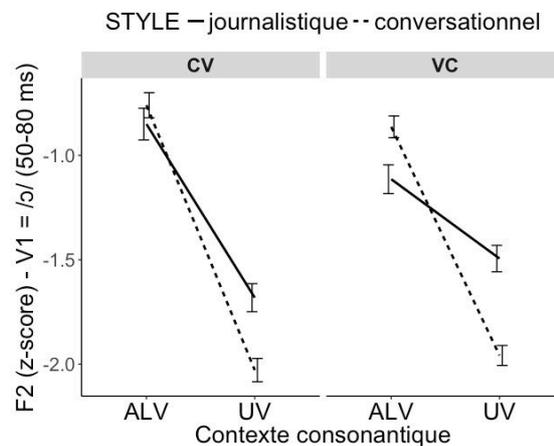


FIGURE 30. F2 (en z-score) de la voyelle /ɔ/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel).

Comme pour la voyelle /a/, nous trouvons que pour la voyelle /ɔ/ le degré de coarticulation C-à-V est moins important en parole journalistique comparé à de la parole conversationnelle. Si cet effet apparaît pour nos deux types de coarticulation (CV et VC), celui-ci est plus fort pour les cas de coarticulation anticipatrice VC.

4.3.5. LA VOYELLE /u/

L'analyse de 1.5k exemplaires de la voyelle /u/ indique un R^2 marginal de 0.36 et un R^2 conditionnel de 0.49 et révèle un effet du CONTEXTE CONSONANTIQUE ($\chi^2(1)=52$, $p < .0001^{***}$) et du STYLE DE PAROLE ($\chi^2(1)=6$, $p = .02^*$). Ainsi, le F2 de la voyelle /u/ est plus élevé en parole journalistique et lorsque la consonne adjacente est une consonne alvéolaire. En revanche, aucun effet du TYPE DE COARTICULATION ne ressort ($\chi^2(1)=0.1$, $p = .8$). Le F2 de /u/ ne varie pas selon qu'il soit en séquence CV ou VC.

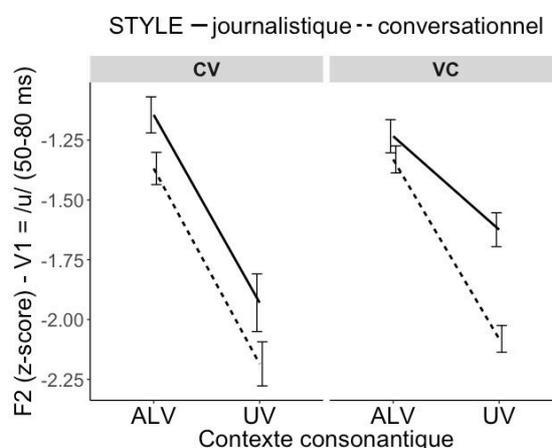


FIGURE 31. F2 (en z-score) de la voyelle /u/ en séquence CV et VC en fonction du contexte consonantique (alvéolaire vs. uvulaire) et du style de parole (journalistique vs. conversationnel).

Notre analyse révèle également une interaction entre le CONTEXTE CONSONANTIQUE et le STYLE DE PAROLE ($\chi^2(1)=16$, $p < .0001^{***}$). Celle-ci indique un degré de coarticulation C-à-V plus faible en parole journalistique comparé à la parole conversationnelle ($\beta = -0.36$, $SE = 0.08$, $t = -4.62$, $p < .0001^{***}$). Enfin, le modèle montre aussi une triple interaction entre nos trois facteurs fixes CONTEXTE CONSONANTIQUE, STYLE et TYPE DE COARTICULATION. Comme on peut le voir Figure 31, l'interaction entre STYLE DE PAROLE et CONTEXTE CONSONANTIQUE, n'apparaît que pour les séquences VC, soit pour les cas de coarticulation anticipatrice.

Pour cette voyelle, comme pour la voyelle /i/, on a pu voir que la coarticulation C-à-V, mesurée par les changements de F2 selon le lieu d'articulation de consonne adjacente, était moins importante en parole journalistique qu'en parole conversationnelle pour les cas de coarticulation anticipatrice (séquences VC) uniquement.

4.3.6. LES EFFETS OBSERVÉS DÉPENDENT-ILS DE LA VOYELLE TESTÉE ?

L'analyse de nos 5 voyelles /i, E, a, ə, u/ révèle un effet principal du CONTEXTE CONSONANTIQUE. Comme attendu, le F2 des voyelles est plus élevé en contexte alvéolaire qu'en contexte uvulaire. Un effet du TYPE DE COARTICULATION est ressorti des analyses des voyelles /E/ et /a/ indiquant des valeurs de F2 plus élevées lorsque ces voyelles sont en séquences CV comparé à VC. Enfin, le F2 des voyelles /i, u, a/ est affecté par le STYLE DE PAROLE et la direction est la même pour les trois voyelles. En parole journalistique, le F2 des voyelles /i, u, a/ est plus élevé qu'en parole conversationnelle. Tous ces résultats sont résumés Table 30.

	CONTEXTE	χ^2	TYPE DE COA	χ^2	STYLE	χ^2
/i/	ALV > UV	26 ^{***}	CV ₁ = V ₁ C	0.4 ^{n.s.}	J > C	5 [*]
/u/	ALV > UV	52 ^{***}	CV ₁ = V ₁ C	0.1 ^{n.s.}	J > C	6 [*]
/E/	ALV > UV	6 [*]	CV ₁ > V ₁ C	11 ^{**}	J = C	0.05 ^{n.s.}
/ə/	ALV > UV	40 ^{***}	CV ₁ = V ₁ C	2 ^{n.s.}	J = C	2 ^{n.s.}
/a/	ALV > UV	65 ^{***}	CV ₁ > V ₁ C	11 ^{***}	J > C	6 [*]

TABLE 30. Test de vraisemblance des modèles mixtes linéaires testant les effets principaux des prédicteurs CONTEXTE CONSONANTIQUE*SEQUENCE*STYLE pour les cinq voyelles testées: /i, E, a, ə, u/. Les valeurs χ^2 (avec un degré de liberté) sont rapportées en tant qu'estimations de la taille de l'effet. « ALV » et « UV » pour alvéolaire et uvulaire, "J" et "C" pour le style journalistique et conversationnel. ALV>UV, par exemple, signifie que F2 de V est plus élevé lorsque V est en contexte alvéolaire qu'en contexte uvulaire. Les valeurs p significatives sont indiquées par le symbole astérisque * (p <.05 = *).

Un résultat fondamental est que nos analyses révèlent une interaction entre le CONTEXTE CONSONANTIQUE et le STYLE DE pour toutes les voyelles testées. Cette interaction révèle

moins de coarticulation en parole journalistique qu'en parole conversationnelle (voir le Table 31 pour un résumé des résultats par voyelle). L'effet du style de parole sur la coarticulation C-à-V ne dépend pas du type de voyelle.

CONTEXTE*STYLE				
	χ^2	β	ET	t
/i/	6*	-0.27	0.09	-3.17
/u/	16***	-0.36	0.08	-4.62
/E/	20***	-0.24	0.04	-5.75
/ɔ/	24***	-0.72	0.09	-7.67
/a/	13***	-0.35	0.06	-5.79

TABLE 31. Test de vraisemblance des modèles mixtes linéaires testant l'interaction entre CONTEXTE CONSONANTIQUE*SEQUENCE*STYLE pour les 5 voyelles: / i, E, a, ɔ, u /. Les valeurs de χ^2 (avec un degré de liberté) sont rapportées en tant qu'estimations de la taille d'effet; Les coefficients β , les erreurs-types (ET) et les valeurs t du modèle sont également rapportés. Les valeurs p significatives sont indiquées par le symbole * (p <.05 = *).

En revanche, l'effet du style sur la coarticulation C-à-V est fonction du TYPE DE COARTICULATION et cela dépend de la voyelle testée. En effet, nos analyses révèlent une triple interaction entre CONTEXTE CONSONANTIQUE, STYLE DE PAROLE et TYPE DE COARTICULATION pour les voyelles /i, u, a, ɔ/. Pour les voyelles fermées /i, u/ le plus faible degré de coarticulation C-à-V en parole journalistique ne ressort que pour les cas de coarticulation anticipative VC. Pour les voyelles (mi)-ouvertes /ɔ, a/ cet effet est présent dans les deux types de coarticulation VC et CV. Cependant, la réduction de la coarticulation en parole journalistique est plus importante en VC qu'en CV. En d'autres termes, pour ces voyelles, le STYLE DE PAROLE semble affecter principalement la coarticulation anticipatrice. Pour la dernière voyelle testée /E/, le TYPE DE COARTICULATION n'affecte pas l'interaction entre CONTEXTE CONSONANTIQUE et STYLE DE PAROLE. On trouve moins de coarticulation en parole journalistique quel que soit le TYPE DE COARTICULATION, anticipatrice (VC) ou persévératrice (CV). Un résumé des résultats est présenté Table 32.

CONSONANTAL CONTEXT*SEQUENCE*STYLE				
	χ^2	β	ET	t
/i/	5*	0.22	0.09	2.41
/u/	5*	0.33	0.14	2.30
/E/	0.09 ^{n.s.}	-0.02	0.07	-0.31
/ɔ/	5*	0.28	0.11	2.49
/a/	7**	0.19	0.07	2.88

TABLE 32. Test de vraisemblance des modèles mixtes linéaires testant l'interaction entre CONTEXTE CONSONANTIQUE*TYPE DE COARTICULATION*STYLE pour les 5 voyelles: / i, E, a, ɔ, u /. Les valeurs de χ^2 (avec un degré de liberté) sont rapportées en tant qu'estimations de la taille d'effet; Les coefficients β , les erreurs-types (ET) et les valeurs t du modèle sont également rapportés. Les valeurs p significatives sont indiquées par le symbole * ($p < .05 = *$).

Enfin, on rajoutera que la réduction de la coarticulation en parole formelle ne se fait pas toujours sur le même contexte et que cela dépend du type de coarticulation et de la voyelle testée (voir les Figure 27 à Figure 31). En effet, pour les voyelles montrant une réduction de la coarticulation sur F2 en parole journalistique pour les séquences CV /E, a, ɔ/, celle-ci se fait toujours et uniquement sur le contexte uvulaire. Pour les séquences VC, la réduction de la coarticulation opère sur les deux contextes /i, E, ɔ/ ou uniquement sur le contexte uvulaire pour les voyelles /u, a/.

Prises ensemble ces trois études montrent que la coarticulation C-à-V en français dépend des caractéristiques phonétiques des consonnes et des voyelles ainsi que des facteurs liés à la structuration prosodique du message linguistique et à la situation de communication. Après avoir passé rapidement en revue les questions relatives aux éventuelles implications de la durée des voyelles dans nos résultats, nous analyserons les résultats montrant que les effets contextuels dépendent des voyelles et de la directionnalité de la coarticulation. Dans un deuxième temps, nous essaierons de montrer à partir des différences observées dans nos études 2 et 3 que la modulation de la coarticulation selon la position prosodique a une fonction linguistique différente de celle observée sous l'effet du style de parole même si toutes les deux conduisent à des résultats semblables : une plus grande résistance des voyelles à la coarticulation en position IPi et en parole formelle. Pour terminer, nous proposerons une réflexion sur les possibles implications de nos résultats dans une perspective d'évolution de la langue en lien avec la préférence des langues pour l'antériorisation de voyelles postérieures par rapport à la postériorisation de voyelles antérieures.

	/i/	/e/	/ɛ/	/a/	/ɔ/	/u/	/o/	/ɔ/
Effets de C sur V : DÉPLACEMENT	C_{UV} : F1 F2, F3&F4 de /i/ C_{ALV} : F1 F2 C_{VEL} : F1 pas d'effet sur F2 les effets contextuels sur F1 sont plus forts pour les voyelles fermées /i, u/ et les voyelles les plus ouvertes /ɛ, a, ɔ/ les effets contextuels sur F2 sont plus forts pour les voyelles postérieures /u, o, ɔ/ et la voyelle /a/							
DISPERSION DISTINCTIVITÉ	$C_{ALV}=C_{VEL}<C_{UV}$ $C_{ALV}=C_{VEL}>C_{UV}$	$C_{ALV}<C_{VEL}=C_{UV}$ $C_{ALV}=C_{VEL}>C_{UV}$	$C_{UV}<C_{ALV}=C_{VEL}$ $C_{UV}>C_{ALV}>C_{VEL}$	$C_{ALV}=C_{VEL}<C_{UV}$ $C_{UV}>C_{ALV}=C_{VEL}$	$C_{ALV}=C_{VEL}=C_{UV}$ $C_{ALV}=C_{VEL}>C_{UV}$	$C_{ALV}=C_{VEL}=C_{UV}$ $C_{ALV}>C_{VEL}>C_{UV}$	$C_{ALV}<C_{VEL}=C_{UV}$ $C_{UV}>C_{ALV}>C_{VEL}$	$C_{ALV}=C_{VEL}<C_{UV}$ $C_{UV}>C_{VEL}>C_{ALV}$
PP (IPi, Wm)*COA								
CV				COA IPi = COA Wm COA IPi = COA Wm				
VC	COA IPi < COA Wm en C _{UV}	COA IPi < COA Wm en C _{UV}	COA IPi < COA Wm en C _{UV} C _{ALV} COA IPi < COA Wm en C _{UV}					COA IPi = COA Wm - COA IPi < COA Wm en C _{UV} C _{ALV}
Style (J, C)*COA F2								
CV	COA J = COA C -	COA J < COA C en C _{UV}	COA J < COA C en C _{UV}		COA J = COA C -			COA J < COA C en C _{UV}
VC	COA J < COA C en C _{UV} C _{ALV}	COA J < COA C en C _{UV} C _{ALV}	COA J < COA C en C _{UV}		COA J < COA C en C _{UV}			COA J < COA C en C _{UV} C _{ALV}
	Effet s\ VC = s\ CV		Effet s\ VC > s\ CV					Effet s\ VC > s\ CV

TABLE 33. Résumé de l'ensemble des résultats obtenus dans cette thèse. COA = coarticulation, PP = position prosodique, IPi = Position initiale de groupe intonatif, Wm = position interne de mot, J = parole journalistique, C = parole conversationnelle, C_{ALV} = contexte alvéolaire, C_{VEL} = contexte vélaire, C_{UV} = contexte uvulaire, s\ = su

5.1. QUESTIONS RELATIVES À LA TEMPORALITÉ DE LA PAROLE

Le style de parole et la position prosodique affectant la temporalité de la parole, la question que l'on peut se poser est de savoir si la modulation de la coarticulation par ces deux facteurs observés ne serait que le produit d'une variation de durée : des segments plus longs en parole plus formelle ou en position I_{Pi} et donc moins coarticulés.

5.1.1. UNE MODULATION DE LA COARTICULATION PAR LE STYLE DE PAROLE MALGRÉ LE CONTRÔLE DE LA DURÉE DES VOYELLES

La question de la durée est au centre des débats concernant les effets du style de parole sur les variations segmentales : Est-ce que les variations observées sous l'effet du style ne sont que le produit des variations de durée ? Dans une version révisée du modèle de voyelle de Lindblom (1963), Moon et Lindblom (1994) ont montré que les déplacements formantiques des voyelles dus au contexte consonantique dépendaient fortement de la durée de la voyelle, et que le lien entre changements spectraux et durée était moins fort en parole claire qu'en forme de citation. Le même constat a également été rapporté pour la coarticulation V-to-C en français (Duez, 1992). En mixtèque, Dicanio et al. (2015) ont trouvé qu'en parole spontanée les segments étaient plus courts et plus fortement influencés par le contexte consonantique. Cependant, d'autres travaux suggèrent que les variations dues au style de parole vont au-delà des effets de durée : les locuteurs peuvent adopter différentes stratégies de production en fonction de la condition de parole (par exemple, Ferguson et Kewley-Port, 2002, Smiljanic et Bradlow, 2005). En réalité, des travaux antérieurs sur les différences de style de parole en français ont montré que les réductions spectrales subies par les voyelles en parole conversationnelle par rapport à la parole lue s'appliquaient principalement aux voyelles courtes, de moins de 40 ms (Rouas, Beppu & Adda-Decker, 2010). De plus, Audibert et al. (2015) ont trouvé plus de réduction de l'espace acoustique, plus de centralisation vocalique et de dispersion intra-catégorie en parole conversationnelle qu'en parole journalistique et parole lue pour les voyelles courtes (jusqu'à 50ms) que pour les voyelles moyennes (jusqu'à 80ms) et longues (jusqu'à 120ms).

Dans l'étude 3, nous avons montré que la coarticulation C-à-V était réduite en parole journalistique à partir de l'analyse de voyelles d'une durée comprise entre 50 et 80 ms. Ces résultats suggèrent donc que l'effet du style de parole sur la coarticulation n'est pas un pur produit des variations de durée. Ces observations viennent s'ajouter à l'ensemble des recherches antérieures montrant des effets du style de parole sur l'acoustique des voyelles lorsque la durée des segments est contrôlée (Rouas et al., 2013 et Audibert et al., 2015). Ainsi, nous avons été en mesure de montrer les effets de style sur le degré de coarticulation des voyelles lorsque la durée des voyelles était contrôlée. À la lumière de ces constatations, il semble donc important de considérer le style comme un facteur de variation en soi en le séparant des conséquences de variations de durée dues à un débit plus rapide en parole relâchée par exemple.

5.1.2. UNE MODULATION DE LA COARTICULATION PAR LA POSITION PROSODIQUE MALGRÉ UN CONTRÔLE DE LA DURÉE SUR UN SOUS-ÉCHANTILLON DE DONNÉES

Une conséquence du renforcement prosodique initial est l'allongement des segments (voir Fougeron & Keating, 1997 pour l'anglais ; Hsu & Jun, 1997 pour le taiwanais). On peut donc se demander si une réduction de la coarticulation des segments en position initiale de constituant prosodique, ne serait pas due à l'augmentation de la durée des segments dans cette position. Cependant, on notera qu'en français l'allongement des segments sous l'effet du renforcement prosodique initial n'est pas systématique (Fougeron, 1998 ; Georgeton, 2014). En effet, Georgeton (2014) rapportait un effet très partiel de la position prosodique sur les variations de durée des voyelles. L'allongement des voyelles en position initiale de Groupe Intonatif comparé à la position initiale de Mot n'apparaissait que pour certaines voyelles et pour certains locuteurs. De plus, une étude portant sur la variation de la coarticulation V-à-V sous l'effet de la frontière prosodique en anglais (V#V) a montré que l'effet du renforcement initial sur la coarticulation, conduisant à une réduction de celle-ci, était indépendant des variations de durée des segments (Cho, 2004). Selon l'auteur les variations de durée peuvent expliquer une partie des effets de renforcement initial, mais elles ne peuvent en rendre compte pleinement.

Les résultats obtenus dans cette thèse montrent que la durée ne permet pas d'expliquer en totalité les variations de coarticulation. En effet, l'analyse préliminaire concernant la modulation de la coarticulation par la position prosodique effectuée sur 1686 /a/ en séquence VC d'une durée de 70 à 80 ms indique que le renforcement initial affecte la coarticulation des segments post-frontaliers indépendamment d'éventuelles variations de durée. Autrement dit, les effets du renforcement initial ne peuvent se résumer à de simples variations de durée causées par les différentes positions étudiées (c.-à-d. allongement des voyelles en position prosodique forte IPi). Les résultats de cette analyse indiquent que les influences contextuelles, mesurées par les changements de F1 et F2 en fonction du contexte consonantique (alvéolaire vs. uvulaire), sont réduites en position prosodique forte (IPi).

Pour résumer, si la coarticulation est en partie déterminée par la durée des segments, les résultats obtenus dans cette thèse ne sont probablement pas attribuables à des variations de durée.

5.2. LA COARTICULATION : UN PHÉNOMÈNE VARIABLE DÉPENDANT DE LA VOYELLE ET DE LA DIRECTION DES EFFETS COARTICULATOIRES

5.2.1. DES EFFETS CONTEXTUELS DÉPENDANTS DES VOYELLES

L'observation des différents résultats de cette thèse résumés Table 33 page 138 suggèrent à la fois des spécificités liées aux voyelles impliquées et des patterns semblables à l'ensemble des voyelles. En effet, dans notre première étude nous avons vu que les déplacements formantiques dépendaient principalement des spécificités phonétiques des consonnes mais que certaines voyelles montraient de plus amples déplacements formantiques que d'autres comme par exemple les voyelles postérieures et la voyelle /a/ sur F2. Au delà des déplacements formantiques, nous avons montré que la coarticulation influait sur la variabilité des voyelles. Comme cela avait déjà été montré pour le mixtèque sur la dimension F2 (Dicano et al., 2015), les voyelles sont moins dispersées en contexte alvéolaire. La distinctivité des voyelles dépend plus fortement de la compatibilité acoustique de l'influence consonantique sur les formants avec le système acoustique des voyelles du français.

5.2.1.1. Questions relatives à la stabilité des voyelles

Les résultats de notre étude préliminaire ne valident totalement aucune théorie de la parole concernant la stabilité intrinsèque de certaines voyelles comme on peut le voir en comparant la Table 34 rapportant les prédictions des modèles et nos résultats présentés Table 35. Cependant, certains résultats corroborent en partie chacune des théories présentées comme nous allons le discuter par la suite.

	Prédictions		
	QT	TDF	DAC
- Stable +	/i, u, a/	/i, u, o, ə, a/	/i/ /e/ /ɛ/ /a/ /u, o, ə/ /x/
	/e, ɛ, x, o, ə/	/e, ɛ, x/	

TABLE 34. Rappel des prédictions des différentes théories de la parole (qt = théorie quantique, tdf = théorie de la dispersion-focalisation, dac = degré de contrainte articulatoire) sur la stabilité des voyelles.

Résultats						
	DF1234	DF1	DF2	DF3	DF4	Distinctivité
- Stable +	/e, x/	/e, o/	/e/	/e,ɛ,a,x/	/ə, x/	/i/ (0.79)
		/i, x/	/i, ɛ/		/e/	/a/ (0.73)
	/i, ɛ, ə, o/			/ə/	/o/	/e/ (0.63)
			/x, a/	/i/		/x/ (0.54)
						/ɛ/ (0.47)
	/a/	/ɛ, ə/	/ə/	/u/	/a, i, ɛ/	/ə/ (0.45)
		/a/	/o/			/u/ (0.44)
/u/	/u/	/u/	/o/	/u/	/o/ (0.30)	

TABLE 35. Résultats de notre étude préliminaire visant à déterminer si certaines voyelles sont plus stables que d'autres (DF1234 = Dispersion multidimensionnelle, DF1 = Dispersion sur F1, DF2 = Dispersion sur F2, DF3 = Dispersion sur F3, DF4 = Dispersion sur F4, Distinctivité à partir du score F1).

iii. *Questions sur la stabilité des voyelles selon leur dispersion intra-catégorie*

Que disent nos résultats vis-à-vis des contraintes articulatoires des voyelles ?

Les prédictions du DAC (Recasens et al., 1997) se retrouvent dans les résultats obtenus sur la dispersion du F2 des voyelles à l'exception de la voyelle /x/ regroupant pourtant les voyelles /ø, œ, ə/. De plus, la voyelle /u/ est toujours la voyelle la plus dispersée (mesure multidimensionnelle et par formant). Or selon les attentes de ce modèle, cette voyelle peu contrainte au niveau du dos de la langue, devait effectivement se trouver en bas de l'échelle de stabilité. En revanche, on ne retrouve pas ce résultat pour les autres voyelles postérieures /o, ɔ/ comme le prédit le DAC. La faible dispersion de la voyelle /e/ observée dans ce travail est en accord avec les prédictions de ce modèle conférant un DAC maximalement élevé aux voyelles antérieures du fait qu'elles soient contraintes articulatoirement au niveau du dos de la langue. Cependant, on se serait attendu à ce que la voyelle /i/ montre une plus faible dispersion acoustique de ses exemplaires que la voyelle /e/ étant donné que le degré de contrainte articulatoire augmente avec l'étroitesse de la constriction. Un autre résultat en accord avec le DAC, concerne l'assez forte variabilité de la voyelle /a/ sur la mesure multidimensionnelle, F1 et F4 notamment. Pour les autres théories, les convergences entre leur prédiction et nos résultats sont plus partielles.

Que disent nos résultats vis-à-vis des voyelles quantiques ?

Comme le prédit la Théorie Quantique (Stevens, 1972 ; 1989) la voyelle /i/ est très peu dispersée acoustiquement sur la dimension F2 même si sa dispersion sur les autres dimensions est considérable. Cependant, on ne retrouve pas ce résultat pour les autres voyelles quantiques /u, a/ dont la variabilité sur cette dimension F2 est considérable.

Que disent nos résultats vis-à-vis des propriétés focales des voyelles ?

La théorie de la Dispersion-Focalisation (Schwartz et al., 1997), confère un statut particulier aux voyelles focales (voyelles caractérisées par un rapprochement très important de deux formants provoquant un renforcement mutuel de leur amplitude et leur perception comme un seul et même pic). Comme on peut le voir Table 35, nos résultats sur la dispersion des exemplaires des voyelles à une valeur de référence de leur catégorie

vocalique, confèrent peu de poids à cette théorie. En effet, la voyelle /i/ focale sur F3-F4 est plus dispersée sur ces dimensions que sur F1 et F2. De même les voyelles /u, o, ə, a/ sont toutes très dispersées sur leur formants focaux F1 et F2.

La question que l'on peut se poser est de savoir si malgré une variabilité acoustique importante sur leurs formants focaux, ces voyelles maintiennent une distance stable entre ces derniers. En effet, la particularité des voyelles focales tenant dans le rapprochement de deux formants dits « focaux », on peut donc imaginer que ces formants focaux varient mais restent proches l'un de l'autre. Ceci étant, nos observations ne vont pas dans ce sens. Comme on peut le voir Figure 32, l'écart-type des distances F2-F1 (lignes pleines) et F4-F3 (lignes en pointillé), reflétant la variabilité dans les distances F2-F1 et F4-F3 respectivement, n'est pas moins important pour les voyelles focales sur les dimensions concernées (/u, o, ə, a/ F1-F2, /i/ F3-F4). On notera cependant que pour le français, il a été suggéré que le plus important pour la voyelle /i/ était d'avoir une distance maximale entre F2 et F3 (Tabain et Perrier, 2005 ; aussi admis par Georgeton, 2014). Nos données ne nous permettent pas de vérifier ce dernier point, mais pourraient expliquer la particularité de la voyelle focale /i/.

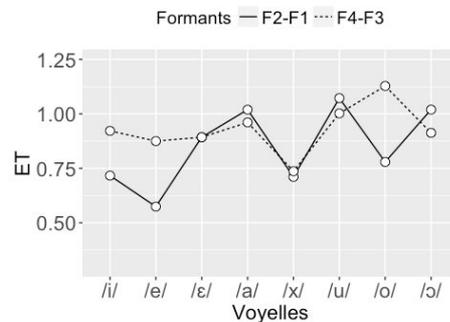


FIGURE 32. Écart-types (ET) des distances entre F2-F1 (ligne pleine) et F3-F4 (ligne en pointillé) pour les différentes voyelles étudiées.

iv. Questions sur la stabilité des voyelles selon leur distinctivité acoustique

Concernant les résultats de la classification automatique, on peut voir que sur cet aspect de stabilité testé, les résultats de la voyelle /i/ corroborent les prédictions de nos différentes théories (QT, TDF et DAC) qui placent toutes la voyelle /i/ comme la plus stable.

Cependant, il est important de noter que notre étude n'inclut pas la voyelle /y/, voisine acoustique de la voyelle /i/. Le bon score de la voyelle /i/, comme le score des autres voyelles, est relatif aux nombres de catégories incluses dans le modèle. Il se peut que, si nous avions inclus la voyelle /y/, le score de la voyelle /i/ eût été moins bon.

Par ailleurs, on a des voyelles postérieures très peu distinctives en accord avec le modèle DAC mais en désaccord avec TDF, puisque ce sont toutes des voyelles focales, et avec la QT théorie pour la voyelle quantique /u/.

La voyelle /a/, autre voyelle quantique et focale, est d'une manière générale très dispersée mais obtient un bon score de distinctivité en accord avec ces deux théories s'opposant au modèle DAC.

La voyelle /e/ reste très stable dans ses réalisations acoustiques mais cela ne lui confère pas un statut particulier en ce qui concerne la distinctivité. Les résultats de la faible variabilité du /e/ en accord avec le modèle DAC se retrouve donc sur le plan de la dispersion acoustique mais pas sur celui de la distinctivité.

v. *Que disent les différences observées entre nos deux mesures de variabilité (dispersion et distinctivité) ?*

La mise en correspondance de nos résultats de dispersion avec ceux de distinctivité et leurs divergences, indique que ces deux aspects de la stabilité ne s'équivalent pas. Une voyelle peut être très dispersée tout en étant très distinctive /i, a/ alors qu'une voyelle peu dispersée peut être peu distinctive /e/. Ceci nous conduit à faire l'hypothèse que ce qui importe dans la distinctivité acoustique ce n'est pas la stabilité dans les réalisations acoustiques des voyelles, mais plutôt le nombre de compétiteurs acoustiques présents dans la zone dans laquelle évolue une voyelle.

L'observation des triangles vocaliques sur les dimensions F1/F2 (Figure 33a), F2/F3 (Figure 33b) et F3/F4 (Figure 33c), confirment cette hypothèse. En effet, on peut voir que l'espace de réalisation de la voyelle /i/ sur F2 (dimension sur laquelle la voyelle est peu variable), ne chevauche que très peu l'espace de dispersion d'une autre voyelle. Cela est sûrement dû au fait que la voyelle /y/, non testée dans cette étude, n'apparaît pas sur ces

espaces vocaliques. Les voyelles /i/ et /y/ se différenciant surtout sur F3, on aurait très certainement trouvé un chevauchement important des ellipses de ces voyelles sur F1 et F2. En revanche, puisqu'elles se distinguent principalement sur F3, en incluant /y/ on aurait sans doute observé peu de chevauchement de ces deux voyelles sur cette dimension. Or, ce qui ressort des triangles vocaliques ci-dessous, c'est justement que le F3 de la voyelle /i/ évolue dans une zone acoustique occupée par aucune autre voyelle. La particularité de cette voyelle tiendrait donc dans son F3 et du fait qu'il n'existe pas de compétiteur acoustique dans la zone où le F3 de /i/ évolue.

Un constat semblable peut être fait pour la voyelle /a/ (deuxième voyelle la plus distinctive). Cette voyelle maximale ouverte et ayant un F1 maximale élevé n'a pas de compétiteur acoustique sur cette dimension. De plus, la particularité de son F1 permet de la garder bien distinctive de la voyelle /x/ dont elle partage la zone acoustique sur F2. Ceci expliquerait pourquoi malgré une forte dispersion, cette voyelle a été bien reconnue par le classifieur automatique.

D'un autre côté, nous avons les voyelles postérieures très peu distinctives et très fortement chevauchées entre elles sur toutes les dimensions (surtout en ce qui concerne /u/ et /o/). Aussi la voyelle /e/ très peu dispersée se trouve toujours dans une zone acoustique dans laquelle évolue plusieurs compétiteurs (/i, ε/ sur F1 et F2, /ε, a/ sur F3 et F4) expliquant pourquoi elle n'a pas obtenu de meilleurs scores de distinctivité.

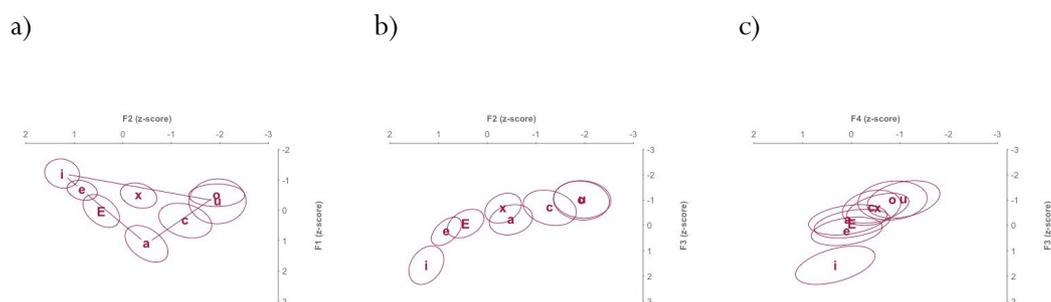


FIGURE 33. Triangles vocaliques sur les dimensions : a) F1/F2, b) F2/F3, c) F3/F4.

La divergence des résultats obtenus sur nos deux métriques de variabilité concernant la voyelle /x/ est intéressante. En effet, cette voyelle a montré une très faible variabilité dans

ses réalisations acoustiques quelle que soit la dimension acoustique considérée contrairement aux prédictions que nous avons faites à partir des différentes théories de la parole et notamment le DAC. Pourtant, son score de distinctivité reste moyen. Nous expliquons ce résultat de par la position (centrale) qu’occupe cette voyelle dans l’espace acoustique. En étant au centre de l’espace vocalique, la voyelle /x/ est relativement proche à la fois des voyelles postérieures et des voyelles antérieures. Le modèle aura donc plus de risque de confondre une de ces dernières voyelles avec un /x/.

5.2.1.2. Un degré de déplacements formantiques dû au contexte consonantique dépendant des voyelles

L’analyse des DÉPLACEMENTS formantiques et de leur direction corrobore la littérature. La Table 36 résume ces résultats conformes à nos attentes.

CONTEXTE	CONSÉQUENCES ACOUSTIQUES
ALVÉOLAIRE	abaissement F1 élévation F2 vers 1800Hz
VÉLAIRE	abaissement F1 peu d’effets sur F2
UVULAIRE	élévation de F1 abaissement de F2 et F3, F4 /i/

TABLE 36. Résultats des déplacements formantiques causés par les différents contextes consonantiques.

Nos observations menées sur le F1 des voyelles traduisent les exigences articulatoires des consonnes. En effet, la consonne uvulaire articulée avec un abaissement considérable de la partie pré-dorsale de la langue et une constriction dorso-uvulaire (Straka, 1965), implique un abaissement de la mâchoire et un conduit oral relativement ouvert responsable des valeurs élevées de F1. Au contraire, les consonnes vélares et alvéolaires contraintes par la réalisation de la constriction tendent à l’abaisser (Recasens et Espinosa, 2009).

Concernant F2, le même constat peut être fait. Le contexte uvulaire ayant une constriction très postérieure agrandit la cavité orale contribuant à l'abaissement du F2. Le contexte alvéolaire, lui, attire le F2 des voyelles vers son locus autour de 1800 Hz. Les faibles effets contextuels des consonnes vélaires sur le F2 des voyelles sont sans doute plus surprenants pour le français puisque jusque-là cela n'avait jamais été montré. Ce résultat s'explique du fait que les occlusives vélaires ont un allophone palatal lorsqu'elles sont entourées d'une voyelle antérieure (voir par exemple Öhman, 1966). Ainsi, ces consonnes s'adaptent au lieu d'articulation de la voyelle environnante subissant plus les effets des voyelles qu'elles n'imposent leurs caractéristiques. Ceci corrobore les résultats obtenus sur le catalan (Recasens, 1985) et l'anglais américain (Fowler & Brancazio, 2000) selon lesquels les consonnes vélaires sont peu résistantes à la coarticulation de lieu d'articulation linguale et de ce fait peu agressives envers les autres segments.

Concernant le F3 de la voyelle /i/, on a vu que celui-ci était abaissé en contexte uvulaire. De plus, cette voyelle montre très peu de variation contextuelle sur F2, en accord avec la théorie quantique (Stevens, 1972, 1989) indiquant que pour le /i/ les effets coarticulatoires, dûs au lieu d'articulation linguale, sont plus marqués sur F3 corroborant les résultats observés par Recasens (1985) sur le catalan sous l'effet du [ɫ] sombre. Enfin, concernant le F4 de cette voyelle, Nos résultats montrent comme pour F3, un abaissement du F4 de /i/.

Les effets coarticulatoires semblent donc déterminés principalement par le contexte consonantique et les particularités articulatoires des consonnes. Cependant, le degré des déplacements formantiques dus au contexte consonantique est fonction de la voyelle. Par exemple, les voyelles postérieures et la voyelle /a/ montrent plus d'effets contextuels sur F2 que les autres voyelles. Le plus fort du contexte consonantique sur F2 pour les voyelles postérieures a aussi été montré en anglais américain (Stevens et House, 1966) et en catalan (Recasens, 1985). Cependant, ces derniers auteurs rapportaient une plus forte élévation du F2 de la voyelle fermée /u/ que des autres voyelles /o, ɔ/ et attribuaient cela au fait que plus une voyelle était arrondie, plus les effets contextuels sur F2 étaient importants,

l'élévation du F2 étant causée par la perte de labialité. Nos résultats ne montrent pas cette différence entre les voyelles postérieures.

Un autre exemple concernant la particularité des voyelles concerne les effets contextuels sur le premier formant. Les voyelles les plus ouvertes /a, ε, ɔ/ (en accord avec des résultats obtenus en anglais américain (Stevens et House, 1966) et en catalan (Recasens, 1985)) et les plus fermées /i, u/ sont plus fortement influencées par le contexte consonantique que les autres voyelles.

5.2.1.3. Effets de la coarticulation sur la dispersion intra-catégorie des voyelles

La grande question de notre première étude était de savoir si, au-delà des déplacements formantiques, la coarticulation affectait la variabilité des voyelles. Nos résultats indiquent clairement que oui. L'analyse des effets contextuels sur la dispersion des voyelles a révélé que celle-ci était plus faible en contexte alvéolaire pour la majorité des voyelles, corroborant les résultats obtenus en mixtèque montrant moins de dispersion des voyelles lorsqu'elles sont précédées par une consonne alvéolaire (Dicanio et al., 2015). En effet seules les voyelles /x/, moins dispersées (avec /e/), et la voyelle /u/ maximalement dispersée ne montrent pas de différence de dispersion selon le contexte consonantique. La voyelle /ε/ est moins dispersée en contexte uvulaire. Pour les voyelles /i, a, ɔ/ les contextes vélaux et alvéolaires ne se distinguent pas, montrant une plus faible dispersion qu'en contexte uvulaire. La Table 37 résume les effets des différents contextes consonantiques sur la dispersion des voyelles (dispersion multidimensionnelle F1234).

Nos résultats concernant la voyelle /i/ indiquent que la dispersion de cette voyelle dépend du degré de conflit existant entre celle-ci et le contexte dans lequel elle se trouve. L'effet du contexte consonantique sur la variabilité de cette voyelle permet de dire, en effet, que la forte variabilité de cette voyelle est causée par le contexte uvulaire. En contextes alvéolaire et uvulaire, cette voyelle présente une faible dispersion intra-catégorie, alors qu'au contact d'une consonne uvulaire, dont la trajectoire articulaire s'oppose à la sienne, cette voyelle devient aussi variable que la voyelle /u/ maximalement dispersée dans notre corpus.

EFFET DU CONTEXTE SUR LA DISPERSION DES VOYELLES	
/i/	ALV = VEL < UV
/e/	ALV < VEL = UV
/ɛ/	UV < ALV = VEL
/a/	ALV = VEL < UV
/ɣ/	ALV = VEL = UV
/u/	ALV = VEL = UV
/o/	ALV < VEL = UV
/ɔ/	ALV = VEL < UV

TABLE 37. Résumé des effets du contexte consonantique (ALV = alvéolaire, VEL = vélaire et UV = uvulaire) sur la dispersion des voyelles.

5.2.1.4. Les effets de la coarticulation sur la distinctivité des voyelles

Concernant la distinctivité des voyelles, nous avons aussi montré un fort impact du contexte consonantique sur les résultats de la classification automatique (voir la Table 38 ci-dessous). Avant toute chose, on rappelle que notre classification repose sur des valeurs statiques de formants bien que l'on sache l'importance des informations dynamiques pour l'identification des sons en contextes. Nous avons fait ce choix afin de limiter au maximum les erreurs de détection de formants dues à une mauvaise segmentation. De plus, notre but n'était pas de créer le meilleur modèle possible pour classer les voyelles, mais de saisir comment les modifications spectrales sous l'effet du contexte consonantique amenaient à des confusions entre les voyelles et quel type de confusions pouvaient être observées.

Les voyelles les plus fermées /i, e, u/ sont plus distinctives en contextes alvéolaire et vélaire alors que les voyelles plus ouvertes /ɛ, a, ɔ/ le sont en contexte uvulaire. La voyelle /o/ diffère des autres voyelles fermées car elle est plus distinctive en contexte uvulaire. Ceci est dû à sa proximité acoustique avec le /u/ notamment sur la dimension F1. La voyelle /o/ se trouve plus distincte du /u/ en contexte uvulaire où son F1 est élevé et son F2 abaissé. De plus, la voyelle /u/ est moins distinctive en contexte uvulaire car l'élévation de son F1 augmente les confusions avec la voyelle /o/. On peut facilement imaginer que le bon score de classification de cette voyelle en contexte alvéolaire, soit dû au fait que la

voyelle /y/ n'ait pas été incluse dans le modèle. Peut-être que, dans ce cas, on aurait relevé plus d'erreurs de classification de /u/ en contexte alvéolaire même si ces deux voyelles se seraient sans doute distinguées sur les dimensions F3 et F4.

	TCC	ALV	VEL	UV
/i/	85 %	90 %	88 %	51 %
/e/	73 %	74 %	70 %	56 %
/ɛ/	46 %	41 %	30 %	67 %
/a/	62 %	57 %	59 %	88 %
/x/	78 %	80 %	77 %	49 %
/u/	41 %	71 %	63 %	20 %
/o/	31 %	22 %	10 %	84 %
/ɔ/	42 %	29 %	38 %	61 %

TABLE 38. Taux de classification correcte (ou score de sensibilité) en pourcentage des voyelles tous contextes confondus (TCC) et par contexte : ALV = alvéolaire, VEL = vélaire et UV = uvulaire.

Nos résultats montrent sans surprise que le contexte consonantique joue aussi sur les erreurs de classification concernant le lieu d'articulation des voyelles. Par exemple, la voyelle /ɔ/ souvent confondue avec la voyelle centrale /x/ l'est principalement en contexte alvéolaire et vélaire qu'en contexte uvulaire. Les confusions de /o/ en /x/ sont aussi nombreuses en contextes alvéolaire et vélaire. En revanche, le /x/ très bien identifié tous contextes confondus est souvent pris pour un /ɔ/ en contexte uvulaire.

Pour résumer, nous avons un effet semblable des contextes alvéolaire et vélaire sur la distinctivité des voyelles. On observe un fort taux de confusion de la voyelle /ɔ/ en /x/ et une meilleure reconnaissance des voyelles fermées, mi-fermées et antérieures. Le contexte uvulaire diminue fortement la distinctivité des voyelles /i, e, x, u/ et améliore celle des voyelles /a, ɛ, o, ɔ/. Cela indique que la distinctivité des voyelles est augmentée lorsqu'elles se trouvent dans un contexte renforçant leurs propriétés inhérentes, c'est à dire lorsque la consonne et la voyelle ne sont pas en conflit en lien avec leur trajectoire

articulatoire. Ainsi, on peut dire que la distinctivité des voyelles en contexte dépend de la compatibilité CV.

5.2.1.5. Une modulation de la coarticulation selon la durée de la voyelle, dépendante du type de voyelle

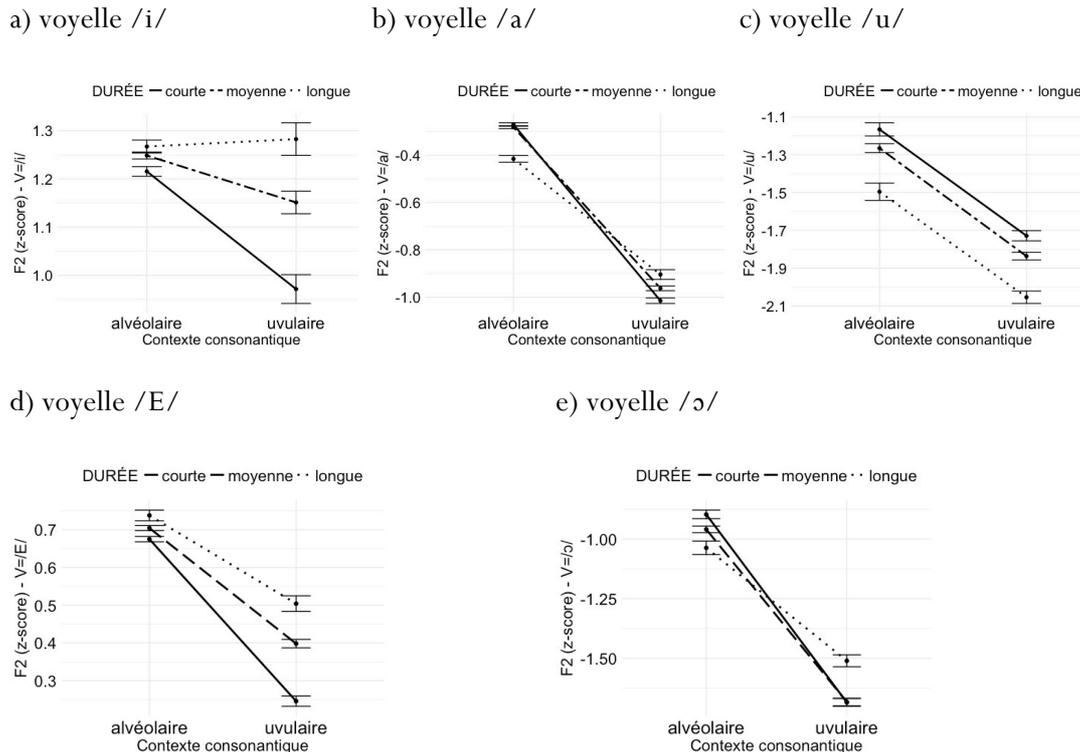


FIGURE 34. F2 (en z-score) des voyelles : a) /i/, b) /a/, c) /u/, d) /E/ (/E/ = /e, ε/), e) /ɔ/, provenant des corpus ESTER et NCCFr, en fonction du contexte consonantique (alvéolaire vs. uvulaire) et de la durée des voyelles (courte = 50 ms, moyenne = 60-80 ms, longue = 90 – 150 ms) selon d'autres études sur le français menées sur des corpus semblables (Gendrot & Adda-Decker, 2005) ou identiques (Audibert et al., 2015).

Une sous-analyse des données de l'étude 3, traitant du style de parole, dont le but était d'identifier précisément comment la coarticulation variait selon la durée des voyelles, a aussi révélé des différences entre voyelles (Guitard-Ivent, 2018). Comme on peut le voir Figure 34, l'étude de 26k exemplaires des voyelles /i, E, a, u, ɔ/ (/E/= /e, ε/), produites par 23 locuteurs masculins, et regroupées en 3 catégories de durée, montre que la durée est effectivement un facteur déterminant pour la coarticulation. Plus les voyelles sont

courtes, plus le degré de coarticulation est fort, à l'exception de la voyelle /u/ dont le degré de coarticulation reste inchangé quelle que soit la durée de la voyelle. Pour cette voyelle les effets contextuels ne se résument pas à des cas de voyelles *undershoot* dus à un manque de temps pour atteindre les cibles articulatoires. Ce résultat interpelle d'autant plus que cette voyelle s'est avérée être la plus variable dans notre première étude effectuée sur le corpus NCCFr de parole conversationnelle.

5.2.2. UNE MODULATION DE LA COARTICULATION DÉPENDANTE DE LA DIRECTIONNALITÉ DE LA COARTICULATION

Les résultats des études 2 et 3 traitent de la modulation de la coarticulation C-à-V par la position prosodique et le style de parole respectivement. Ils indiquent des différences entre la coarticulation anticipatrice VC et persévératrice CV : la modulation de la coarticulation par la position prosodique et le style de parole a lieu principalement pour les cas de coarticulation anticipatrice VC. L'objet de cette section est de discuter de ces résultats et de leurs implications.

L'observation de 7k /a/ en séquences V_1C et CV_1 (où « C » était soit une consonne ALVéolaire, soit une consonne uvulaire) dans deux positions prosodiques IP_i et W_m , montre clairement que les effets du renforcement initial sur le degré de coarticulation C-à-V est fonction du type de séquence étudiée. Alors qu'une réduction de la coarticulation en position IP_i est observée en séquence V_1C , (sur F1 comme sur F2), aucun effet de la position prosodique sur le degré de coarticulation C-à-V n'est ressorti de l'analyse des séquences CV_1 . Ce dernier résultat vient s'ajouter à la lignée des études n'indiquant aucune modification du degré de coarticulation au sein de la syllabe post-frontalière CV_1 (voir Meynadier, 2003 pour le français et Byrd, 2000 pour l'anglais). Dans sa thèse de doctorat par exemple Meynadier (2003), à partir de données Electro-palatographiques, ne montrait aucune modification du chevauchement des gestes de la consonne et de la voyelle au sein de la syllabe post-frontalière (pouvant être /ka/, /ta/ ou /la/) selon la force de la frontière prosodique (Groupe Intonatif, Groupe Accentuel, absence de frontière) en français. Ce résultat indique qu'il n'y a pas de modulation de la coarticulation entre la consonne et la

voyelle suivant la frontière de constituant prosodique (#CV). Par contre, en anglais Cho et al. (2017) ont trouvé une réduction de la coarticulation nasale C-à-V lorsque les séquences CV se trouvaient en position initiale de Groupe Intonatif comparé à la position initiale de mot.

À notre connaissance, aucune étude ne s'était auparavant intéressée aux séquences VC post-frontalières, c'est-à-dire lorsque la voyelle se trouve directement après la frontière du constituant prosodique. Cependant, dans son étude de 2004, portant sur la coarticulation V-à-V en anglais lorsque les voyelles étaient séparées par une frontière prosodique, Cho montrait une réduction du chevauchement inter-gestuel des voyelles transfrontalières en frontière prosodique forte, et rapportait que les modifications étaient surtout portées par la voyelle initiale. Nos résultats sur #VC ainsi que ceux de Cho (2004) sur V#V montrent que les voyelles strictement initiales de constituant prosodique de niveau hiérarchique élevé sont plus résistantes à la coarticulation. L'hypothèse d'une fonction démarcative du renforcement initial est souvent avancée. Il convient de se demander si la plus grande résistance des voyelles initiales à la coarticulation, autre conséquence du renforcement initial, serait une autre façon de marquer la frontière prosodique et contribuerait à la préservation de l'identité phonétique des segments. Par préservation nous entendons préservation de l'identité du segment, préservation vis-à-vis des influences contextuelles. Nos résultats ne nous permettent pas d'en dire plus sur la question. Cette hypothèse devra être examinée plus en détail pour avoir une idée plus précise du sens de cette résistance à la coarticulation des voyelles en position initiale de constituant prosodique.

Plusieurs explications très probablement cumulatives permettent de rendre compte des différences observées entre nos deux types de coarticulation (anticipatrice et persévératrice) :

- a. La position de la voyelle vis-à-vis de la frontière prosodique pourrait donc expliquer les différences observées entre nos deux types de coarticulation VC et CV. En effet, dans les cas de coarticulation au sein de la suite VC, en position I_{Pi}, la voyelle coarticulée se trouve directement à la frontière du constituant prosodique, alors que dans les cas de coarticulation persévératrice CV, c'est la consonne qui se trouve dans

cette position. L'hypothèse de la position du segment vis-à-vis de la frontière prosodique est centrale dans le modèle du π -gesture (Byrd & Saltzman, 2003). Dans le cadre de la Phonologie Articulatoire, ces auteurs ont proposé un modèle dans le but d'expliquer les variations segmentales en frontière de constituants. Au cœur de ce modèle se trouve un geste prosodique abstrait et non défini en termes de constriction articulatoire, le « π -gesture », qui gouvernerait la vitesse des mouvements articulatoires et modulerait l'horloge contrôlant l'activation articulatoire des gestes de constrictions sous son influence. Ainsi, aux frontières de constituants prosodiques, l'horloge est ralentie entraînant dans un même temps un ralentissement des gestes et une réduction de leur chevauchement. Comme le π -gesture est ancré temporellement aux frontières de constituants, son effet est plus fort pour les segments situés en bordure de constituants prosodique et diminue progressivement en s'éloignant du bord.

Dans notre cas, il semblerait que la réduction du chevauchement inter-gestuel due au ralentissement de l'horloge sous l'effet du π -gesture, ait une application strictement locale sur le premier segment dont l'activation sera plus longue causant moins de chevauchement avec le son voisin. Si c'est une séquence #VC, alors le π -gesture affectera la voyelle initiale, et si c'est une séquence #CV, ce dernier s'appliquera sur la consonne. Cela expliquerait pourquoi l'effet de la position prosodique (indépendamment de la coarticulation) soit peu flagrant sur V en séquence #CV en français (Fougeron, 1998). Cependant, il semblerait qu'il existe des différences entre les langues puisque d'autres études indiquent des effets du renforcement initial sur V en séquence #CV en anglais (Cho, 2006) et en coréen (Cho, Lee & Kim, 2011).

- b. Une question qu'il convient de se poser quant à nos résultats de l'étude 2, concernant la modulation de la coarticulation selon la position prosodique, est de savoir si la réduction de la coarticulation des segments en position prosodique IPi n'est pas attribuable à la présence de pause dans le signal. En effet, afin de rendre la procédure d'identification des positions prosodiques facilement automatisable, nous avons pris le parti de ne considérer que les frontières de Groupes Intonatifs

précédées d'une pause. Toutes les séquences extraites en position initiales de Groupe Intonatif étaient donc précédées d'une pause. Or, on sait que la pause est un temps utilisé par le locuteur pour planifier la phrase à venir (voir Fletcher, 2010 citant les travaux de Goldman-Eisler, 1968 ; Butcher, 1981 ; Levelt, 1989). Si les locuteurs utilisent la pause pour planifier ce qu'ils vont dire et donc préparer leurs gestes articulatoires, voire les anticiper lorsque c'est possible (c'est-à-dire lorsque deux gestes n'entrent pas en conflit), on peut imaginer que les accommodations contextuelles entre segments adjacents seront moins importantes. Considérons une séquence CV dont la distance entre C et V est grande, et donc les risques d'accommodations contextuelles importants. Si le locuteur a le temps de planifier sa séquence, il pourra par exemple augmenter sa précision ou sa vitesse articulatoire et atteindre sa cible linguale déjà durant la pause ce qui contribuerait à réduire les accommodations contextuelles.

Il n'y a pas de doute sur l'existence des effets du Renforcement Prosodique Initial au delà des effets de pause. En effet, les effets du Renforcement Initial sont cumulatifs et apparaissent sur des niveaux prosodiques qui ne sont pas démarqués par une pause (Groupe Accentuel, Mot, Syllabe) comme le montre par exemple Fougeron (2001). De plus, le Groupe Intonatif, n'est pas systématiquement précédé ou suivi d'une pause. Or, les effets attribuables à la position initiale de Groupe Intonatif n'ont jamais été remis en question. Cependant, nos données ne nous permettent pas vraiment de dissocier les effets de la position prosodique des effets de la pause. Il serait intéressant d'analyser les effets de la position prosodique sur le degré de coarticulation C-à-V en comparant des séquences en position initiale de Groupes Intonatifs selon que ces derniers soient ou non précédés d'une pause. Ainsi on pourrait voir l'effet réel de la pause et celui de la position prosodique. Pour l'instant, on privilégiera la prudence dans l'interprétation des résultats.

- c. Outre la position de la voyelle vis-à-vis de la frontière de constituant prosodique, le type de coarticulation lui-même pourrait expliquer que seule la coarticulation anticipatrice VC soit modulée par la position prosodique. Dans la lignée de nos résultats de l'étude 2 sur la position prosodique, nos résultats obtenus dans l'étude 3

traitant de la modulation de la coarticulation selon le style de parole (CHAPITRE 4), indiquent que les deux types de coarticulation testés ne sont pas affectés de la même manière par le style de parole. Nous avons montré, à partir de 22k exemplaires des voyelles /i, E, a, u, ə/ (/E/= /e, ε/) en contexte ALV ou UV, que la coarticulation C-à-V était réduite en parole journalistique plus formelle et que cela apparaissait principalement en séquence VC (coarticulation anticipatrice). En effet, seule la voyelle /E/ présente une réduction de la coarticulation en parole formelle de même ampleur en séquence VC et CV. Pour toutes les autres voyelles, la modulation de la coarticulation selon le style de parole est plus forte en séquence VC (/a, ə/) voire n'apparaît qu'en séquence VC (/i, u/).

Le fait que seule la coarticulation anticipatrice soit modulée par la position prosodique et le style de parole, est particulièrement intéressant et renforce l'idée que l'on a affaire à deux types de coarticulations résultant de mécanismes différents : la coarticulation persévératrice CV, due à l'inertie des articulateurs mis en jeu pour l'articulation de C, qui étendent leur influence sur la voyelle V suivante, alors que la coarticulation anticipatrice, elle, peut être aussi due à la planification anticipée de l'articulation de la consonne à venir pendant la voyelle. L'inertie des articulateurs est un phénomène biomécanique que l'on ne peut pas vraiment contrôler, contrairement à l'anticipation qui, elle, est contrôlée, donc modifiable en fonction du style ou de la position prosodique.

En d'autres termes, si on adopte une vision de Phonology Articulatoire : le modèle «oscillateur couplé» permet de rendre compte des différences observées entre les séquences VC et CV. Étant caractérisée par une coordination de synchronisation inter-articulateur stable, il est concevable qu'une séquence CV (couplée en phase) montre moins de sensibilité à la variation du style de parole ou de la position prosodique. Au contraire, sur la base d'un couplage anti-phase, dans une séquence syllabique VC, ou du non couplage V.C (pour rappel l'appartenance syllabique de V et C est variable dans nos données), la coordination des segments est moins stable et peut donc être plus sujette à des sources potentielles de variation comme la position prosodique ou le style de parole (Kelso et al., 1986, Tuller et Kelso, 1991 ; Nam,

Goldstein & Saltzman, 2010). La modulation de la coarticulation par la position prosodique ou le style de parole n'apparaît pas dans le couplage en phase.

5.3. UNE PLUS GRANDE RÉSISTANCE DES VOYELLES À LA COARTICULATION EN PAROLE FORMELLE ET EN IPI : QUELLE(S) FONCTION(S) LINGUISTIQUE(S) ?

Dans cette thèse nous avons montré que les voyelles, quelles qu'elles soient, en position initiale de Groupe Intonatif et celles produites en parole journalistique formelle sont plus résistantes à la coarticulation. Cependant, l'observation de plusieurs résultats nous conduit à émettre l'hypothèse d'une fonction linguistique différente de cette plus grande résistance à la coarticulation des voyelles selon qu'elle dépende de la position prosodique (IPI) ou du style de parole (journalistique).

Les effets du style de parole sont testés sur des locuteurs différents alors que ceux de la prosodie impliquent les mêmes locuteurs. Nous ne pouvons pas savoir dans quelle mesure cela a un impact sur les résultats divergents entre nos deux études. Dans la discussion qui va suivre, nous ne discuterons pas du locuteur, mais nous ne pouvons pas écarter l'implication de ce facteur dans l'obtention des différences observées dans nos deux études.

		/i/	/e/	/E/	/a/	/u/	/ɔ/
PP	CV						
	F1				IPI > Wm		
	F2				ns.		
	VC						
	F1	ns.	ns.		IPI > Wm		ns.
	F2	IPI > Wm	IPI > Wm		ns.		ns.
STYLE	CV & VC						
	F2	J > C		ns.	J > C	J > C	ns.

TABLE 39. Résumé des résultats de l'effet de la position prosodique (PP) et du style de parole (STYLE) sur l'acoustique des voyelles testées /i, e, a, ɔ/ pour PP (IPI = position initiale de groupe intonatif, Wm = position interne de mot) et /i, ε, a, u, ɔ/ (/ε/ = /e, E/) pour le STYLE (J = parole journalistique, C = parole conversationnelle, J > C indique que le F2 de la voyelle en question est plus élevé en parole journalistique).

1. Nos données indiquent que la plus forte résistance des voyelles à la coarticulation en position prosodique IPi pourrait contribuer à marquer la frontière prosodique et ainsi à aider les auditeurs à analyser le signal de parole en facilitant la segmentation lexicale.

a. Dans nos données, l'effet de la position prosodique sur l'acoustique des voyelles dépend de leurs propriétés phonétiques. Les résultats indiquent qu'en position IPi, les caractéristiques distinctives des voyelles sont renforcées comme on peut le voir Table 39 pannels du haut :

- Seul le premier formant de la voyelle /a/ est plus élevé en position IPi en séquence VC comme CV. Les valeurs de F1 des autres voyelles ne varient pas selon la position prosodique. Ainsi, l'élévation du F1 de /a/ en position IPi reflèterait plutôt un renforcement des caractéristiques distinctives de cette voyelle qu'une maximisation du contraste de sonorité (syntagmatique) CV dans cette position prosodique.
- le F2 des voyelles antérieures /i/ et /e/ est plus élevé en position IPi comparé à la position Wm. Ces voyelles sont plus antérieures lorsqu'elles occupent la position initiale de Groupe Intonatif.

b. Le contexte sur lequel s'effectue la réduction des effets coarticulatoires en position IPi dépend aussi des caractéristiques de la voyelle :

- les voyelles les plus fermées /i, e/ montrent une réduction de la coarticulation sur le contexte uvulaire dont les conséquences acoustiques sont d'élever F1. La voyelle ouverte /a/ montre une réduction de la coarticulation dans les deux contextes, alvéolaire et uvulaire. Aucune différence significative de la coarticulation dans les deux positions prosodiques n'a été révélée pour le F1 de /ɔ/, mais on observe une tendance à la réduction de la coarticulation en position IPi et celle-ci concerne le contexte alvéolaire (abaissant le F1 des voyelles).
- la réduction de la coarticulation en position IPi se fait sur le contexte uvulaire, abaissant F2, pour les voyelles antérieures /i, e/ caractérisées par un F2 maximalelement élevé, alors que pour la voyelle postérieure /ɔ/ la réduction de la coarticulation en position IPi concerne les deux contextes alvéolaire et uvulaire.

Pour la voyelle acoustiquement centrale, /a/ on observe une réduction de la coarticulation en contexte uvulaire.

→ Pour les voyelles antérieures /i, e/ la réduction de la coarticulation en position IPI opère uniquement sur le contexte uvulaire avec la voyelle en conflit (sur le contexte dont la trajectoire articulaire s'oppose à celle de la voyelle). Ce résultat montre que la modulation de la coarticulation par la position prosodique est plutôt locale. Elle tend à diminuer les effets coarticulatoires du contexte pouvant diminuer la distinctivité des voyelles. Cependant, cela ne se retrouve pas pour les voyelles plus ouvertes /ɔ, a/.

Ces résultats pris ensemble suggèrent que la réduction de la coarticulation permettrait le renforcement des caractéristiques phonétiques des voyelles dans certaines positions prosodiques. La plus grande résistance des voyelles à la coarticulation en position IPI s'ajouterait à un renforcement des contrastes syntagmatiques et paradigmatiques en position prosodique forte pour marquer la frontière.

2. Nos résultats concernant la modulation de la coarticulation selon le style de parole semblent, eux, indiquer que la plus grande résistance à la coarticulation des voyelles produites en parole formelle serait plutôt la conséquence d'une hyperarticulation en parole journalistique. La réduction de la coarticulation en parole formelle serait un moyen d'augmenter l'intelligibilité de la parole et répondrait à une adaptation du locuteur à la situation de communication.

- a. Nos résultats concernant l'effet du style de parole sur l'acoustique des voyelles montre une élévation du second formant des voyelles /i, u, a/ en parole journalistique. Le style de parole affecte l'acoustique de l'ensemble du système vocalique dans la même direction.

- b. Contrairement à la modulation de la coarticulation par la position prosodique, le contexte sur lequel s'effectue la réduction des effets coarticulatoires en parole formelle, ne dépend pas de la voyelle :
- pour la voyelle /i/ et les voyelles /E, ə/ en séquence VC (pour rappel nous n'avons pas trouvé de modulation de la coarticulation de la voyelle /i/ en séquence CV), la réduction de la coarticulation en parole journalistique se fait à la fois sur le contexte alvéolaire et le contexte uvulaire.
 - pour la voyelle /a/, la voyelle /u/ en séquence VC (aucun effet ne ressort en CV) et les voyelles /E, ə/ en séquence CV la réduction de la coarticulation opère sur le contexte uvulaire.
- L'effet du style sur la coarticulation est plutôt global dans le sens où la réduction de la coarticulation ne se fait pas forcément sur le contexte consonantique en conflit avec la voyelle. Il peut s'opérer tant sur les deux contextes, que sur le seul contexte dont la trajectoire peut être ou non comparable avec celle de la voyelle.

Ces résultats montrent que le style de parole affecte de la même manière l'acoustique de l'ensemble du système vocalique contrairement à la position prosodique dont les effets dépendent des caractéristiques phonétiques des voyelles. De plus, la modulation de la coarticulation par le style de parole se fait plutôt de manière globale alors que celle observée sous l'effet de la position prosodique se fait principalement sur le contexte dont la trajectoire articulaire s'oppose à celle de la voyelle, et est, en ce sens, plutôt locale.

La modulation de la coarticulation sous l'effet du style parole indépendante des caractéristiques phonétiques des voyelles serait le reflet de l'adaptation du locuteur, de l'intelligibilité de sa parole selon les contraintes relatives à la situation de communication.

5.4. RÉFLEXION SUR LES POSSIBLES IMPLICATIONS DE CERTAINS RÉSULTATS SUR LES CHANGEMENTS DE SONS

Rien n'est permanent sauf le changement
Héraclite d'Éphèse

Les variations phonétiques dans la parole en synchronie n'ont pas toujours un lien direct avec les changements de sons diachroniques dans les langues, mais la relation fondamentale entre les deux ne peut être négligée :

« Virtually all writers accept that sound change is happening at all times including the present. Where they disagree is whether or not sound change can be detected as it happens. I claim that although sound change may not be easily detected 'on the hoof', it and its phonetic pre-conditions can be observed in the lab. The basis for this claim is very simple: a detailed examination of speech phenomena, both in production and perception, reveals extensive variation ("change") that parallels sound change. The points of similarity are so close, I maintain, that one must conclude that somehow one is dealing with the same thing. There are differences between synchronic variation and sound change but these can be accounted for without denying the essential link between the two. » (Ohala, 1990 : 2).

De nombreux travaux, à partir d'expérimentations phonétiques et des connaissances portant sur les tendances universelles d'évolution des langues, essaient de trouver la cause de différents changements de sons. Par exemple, Recasens (2014) s'est particulièrement intéressé aux effets que pouvaient avoir la coarticulation sur les changements de sons observés dans les langues romanes. L'auteur montre comment selon les caractéristiques articulatoires des voyelles et des consonnes et la directionnalité de la coarticulation, les variations contextuelles ont pu déboucher sur des changements de sons tels que les élisions, insertions et substitutions (assimilations et dissimilations) dans cette famille linguistique.

Ohala (1981, 1993) propose un modèle permettant d'expliquer comment des variations phonétiques répondant à des contraintes physiologiques comme la coarticulation, peuvent conduire à des changements linguistiques, et donc avoir des conséquences sur la phonologie des langues, à partir de la relation locuteur-auditeur et d'erreurs de perception. Selon cet auteur, les changements de sons arriveraient lorsque l'auditeur fait une mauvaise analyse du signal de parole et qu'il ne parvient pas à compenser perceptivement les effets de la

coarticulation. L'auditeur peut associer par erreur une ou des caractéristiques phonétiques d'un phonème à celles du phonème voisin du fait de la non séquentialité de la parole.

Un des postulats de ce modèle est que les auditeurs font une sorte de normalisation perceptive permettant de passer outre les variations contextuelles omniprésentes dans la parole. Par exemple, un /i/ en contexte uvulaire aura un F1 plus élevé que ce que l'on peut attendre d'un /i/, mais l'auditeur est capable de passer outre cette variation qui s'explique par le fait que l'élévation du premier formant est due à l'effet de la consonne uvulaire adjacente. Ainsi, l'auditeur ne confondra pas le /i/ en contexte uvulaire avec une voyelle plus ouverte telle que la voyelle /e/. Le problème survient lorsque l'auditeur n'arrive pas à compenser la variation contextuelle. L'auditeur entend un /e/ à la place d'un /i/. Sa mauvaise perception le conduit à modifier ses propres productions qui pourront être imitées par d'autres auditeurs-locuteurs. À la base des changements de sons, serait donc la décontextualisation des variations coarticulatoires, en produisant un /e/ à la place d'un /i/ dans d'autres contextes n'impliquant pas une élévation du premier formant comme dans *mite* [mit].

Les résultats de Harrington et al. (2008) sur l'antériorisation du /u/ donnent du poids au modèle développé par Ohala : à partir d'une étude mêlant production et perception en anglais britannique, connu pour être affecté par un processus d'antériorisation de voyelles postérieures fermées (en cours depuis une cinquantaine d'années), les auteurs ont montré que les jeunes locuteurs produisaient des /u/ plus antérieurs que les locuteurs âgés et que l'effet du contexte sur l'antériorisation de /u/ était moins importants pour les premiers. De plus, les locuteurs jeunes percevaient plus souvent le /u/ comme un /i/ et ils parvenaient moins facilement à compenser perceptivement l'antériorisation de /u/ due aux effets contextuels que les sujets âgés. Ainsi les auteurs concluaient que ce type de changement de son pouvait être lié à une mauvaise compensation des effets coarticulatoires des auditeurs.

Plusieurs résultats observés dans cette thèse, et notamment dans notre première étude, montrent un pattern de variations particulier de la voyelle /u/, dont la tendance universelle est à l'antériorisation (Labov, 1994), et plus largement de l'ensemble des

voyelles postérieures qu'il nous semble intéressant de discuter dans une perspective diachronique :

- a. Les voyelles postérieures présentent des déplacements de F2 dus au contexte consonantique plus importants que les autres voyelles.
- b. Nous avons aussi montré que ces voyelles présentaient les plus mauvais taux de reconnaissance correcte (38 %) obtenus à partir de notre classification automatique. Ce résultat indique une plus faible distinctivité acoustique de ces voyelles postérieures.
- c. Notre classification révèle aussi un taux important de confusions de /ɔ/ en /x/ en contexte alvéolaire. Or, un processus en cours d'antériorisation du /ɔ/ a déjà été décrit en français (Martinet, 1958 ; Boula de Mareuil et al., 2010). Les derniers auteurs indiquaient aussi que les cas d'antériorisations observés avaient principalement lieu lorsque la voyelle apparaissait en contexte alvéolaire renforçant l'idée d'un processus d'origine coarticulatoire.
- d. La voyelle /u/ montre une plus forte dispersion acoustique intra-catégorie que les autres voyelles sur la mesure dimensionnelle, ainsi que sur chaque formant, pris séparément. De plus, contrairement à d'autres voyelles (telle que le /i/), sa dispersion n'est pas attribuable à certains contextes consonantiques. La dispersion du /u/ reste la même quel que soit le contexte dans lequel elle se trouve (alvéolaire, vélaire ou uvulaire).
- e. Les résultats d'une sous-analyse de l'effet de la durée sur le degré de coarticulation C-à-V rapportés page 153, montrent que contrairement aux autres voyelles, le degré de coarticulation de la voyelle /u/ reste le même quelle que soit sa durée. Ce résultat nous a conduit à dire que pour cette voyelle les effets contextuels ne se résument pas à des cas de voyelles *undershoot* dus à un manque de temps pour atteindre les cibles articulatoires.

Nos résultats montrent un pattern de variations particulier des voyelles postérieures. On peut se demander si cette asymétrie entre voyelles antérieures et voyelles postérieures vis-à-vis de la variation (voyelles postérieures plus variables que les autres) peut être à l'origine

de la préférence des changements de sons pour l'antériorisation des voyelles postérieures fermées par rapport à la postériorisation des voyelles antérieures fermées, phénomène beaucoup plus marginal. Si cette tendance universelle concerne les voyelles postérieures fermées, nos résultats nous conduisent à nous interroger sur l'ensemble des voyelles postérieures.

Les résultats *a*, et *c* présentés ci-dessus, vont dans le sens d'une origine coarticulatoire du phénomène : les effets coarticulatoires sont plus importants sur les voyelles postérieures que pour les autres voyelles et peuvent causer plus de confusions des voyelles postérieures avec des voyelles antérieures, comme l'illustre le fort taux de confusions de /ɔ/ en /x/. Ce résultat est d'autant plus intéressant qu'un changement en cours a été décrit depuis déjà une soixantaine d'années (Martinet, 1958) et qu'une étude récente montre que les antériorisations sont principalement observées lorsque la voyelle /ɔ/ se trouve en contexte antériorisant (Boula de Mareuil et al., 2010).

Cependant, les résultats *b* et *d*, indiquent que cette plus forte variabilité des voyelles postérieures et notamment /u/ va au-delà des variations contextuelles. Le plus fort taux de confusions des voyelles postérieures obtenu à partir de notre classification automatique pourrait s'expliquer parce que l'espace acoustique est plus réduit à l'arrière de la cavité buccale, entraînant plus de confusions entre les voyelles postérieures comme nous l'avons observé entre /o/ et /u/ principalement. Ceci étant, cela n'explique pas la forte dispersion acoustique de la voyelle /u/ dans tous les contextes testés. De plus amples analyses seraient nécessaires pour trouver l'origine de cette forte variabilité du /u/ dans nos données. Cependant, l'observation de ce résultat en parallèle avec le résultat *e* suggère une prédisposition de cette voyelle /u/ à subir des changements linguistiques. En plus d'être fortement variable, son degré de coarticulation avec la consonne environnante reste inchangé quelle que soit sa durée contrairement aux autres voyelles. Cette particularité de la voyelle /u/ pourrait venir s'ajouter aux causes coarticulatoires d'antériorisation des voyelles postérieures fermées.

L'observation de la distribution combinatoire des sons du français suggère une distribution particulière des consonnes alvéolaires (antériorisantes) dans cette langue, pouvant aussi

avoir des implications sur la préférence des langues à l'antériorisation des voyelles postérieures par rapport à la postériorisation des voyelles antérieures. Adda-Decker (2006) à partir de l'analyse de la distribution des voyelles du français en contexte menée à partir de 3 grands corpus oraux du français montre que :

- a. le contexte alvéolaire est le plus représenté en tant que contexte gauche (CV) et apparaît dans un plus grand nombre de combinaisons CVC en contexte droit ce qui en fait le contexte « le plus varié » dans cette position.
- b. concernant la combinaison consonne voyelle, lorsque la consonne de gauche est une consonne alvéolaire, on aura plus de probabilité de trouver aussi une consonne alvéolaire en contexte droit pour les voyelles antérieures comme postérieures. Alors que, lorsque la consonne de gauche est un /R/ (contexte postériorisant), les voyelles antérieures /i, e, ε/ se trouveront plus souvent suivies d'une consonne alvéolaire. Ainsi, les voyelles postérieures auront plus de risque de se trouver dans un contexte antériorisant que les voyelles antérieures en contexte postériorisant comme on peut le voir Figure 35.

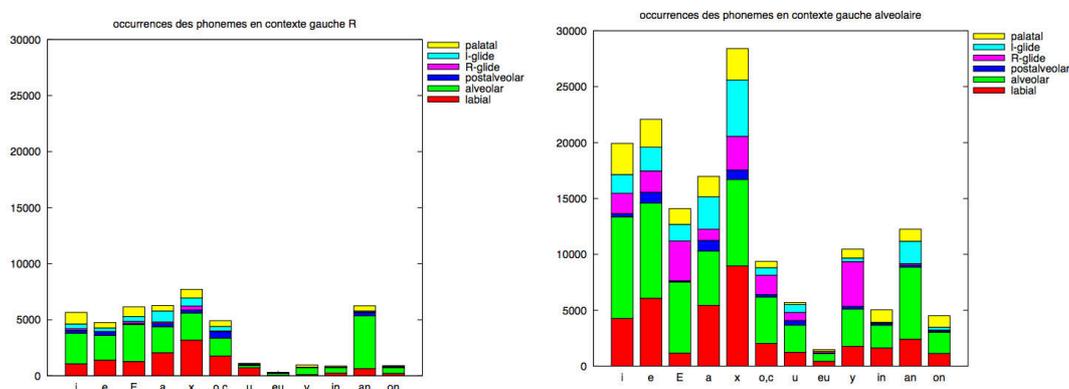


FIGURE 35. Fréquences d'occurrence des voyelles en contextes gauche et droit de labiales, d'alvéolaires, de postalvéolaires, de /R/, de la liquide /l/ et de palato-vélaires mesurées sur le corpus journalistique. a) : le contexte gauche correspond à la consonne /R/ ; b) : le contexte gauche correspond aux alvéolaires. Pour chaque contexte gauche les histogrammes indiquent la distribution des voyelles sur un des trois corpus étudiés en fonction des 6 classes de consonnes en contexte droit. Graphiques tirés de Adda-Decker (2006).

Ces observations faites à partir de données brutes font écho à celles de Harrington (2007) montrant qu'en anglais britannique du sud, 70% des /u/ présents dans la base de données CELEX (Baayen, Dijkstra & Schreuder, 1997) sont précédés d'une consonne ayant un locus élevé. Selon Harrington (2008) cela pourrait aussi expliquer les cas d'antériorisation du /u/ dans cette langue sur la base d'un « *episodic model of perception and production in which the frequency and probability with which /u/ occurs in a certain context is the trigger for sound change.* » (Harrington, 2008 : 2833)

Le fait que les contextes antériorisants soient plus florissants en français que les contextes postériorisants pourrait indiquer une prédisposition de la langue à des phénomènes d'antériorisation des voyelles postérieures. Plus généralement, on peut se demander si cela est à l'origine de la préférence des langues pour les phénomènes d'antériorisation que de postériorisation. Pour cela, il serait intéressant de rechercher dans d'autres langues ce type de distribution, afin d'établir des tendances linguistiques. Une question à laquelle il serait particulièrement important de répondre est de savoir quelle est la probabilité de trouver un /u/ en contexte antériorisant dans les langues affectées par un processus d'antériorisation de voyelles postérieures fermées.

À l'heure actuelle nos résultats ne nous permettent pas de dire si les variations observées dans cette thèse peuvent avoir des répercussions en diachronie. On ne sait d'ailleurs pas si elles ont des conséquences perceptives. Cependant, le pattern de variations particulier des voyelles postérieures observé sur un grand corpus fait écho aux travaux traitant de l'antériorisation des voyelles postérieures. Ces résultats offrent des pistes de réflexions intéressantes dans une perspective d'évolution de la langue, qui devront faire l'objet de plus amples investigations.

CONCLUSION

Le recours à de grands corpus de parole pour l'étude de la variation offre des perspectives prometteuses pour l'étude de la variation dans la parole. Notamment, parce qu'ils permettent d'aller tester des théories de la parole à grande échelle à partir de données enregistrées dans des conditions plus écologiques. Nous avons par exemple vu que la voyelle /i/ était bien plus stable sur la dimension F2 comme le prédit la théorie quantique alors que les voyelles postérieures sont, elles, plus variables que les autres voyelles en accord avec le modèle DAC.

Plus précisément, nos résultats montrent qu'un même phénomène de variation, la coarticulation dans notre cas, dépend des influences de différents facteurs répondant à des contraintes diverses. Avoir la possibilité d'étudier un même phénomène de variation sous différents angles permet de se faire une image plus complète de ce dernier et de mieux comprendre le signal de parole. À partir des théories existantes, différentes hypothèses ont pu être testées et à partir de cela nous avons pu émettre des suggestions quant aux fonctions linguistiques de la modulation de la coarticulation par les facteurs testés. Si à l'heure actuelle ce ne sont que des hypothèses qui mériteront d'être étudiées plus en profondeur, ce genre d'étude permet d'appréhender la variation différemment, et de soulever de nouvelles questions à partir des théories existantes, notamment en lien avec l'évolution de la langue.

Les études sur grands corpus ne peuvent se passer des expérimentations menées en condition de laboratoire. Il est plus important que jamais de rapprocher les deux approches pour tirer au mieux profit des nouvelles technologies. Les données expérimentales sont fondamentales pour l'obtention de données articulatoires et permettent d'avoir recours à des mesures très précises permettant une analyse plus fine du signal de parole sur des données équilibrées. De plus, l'étude sur grand corpus offre la possibilité de tester à grande échelle des résultats obtenus en condition de laboratoire sur de la parole plus naturelle.

Cependant, le recours à ce genre de corpus ne permet pas toujours une analyse si fine du signal de parole (par exemple, nous avons préféré des mesures statiques de formants plutôt que des équations du locus pour pallier les éventuels problèmes de segmentation) et le déséquilibre dans les données peut s'avérer problématique dès que l'on commence à contrôler plusieurs facteurs.

BIBLIOGRAPHIE

- ADDA-DECKER, M. (2006). De la reconnaissance automatique de la parole à l'analyse linguistique de corpus oraux. *Actes des Journées d'étude sur la Parole*, 389-400.
- AGWUELE, A., SUSSMAN, H. M., & LINDBLOM, B. (2008). The effect of speaking rate on consonant vowel coarticulation. *Phonetica*, 65(4), 194-209.
- ALI, L., GALLAGHER, T., GOLDSTEIN, J., & DANILOFF, R. (1971). Perception of coarticulated nasality. *The Journal of the Acoustical Society of America*, 49(2B), 538-540.
- ASSMANN, P. F. (1995). The role of formant transitions in the perception of concurrent vowels. *The Journal of the Acoustical Society of America*, 97, 575-584.
- AUDIBERT, N., & FOUGERON, C. (2012). Distorsions de l'espace vocalique : quelles mesures ? Application à la dysarthrie. *Actes des 29èmes Journées d'Etudes sur la Parole/JEP-TALN-RECITAL 2012*, 217-224.
- AUDIBERT, N., FOUGERON, C., GENDROT, C., & ADDA-DECKER, M. (2015). Duration- vs. style-dependent variation: a multiparametric investigation. *Proceedings of the International Conference of Phonetics Science 2015 (ICPhS)*, Glasgow. Retrieved from: <https://www.internationalphoneticassociation.org/icphsproceedings/ICPhS2015/Papers/ICPHS0753.pdf>
- BAAYEN, R., DIJKSTRA, T., & SCHREUDER, R. (1997). The CELEX Lexical Database. *Linguistic Data Consortium*, University of Pennsylvania, Philadelphia.
- BARD, E. G., & ANDERSON, A. H. (1983). The unintelligibility of speech to children. *Journal of Child Language*, 10(2), 265-292.
- BARTON, K. (2014). MuMIn: multi-model inference, R Package Version 1.9. 13. Available at: <http://cran.r-project.org/web/packages/MuMIn/>. Last accessed, 10.
- BATES, D., MAECHLER, M., BOLKER, B., & WALKER, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1-48.
- BEDDOR, P. S. (2009). A coarticulatory path to sound change. *Language*, 85(4), 785-821.
- BEDDOR, P. S., HARNSBERGER, J. D., & LINDEMANN, S. (2002). Language-specific patterns of vowel-to-vowel coarticulation: acoustic structures and their perceptual correlates. *Journal of Phonetics*, 30 (4), 591-627.
- BELL BERTI, F., & HARRIS, S. K. (1981). A temporal model of speech production. *Phonetica* 30, 41-55.

- BLADON, R. A., & AL-BAMERNI, A. (1976). Coarticulation resistance in English /l/. *Journal of Phonetics*, 4, 137-150.
- BOERSMA, P., & WEENINK, D. (2014). Praat: doing phonetics by computer (Version 5.4.04) [Computer program]. Retrieved from: <http://www.praat.org/>
- BOLINGER, D. (1989). *Intonation and its uses: Melody in grammar and discourse*. Stanford: Stanford University Press.
- BOMBIEN, L., MOOSHAMMER, C., HOOLE, P., RATHCKE, T., & KÜHNERT, B. (2007). Articulatory strengthening in initial German /kl/ clusters under prosodic variation. *Proceedings of the International Conference of Phonetics Science (ICPhS)*, 457-460. Saarbrücken.
- BOULA DE MAREÛIL, P. (2014). Qu'est-ce qu'un (phono) style? *Cahiers de linguistique française*, 31, 9-19.
- BOULA DE MAREÛIL, P., ADDA-DECKER, M., & WOEHLING, C. (2010). Antériorisation/aperture des voyelles /O/~o/ en français du Nord et du Sud. *Actes des Journées d'Étude sur la Parole*, 81-84, Mons.
- BRADLOW, A. R. (2002). Confluent talker-and listener-oriented forces in clear speech production. *Laboratory phonology*, 7, 241-272.
- BROWMAN, C. P., & GOLDSTEIN, L. (1990). Tiers in articulatory phonology, with some implications for casual speech. *Laboratory phonology, 1: Between the grammar and physics of speech*, 341-376.
- BROWMAN, C. P., & GOLDSTEIN, L. (1992). Articulatory phonology: An overview. *Phonetica*, 49(3-4), 155-180.
- BUTCHER, A. (1981). Aspects of the speech pause: Phonetics correlates and communicative functions. *Aipuk* (Arbeitsberichte Institut für Phonetik Kiel), 15.
- BUTCHER, A., & WEILHER, E. (1976). An electropalatographic investigation of coarticulation in vcv sequences. *Journal of Phonetics*, 4, 59-74.
- BYBEE, J. L. (2001) *Phonology and language use*. Cambridge, UK: Cambridge University Press.
- BYRD, D. (2000). Articulatory vowel lengthening and coordination at phrasal junctures. *Phonetica*, 57(1), 3-16.
- BYRD, D., & CHOI, S. (2010). At the juncture of prosody, phonology, and phonetics—The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory Phonology*, 10, 31-59.

- BYRD, D., KAUN, A., NARAYANAN, S., & SALTZMAN, E. (2000). Phrasal signatures in articulation. *Laboratory phonology*, 5, 70-87.
- BYRD D., & SALTZMAN, E. (2003). The elastic phrase: Modeling the dynamics of boundary-adjacent lengthening. *Journal of Phonetics*, 31, 149–180.
- CARNEY, P. J., & MOLL, K. L. (1971). A cinefluorographic investigation of fricative consonant-vowel coarticulation. *Phonetica*, 23(4), 193-202.
- CHAFE, W. L. (1974). Language and consciousness. *Language*, 111-133.
- CHO, T. (1998). Domain-initial strengthening in the prosodic hierarchy: an EPG study. *Proceedings of 11th International Conference on Korean Linguistics (ICKL 11)*, University of Hawaii at Manoa.
- CHO, T. (2001). Effects of morpheme boundaries on intergestural timing: evidence from Korean. *Phonetica*, 58, 129-162.
- CHO, T. (2004). Prosodically conditioned strengthening and vowel-to-vowel coarticulation in English. *Journal of Phonetics*, 32(2), 141-176.
- CHO, T. (2006). Manifestation of prosodic structure in articulation: Evidence from lip kinematics in English. *Laboratory phonology*, 8, 519-548.
- CHO, T. (2011). Laboratory phonology. *The continuum companion to phonology*, 343-368.
- CHO, T., & KEATING, P. A. (2001). Articulatory and acoustic studies on domain-initial strengthening in Korean. *Journal of Phonetics*, 29(2), 155-190.
- CHO, T., & KEATING, P. (2009). Effects of initial position versus prominence in English. *Journal of Phonetics*, 37(4), 466-485.
- CHO, T., KIM, D., & KIM, S. (2017). Prosodically-conditioned fine-tuning of coarticulatory vowel nasalization in English. *Journal of Phonetics*, 64, 71-89.
- CHO, T., LEE, Y., & KIM, S. (2011). Communicatively driven versus prosodically driven hyper-articulation in Korean. *Journal of Phonetics*, 39(3), 344–361.
- CHO, T., LEE, Y., & KIM, S. (2014). Prosodic strengthening on the /s/-stop cluster and the phonetic implementation of an allophonic rule in English. *Journal of Phonetics*, 46, 128-146.
- CHO, T., YOON, Y., & KIM, S. (2014). Effects of prosodic boundary and syllable structure on the temporal realization of CV gestures. *Journal of Phonetics*, 44, 96-109.

- COHN, A. (1990). *Phonetic and Phonological Rules of Nasalization*. (Thèse de Doctorat, UCLA aussi publiée comme *UCLA Working Papers in Phonetics* 76).
- CLUMECK, H. (1976). Patterns of soft palate movement in six languages. *Journal of Phonetics*, 4, 337-351.
- COLE, J., KIM, H., CHOI, H., & HASEGAWA-JOHNSON, M. (2007). Prosodic effects on acoustic cues to stop voicing and place of articulation: Evidence from Radio News speech. *Journal of Phonetics*, 35, 180-209.
- CUNNINGS, I. (2012). An overview of mixed-effects statistical models for second language researchers. *Second Language Research*, 28(3), 369-382.
- DELAIS-ROUSSARIE, E., POST, B., AVANZI, M., BUTHKE, C., DI CRISTO, A., FELDHAUSEN, I., ... & SICHEL-BAZIN, R. (2015). Intonational phonology of French: Developing a ToBI system for French. In Frota, S., & Prieto, P. (Eds), *Intonation in Romance*, Oxford University Press.
- DICANIO, C., NAM, H., AMITH, J. D., GARCÍA, R. C., & WHALEN, D.H. (2015). Vowel variability in elicited versus spontaneous speech: Evidence from Mixtec. *Journal of Phonetics* 48, 45-59.
- DUEZ, D. (1992). Second formant locus-nucleus patterns: An investigation of spontaneous French speech. *Speech Communication*, 11(4-5), 417-427.
- ECKERT, P. (2003). Language and adolescent peer groups. *Journal of Language and Social Psychology* 22(1), 112-118.
- ECKERT, P. (2005). Variation, convention and social meaning. Communication présentée à *Annual Meeting of the Linguistic Society of America*, 1-35, Oakland.
- ESKÉNAZI, M. (1993). Trends in speaking styles research. *3rd European Conference on Speech Communication and Technology*, 501-509, Berlin.
- FANT, G. (1960). *Acoustic Theory of Speech Production*. The Hague : Mouton.
- FANT, G. (1973). *Speech sounds and features*.
- FARNETANI, E. (1990). VCV lingual coarticulation and its spatiotemporal domain. In *Speech production and speech modeling*, 93-130. Springer, Dordrecht.
- FARNETANI, E., & RECASENS, D. (2010). Coarticulation and connected speech processes. In William J. Hardcastle, John Laver & Fiona E. Gibbon (Eds), *The Handbook of Phonetic Sciences*, 316-32. Chichester : Wiley-Blackwell.

- FERGUSON, S. H., & KEWLEY-PORT, D. (2002). Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners. *The Journal of the Acoustical Society of America*, 112(1), 259-271.
- FERGUSON, S. H., KEWLEY-PORT, D. (2007). Talker Differences in Clear and Conversational Speech: Acoustic Characteristics of Vowels. *Journal of Speech, Language, and Hearing Research*, 50, 1241-1255.
- FLETCHER, J. (2010). The prosody of speech: Timing and rhythm. In William J. Hardcastle, John Laver & Fiona E. Gibbon (Eds), *The handbook of phonetic sciences*, 521-602. Chichester : Wiley-Blackwell.
- FOUGERON, C. (1998). *Variations articulatoires en début de constituants prosodiques de différents niveaux en français*. (Thèse de doctorat, ANRT Université Sorbonne Nouvelle - Paris 3).
- FOUGERON, C. (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics*, 29(2), 109-135.
- FOUGERON, C., & KEATING, P. A. (1997). Articulatory strengthening at edges of prosodic domains. *The journal of the acoustical society of America*, 101(6), 3728-3740.
- FOWLER, C. A. (1981). Production and perception of coarticulation among stress-unstressed vowels. *Journal of Speech and Hearing Research*, 24, 127-139.
- FOWLER, C. A. (1984). Segmentation of coarticulated speech in perception. *Perception & Psychophysics*, 36(4), 359-368.
- FOWLER, C. A., & BRANCAZIO, L. (2000). Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech*, 43(1), 1-41.
- FOWLER, C. A., & HOUSUM, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of memory and language*, 26(5), 489-504.
- FOX, J., WEISBERG, S., FRIENDLY, M., HONG, J., ANDERSEN, R., FIRTH, D., ... & FOX, M. J. (2018). Package 'effects'.
- GALLIANO, S., GEOFFROIS, E., MOSTEFA, D., CHOUKRI, K., BONASTRE, J.-F., & GRAVIER, G. (2005). ESTER Phase II Evaluation Campaign for the Rich Transcription of French Broadcast News. *Proceedings of Interspeech*, Lisbonne, Portugal.
- GAUVAIN, J.L., LAMEL, L. AND ADDA, G. (2002). The Limsi Broadcast News Transcription System. *Speech Communication*, 37(1-2), 89-108.

- GAY, T. (1978). Effect of speaking rate on vowel formant movements. *The journal of the Acoustical society of America*, 63(1), 223-230.
- GENDROT, C., & ADDA-DECKER, M. (2005). Impact of duration on F1/F2 formant values of oral vowels: an automatic analysis of large broadcast news corpora in French and German. *Proceedings of Eurospeech*, 2453-2456, Lisbonne, Portugal.
- GENDROT, C., & GERDES, K. (2011). Prosodic hierarchy and spectral realization of vowels in French. In H. -Y. Yoo, & E. Delais-Roussarie (Eds.), IDP 2009, 191-205. Paris, France.
- GENDROT, C., GERDES, K., & ADDA-DECKER, M. (2011). Impact of prosodic position on vocalic space in German and French. *Proceedings of International Conference of Phonetics Science*, 731-734, Hong-Kong, Chine.
- GEORGETON, L. (2014). *Renforcement des voyelles orales du français en position initiale de constituants prosodiques : interaction avec les contrastes phonologiques*. (Thèse de doctorat Université Sorbonne Nouvelle - Paris 3).
- GEORGETON, L., & FOUGERON, C. (2014). Domain-initial strengthening on French vowels and phonological contrasts: evidence from lip articulation and spectral variation. *Journal of Phonetics*, 44, 83-95.
- GOFFMAN, L., SMITH, A., HEISLER, L., & HO, M. (2008). The breadth of coarticulatory units in children and adults. *Journal of Speech, Language and Hearing Research*, 51, 1242-1437.
- GOLDMAN-EISLER, F. (1968). *Psycholinguistics: Experiments in Spontaneous Speech*. London: Academic Press.
- GRAVIER, G., BONASTRE, J-F., GEOFFROIS, E., GALLIANO, S., TAIT, K. MC., & CHOUKRI, K. (2006). Corpus description of the ESTER evaluation campaign for the rich transcription of French broadcast news. *Proceedings of European Conference on Speech Communication and Technology*, 139-142.
- GUIARD-IVENT, F. (2018). Effets de la durée vocalique et du locuteur sur le degré de coarticulation C-à-V en français : étude sur grands corpus. *Actes des 32^{ème} Journées d'Étude sur la Parole (JEP)*, Aix-en-provence.
- HARDCASTLE, W. J. (1985). Some phonetic and syntactic constraints on lingual coarticulation during /kl/sequences. *Speech Communication*, 4(1-3), 247-263.
- HARRINGTON, J. (2007). Evidence for a relationship between synchronic variability and diachronic change in the Queen's annual Christmas broadcasts. In J. Cole and J. Hualde (Eds), *Laboratory Phonology* 9, 125-143, Mouton de Gruyter, Berlin.

- HARRINGTON, J., KLEBER, F., & REUBOLD, U. (2008). Compensation for Coarticulation, /u/-Fronting, and Sound Change in Standard Southern British: An acoustic and perceptual study. *Journal of the Acoustical Society of America*, 123, 2825-2835.
- HAWKINS, S., & WARREN, P. (1994). Phonetic influences on the intelligibility of conversational speech. *Journal of Phonetics*.
- HSU, C. S. K., & JUN, S. A. (1998). Prosodic strengthening in Taiwanese: Syntagmatic or paradigmatic?. *UCLA working papers in phonetics*, 69-89.
- HUET, K., & HARMEGNIES, B. (2000). Contribution à la quantification du degré d'organisation des systèmes vocaliques. *Actes des XXIIIe Journées d'Étude sur la Parole*, 225-228.
- HUNNICUTT, S. (1985). Intelligibility versus redundancy-conditions of dependency. *Language and Speech*, 28(1), 47-56.
- JOHNSON, K., FLEMMING, E., & WRIGHT, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language*, 69, 505-528.
- KAHN J. (2011). *Parole de locuteur : performance et confiance en identification biométrique vocale*. (Thèse de doctorat, ANRT Université Sorbonne Nouvelle - Paris 3).
- KATZ, W. F., KRIPKE, C., & TALLAL, P. (1991). Anticipatory coarticulation in the speech of adults and young children: Acoustic, perceptual, and video data. *Journal of Speech, Language, and Hearing Research*, 34(6), 1222-1232.
- KEATING, P. A. (1985). Universal phonetics and the organization of grammars. *UCLA Working Papers in Phonetics*, 59, 35-49.
- KEATING, P. A. (1988). Underspecification in phonetics. *Phonology*, 5(2), 275-292.
- KEATING, P. A. (1990). The window model of coarticulation: Articulatory evidence. In J. Kingston, & M. E. Beckman (Eds.), *Papers in laboratory phonology I: Between the grammar and the physics of speech*, 451-470. Cambridge: Cambridge University Press.
- KEATING, P., CHO, T., FOUGERON, C., & HSU, C. (2003). Domain-initial articulatory strengthening in four languages. *Laboratory Phonology*, 6, 143-161.
- KELSO, J. A., SALTZMAN, E. L., & TULLER, B. (1986). The dynamical perspective on speech production: Data and theory. *Journal of Phonetics*, 14 (1), 29-59.
- KENT, R. (1983). The segmental organization of speech. In P. MacNeilage (ed.), *The production of Speech*, 57-89, New York: Springer.

- KIM, S., & CHO, T. (2012). Prosodic strengthening in the articulation of English /æ/. *Studies in Phonetics, Phonology and Morphology*, 18(2), 321-337.
- KRULL D. (1989). Second formant locus patterns and consonant-vowel coarticulation in spontaneous speech. *Perilus*, 10, 87-108.
- KÜHNERT, B., & NOLAN, F. (1999). The origin of coarticulation. *Coarticulation: Theory, data and techniques*, 7-30.
- KUZNETSOVA, A., BROCKHOFF, P.B., & CHRISTENSEN, R. H. B. (2013). Lmertest: Tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2. 0-29.
- LABOV, W. (1976). *Sociolinguistique*. Paris : Éditions de Minuit.
- LABOV, W. (1994). *Principles of language change: Internal factors*.
- LAMEL, L., GAUVAIN, J., & ESKENAZI, M. (1991). BREF, a Large Vocabulary Spoken Corpus for French. *Proceedings of the European Conference on Speech Communication and Technology (EUROSPEECH)*, Genova, Italy.
- LANE, H., & TRANEL, B. (1971). The Lombard sign and the role of hearing in speech. *Journal of Speech, Language, and Hearing Research*, 14(4), 677-709.
- LANE, H., TRANEL, B., & SISSON, C. (1970). Regulation of voice communication by sensory dynamics. *The Journal of the Acoustical Society of America*, 47(2B), 618-624.
- LEHNERT-LEHOULLIER, H., MCDONOUGH, J., & MCALEAVEY, S. (2010). Prosodic strengthening in American English domain-initial vowels. *Proceedings of Speech Prosody 2010-Fifth International Conference*.
- LENTH, R., LOVE, J., & LENTH, M. R. (2018). Package 'lsmeans'. *The American Statistician*, 34(4), 216-221.
- LEVELT, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- LIEBERMAN, P. (1963). Some effects of semantic and grammatical context on the production and perception of speech. *Language and speech*, 6(3), 172-187.
- LINDBLOM, L. (1963). On vowel reduction. *The Royal Institute of Technology, Speech Transmission Laboratory*, 29.
- LINDBLOM, B. (1990). Explaining phonetic variation. A sketch of the H&H theory. In A. Marchal and W. Hardcastle (Eds), *Speech Production and Speech Modelling*, 403-439, Dordrecht: Kluwer.

- LOBANOV, B. M. (1971). Classification of Russian vowels spoken by different speakers. *The Journal of the Acoustical Society of America*, 49(2b), 606-608.
- LOMBARD, E. (1911). Le signe de l'élévation de la voix, *Annales des Maladies de l'Oreille et du Larynx*, 37, 101-119.
- MA, L., PERRIER, P., & DANG, J. (2006). Anticipatory Coarticulation in Vowel-Consonant-Vowel sequences: A crosslinguistic study of French and Mandarin speakers. *Proceedings of the 7th International Seminar on Speech Production*, 151-158, Ubatuba.
- MANUEL, S. Y. (1990). The role of contrast in limiting vowel-to-vowel coarticulation in different languages. *The Journal of the Acoustical Society of America*, 88(3), 1286-1298.
- MANUEL, S. Y., & KRAKOW, R. A. (1984). Universal and language particular aspects of vowel-to-vowel coarticulation. *Haskins Laboratories Status Report on Speech Research*, 77(78), 69-78.
- MARTINET, A. (1958). C'est jeuli, le Mareuc!. *Romance philology*, 11, 345-355.
- MATTHIES, M., PERRIER, P., PERKELL, J. S. & ZANDIPOUR, M. (2001). Variation in anticipatory coarticulation with changes in clarity and rate. *Journal of Speech Language and Hearing Research* 44, 340-353.
- MEUNIER, C. (1994). *Les groupes de consonnes: problématique de la segmentation et variabilité acoustique* (Thèse de doctorat, Université Aix-Marseille 1).
- MEUNIER, C. (2005). Invariant et variabilité phonétique. In: Nguyen, N. (ed). *Forme et Substance*, 349-374.
- MEUNIER, C., & ESPESSE, R. (2012). Vowel reduction in conversational speech in French: The role of lexical factors. *Journal of Phonetics* 39(3), 271-278.
- MEUNIER C., & FLOCCIA, C. (1997). La spécificité des types de parole pour la perception de la voyelle, *Actes des 1ères Journées d'Études Linguistiques, La voyelle dans tous ses états*, 158-163, Nantes.
- MEUNIER C., & FLOCCIA C. (1999). Syllabe ou mot: quelle unité permet d'identifier les catégories phonétiques ? , *Actes des 2èmes Journées d'Études Linguistiques Syllabes*, 87-92, Nantes.
- MEYNADIER, Y. (2003). *Interaction entre prosodie et (co)articulation linguopalatale en français*. These de Doctorat (Université Aix-Marseille).
- MOON, S. J., & LINDBLOM, B. (1994). Interaction between duration, context, and speaking style in English stressed vowels. *The Journal of the Acoustical society of America*, 96(1), 40-55.

- NAM, H., GOLDSTEIN, L. & SALTZMAN, E. (2010). Self-organization of syllable structure: A coupled oscillator model. In Pellegrino, F. (Ed.) *Approaches to phonological complexity*, 16, 299-328, Walter de Gruyter, Berlin.
- NGUYEN, N. (2001). Rôle de la coarticulation dans la reconnaissance des mots. *Année psychologique*, 101, 125-154.
- NGUYEN, N., & ADDA-DECKER, M. (2013). *Méthodes et outils pour l'analyse phonétique des grands corpus oraux*. Hermes Science Publications.
- NITTROUER, S., STUDDERT-KENNEDY, M. & MCGOWAN, R. (1989). The emergence of phonetic segments: evidence from the spectral structure of fricative-vowel syllables spoken by children and adults. *Journal of Speech and Hearing Research* 32, 120-132.
- NITTROUER, S. & WHALEN, D. (1989). The perceptual effects of child-adult differences in fricative vowel coarticulation. *Journal of the Acoustical Society of America* 86, 1266-1276.
- NOIRAY, A., MÉNARD, L., & ISKAROUS, K. (2013). The development of motor synergies in children: Ultrasound and acoustic measurements. *The Journal of the Acoustical Society of America*, 133(1), 444-452.
- NOLAN F. (1983). *The Phonetic Bases of Speaker Recognition*. Cambridge : Cambridge Univ. Press.
- NOLAN, F. (1985). Idiosyncrasy in coarticulatory strategies. *Cambridge Papers in Phonetics and Experimental Linguistics* 4, 1-9.
- OHALA, J. J. (1981). The Listener as a Source of Sound Change. In Carrie Masek, Roberta Hendrick & Mary Frances Miller (Eds), *Papers from the Parasession on Language and Behavior*, 178-203. Chicago: Chicago Linguistic Society.
- OHALA, J. J. (1990). The phonetics and phonology of aspects of assimilation. In J. Kingston and M. Beckman (Eds), *Papers in Laboratory Phonology I: Between the Grammar and the Physics of Speech*, 258-275, Cambridge University Press, Cambridge.
- OHALA, J. J. (1993). The phonetics of sound change. In C. Jones Longman (Ed), *Historical Linguistics: Problems and Perspectives*, 237-278, London.
- ÖHMAN, S. E. (1966). Coarticulation in VCV utterances: Spectrographic measurements. *The Journal of the Acoustical Society of America*, 39(1), 151-168.
- ONAKA, A. (2003). Domain-initial strengthening in Japanese: An acoustic and articulatory study. In *Proceedings of the 15th international congress of phonetic sciences*, 2091-2094, Barcelona, Spain.

- PICHENY, M. A., DURLACH, N. I., & BRAIDA, L. D. (1986). Speaking clearly for the hard of hearing. II: Acoustic characteristics of clear and conversational speech. *Journal of Speech and Hearing Research*, 29, 434-446.
- R DEVELOPMENT CORE TEAM. (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0, retrieved from <http://www.R-project.org/>.
- RECASENS, D. (1983). Place cues for nasal consonants with special reference to Catalan. *The Journal of the Acoustical Society of America*, 73(4), 1346-1353.
- RECASENS, D. (1984). Vowel - to - vowel coarticulation in Catalan VCV sequences. *The Journal of the Acoustical Society of America*, 76(6), 1624-1635.
- RECASENS, D. (1985). Coarticulatory patterns and degrees of coarticulatory resistance in Catalan CV sequences. *Language and Speech*, 28(2), 97-114.
- RECASENS, D. (2014). *Coarticulation and Sound Change in Romance*. In John Benjamins Publishing Company (ed).
- RECASENS, D. (2018). Coarticulation. Oxford Research Encyclopedia of Linguistics. Retrieved from: <http://linguistics.oxfordre.com/view/10.1093/acrefore/9780199384655.001.0001/acrefore-9780199384655-e-416>.
- RECASENS, D., & ESPINOSA, A. (2009). An articulatory investigation of lingual coarticulatory resistance and aggressiveness for consonants and vowels in Catalan. *The Journal of the acoustical society of America*, 125(4), 2288-2298.
- RECASENS, D., PALLARÈS, M. D., & FONTDEVILA, J. (1997). A model of lingual coarticulation based on articulatory constraints. *The Journal of the Acoustical Society of America*, 102(1), 544-561.
- RENWICK, M. E., & CASSIDY, C. N. (2015). Detecting palatalization in spontaneous spoken English. *In Proceedings of Meetings on Acoustics 169ASA*, 23(1).
- ROUAS, J.L., BEPPI, M., & ADDA-DECKER, M. (2010). Comparison of Spectral Properties of Read, Prepared and Casual Speech in French. *International Conference on Language Resources and Evaluation (LREC 2010)*, 606-611, Malta.
- SCARBOROUGH, R. A. (2004). *Coarticulation and the structure of the lexicon*. (Thèse de doctorat, University of California, Los Angeles).
- SCARBOROUGH, R. A., & ZELLOU, G. (2013). Clarity in communication: "Clear" speech authenticity and lexical neighborhood density effects in speech production and perception. *The Journal of the Acoustical Society of America*, 134(5), 3793-3807.

- SCHAFER, A. J., SPEER, S. R., & WARREN, P. (2005). Prosodic influences on the production and comprehension of syntactic ambiguity in a game-based conversation task. In J. C. Trueswell & M. K. Tanenhaus (eds.), *Approaches to Studying World-situated Language Use*, 209-25, Cambridge, MA: MIT Press.
- SCHWARTZ, J.L., BOË, L. J., VALLÉE, N., & ABRY, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics* 25, 255-286.
- SHI, R., GICK, N., KANWISCHER, D., & WILSON, I. (2005). Frequency and category factors in the reduction and assimilation of function words: EPG and acoustic measures. *Journal of Psycholinguistic Research*, 34, 315–364.
- SMILJANIĆ, R., & BRADLOW, A.R (2005). Production and perception of clear speech in Croatian and English. *The Journal of the Acoustical Society of America*, 118(3), 1677-1688.
- STABLE. (déf. 2) *Dictionnaire Larousse en ligne*. Paris: Éditions Larousse, Inc. *Larousse.fr*, retrouvé à <http://www.larousse.fr/dictionnaires/francais/stable/74409?q=stable#73571>
- STABLE. (def B.2.b) Centre National des Ressources Textuelles (CNRTL). *cnrtl.fr*, <http://www.cnrtl.fr/definition/stable>
- STEVENS, K. N. (1972). *The quantal nature of speech: Evidence from articulatory-acoustic data*.
- STEVENS, K. N., & HOUSE, A. S. (1963). Perturbations of vowel articulations by consonantal context: An acoustical study. *Journal of Speech and Hearing Research*, 6, 111-128.
- STEVENS, K. N., HOUSE, A. S., & PAUL, A. P. (1966). Acoustical description of syllabic nuclei: An interpretation in terms of a dynamic model of articulation. *The Journal of the Acoustical Society of America*, 40(1), 123-132.
- STEVENS, K. N., & KEYSER, S. J. (1989). Primary features and their enhancement in consonants. *Language*, 81-106.
- STRAKA, G. (1965). Contribution à l'histoire de la consonne R en français. *Neuphilologische Mitteilungen* 66. 572-606.
- STYLE. *Dictionnaire Larousse en ligne*. Paris: Éditions Larousse, Inc. *Larousse.fr*, retrouvé à <http://www.larousse.fr/dictionnaires/francais/style/74959?q=style#74108>
- SU L. S., LI K. P., FU K. S. (1974). Identification of speakers by use of nasal coarticulation. *Journal of the Acoustical Society of America*, 56, 1867-1882.

- TABAIN, M. (2003). Effects of prosodic boundary on /aC/ sequences: acoustic results. *Journal of Acoustic Society of America (JASA)*, 113, 516-531.
- TABAIN, M. & PERRIER, P. (2005). Articulation and acoustics of /i/ in pre-boundary position in French. *Journal of Phonetics* 33, 77-100.
- TORREIRA, F., ADDA-DECKER, M., & ERNESTUS, M. (2010). The Nijmegen corpus of casual French. *Speech Communication*, 52, 201–212.
- TRAUNMÜLLER, H. (1997). Auditory scales of frequency representation. [Online: <http://www.ling.su.se/staff/hartmut/bark.htm>]
- TULLER, B., & KELSO, J. A. S. (1991). The production and perception of syllable structure. *Journal of Speech and Hearing Research*, 34, 501-508.
- VAISSIÈRE, J. (1985). Étude des variations allophoniques de la voyelle /a/ et ses conséquences pour la reconnaissance automatique de la parole. *XIV Journées d'Étude sur la Parole*, 304-307, Paris.
- VAISSIÈRE, J. (1987). Effect of phonetic context and timing on the F-pattern of the vowels in continuous speech. *Proceedings of 11th Congress of Phonetic Sciences*, 43-46. Tallin.
- VAISSIÈRE, J. (1996). From Latin to modern French: on diachronic changes and synchronic variations. *Arbeitsberichte, U. Kiel*, 31, 61-74.
- VAISSIÈRE, J. (2006). *La phonétique. Que sais-je*. Presses Universitaires de France.
- VAISSIÈRE, J. (2009). Articulatory modeling and the definition of acoustic-perceptual targets for reference vowels. *The Chinese Phonetics Journal*, 2, 22-33.
- VENABLES, W. N. & RIPLEY, B. D. (2002). *Modern Applied Statistics with S*. Fourth Edition. Springer, New York. ISBN 0-387-95457-0
- WHALEN, D. H. (1984). Subcategorical phonetic mismatches slow phonetic judgments. *Perception & psychophysics*, 35(1), 49-64.
- WICKHAM, H., FRANCOIS, R., HENRY, L., & MÜLLER, K. (2017). Package 'dplyr': A Grammar of Data Manipulation.
- WOOD, S. (1979). A radiographic analysis of constriction locations for vowels. *Journal of Phonetics*, 7(1), 25-43.
- ZHARKOVA, N., HEWLETT, N., & HARDCASTLE, W. J. (2011). Coarticulation as an indicator of speech motor control development in children: An ultra- sound study. *Motor Control* 15, 118-140.

- ZHARKOVA, N., HEWLETT, N., & HARDCASTLE, W. J. (2012). An ultrasound study of lingual coarticulation in /sV/ syllables produced by adults and typically developing children. *Journal of International Phonetic Association*, 42(2), 193-208.

ANNEXES

ANNEXE I. DÉTAIL DU MATÉRIEL LINGUISTIQUE UTILISÉ DANS L'ÉTUDE I

- A. RÉSUMÉ DU MATÉRIEL DISPONIBLE PAR LOCUTEUR EN FONCTION DE LA DURÉE DE CHAQUE VOYELLE (C=COURTE 50ms, M=MOYENNE 60-80ms ET L=LONGUE 90-150ms) PAR CONTEXTE CONSONANTIQUE (ALV = ALVÉOLAIRE, UV = UVULAIRE OU VEL=VÉLAIRE).

Pages (191 à 194)

- B. VALEURS MOYENNES (M) ET ÉCART-TYPES (ET) DE CHAQUE FORMANT (F1, F2, F3, F4) EN Z-SCORE PAR VOYELLE ET PAR LOCUTEUR.

Pages (195 à 202)

- C. GRAPHIQUE DE DENSITÉ ILLUSTRANT LA DISTRIBUTION DE LA DURÉE (50-150 MS) DES 8 VOYELLES TESTÉES EN FONCTION DU CONTEXTE CONSONANTIQUE (ROSE : ALVÉOLAIRE, VERT : UVULAIRE ET BLEU : VÉLAIRE).

Page 203

		/i/			/e/			/ɛ/			/a/			/ɔ/			/u/			/o/			/ɔ/		
		ALV	UV	VEL																					
03_12_07_ nb1_1_16	C	27	4	4	35	1	1	25	10	10	57	19	3	17	1	1	3	12	4	3	1		12	10	5
	M	56	1	6	64	1		35	11	5	75	21	15	41	2	1	12	30	6	6	4	8	26	14	8
	L	17	1	4	12		2	7	9	1	13	3	3	4		6	2	8	2	3	1	4	6	6	1
04_12_07_ nb1_2_16	C	19	3	3	45	5	1	30	16	12	80	19	10	16			6	8	3	5		5	4	9	11
	M	51	5	10	43	3	1	39	16	10	97	28	26	14	2	2	6	14	1	10	2	7	15	10	8
	L	11	4	3	6	1		12	3	2	22	8	10	2				7		6		3	3	4	
04_12_07_ nb2_1_16	C	27	4	1	57	5	2	30	19	9	97	22	8	25		4	10	6	4	6	2	8	18	13	13
	M	37	6	7	66	3	4	41	15	9	121	21	25	25		7	12	18	1	8	2	9	16	7	15
	L	12	2	3	15	1	1	8	6	5	24	7	8	3	2	3	2	8		3	5	1	13	3	
04_12_07_ nb3_2_16	C	42	2	6	80	4		38	24	14	130	14	14	18		6	2	10		4	6	4	10	8	6
	M	40	4	4	54	2	6	36	16	6	148	20	18	30		18	12	24	6	10	4	6	6	18	4
	L	4		2	10			8	8		16		6	6		4		10				2		4	
05_12_07_ nb1_1_16	C	26	6	3	55		4	55	16	8	72	20	10	34	3	3	10	7	6	9		9	20	17	5
	M	39	2	7	96	4	5	47	11	11	135	33	26	55	3	11	17	16	3	23	5	8	25	12	9

	L	13	1	3	10	1	2	9	4		25	10	2	11	5	5		6	1	7	1		4	3	
05_12_07_ nb1_2_16	C	28	3	1	33	2		45	16	5	82	25	6	46	2	1	3	15		7	1	4	10	21	7
	M	47	4	2	87	4	4	68	28	9	176	44	12	42	2	3	12	11	2	8	2	3	23	29	5
	L	20	4	3	20	1	2	25	6	3	42	12	4	11	3	1	3	13	1		3	1	10	7	2
14_11_07_ nb1_2_16	C	22	4	2	29		1	64	9	6	61	19	6	22		1	4	10	2	4		1	7	5	10
	M	23	4	5	59		2	83	16	11	202	29	22	44	1	7	5	14		7	4	3	8	7	17
	L	10	1	1	20	2	1	14	5	1	39	17	1	10				7		1	2	1	3	1	1
14_11_07_ nb2_1_16	C	15	8	3	54	1	2	60	9	16	158	15	12	41	1	5	9	13	3	5		3	13	11	20
	M	34	2	5	71	2	1	64	12	12	204	17	15	41	6	6	7	7	3	7	4	5	15	9	11
	L	5		2	17		1	13	4		39	2	4	10		8	3	4		2	2	2	2	3	1
14_11_07_ nb2_2_16	C	24	2	6	46	3	2	62	14	5	101	15	12	20	3	2	6	16	5	1		3	9	9	12
	M	44	10	13	101	3	5	124	18	8	239	21	27	74	4	7	12	21	6	13	2	6	25	22	34
	L	19	3	3	22	1	3	29	7	4	56	3	4	13	3		5	11	1	5	1	1	11	5	1
26_11_07_ nb3_1_16	C	22	4	5	52	6	1	64	15	8	64	29	6	27	7	3	1	15	1	4	1		4	10	17
	M	36	10	8	99	6	7	96	25	21	228	78	17	51	6	7	6	25	6	7	9	6	20	7	25
	L	25	10	5	55			42	16	9	67	25	12	12	5	5	5	8	1		7	1	7	9	5

26_11_07_ nb3_2_16	C	34	8	12	75		2	56	14	19	142	29	15	44	2	2	4	15	3	8	3	4	7	17	23
	M	52	24	16	99	2	5	104	22	24	188	38	30	56	2	2	9	16	2	9	6	9	19	20	21
	L	40	11	5	42	4	2	29	15	11	73	20	7	15	2	4	7	6		3	5	3	2	8	2
28_11_07_ nb2_2_16	C	37	7	7	55	14	6	62	51	24	133	36	10	40		3	8	12	1	6	1	7	27	15	14
	M	66	13	10	116	15	6	94	45	27	278	66	28	51	1	6	10	17	5	18	3	21	40	13	15
	L	20	7	3	29	2	3	20	9	2	55	21	8	12	2		4	8	3	6		2	8	5	
29_11_07_ nb2_1_16	C	17	1	2	57	3	2	42	24	8	75	24	5	26		3	6	8	4	5	4	3	3	3	10
	M	28		9	110		5	70	39	8	178	35	13	36	3	1	12	10	4	7	6	6	12	10	11
	L	6	2	4	35		3	24	10	5	41	7	6	6	2	2	6	3	2	3	4	3	4	6	4
29_11_07_ nb2_2_16	C	18	2	3	44	6	4	37	13	11	76	21	5	23	4	3	6	6	1	6	3	5	6	7	7
	M	32	2	6	45	8	6	39	27	10	144	26	19	47	2	3	8	11	5	11	8	3	22	6	7
	L	7	1	2	7	2	1	9	7		25	9	2	10	1		4	11		3	3	2	5	1	
30_11_07_ nb1_1_16	C	14	3	1	53	3	2	34	7	8	81	15	9	42	1	7	2	4	2	4		1	5	10	15
	M	39	5	5	92	7	1	43	26	9	131	29	18	46	3	4	4	12	1	5	5	3	13	16	12
	L	11	2	6	28	1	1	9	13	3	41	22	3	7	7	2	2	8	3	8	1	2	7	3	3
30_11_07_	C	12	3	4	48	2		23	18	7	56	6	1	11		4	7	5		5		3	3	1	11

nbl_2_16	M	37	6	8	91	4	1	56	32	20	148	27	13	32	1	6	8	10	1	11	3	5	14	5	18
	L	12	6	1	21		2	13	21	8	40	16	6	8	1		4	5		4	1	1	8	3	2
Total voyelles		1277	217	234	2460	135	113	2037	777	424	4775	1043	542	1277	95	179	286	541	105	296	127	197	550	442	428

Locuteur	Voyelle	N	F1 (M)	F1 (ET)	F2 (M)	F2 (ET)	F3 (M)	F3 (ET)	F4 (M)	F4 (ET)
03_12_07_nb1_1_16	/i/	120	-1,26	0,66	1,26	0,38	1,10	0,88	0,45	0,83
	/e/	116	-0,35	0,41	0,68	0,32	-0,14	0,62	0,20	0,75
	/ɛ/	113	0,23	0,67	0,38	0,44	-0,15	0,72	0,27	1,05
	/a/	209	1,14	0,72	-0,35	0,49	0,09	0,79	0,16	0,88
	/ɔ/	73	-0,19	0,55	-0,36	0,35	-0,48	0,67	-0,51	0,77
	/u/	79	-0,20	1,00	-2,00	0,56	-1,05	0,96	-1,24	0,77
	/o/	30	-0,46	0,48	-1,76	0,58	-0,86	0,81	-1,20	0,66
	/ɔ̃/	88	0,39	0,69	-1,13	0,62	-0,37	0,91	-0,35	0,82
04_12_07_nb1_2_16	/i/	109	-1,20	0,52	1,39	0,30	2,02	0,75	-0,21	0,87
	/e/	105	-0,70	0,41	0,76	0,26	-0,02	0,58	0,07	0,86
	/ɛ/	140	0,15	0,62	0,42	0,35	-0,25	0,61	-0,21	0,87
	/a/	300	1,07	0,64	-0,53	0,45	-0,07	0,63	0,25	0,91
	/ɔ/	36	-0,59	0,43	-0,29	0,39	-0,80	0,58	-0,69	0,92
	/u/	45	-0,14	0,80	-1,77	0,69	-0,64	0,76	-0,62	0,95
	/o/	38	-0,55	0,37	-1,93	0,49	-1,02	0,68	-0,56	0,65
	/ɔ̃/	64	0,55	0,62	-1,31	0,60	-0,31	0,91	-0,05	0,82
04_12_07_nb2_1_16	/i/	99	-1,22	0,54	1,41	0,36	1,70	0,97	0,20	0,96

	/e/	154	-0,67	0,37	0,81	0,34	0,30	0,68	-0,03	0,80
	/ɛ/	142	0,15	0,65	0,36	0,38	-0,07	0,66	-0,06	0,90
	/a/	333	1,18	0,60	-0,44	0,43	-0,40	0,65	0,05	0,96
	/x/	69	-0,38	0,48	-0,24	0,45	-0,56	0,55	-0,67	0,67
	/u/	61	-0,24	0,98	-1,72	0,70	-0,94	0,74	-0,63	1,15
	/o/	44	-0,56	0,47	-1,91	0,61	-0,88	0,77	-0,61	1,03
	/ɔ/	98	0,40	0,56	-1,09	0,56	-0,80	0,66	-0,36	0,59
04_12_07_nb3_2_16	/i/	104	-1,38	0,37	1,11	0,38	1,95	0,68	0,39	0,93
	/e/	156	-0,48	0,40	0,83	0,27	0,41	0,47	0,28	0,80
	/ɛ/	150	0,07	0,70	0,40	0,43	-0,15	0,49	0,01	0,94
	/a/	366	1,00	0,69	-0,37	0,43	-0,29	0,45	0,20	0,84
	/x/	82	-0,43	0,42	-0,12	0,32	-0,65	0,56	-0,68	0,74
	/u/	64	0,01	0,95	-2,08	0,67	-0,77	0,51	-1,15	0,85
	/o/	36	-0,20	0,54	-1,83	0,70	-0,86	0,94	-1,13	1,07
05_12_07_nb1_1_16	/ɔ/	56	0,73	0,75	-1,31	0,86	-0,57	0,56	-0,61	0,58
	/i/	100	-1,14	0,34	1,24	0,39	1,69	0,78	0,40	1,11
	/e/	177	-0,56	0,31	0,86	0,33	0,13	0,56	0,17	0,65

	/ɛ/	161	0,05	0,59	0,41	0,40	-0,08	0,50	0,06	0,81
	/a/	333	1,24	0,70	-0,60	0,54	-0,28	0,56	0,02	0,89
	/x/	130	-0,46	0,43	-0,29	0,37	-0,81	0,50	-0,68	0,58
	/u/	66	-0,35	0,85	-1,90	0,63	-1,01	0,60	-1,46	0,88
	/o/	62	-0,55	0,33	-1,64	0,67	-0,91	0,49	-0,90	0,64
	/ɔ/	95	0,24	0,71	-1,37	0,65	-0,82	0,53	-0,69	0,68
	/i/	112	-1,26	0,37	1,27	0,41	1,80	0,58	0,55	0,90
	/e/	153	-0,71	0,29	0,88	0,31	0,37	0,58	0,10	0,82
	/ɛ/	205	0,16	0,60	0,38	0,39	0,00	0,47	0,04	0,78
	/a/	403	1,18	0,63	-0,40	0,52	-0,33	0,44	0,15	0,86
	/x/	111	-0,50	0,51	-0,36	0,40	-0,76	0,66	-0,58	0,71
	/u/	60	-0,21	0,85	-2,10	0,67	-1,35	0,75	-1,37	0,87
	/o/	29	-0,66	0,35	-1,97	0,67	-1,34	0,63	-0,88	0,76
05_12_07_nbl_2_16	/ɔ/	114	0,29	0,59	-1,44	0,66	-0,80	0,65	-0,57	0,72
	/i/	72	-1,21	0,54	1,36	0,40	0,68	0,95	0,11	0,75
14_11_07_nbl_2_16	/e/	114	-0,69	0,28	0,88	0,39	0,19	0,66	0,04	0,82
	/ɛ/	209	-0,15	0,59	0,49	0,41	0,01	0,77	-0,06	0,97

	/a/	396	1,11	0,68	-0,54	0,54	-0,06	0,98	0,02	0,99
	/x/	85	-0,68	0,36	-0,38	0,44	-0,75	0,84	-0,69	0,71
	/u/	42	0,04	0,83	-1,94	0,68	-0,97	1,11	-1,02	0,63
	/o/	23	-0,72	0,35	-1,87	0,80	-1,01	1,49	-0,86	0,67
	/ɔ/	59	0,15	0,69	-1,39	0,55	-0,38	0,89	-0,34	0,69
14_11_07_nb2_1_16	/i/	74	-1,17	0,55	1,47	0,42	1,34	0,86	0,48	0,99
	/e/	149	-0,79	0,33	0,97	0,35	0,24	0,58	0,11	0,77
	/ɛ/	190	-0,21	0,58	0,58	0,45	-0,08	0,68	0,06	0,78
	/a/	466	1,09	0,62	-0,60	0,45	-0,17	0,74	0,00	0,93
	/x/	118	-0,59	0,44	-0,27	0,36	-0,69	0,77	-0,49	0,71
	/u/	49	-0,30	0,99	-1,76	0,73	-1,34	0,89	-1,15	0,75
	/o/	30	-0,54	0,39	-1,64	0,68	-1,23	0,94	-1,34	0,45
	/ɔ/	85	0,35	0,67	-1,21	0,54	-0,64	0,77	-0,69	0,53
14_11_07_nb2_2_16	/i/	124	-1,12	0,56	1,16	0,42	1,59	0,71	0,39	0,88
	/e/	186	-0,52	0,34	0,82	0,38	0,04	0,62	0,01	0,93
	/ɛ/	271	-0,18	0,54	0,51	0,43	-0,14	0,68	-0,05	0,87
	/a/	478	1,20	0,68	-0,59	0,50	-0,16	0,68	-0,13	0,90
	/x/	126	-0,44	0,45	-0,35	0,39	-0,84	0,63	-0,50	0,75

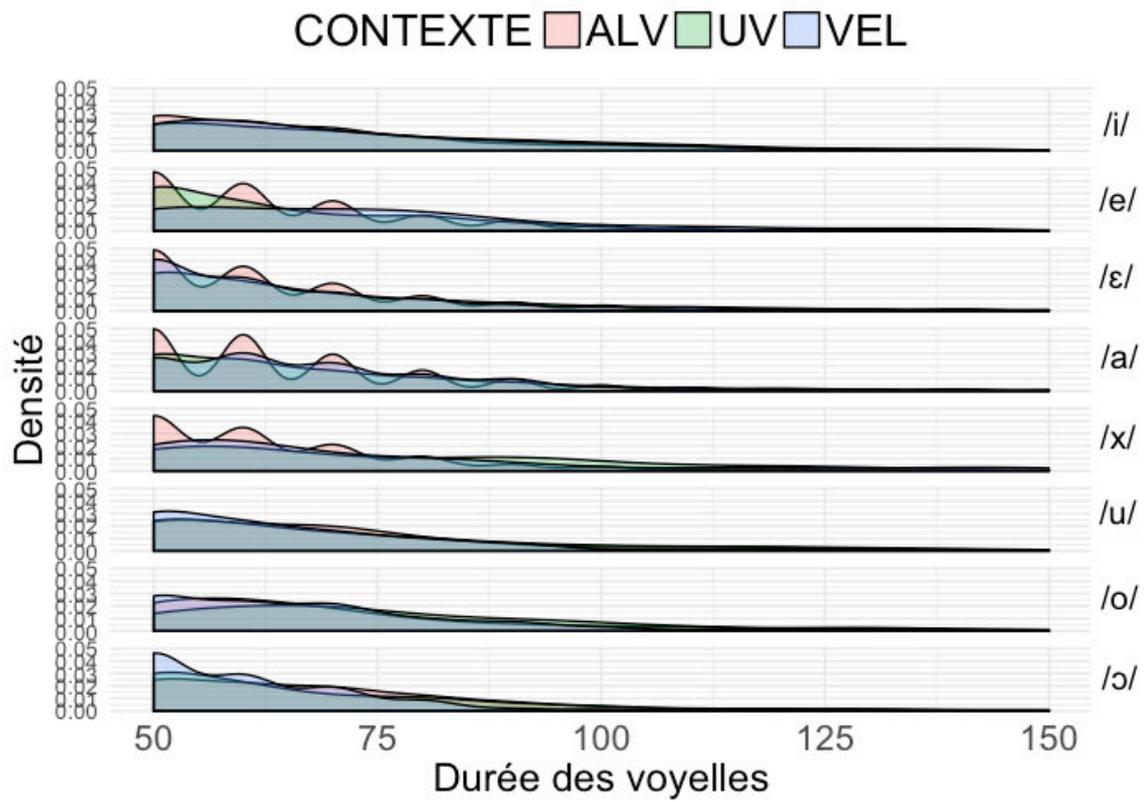
	/u/	83	-0,20	0,82	-2,00	0,68	-1,11	0,89	-0,90	0,78
	/o/	32	-0,60	0,27	-1,87	0,76	-1,36	0,92	-0,50	0,64
	/ɔ/	128	0,54	0,57	-1,34	0,68	-0,56	0,92	-0,34	0,70
26_11_07_nb3_1_16	/i/	125	-1,20	0,55	1,10	0,43	1,49	0,78	0,64	0,98
	/e/	226	-0,70	0,38	0,80	0,40	0,25	0,67	0,03	0,82
	/ɛ/	296	-0,04	0,59	0,44	0,47	0,01	0,75	-0,06	0,86
	/a/	526	1,07	0,74	-0,42	0,59	-0,11	0,68	0,11	0,98
	/ɤ/	123	-0,51	0,52	-0,40	0,58	-0,77	0,74	-0,44	0,77
	/u/	68	-0,49	0,68	-2,18	0,50	-1,31	0,61	-0,90	0,72
	/o/	35	-0,39	0,49	-2,12	0,72	-1,09	0,68	-0,87	1,04
	/ɔ/	104	0,18	0,58	-1,29	0,59	-0,96	0,87	-0,38	0,63
26_11_07_nb3_2_16	/i/	202	-1,15	0,53	1,16	0,40	1,58	0,79	0,15	0,91
	/e/	231	-0,69	0,34	0,76	0,37	0,24	0,57	0,21	0,82
	/ɛ/	294	0,13	0,62	0,39	0,47	-0,03	0,56	0,04	0,96
	/a/	542	1,11	0,66	-0,56	0,49	-0,32	0,60	0,02	0,86
	/ɤ/	129	-0,46	0,44	-0,42	0,46	-0,75	0,56	-0,69	0,79
	/u/	62	-0,40	0,64	-1,99	0,76	-1,00	0,69	-1,21	0,73
	/o/	50	-0,35	0,43	-2,05	0,66	-1,23	0,67	-1,14	0,96

	/ɔ/	119	0,38	0,62	-1,38	0,58	-0,83	0,69	-0,43	0,51
28_11_07_nb2_2_16	/i/	170	-1,21	0,64	1,06	0,40	1,84	0,76	0,44	0,98
	/e/	246	-0,66	0,41	0,93	0,40	0,53	0,61	-0,05	0,98
	/ɛ/	334	0,05	0,57	0,43	0,43	-0,01	0,52	-0,08	0,94
	/a/	635	1,10	0,69	-0,41	0,50	-0,49	0,50	0,23	0,98
	/x/	115	-0,45	0,42	-0,38	0,43	-0,64	0,46	-0,34	0,77
	/u/	68	-0,35	0,69	-2,18	0,71	-0,77	0,49	-1,14	0,86
	/o/	64	-0,57	0,39	-2,38	0,53	-1,15	0,48	-0,54	0,77
	/ɔ/	137	0,22	0,68	-1,29	0,60	-0,97	0,52	-0,36	0,62
29_11_07_nb2_1_16	/i/	69	-1,20	0,50	1,33	0,35	0,86	0,80	0,44	0,71
	/e/	215	-0,82	0,39	0,94	0,36	0,21	0,71	0,21	0,85
	/ɛ/	230	0,12	0,66	0,41	0,46	-0,01	0,78	0,15	0,92
	/a/	384	0,97	0,76	-0,45	0,55	-0,10	0,91	-0,12	0,98
	/x/	79	-0,60	0,55	-0,23	0,43	-0,42	1,02	-0,12	0,68
	/u/	55	-0,40	0,88	-1,88	0,56	-1,16	1,25	-0,96	0,93
		/o/	41	-0,37	0,50	-2,12	0,55	-0,58	1,51	-0,74
	/ɔ/	63	0,38	0,69	-1,35	0,60	-0,39	1,13	0,05	0,64

29_11_07_nb2_2_16	/i/	73	-1,10	0,72	1,38	0,38	1,35	0,91	0,38	0,79
	/e/	123	-0,66	0,36	0,81	0,35	0,03	0,72	0,03	0,78
	/ɛ/	153	0,07	0,68	0,50	0,42	-0,26	0,72	-0,21	0,77
	/a/	327	1,02	0,80	-0,31	0,50	0,08	0,70	0,16	0,89
	/x/	93	-0,49	0,49	-0,45	0,41	-0,54	0,82	-0,60	0,85
	/u/	52	-0,24	0,84	-1,73	0,70	-1,15	1,12	-1,03	0,72
	/o/	44	-0,44	0,42	-1,82	0,67	-0,68	1,18	-0,98	0,70
	/ɔ/	61	0,19	0,77	-1,00	0,61	-0,36	1,14	-0,37	0,82
30_11_07_nb1_1_16	/i/	86	-1,24	0,66	1,33	0,43	1,62	0,93	0,34	0,96
	/e/	188	-0,67	0,39	0,85	0,34	0,47	0,65	-0,03	0,85
	/ɛ/	152	-0,04	0,61	0,46	0,44	-0,06	0,60	-0,02	0,77
	/a/	349	1,10	0,76	-0,58	0,49	-0,42	0,64	0,02	0,87
	/x/	119	-0,42	0,49	-0,19	0,43	-0,41	0,53	-0,57	0,52
	/u/	38	-0,70	0,57	-1,86	0,66	-0,89	0,57	-1,18	0,63
	/o/	29	-0,50	0,38	-1,87	0,63	-0,87	0,47	-0,66	0,63
	/ɔ/	84	0,35	0,61	-1,34	0,63	-0,93	0,55	-0,36	0,69
30_11_07_nb1_2_16	/i/	89	-1,30	0,53	1,46	0,37	1,86	0,88	0,14	0,87
	/e/	169	-0,65	0,32	0,73	0,28	0,17	0,60	0,15	0,84

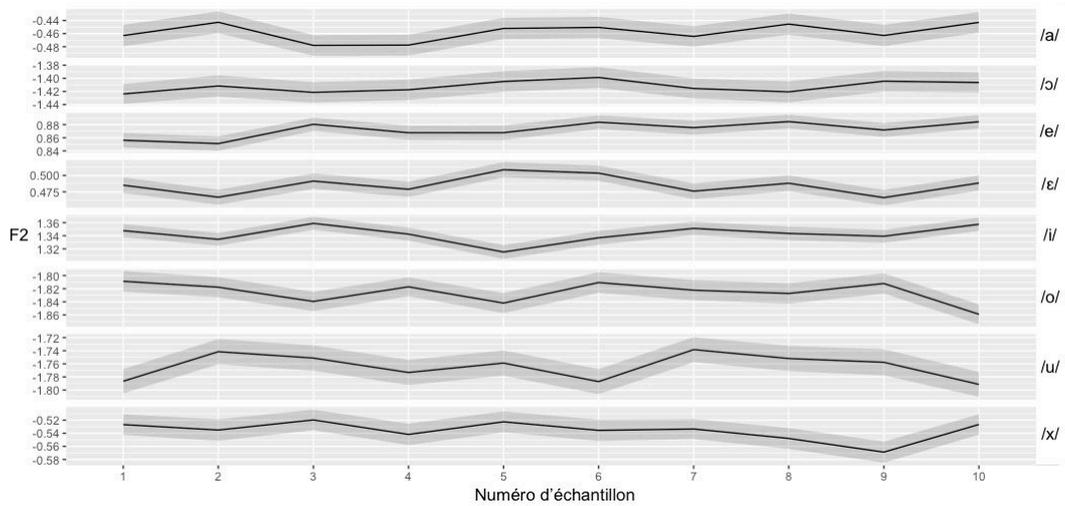
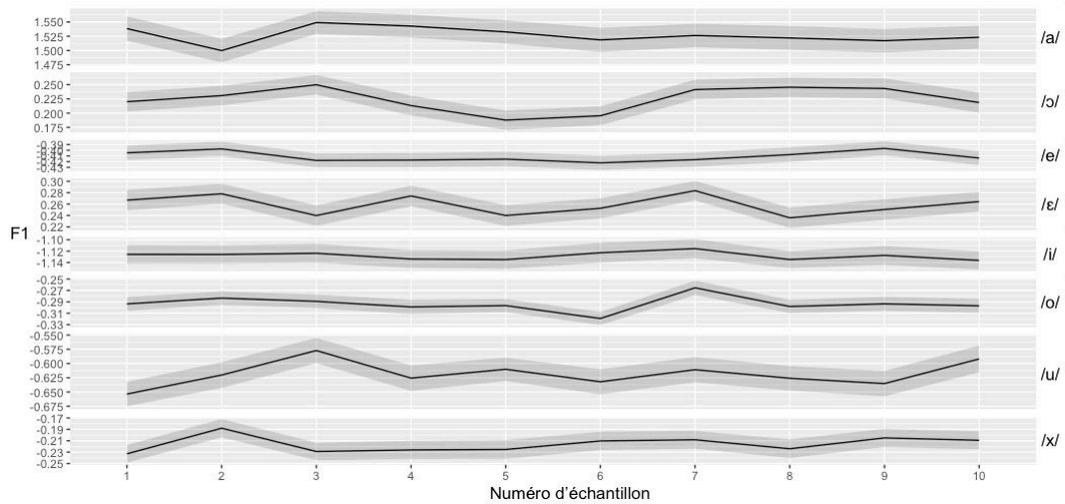
	/ɛ/	198	0,05	0,61	0,52	0,40	-0,09	0,68	-0,11	0,90
	/a/	313	1,19	0,61	-0,59	0,42	-0,21	0,62	0,03	0,80
	/x/	63	-0,55	0,50	-0,41	0,41	-0,85	0,61	-0,55	0,50
	/u/	40	-0,64	0,73	-1,81	0,70	-0,80	1,03	-1,11	0,81
	/o/	33	-0,46	0,36	-1,84	0,40	-0,75	0,82	-0,86	0,63
	/ɔ/	65	0,20	0,55	-1,18	0,50	-0,73	0,68	-0,64	0,58

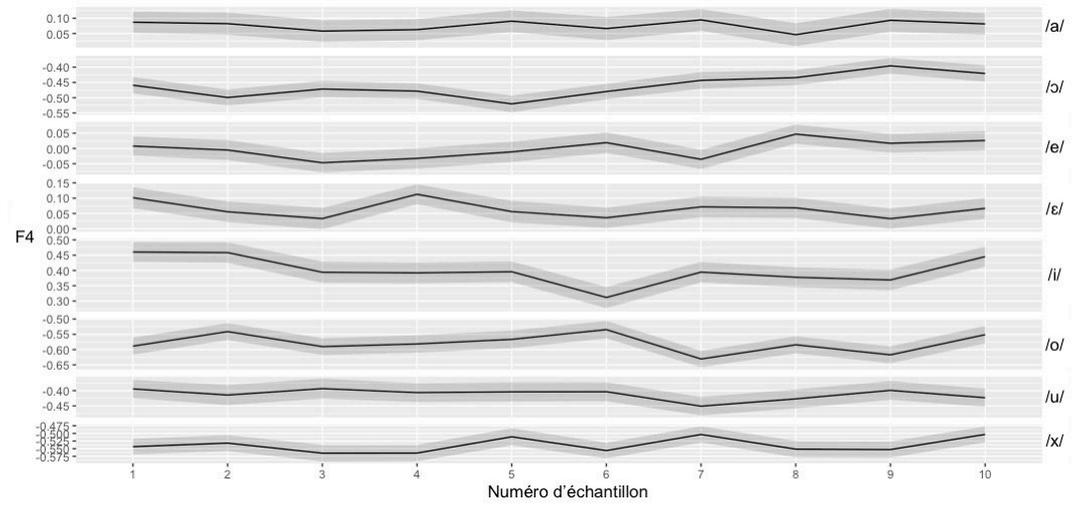
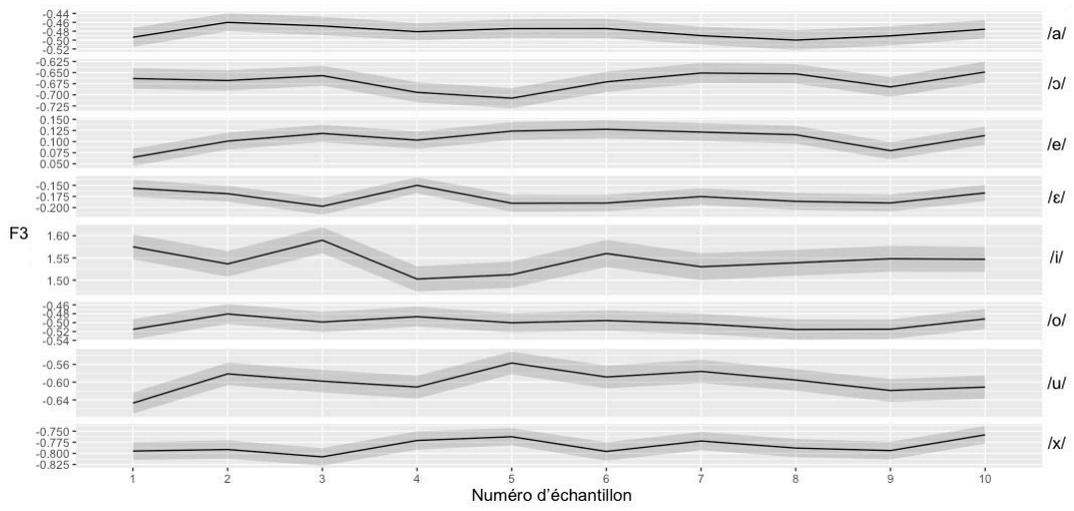
Distribution de la durée des voyelles de 50 à 150 ms en fonction du contexte consonantique alvéolaire (ALV = rose), uvulaire (UV = vert) et vélaire (VEL = bleu)



ANNEXE II. CONTENU DE L'ÉCHANTILLON DE RÉFÉRENCE : CORPUS BREF DE PAROLE LUE

A. VALEURS MOYENNES DES FORMANTS (EN Z-SCORE) AVEC INTERVALLE DE CONFIANCE A +/- UNE ERREUR STANDARD POUR CHAQUE VOYELLE PAR TIRAGE ALÉATOIRE (DE 1 À 10) EN ABSCISSE.





B. DÉTAIL DU CONTENU DE CHAQUE ÉCHANTILLON ALÉATOIRE TIRÉ (T). POUR CHAQUE VOYELLE (V) : VALEUR MOYENNE (M), ÉCART-TYPE (ET), ERREUR STANDARD (ES) ET INTERVALLE DE CONFIANCE 2,5% ET 97, 5% POUR CHACUN DES 4 FORMANTS TESTÉS (F1 ET F2 TABLE 1 ET F3 ET F4 TABLE 2)

TABLE 1.

T	V	F1	F1	F1	F1	F1	F2	F2	F2	F2	F2
		(M)	(ET)	(ES)	(2,5%)	(97,5%)	(M)	(ET)	(ES)	(2,5%)	(97,5%)
1	/a/	1,54	0,60	0,02	0,40	2,58	-0,46	0,46	0,02	-1,32	0,38
2	/a/	1,50	0,59	0,02	0,44	2,64	-0,44	0,47	0,02	-1,35	0,37
3	/a/	1,55	0,56	0,02	0,45	2,58	-0,48	0,44	0,02	-1,31	0,31
4	/a/	1,54	0,56	0,02	0,51	2,54	-0,48	0,45	0,02	-1,34	0,35
5	/a/	1,53	0,57	0,02	0,44	2,57	-0,45	0,46	0,02	-1,31	0,36
6	/a/	1,52	0,60	0,02	0,35	2,63	-0,45	0,46	0,02	-1,34	0,38
7	/a/	1,53	0,58	0,02	0,41	2,54	-0,46	0,44	0,02	-1,32	0,32
8	/a/	1,52	0,59	0,02	0,40	2,61	-0,45	0,46	0,02	-1,25	0,36
9	/a/	1,52	0,59	0,02	0,36	2,54	-0,46	0,46	0,02	-1,33	0,34
10	/a/	1,52	0,56	0,02	0,44	2,60	-0,44	0,45	0,02	-1,31	0,38
1	/ɔ/	0,22	0,48	0,02	-0,63	1,06	-1,42	0,43	0,02	-2,16	-0,57
2	/ɔ/	0,23	0,48	0,02	-0,62	1,13	-1,41	0,46	0,02	-2,16	-0,46
3	/ɔ/	0,25	0,48	0,02	-0,62	1,13	-1,42	0,44	0,02	-2,20	-0,52
4	/ɔ/	0,21	0,49	0,02	-0,63	1,13	-1,42	0,44	0,02	-2,17	-0,53
5	/ɔ/	0,19	0,48	0,02	-0,68	1,04	-1,40	0,45	0,02	-2,16	-0,53
6	/ɔ/	0,20	0,47	0,02	-0,64	1,03	-1,40	0,45	0,02	-2,25	-0,53
7	/ɔ/	0,24	0,49	0,02	-0,63	1,13	-1,42	0,43	0,02	-2,21	-0,53
8	/ɔ/	0,25	0,48	0,02	-0,63	1,19	-1,42	0,45	0,02	-2,17	-0,52
9	/ɔ/	0,24	0,50	0,02	-0,64	1,11	-1,40	0,44	0,02	-2,16	-0,54
10	/ɔ/	0,22	0,50	0,02	-0,60	1,15	-1,41	0,44	0,02	-2,18	-0,50
1	/e/	-0,40	0,36	0,01	-1,03	0,32	0,86	0,31	0,01	0,16	1,39
2	/e/	-0,40	0,36	0,01	-1,04	0,32	0,85	0,31	0,01	0,25	1,40

3	/e/	-0,42	0,34	0,01	-1,04	0,24	0,88	0,29	0,01	0,28	1,42
4	/e/	-0,42	0,36	0,01	-1,02	0,34	0,87	0,31	0,01	0,25	1,49
5	/e/	-0,42	0,37	0,01	-1,08	0,34	0,87	0,31	0,01	0,24	1,43
6	/e/	-0,42	0,35	0,01	-1,04	0,33	0,88	0,29	0,01	0,30	1,40
7	/e/	-0,42	0,35	0,01	-1,03	0,34	0,88	0,30	0,01	0,24	1,40
8	/e/	-0,41	0,36	0,01	-1,05	0,36	0,88	0,29	0,01	0,28	1,41
9	/e/	-0,40	0,35	0,01	-1,03	0,34	0,87	0,30	0,01	0,24	1,44
10	/e/	-0,41	0,35	0,01	-1,07	0,34	0,88	0,29	0,01	0,26	1,41
1	/ɛ/	0,27	0,50	0,02	-0,66	1,25	0,49	0,33	0,01	-0,14	1,14
2	/ɛ/	0,28	0,49	0,02	-0,68	1,19	0,47	0,32	0,01	-0,18	1,10
3	/ɛ/	0,24	0,51	0,02	-0,72	1,23	0,49	0,33	0,01	-0,15	1,11
4	/ɛ/	0,27	0,51	0,02	-0,64	1,25	0,48	0,32	0,01	-0,16	1,10
5	/ɛ/	0,24	0,51	0,02	-0,69	1,18	0,51	0,33	0,01	-0,14	1,13
6	/ɛ/	0,25	0,50	0,02	-0,67	1,17	0,50	0,33	0,01	-0,19	1,13
7	/ɛ/	0,28	0,48	0,02	-0,67	1,18	0,48	0,34	0,01	-0,21	1,13
8	/ɛ/	0,24	0,50	0,02	-0,75	1,26	0,49	0,34	0,01	-0,19	1,13
9	/ɛ/	0,25	0,51	0,02	-0,70	1,20	0,47	0,33	0,01	-0,22	1,07
10	/ɛ/	0,26	0,48	0,02	-0,73	1,14	0,49	0,33	0,01	-0,15	1,12
1	/i/	-1,13	0,46	0,02	-1,95	-0,10	1,35	0,28	0,01	0,76	1,92
2	/i/	-1,13	0,45	0,02	-1,93	-0,17	1,33	0,27	0,01	0,80	1,88
3	/i/	-1,12	0,47	0,02	-1,93	-0,05	1,36	0,28	0,01	0,81	1,92
4	/i/	-1,13	0,45	0,02	-1,88	-0,11	1,34	0,28	0,01	0,80	1,92
5	/i/	-1,13	0,46	0,02	-1,98	-0,21	1,32	0,30	0,01	0,76	1,92
6	/i/	-1,12	0,49	0,02	-1,98	-0,01	1,34	0,30	0,01	0,70	1,90
7	/i/	-1,12	0,47	0,02	-1,91	-0,08	1,35	0,29	0,01	0,75	1,90
8	/i/	-1,13	0,42	0,01	-1,87	-0,11	1,34	0,29	0,01	0,74	1,88
9	/i/	-1,13	0,47	0,02	-2,01	-0,09	1,34	0,28	0,01	0,79	1,87
10	/i/	-1,14	0,45	0,02	-1,88	-0,03	1,36	0,28	0,01	0,79	1,90

1	/o/	-0,29	0,34	0,01	-0,90	0,39	-1,81	0,45	0,02	-2,56	-0,85
2	/o/	-0,28	0,33	0,01	-0,86	0,40	-1,82	0,43	0,02	-2,58	-0,92
3	/o/	-0,29	0,34	0,01	-0,91	0,39	-1,84	0,41	0,01	-2,52	-0,88
4	/o/	-0,30	0,35	0,01	-0,90	0,42	-1,82	0,41	0,01	-2,45	-0,93
5	/o/	-0,30	0,32	0,01	-0,89	0,39	-1,84	0,43	0,02	-2,52	-0,84
6	/o/	-0,32	0,32	0,01	-0,91	0,40	-1,81	0,44	0,02	-2,51	-0,87
7	/o/	-0,27	0,36	0,01	-0,90	0,50	-1,82	0,43	0,02	-2,56	-0,91
8	/o/	-0,30	0,33	0,01	-0,89	0,37	-1,83	0,44	0,02	-2,59	-0,93
9	/o/	-0,29	0,34	0,01	-0,93	0,40	-1,81	0,43	0,02	-2,54	-0,86
10	/o/	-0,30	0,34	0,01	-0,95	0,38	-1,86	0,43	0,02	-2,56	-0,88
1	/u/	-0,65	0,60	0,02	-1,60	0,86	-1,79	0,52	0,02	-2,68	-0,70
2	/u/	-0,62	0,63	0,02	-1,61	0,93	-1,74	0,54	0,02	-2,64	-0,65
3	/u/	-0,58	0,62	0,02	-1,52	0,98	-1,75	0,55	0,02	-2,59	-0,58
4	/u/	-0,63	0,63	0,02	-1,62	0,91	-1,77	0,54	0,02	-2,67	-0,64
5	/u/	-0,61	0,58	0,02	-1,51	0,72	-1,76	0,55	0,02	-2,67	-0,60
6	/u/	-0,63	0,62	0,02	-1,60	0,83	-1,79	0,54	0,02	-2,69	-0,64
7	/u/	-0,61	0,62	0,02	-1,57	0,81	-1,74	0,54	0,02	-2,67	-0,58
8	/u/	-0,63	0,61	0,02	-1,55	0,83	-1,75	0,54	0,02	-2,71	-0,69
9	/u/	-0,64	0,63	0,02	-1,60	0,95	-1,76	0,56	0,02	-2,63	-0,59
10	/u/	-0,59	0,66	0,02	-1,57	0,95	-1,79	0,54	0,02	-2,63	-0,59
1	/x/	-0,23	0,44	0,02	-0,96	0,72	-0,53	0,44	0,02	-1,44	0,25
2	/x/	-0,19	0,45	0,02	-0,94	0,74	-0,54	0,47	0,02	-1,45	0,30
3	/x/	-0,23	0,43	0,02	-0,96	0,69	-0,52	0,44	0,02	-1,37	0,29
4	/x/	-0,23	0,45	0,02	-0,95	0,75	-0,54	0,46	0,02	-1,41	0,30
5	/x/	-0,23	0,46	0,02	-0,96	0,75	-0,52	0,45	0,02	-1,45	0,26
6	/x/	-0,21	0,46	0,02	-1,01	0,71	-0,54	0,46	0,02	-1,41	0,29
7	/x/	-0,21	0,44	0,02	-0,95	0,68	-0,53	0,43	0,02	-1,37	0,19
8	/x/	-0,22	0,46	0,02	-1,00	0,75	-0,55	0,45	0,02	-1,43	0,24
9	/x/	-0,20	0,45	0,02	-0,93	0,71	-0,57	0,45	0,02	-1,43	0,28

10	/x/	-0,21	0,45	0,02	-0,92	0,73	-0,53	0,44	0,02	-1,41	0,23
----	-----	-------	------	------	-------	------	-------	------	------	-------	------

Table 2

T	V	F3	F3	F3	F3	F3	F4	F4	F4	F4	F4
		(M)	(ET)	(ES)	(2,5%)	(97,5%)	(M)	(ET)	(ES)	(2,5%)	(97,5%)
1	/a/	-0,49	0,61	0,02	-1,85	0,52	0,09	0,98	0,03	-1,94	1,87
2	/a/	-0,46	0,57	0,02	-1,64	0,62	0,08	1,03	0,04	-2,00	2,03
3	/a/	-0,47	0,59	0,02	-1,62	0,61	0,06	1,00	0,04	-2,06	1,82
4	/a/	-0,48	0,55	0,02	-1,57	0,58	0,06	0,98	0,03	-1,79	2,01
5	/a/	-0,47	0,59	0,02	-1,77	0,64	0,09	1,01	0,04	-1,97	1,94
6	/a/	-0,47	0,61	0,02	-1,78	0,62	0,07	1,07	0,04	-2,24	2,20
7	/a/	-0,49	0,57	0,02	-1,61	0,63	0,09	1,01	0,04	-2,13	1,90
8	/a/	-0,50	0,63	0,02	-1,77	0,70	0,05	1,05	0,04	-2,30	1,94
9	/a/	-0,49	0,61	0,02	-1,94	0,56	0,09	1,05	0,04	-2,24	2,03
10	/a/	-0,48	0,59	0,02	-1,81	0,63	0,08	1,00	0,04	-2,10	1,91
1	/ɔ/	-0,66	0,66	0,02	-1,94	0,70	-0,46	0,75	0,03	-1,99	0,95
2	/ɔ/	-0,67	0,65	0,02	-1,87	0,70	-0,50	0,74	0,03	-2,09	0,88
3	/ɔ/	-0,66	0,64	0,02	-1,81	0,66	-0,47	0,76	0,03	-2,04	0,89
4	/ɔ/	-0,69	0,64	0,02	-1,96	0,63	-0,48	0,72	0,03	-1,93	0,80
5	/ɔ/	-0,71	0,64	0,02	-1,86	0,58	-0,52	0,76	0,03	-2,11	0,96
6	/ɔ/	-0,67	0,65	0,02	-1,85	0,64	-0,48	0,71	0,03	-1,94	0,90
7	/ɔ/	-0,65	0,64	0,02	-1,92	0,60	-0,44	0,77	0,03	-1,96	1,01
8	/ɔ/	-0,65	0,62	0,02	-1,72	0,63	-0,43	0,69	0,02	-1,74	0,89
9	/ɔ/	-0,68	0,62	0,02	-1,86	0,60	-0,40	0,74	0,03	-1,76	0,98
10	/ɔ/	-0,65	0,66	0,02	-1,87	0,64	-0,42	0,75	0,03	-1,85	1,05
1	/e/	0,06	0,55	0,02	-1,08	1,11	0,01	0,88	0,03	-1,68	1,77

2	/e/	0,10	0,54	0,02	-0,97	1,15	-0,01	0,92	0,03	-1,91	1,76
3	/e/	0,12	0,55	0,02	-0,91	1,19	-0,05	0,90	0,03	-1,91	1,56
4	/e/	0,10	0,56	0,02	-1,08	1,16	-0,03	0,94	0,03	-2,08	1,76
5	/e/	0,12	0,57	0,02	-1,02	1,23	-0,01	0,94	0,03	-1,94	1,72
6	/e/	0,13	0,58	0,02	-1,00	1,21	0,02	0,95	0,03	-1,93	1,80
7	/e/	0,12	0,56	0,02	-0,96	1,26	-0,04	0,88	0,03	-1,86	1,60
8	/e/	0,12	0,57	0,02	-1,10	1,11	0,05	0,88	0,03	-1,78	1,63
9	/e/	0,08	0,55	0,02	-0,97	1,12	0,02	0,86	0,03	-1,67	1,70
10	/e/	0,11	0,58	0,02	-1,07	1,21	0,03	0,89	0,03	-1,82	1,68
1	/ɛ/	-0,16	0,55	0,02	-1,35	0,77	0,10	0,96	0,03	-1,83	1,90
2	/ɛ/	-0,17	0,50	0,02	-1,19	0,67	0,06	0,95	0,03	-1,96	1,84
3	/ɛ/	-0,20	0,54	0,02	-1,32	0,77	0,03	0,99	0,04	-2,19	1,82
4	/ɛ/	-0,15	0,50	0,02	-1,21	0,76	0,11	0,91	0,03	-1,82	1,91
5	/ɛ/	-0,19	0,54	0,02	-1,34	0,75	0,06	0,99	0,04	-1,92	1,76
6	/ɛ/	-0,19	0,52	0,02	-1,32	0,72	0,04	0,94	0,03	-1,83	1,85
7	/ɛ/	-0,18	0,53	0,02	-1,27	0,77	0,07	0,95	0,03	-1,90	1,84
8	/ɛ/	-0,19	0,54	0,02	-1,29	0,78	0,07	0,93	0,03	-1,82	1,95
9	/ɛ/	-0,19	0,53	0,02	-1,26	0,75	0,03	0,93	0,03	-1,78	1,76
10	/ɛ/	-0,17	0,52	0,02	-1,20	0,82	0,07	0,96	0,03	-1,84	1,92
1	/i/	1,58	0,79	0,03	-0,18	2,93	0,46	0,91	0,03	-1,38	2,28
2	/i/	1,54	0,81	0,03	-0,32	2,87	0,46	0,95	0,03	-1,35	2,27
3	/i/	1,59	0,83	0,03	-0,18	2,99	0,39	0,98	0,03	-1,62	2,22
4	/i/	1,50	0,80	0,03	-0,24	2,89	0,39	0,96	0,03	-1,34	2,44
5	/i/	1,51	0,84	0,03	-0,26	2,90	0,40	0,95	0,03	-1,42	2,34
6	/i/	1,56	0,86	0,03	-0,21	3,02	0,31	0,98	0,03	-1,64	2,21
7	/i/	1,53	0,86	0,03	-0,28	2,93	0,39	0,94	0,03	-1,43	2,38
8	/i/	1,54	0,82	0,03	-0,23	2,95	0,38	0,91	0,03	-1,52	2,10
9	/i/	1,55	0,83	0,03	-0,11	3,00	0,37	0,95	0,03	-1,43	2,26

10	/i/	1,55	0,80	0,03	-0,21	2,85	0,45	0,93	0,03	-1,32	2,27
1	/o/	-0,52	0,64	0,02	-1,71	0,85	-0,59	0,79	0,03	-2,09	1,02
2	/o/	-0,48	0,63	0,02	-1,75	0,78	-0,54	0,78	0,03	-1,96	1,02
3	/o/	-0,50	0,65	0,02	-1,75	0,94	-0,59	0,77	0,03	-2,13	0,80
4	/o/	-0,49	0,63	0,02	-1,67	0,80	-0,58	0,81	0,03	-2,06	0,89
5	/o/	-0,50	0,61	0,02	-1,66	0,75	-0,57	0,81	0,03	-2,19	0,98
6	/o/	-0,50	0,65	0,02	-1,70	0,85	-0,53	0,80	0,03	-2,12	0,89
7	/o/	-0,50	0,65	0,02	-1,69	0,80	-0,63	0,76	0,03	-1,98	0,84
8	/o/	-0,52	0,64	0,02	-1,76	0,86	-0,58	0,78	0,03	-1,98	0,93
9	/o/	-0,51	0,62	0,02	-1,68	0,75	-0,62	0,78	0,03	-1,98	0,92
10	/o/	-0,49	0,64	0,02	-1,66	0,78	-0,55	0,81	0,03	-2,02	0,93
1	/u/	-0,65	0,67	0,02	-1,86	0,88	-0,39	0,84	0,03	-2,00	1,32
2	/u/	-0,58	0,72	0,03	-1,97	0,94	-0,41	0,93	0,03	-2,05	1,43
3	/u/	-0,60	0,71	0,03	-1,93	0,97	-0,39	0,92	0,03	-2,04	1,37
4	/u/	-0,61	0,73	0,03	-1,93	0,89	-0,41	0,87	0,03	-2,04	1,44
5	/u/	-0,56	0,72	0,03	-1,86	0,89	-0,40	0,89	0,03	-2,13	1,29
6	/u/	-0,59	0,73	0,03	-1,84	0,99	-0,40	0,87	0,03	-1,96	1,37
7	/u/	-0,58	0,74	0,03	-1,90	1,01	-0,45	0,86	0,03	-2,21	1,26
8	/u/	-0,60	0,68	0,02	-1,76	0,87	-0,43	0,87	0,03	-2,11	1,27
9	/u/	-0,62	0,72	0,03	-1,84	1,04	-0,40	0,87	0,03	-2,04	1,32
10	/u/	-0,61	0,74	0,03	-1,97	0,91	-0,42	0,84	0,03	-2,11	1,32
1	/x/	-0,79	0,57	0,02	-1,86	0,36	-0,54	0,72	0,03	-1,99	0,91
2	/x/	-0,79	0,59	0,02	-1,88	0,41	-0,53	0,74	0,03	-1,99	0,93
3	/x/	-0,81	0,56	0,02	-1,87	0,31	-0,56	0,76	0,03	-2,06	0,99
4	/x/	-0,77	0,58	0,02	-1,80	0,42	-0,56	0,75	0,03	-1,97	1,03
5	/x/	-0,76	0,55	0,02	-1,75	0,41	-0,51	0,79	0,03	-1,95	1,11
6	/x/	-0,80	0,59	0,02	-1,86	0,38	-0,56	0,72	0,03	-1,92	0,83
7	/x/	-0,77	0,58	0,02	-1,86	0,34	-0,50	0,75	0,03	-1,96	0,96
8	/x/	-0,79	0,58	0,02	-1,82	0,40	-0,55	0,75	0,03	-1,99	0,91

9	/x/	-0,79	0,57	0,02	-1,86	0,32	-0,55	0,74	0,03	-1,95	0,97
10	/x/	-0,76	0,56	0,02	-1,72	0,36	-0,50	0,74	0,03	-1,93	1,10

ANNEXE III. DÉTAIL DU CLASSIFIEUR AUTOMATIQUE CONSTRUIT À PARTIR DE L'ÉCHANTILLON DE RÉFÉRENCE SUR LE CORPUS DE PAROLE LUE (BREF)

Formule :

lda(Catégorie de voyelle \sim f1_zscore + f2_zscore + f3_zscore + f4_zscore, data = Échantillon d'entraînement)

Probabilité a priori des groupes :

/i/	/e/	/ɛ/	/a/	/x/	/u/	/o/	/ɔ/
0.1250567	0.1251349	0.1250723	0.1250567	0.1251349	0.1242903	0.1251349	0.1251193

Moyennes des groupes :

	f1_zscore	f2_zscore	f3_zscore	f4_zscore
/i/	-1.1288174	1.3424430	1.5432240	0.400004253
/e/	-0.4109690	0.8723427	0.1068837	-0.001433618
/ɛ/	0.2583889	0.4852647	-0.1773032	0.063449133
/a/	1.5266708	-0.4582752	-0.4809231	0.076269284
/x/	-0.2157975	-0.5358252	-0.7834897	-0.538286967
/u/	-0.6234873	-1.7695769	-0.6072896	-0.411574051
/o/	-0.2934990	-1.8256019	-0.5001616	-0.578599741
/ɔ/	0.2245137	-1.4125974	-0.6697242	-0.460202951

Coefficients des fonctions linéaires discriminantes :

	LD1	LD2	LD3	LD4
f1_zscore	-0.12312442	2.0209261	0.59408272	0.2564492
f2_zscore	-2.34043312	0.5406904	-0.63304365	0.1200166
f3_zscore	-0.39766364	-0.6430704	1.38758079	0.4441347
f4_zscore	0.05741779	0.1514055	0.06662624	-1.1826280

Proportion de la variance expliquée par chacune des fonctions linéaires discriminantes :

LD1	LD2	LD3	LD4
0.7293	0.2456	0.0241	0.0009

Coarticulation C-à-V en français : interaction avec le type de voyelle, la position prosodique et le style de parole

Cette thèse étudie la coarticulation C-à-V en français et son interaction avec d'autres sources de variations dans le but de mieux comprendre ce qui la module et ce qui gouverne la variation dans la parole. Pour cela, à partir de grands corpus de parole, nous avons testé comment la coarticulation C-à-V était fonction : 1) des caractéristiques articulatoires des consonnes et voyelles impliquées à partir de 18.5k voyelles /i, e, ε, a, x, u, o, ə/ (/x/=/ø, œ, ə/) en contexte ALVéolaire, UVulaire et VÉLaire ; 2) de la position prosodique occupée par les voyelles, en comparant le degré de coarticulation de 17k séquences CV et VC, V=/i, e, a, ə/ et C=ALV|UV, en position initiale de groupe intonatif, avec celui de séquences semblables en position interne de mot ; et 3) du style de parole, en analysant le degré de coarticulation dans 22k séquences CV et VC, V= /i, E, a, u, ə/ (/E/= /e, ε/) et C=ALV|UV, issues de parole journalistique et conversationnelle. Cette thèse montre qu'en plus de dépendre des caractéristiques articulatoires des segments, la coarticulation est aussi modulée par des facteurs linguistiques, liés à l'organisation prosodique du message, et des facteurs communicationnels dépendant de la situation de communication. Cependant, certains résultats suggèrent que la modulation de la coarticulation par la position prosodique et le style de parole, ont des fonctions linguistiques différentes dont les implications sur la variation dans la parole seront discutées. Enfin, une réflexion sur les changements de sons en lien avec la préférence universelle pour l'antériorisation des voyelles postérieures fermées sera proposée à partir des différences observées entre les voyelles.

Mots clés : grand corpus, acoustique, coarticulation C-à-V, position prosodique, style de parole, français, voyelles.

C-to-V coarticulation in French: interaction with vowel type, prosodic position and speech style

This dissertation examines C-to-V coarticulation in French and its interaction with others sources of variation in order to better understand what modulates and governs variation in speech. Based on data from large speech corpora, we tested how C-to-V coarticulation is a function of: 1) the articulatory properties of the tested segments, i.e. 18.5k vowels /i, e, ε, a, x, u, o, ə/ (/x/=/ø, œ, ə/) in ALVeoalr, UVular et VELar contexts; 2) the prosodic position occupied by the vowels, comparing the degree of coarticulation of 17k CV and VC sequences V=/i, e, a, ə/ and C=ALV|UV in initial position of intonational phrases, to similar sequences in internal word position; 3) the speech style, by analyzing the degree of coarticulation in 22k CV and VC sequences, V = /i, E, a, u, ə/ (/E/ = /e, ε/) and C = ALV|UV, in journalistic and conversational speech. The thesis demonstrates that coarticulation, in addition to being dependent on the articulatory characteristics of segments, is also modulated by other linguistic factors, related to the prosodic organization of the message, and communicational factors depending on the communication situation. Indeed, vowels are more resistant to coarticulation in strong prosodic positions as in formal speech. However, some results suggest that the modulation of coarticulation by prosodic position and speech style have different linguistic functions whose implications for speech variation will be discussed. Finally, a reflection on sound changes related to the universal preference for the anteriorization of back closed vowels will be proposed from the observed differences between the vowels.

Key words: large corpora, acoustics, C-to-V coarticulation, prosodic position, speech style, French, vowels