



**HAL**  
open science

# Étude acoustique et perceptive de l'harmonie vocalique en français

Juliusz Cecelewski

► **To cite this version:**

Juliusz Cecelewski. Étude acoustique et perceptive de l'harmonie vocalique en français. Linguistique. 2023. dumas-04136664

**HAL Id: dumas-04136664**

<https://dumas.ccsd.cnrs.fr/dumas-04136664v1>

Submitted on 21 Jun 2023

**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.



Distributed under a Creative Commons Attribution 4.0 International License

Université Sorbonne Nouvelle  
Institut de Linguistique et Phonétique Générales et Appliquées (ILPGA)  
Laboratoire de Phonétique et Phonologie, UMR 7018

# Étude acoustique et perceptive de l'harmonie vocalique en français

---

MEMOIRE DE MASTER 2 RECHERCHE MENTION SCIENCES DU LANGAGE

Parcours : Phonétique et Phonologie

Présenté par

**Juliusz CECELEWSKI**

Sous la direction de Cédric Gendrot

Année universitaire 2022-2023

Soutenu le 5 juin 2023 devant le jury composé de :

*Présidente :*

Martine ADDA-DECKER

*Membres :*

Rachid RIDOUANE

Cédric GENDROT

Claire PILLOT-LOISEAU

Jane STUART-SMITH



## Remerciements

En tout premier lieu, je tiens à remercier mon directeur, Cédric Gendrot, pour sa grande disponibilité, ses conseils avisés et l'art avec lequel il a su me guider. Sa présence infaillible et ses encouragements continus m'ont ouvert bien des portes.

Ce travail n'aurait pu être mené à bien sans la générosité et la bienveillance dont Martine Adda-Decker et Philippe Boula de Mareüil du LISN ont fait preuve à mon égard. Leur expertise et leurs conseils avisés ont été des guides importants dans la réalisation de ce travail de recherche.

Mes vifs remerciements vont aussi à Claire Pillot-Loiseau et Rachid Ridouane, qui m'ont fait l'honneur de bien vouloir évaluer mon travail et qui ont eu l'amabilité d'accepter de participer à ce jury.

I extend my heartfelt gratitude to Jane Stuart-Smith, whose groundbreaking scientific work has been an intellectual inspiration for me throughout this study. I am deeply appreciative of her willingness to participate as an invited member of the jury.

## RÉSUMÉ

Dans la tradition phonologique française, le terme d'*harmonie vocalique* (HV) désigne un phénomène d'assimilation régressive qui affecte les voyelles moyennes  $V_1 \in \{e \ \varepsilon \ o \ \circ \ \emptyset \ \text{œ}\}$  en syllabe pénultième sous l'influence de la voyelle finale.

Une étude acoustique de trois corpus (BREF, ESTER et NCCFr) a permis d'apporter un nouvel éclairage sur l'impact du style de parole sur l'HV. Nous montrons que la mécanique exacte de cette interaction est plus complexe que ce qui était supposé et posons la nécessité de distinguer une *HV fermante* et une *HV ouvrante* (c'est-à-dire agissant, respectivement, sur les voyelles moyennes mi-ouvertes et mi-fermées), afin de prendre en compte la variabilité de ce phénomène selon la  $V_1$  concernée.

De plus, une expérience à base d'imitation a été conçue en vue de fournir une validation supplémentaire de l'hypothèse selon laquelle la diminution significative de l'HV en parole journalistique entre les années 1940-50 et les périodes plus récentes ne peut être attribuée au style emphatique des speakers d'époque. Nous démontrons qu'il s'agit d'un changement phonétique propre au français standard, en continuité avec la tendance récente à réduire le degré d'HV.

En outre, une expérience de perception a été réalisée dans le paradigme de la complétion de mot, qui a permis de montrer que l'information acoustique véhiculée par l'HV peut être exploitée par l'auditeur pour faciliter l'accès au lexique.

Enfin, nous postulons que l'HV en français correspond à une *coarticulation fossilisée*, à mi-chemin entre un processus phonologique et une coarticulation régie par les lois phonétiques. La généalogie du terme d'*harmonie vocalique* dans la tradition descriptive française est évoquée en guise d'éclairage supplémentaire.

---

**MOTS-CLÉS : harmonie vocalique, coarticulation fossilisée, changement phonétique, voyelles moyennes, phonétique acoustique, phonétique perceptive**

---

## ABSTRACT

In the French phonological tradition, the term 'vowel harmony' (VH) refers to a regressive assimilation phenomenon that affects the mid vowels  $V_1 \in \{e \ \varepsilon \ o \ \circ \ \emptyset \ \text{œ}\}$  in the penultimate syllable under the influence of the final vowel.

An acoustic study of three speech corpora (BREF, ESTER and NCCFr) sheds new light on the impact of speech style on VH. We demonstrate that the exact mechanics of this interaction are more complex than previously assumed and argue for the need to distinguish between 'closing VH' and 'opening VH', which act on open-mid and close-mid vowels, respectively. This distinction is important for accounting for the variability of this phenomenon depending on the  $V_1$  involved.

Furthermore, an imitation-based experiment was designed to provide additional validation of the hypothesis that the significant decrease in VH in journalistic speech between the 1940s-1950s and more recent periods cannot be attributed solely to the emphatic style of speakers from that era. We demonstrate that it is a phonetic change specific to Standard French, in line with the recent trend of reducing the degree of VH.

Additionally, a perception experiment was conducted using a word completion paradigm, which showed that the acoustic information conveyed by VH can be exploited by the listener to facilitate lexical access.

Finally, we postulate that VH in French corresponds to 'fossilized coarticulation', situated midway between a phonological process and coarticulation governed by phonetic laws. The genealogy of the term 'vowel harmony' in the French descriptive tradition is discussed for further illumination.

---

**KEY-WORDS: vowel harmony, fossilized coarticulation, sound change, mid-vowels, acoustic phonetics, auditory phonetics**

---

## Table des matières

1. Introduction .....	- 1 -
2. Études antérieures sur l'harmonie vocalique	
2.1. Travaux historiques et témoignages anciens.....	- 4 -
2.2. « <i>Is there vowel harmony in French ?</i> » – Études sur des corpus enregistrés en laboratoire .....	- 5 -
2.3. « Que nous apprennent les gros corpus sur l'harmonie vocalique en français ? ».....	- 9 -
2.4. Harmonie vocalique en perspective diachronique .....	- 13 -
2.5. Hypothèses et objectifs .....	- 14 -
3. Harmonie vocalique dans les grands corpus	
3.1. Méthode	
3.1.1. Corpus.....	- 16 -
3.2. Résultats	
3.2.1. Harmonie vocalique sur /E, O/ dans les corpus BREF, ESTER et NCCFr.....	- 19 -
3.2.2. Harmonie vocalique ouvrante sur /e, o/ dans les corpus BREF, ESTER et NCCFr.....	- 21 -
3.2.3. Harmonie vocalique fermante sur /ɛ, ə/ dans les corpus BREF, ESTER et NCCFr .....	- 23 -
3.2.4. Harmonie vocalique versus assimilation de voyelle à voyelle dans des syllabes initiales prétoniques internes .....	- 24 -
3.2.5. Évolution d'harmonie vocalique : étude en temps apparent du corpus BREF.....	- 27 -
4. Expérience de production	
4.1. Méthode	
4.1.1. Corpus.....	- 30 -
4.1.2. Sur la <i>loi de position</i> en français standard : stratégie de transcription.....	- 31 -
4.1.3. Locuteurs et enregistrements .....	- 33 -
4.1.4. Mesures acoustiques.....	- 34 -
4.2. Résultats	
4.2.1. Système des voyelles orales dans les productions des six locuteurs .....	- 34 -
4.2.2. Harmonie vocalique ouvrante sur /e, o/.....	- 35 -
4.2.3. Harmonie vocalique fermante sur /ɛ, ə/.....	- 40 -
4.2.4. Harmonie vocalique sur /E, œ, O, A/ .....	- 44 -
5. Expérience de perception	
5.1. Coarticulation, métaphonie et harmonie vocalique dans l'optique de la phonétique auditive	
5.1.1. Coarticulation de voyelle à voyelle et la perception de la parole.....	- 47 -
5.1.2. Harmonie vocalique et traitement perceptif du signal.....	- 48 -
5.1.3. Métaphonie et la reconnaissance de mots .....	- 49 -
5.1.4. Deux expériences de perception sur l'harmonie vocalique en français .....	- 51 -
5.2. Objectifs et hypothèses .....	- 52 -
5.3. Méthode	
5.3.1. Stimuli .....	- 53 -
5.3.2. Tâche de complétion de mot.....	- 55 -

<b>5.4. Résultats</b>	
<b>5.4.1. Participants.....</b>	<b>- 56 -</b>
<b>5.4.2. Effet des caractéristiques acoustiques des stimuli sur la décision des participants.....</b>	<b>- 56 -</b>
<b>5.4.3. Étude de l'effet du degré d'harmonie vocalique sur la capacité à anticiper la fin du mot .....</b>	<b>- 57 -</b>
<b>6. Discussion</b>	
<b>6.1. Conclusions générales.....</b>	<b>- 65 -</b>
<b>6.2. Évolution d'harmonie vocalique entre 1940 et 2022 .....</b>	<b>- 67 -</b>
<b>6.3. Entre <i>harmonisation</i> phonétique et <i>harmonie</i> phonologique.....</b>	<b>- 70 -</b>
<b>6.4. Harmonie vocalique et l'aménagement du système vocalique français .....</b>	<b>- 74 -</b>
<b>6.5. Vers une perte des distinctions des voyelles à double timbre ? .....</b>	<b>- 76 -</b>
<b>7. Conclusion .....</b>	<b>- 78 -</b>
<b>8. Bibliographie.....</b>	<b>- 80 -</b>
<b>9. Annexes.....</b>	<b>- 88 -</b>

# 1. Introduction

Lorsqu'en 1989, Fónagy écrivait, en préambule à une synthèse impressionniste consacrée à des innovations phonétiques récentes, que : « Un professeur de français revenant à Paris de Prague ou de Kyoto, après une absence de vingt ans, serait probablement déconcerté devant certains néologismes phonétiques » (Fónagy, 1989), il évoquait, sans guère y prêter d'attention, une question complexe et, à notre sens, relativement peu considérée dans la littérature, qui est celle de savoir quel empan temporel minimal doit s'étendre entre deux observations pour que l'écart significatif mis en évidence par une expérience robuste puisse être considéré comme un changement signifiant et non plus comme relevant d'un « *pool of synchronic variation* » (Ohala, 1993), ou encore comme un « néologisme phonétique », pour reprendre le terme imagé utilisé par Fónagy justement – nous semble-t-il – pour contourner celui de « changement ». Autant l'espace de près d'un siècle, tout aussi négligeable dans l'optique de la phonologie diachronique traditionnelle qu'inatteignable, ou guère, pour la phonétique expérimentale pour des raisons matérielles évidentes, paraît désormais être un cadre temporel possible à embrasser expérimentalement grâce notamment à des méthodologies hybrides, études en temps réel et en temps apparent (Sóskuthy & Stuart-Smith, 2020), autant, un empan chronologique plus bref, d'à peine quelques décennies, laisse place à davantage d'ambiguïté, en situant les faits observés à mi-chemin entre une tendance périssable et un véritable changement.

Une ambiguïté qui ne devient que plus prégnante dès lors que l'évolution que l'on vise à quantifier expérimentalement concerne un objet phonétique lui-même hautement variable, et dont l'image acoustique est la résultante de tout un faisceau de facteurs intriqués dont chacun se caractérise par une dynamique propre. Ainsi, autant le vocalisme français peut être considéré comme riche, voire « saturé » (Vallée *et al.*, 1999), avec une quinzaine de segments distinctifs, « [les] voyelles moyennes occupent une position unique [en son sein] et leur instabilité est bien attestée » (Landick, 2004). Aux lois de position et plusieurs autres règles énoncées par les phonologues s'ingéniant à décrire la distribution des paires *e~ε*, *o~ɔ*, *ø~æ* s'ajoute un phénomène d'assimilation anticipatoire du degré d'aperture qui affecte les voyelles moyennes en syllabe pénultième sous l'influence de la voyelle finale de mot. Un processus qui n'a point échappé à l'attention des pionniers de la phonétique expérimentale, à une période où le terme inexact d'*harmonie vocalique* qui allait lui être attribué plus tard (cf. Dell, 1973) n'était pas encore énoncé :

(...) ou bien une syllabe tonique devient atone, soit dans la dérivation (tête, têtu) et la conjugaison (pêche, pêcher) (...). La voyelle, ainsi modifiée par rapport à l'accent, subit un changement de timbre : (...)<sup>1</sup> *ò*<sup>1</sup> et *ó* en *o* : (*kòr* « corps », *korsè* « corset ») ; (...) *è* se change en *é* (*tèt*, *tétu* « têtu » ; *mè* et *mé wi* « mais oui »). Et *æ* en *é* (*plæ̀r* et *plé`ré* « pleurer ») (Rousselot & Laclotte, 1902)

---

<sup>1</sup> Rousselot & Laclotte discernent, en partant d'une analyse phonétique fine et subtile, cinq catégories de voyelles moyennes dont trois « bien nettes » (1902), en suivant les variations du degré d'aperture, qu'ils transcrivent : « è, è, ò » pour la série ouverte, lorsque la voyelle moyenne précède un r en syllabe fermée, e.g. *paire*, *heure*, *encore*, « e, œ, o » pour la série moyenne, e.g. *église*, *jeune*, *votre*, enfin « é, é, ó » pour la série fermée, e.g. *chanté*, *jeûne*, *gigot* (Rousselot & Laclotte, 1902).

La motivation initiale du travail de recherche dont les résultats seront présentés ici était de s'inscrire dans la continuité de notre étude consacrée à l'évolution de l'harmonie vocalique en français depuis 1940, en perspective diachronique et en temps réel (Cecelewski, 2022), notamment en apportant une démonstration complémentaire à l'appui de l'hypothèse, énoncée dans Cecelewski *et al.* (à paraître), selon laquelle une forme d'assimilation particulièrement forte en parole journalistique dans les années 1940 et 1950 ne saurait être due, ou du moins ne le saurait-elle que partiellement, au style emphatique, ou style « Gaumont-Pathé » (cf. Boula de Mareüil *et al.*, 2008, 2012a, 2012b), propre à la parole des speakers radiophoniques d'alors.

Une expérience de production a ainsi été conçue en vue de reproduire, dans les conditions de laboratoire, l'influence éventuelle de ce style de parole particulier sur la réalisation phonétique de l'harmonie vocalique, en comparant des productions de six locuteurs en *voix modale* et en *voix projetée*.

Une autre motivation sous-tendait cette première expérience avec, toujours, une perspective diachronique pour toile de fond, qui était d'étendre l'analyse du phénomène d'harmonie vocalique à des données plus récentes, récoltées à près de deux décennies d'intervalle après le dernier corpus analysé et, avec toutes les précautions que cela impose, de rallonger davantage l'empan temporel initialement considéré.

Nous avons également saisi cette opportunité pour inclure dans ce nouveau corpus des contextes en /ø, œ/ qui, quoique mentionnés aux côtés de /e, ε, o, ɔ/ comme sujets à l'harmonie vocalique dans une partie de la littérature historique (Straka, 1950, entre autres), n'ont pas fait jusqu'à présent l'objet d'une étude empirique.

Ainsi, en partant d'une optique diachronique initiale, au fur et à mesure de nos résultats, de nouvelles questions de recherche apparaissaient, nous amenant à adopter un point de vue plus synchronique, avec pour vocation de compléter au mieux l'image d'harmonie vocalique là où elle s'avérait incomplète, et où il nous était possible d'apporter une précision grâce à un choix méthodologique différent par rapport aux études antérieures.

C'est ainsi que nous nous sommes proposé d'enquêter sur la directionnalité d'harmonie vocalique, pour distinguer une tendance à l'ouverture d'une voyelle moyenne pénultième mi-fermée sous l'effet d'une voyelle finale ouverte (ou mi-ouverte) de l'effet inverse que subit une voyelle moyenne mi-ouverte en syllabe pénultième sous l'influence d'une voyelle finale fermée (ou mi-fermée). Nous essaierons de montrer que cette distinction, issue de l'outillage définitionnel communément admis dans la classification de phénomènes d'harmonie vocalique dans les langues du monde (Ultan, 1973 ; Ohala, 1994) alors qu'elle avait probablement été jugée superflue ou inadaptée à l'harmonie vocalique en français et, par conséquent, n'a pas été exploitée dans les études antérieures (Fagyal *et al.*, 2003 ; Nguyen & Fagyal, 2008 ; Turco *et al.*, 2016a, 2016b) s'avère être plus qu'un artefact théorique, dans la proportion où le défaut de symétrie entre l'*harmonie vocalique fermante* et



L'harmonie vocalique ouvrante mérite à nos yeux d'être considéré comme l'une des caractéristiques inhérentes à ce phénomène, comme nous tâcherons de l'illustrer ici.

En plus de l'étude du corpus contrôlé enregistré en laboratoire, une analyse comparative du degré d'harmonie vocalique dans trois grands corpus de parole lue (BREF), journalistique (ESTER) et spontanée (NCCFr), a été menée pour apporter cette précision méthodologique. Elle a également permis d'intégrer dans notre analyse un examen plus étendu de la variation du degré d'harmonie vocalique en fonction du registre ou style de parole, séparément pour chacune des quatre voyelles moyennes /e, ε, o, ə/.

Une autre question, que nous n'avons sans doute pas explorée sous toutes ses facettes, mais dont la pertinence nous semble indéniable, a été de savoir s'il était pertinent de conférer à l'harmonie vocalique le statut d'un processus phonologique. En effet, bien que la nature par excellence graduelle et hautement variable d'harmonie vocalique ait été confirmée expérimentalement dès la première étude empirique contemporaine de Fagyal *et al.*, (2003), il est difficile de trouver une description phonologique datant d'avant le début du XX<sup>e</sup> siècle qui n'ait pas fait mention de ce phénomène, à commencer par l'ouvrage classique de Dell (1973), dans lequel elle trouve sa place parmi les règles dérivationnelles régissant la réalisation des voyelles moyennes du français.

Cette question sera, dans notre travail, scindée en deux autres, moins directes dans leur formulation, mais qui, dans le fond, ne viseront qu'à désambiguïser la réponse à cette première, à savoir : l'harmonie vocalique se distingue-t-elle d'autres formes de coarticulation de voyelle à voyelle, dont le statut *non phonologique* laisse planer moins de doute ? Notre démarche ne sera pas ici de chercher à démontrer que l'harmonie vocalique n'est pas un processus catégoriel ou binaire : chose que nous pressentons au moins depuis Grammont (1914) et dont nous avons une preuve empirique robuste depuis Fagyal *et al.* (2003). Il ne s'agira pas non plus d'explorer la mécanique complexe et subtile qui sous-tend la coarticulation elle-même. Notre approche sera plutôt celle de déceler la nature propre d'harmonie vocalique en la comparant à des phénomènes de coarticulation avec qui elle partage des mécanismes phonétiques, mais qui ne font pas partie de sa définition phonologique.

La dernière question irrésolue que nous avons abordée, suggérée initialement par Nguyen *et al.*, (2004) et formulée plus explicitement par Turco *et al.*, (2016a) était de savoir dans quelle mesure les auditeurs exploitent cognitivement l'harmonie vocalique pour améliorer l'accès lexical. Une expérience de perception s'appuyant sur le paradigme de complétion de mot a ainsi été menée pour essayer de combler cette lacune dans nos connaissances sur le rôle perceptuel de l'harmonie vocalique, et dont les résultats apporteront un élément de réponse complémentaire pour discuter sur la place éventuelle de l'harmonie vocalique dans le système phonologique du français.

Notre travail se déclinera dès lors en deux parties. Dans un premier temps, nous aborderons des aspects acoustiques de la réalisation d'harmonie vocalique. Une analyse de ce phénomène dans les corpus BREF, ESTER et NCCFr préludera à une étude acoustique d'un corpus bien contrôlé enregistré

en laboratoire. Dans un second temps, nous présenterons les résultats d'une étude perceptive sur l'harmonie vocalique en français. Chacune des deux parties sera précédée d'une revue de résultats apportés par des études empiriques antérieures sur l'harmonie vocalique dans la dimension acoustique et, séparément, dans la dimension perceptive. Une synthèse des résultats obtenus sera enfin proposée, en vue, sinon d'approfondir et d'éclairer certains aspects acoustiques et perceptifs d'harmonie vocalique en français, au moins d'attester d'un réel intérêt scientifique que représente l'analyse empirique fine de ce phénomène complexe et éphémère, en l'inscrivant dans un cadre problématique plus large de l'évolution des voyelles moyennes en français.

## **2. Études antérieures sur l'harmonie vocalique**

### **2.1. Travaux historiques et témoignages anciens**

Dans Cecelewski (2022), nous avons réalisé un survol historique de l'harmonie vocalique en français. À ce titre, nous avons choisi de faire dialoguer des travaux pionniers en phonétique, des témoignages d'antan ainsi que des études impressionnistes. Dans le même temps, nous nous étions efforcés d'établir une problématique permettant non seulement de fédérer ces différents travaux, mais également de mettre en avant leurs perspectives communes et leurs divergences.

Abordé sous différentes termes tels que *dilation*, *harmonisation*, *inflexion* ou encore *brisure vocalique*, l'harmonie vocalique était ainsi décrite comme un processus faisant varier, selon l'acception, une, deux ou trois voyelles moyennes sous l'effet d'un impact distant de la voyelle finale de mot. Bien que les voyelles moyennes antérieures non-arrondies /e, ε/ sont souvent considérées comme les représentantes prototypiques de cette harmonie, les voyelles moyennes postérieures /o, ɔ/ n'ont pas pour autant été ignorées par les premiers phonéticiens. Les voyelles moyennes antérieures arrondies, en l'occurrence /ø/ et /œ/, n'ont suscité, quant à elles, qu'un intérêt restreint de la part des chercheurs étudiés (Straka, 1950 ; Selkirk, 1972 ; Tranel, 1987). Il importe également de noter une certaine discordance des études impressionnistes quant au répertoire exact des voyelles finales provoquant l'assimilation, à sa dimension synchronique, chez les uns (Fouché, 1956) et diachronique chez d'autres (Grammont, 1914), enfin à sa nature, tantôt phonologique, systémique et régulière, tantôt phonétique, due à l'imprécision de la coordination motrice des articulateurs dans un style d'élocution peu soutenu.

Nous avons ainsi entrepris de retracer les premières mentions de la variabilité des voyelles moyennes en fonction de la voyelle suivante, en remontant aux traités de prononciation du XVIII<sup>e</sup> siècle, et en examinant les évolutions terminologiques, descriptives et théoriques de l'harmonie vocalique depuis lors. Nous avons accordé une attention particulière aux premiers phonéticiens impressionnistes du début du XX<sup>e</sup> siècle, ainsi qu'aux contributions majeures de la phonologie moderne émergente, jusqu'aux premières études expérimentales des années 1970 (Boumendil-Lucot, 1977 ; Robert, 1977).

Nous avons également présenté un aperçu historique du phénomène d'harmonie vocalique, en soulignant que cette appellation n'impliquait pas une *harmonie vocalique* au sens strict (Ultan, 1973). Bien que phonétiquement similaire à la *métaphonie* fréquente dans les langues romanes (Calabrese, 2011), la nature morphophonologique de l'harmonie vocalique en français demeurerait incertaine. Si elle se rapprochait également, par sa nature phonétique, du *umlaut* germanique (Grosvald, 2010), sa synchronicité ne correspondait pas au statut de loi diachronique de ce dernier. Il est à supposer que de nos jours, elle serait considérée comme une forme fossilisée et renforcée de *coarticulation* à distance (*transconsonantique*) si seulement le terme d'harmonie vocalique ne s'était pas déjà ancré dans la tradition phonologique française bien avant que le terme de *coarticulation* ne fût forgé et diffusé (cf. Kühnert et Nolan, 1999, pour un aperçu historique).

En revanche, dans Cecelewski (2022), nous avons accordé moins d'importance aux études phonétiques contemporaines sur l'harmonie vocalique du français. Dès le début du XXI<sup>e</sup> siècle, celles-ci ont pourtant fourni des données empiriques éclairant la nature acoustique et perceptive de ce phénomène. L'objectif de ce chapitre est donc de fournir une synthèse aussi exhaustive que possible de ces travaux, dont les conclusions nous conduiront ensuite à formuler les hypothèses de notre propre travail de recherche.

## **2.2. « *Is there vowel harmony in French ?* » – Études sur des corpus enregistrés en laboratoire**

Il y a tout juste vingt ans, la première étude expérimentale ayant véritablement posé la question de savoir si une forme d'harmonie vocalique était présente en français a été celle de Fagyal *et al.* (2003) qui allait précéder l'étude complémentaire (et plus complète) de Nguyen & Fagyal (2008). L'objectif de Fagyal *et al.* (2003) a été d'apporter des preuves empiriques à l'appui du phénomène d'assimilation des voyelles moyennes  $V_1 \in \{e, \varepsilon, o, \circ\}$  et, plus généralement, de contribuer à combler le « *general lack of evidence on actual realizations of mid vowels in different dialects of French* » (Fagyal *et al.*, 2003). Mais l'enjeu était également d'aborder la question de l'adéquation du terme d'*harmonie vocalique* utilisé pour décrire ce phénomène d'assimilation qui, comme l'ont suggéré les auteurs de l'étude, appellerait une analyse en termes d'umlaut ou de métaphonie, à condition de pouvoir être relié à des environnements morphologiques spécifiques. La question de l'impact de l'interface morphophonologique a ainsi été posée, sans pour autant être mise en examen de façon systématique. Néanmoins, les auteurs se sont proposés de comparer l'harmonie vocalique dans deux variétés géographiques différentes, en construisant le corpus de façon à rendre compte d'une éventuelle variation diatopique.

Ont été considérées les productions des trois femmes âgées de 28 à 34 ans, dont deux représentant une variété « *typical of the Île-de-France variety of Northern Metropolitan French* » (Fagyal *et al.*, 2003) et la troisième une variété de français du Sud parlée à Aix-en-Provence. Dans

cette étude préliminaire, l'harmonie vocalique a été définie en termes d'assimilation d'antériorité et non de hauteur. C'est pourquoi l'analyse acoustique a été limitée à la hauteur du second formant. Ce choix méthodologique a également imposé de distinguer plusieurs sous-groupes de paires harmoniques en fonction non seulement du degré d'aperture, mais également de celui de l'antériorité des voyelles-cibles afin d'analyser séparément les contextes suivants :  $V_1-e/\epsilon$ ,  $V_1-\emptyset/\text{œ}$ ,  $V_1-i/a$ ,  $V_1-o/\text{ɔ}$ ,  $V_1-e/\epsilon$  et  $V_1-e/\epsilon$ . En revanche, la distinction entre les voyelles mi-fermées et mi-ouvertes en position prétonique n'a pas été respectée, et la question de la directionnalité du processus d'assimilation (*harmonie fermante* versus *ouvrante*), quoique mentionnée dans l'introduction, n'a pas été examinée. En effet, la stratégie de transcription adoptée a consisté à considérer les voyelles moyennes en syllabe ouverte comme étant toutes phonologiquement mi-fermées, en suivant la *loi de position*.

Les résultats ont indiqué une tendance à l'assimilation d'antériorité dans près de 60% à 90% des cas étudiés, entraînant un rehaussement ou un abaissement du second formant allant de 20 Hz jusqu'à 120 Hz pour les résultats significatifs. Les contextes linguistiques pour lesquels les résultats ne se sont pas avérés significatifs ont concerné les sous-classes suivantes : *dévo-t-dévot-e* (e-o/ɔ), *prê-teur-prê-teuse* (e-ø/œ) et *noti-ce-nota* (o-i/a), révélant une faible amplitude d'assimilation de F2 de moins de 5 Hz. La variation dialectale n'a pas pu faire l'objet d'une analyse rigoureuse, étant donné que la locutrice, originaire du Sud de la France, manifestait une tendance à perdre les traits phonétiques spécifiques à la variété du français d'Aix-de-Provence durant la lecture du corpus.

Ce dernier aspect allait être mis en examen de façon plus systématique dans l'étude de Nguyen & Fagyal (2008) qui reprend de manière plus perfectionnée et complète le protocole de l'étude précédente. Six locuteurs (trois hommes et trois femmes) ont été enregistrés dont deux sujets originaires de Salon-de-Provence. Il importe de garder à l'esprit une difficulté de comparer des patrons d'harmonie vocalique dans deux variétés géographiques différentes. En effet, celles-ci découlent des distributions différentes des voyelles moyennes : d'une part, phonémiques dans la variété standard et, d'autre part, allophoniques, distribuées au gré de la *loi de position*, dans la variété du Nord. Les voyelles censées entraîner l'harmonisation étaient ainsi susceptibles d'avoir des timbres différents selon la variété, notamment dans les cas des paires du type *prê-teur-prê-teuse* et *poseur-poseuse*, pouvant être produites [ $p\beta Et\text{œ}\beta$ -  $p\beta Et\phi:z$  ;  $pOz\text{œ}\beta$ - $pOz\phi:z$ ] dans la variété du Nord et [ $p\beta Et\text{œ}\beta$ -  $p\beta Et\text{œ}z$  ;  $pOz\text{œ}\beta$ - $pOz\text{œ}z$ ] en français du Sud. Cette divergence notable dans les productions phonétiques a vraisemblablement constitué une raison suffisante pour ne pas généraliser de manière uniforme les aspects phonologiques pour les deux variétés. En conséquence, chaque sous-catégorie des paires harmoniques a été examinée séparément, dans une perspective plus minutieuse.

Les résultats ont indiqué que les voyelles initiales étaient généralement plus fermées devant une voyelle finale mi-fermée. Les voyelles initiales étaient aussi généralement plus fermées et antérieures devant /i/ et plus ouvertes et postérieures devant /a/. Ces variations ont affecté à la fois les voyelles antérieures et postérieures de la première syllabe, et étaient indépendantes de la présence ou de l'absence d'une frontière morphémique entre les deux syllabes et des fréquences des mots porteurs.

L'assimilation était plus systématique chez certains locuteurs et moins régulière chez d'autres, la frontière séparant les deux variétés géographiques s'étant, quant à elle, révélée particulièrement graduelle créant un spectre de réalisations plutôt que deux schémas clairement délimités : « *at the strongest endpoint of the continuum we find a prototypical Parisian speaker who can be opposed to a prototypical Southern speaker* » (Nguyen & Fagyal, 2003). L'étude a donc mis en évidence la réalité empirique de l'harmonie vocalique présente tant chez les locuteurs du français standard que chez ceux de la variété méridionale.

Outre la variation dialectale (qui s'est révélée être un artefact non approprié à la description de l'harmonie vocalique en français), deux autres aspects ont été abordés en vue de mettre à l'essai certaines des définitions traditionnelles, à savoir l'impact potentiel de la composition morphologique des mots et celui de la fréquence lexicale sur le degré d'assimilation. Concernant ce premier point, certaines descriptions antérieures soutenaient que l'harmonie vocalique était un phénomène limité à des contextes bi-morphémiques qui faisait ainsi varier la voyelle radicale sous l'influence de la voyelle du suffixe, comme dans l'exemple *cède* → *céd-er~céd-ant~céd-a* cité par Dell (1973). En contrepartie, l'harmonie ne serait pas présente dans les mots mono-morphémiques tels que *récolte* ou *prétend* (Nguyen & Fagyal, 2008). Dans le cas d'un phénomène que nous savons désormais de nature *phonétique* plutôt que *phonologique*, cette différenciation paraît d'emblée artificielle. Néanmoins, si elle se révélait juste, alors l'harmonie vocalique en français partagerait un trait typologique commun avec les phénomènes de métaphonie tels que définis pour un nombre conséquent de langues et dialectes romans (Calabrese, 2011). Toutefois, ce problème apparemment clair et bien posé, devient plus complexe que prévu dans le cas des mots dont la composition morphologique varie en fonction de la catégorie morphosyntaxique (p.ex. *dé-fait-e* Adj. ; *dé-faite* N.), de la transparence étymologique des mots. À cet égard, nous pouvons songer à des emprunts récents comme *motel* (bi-morphémique en anglais, *mo(tor)+(ho)tel*, mais monomorphémique en français) ou au contraire, plus anciens, comme *notaire*, issu du latin *notare* mais resegmenté en français en *not-aire*. Pour remédier à ces incertitudes, trois annotateurs professionnels ont été sollicités afin de déterminer le nombre de morphèmes dans les formes parlées des mots cibles utilisés dans l'étude. Ainsi, le degré d'harmonie vocalique dans les paires des mots bi-morphémiques du type *prêteur-prêteuse* a été comparé à celui dans les mots monomorphémiques comme *été-éther*, conduisant les chercheurs à la conclusion qu'une faible corrélation entre la magnitude de l'effet assimilatoire et la structure morphologique était présente dans les données analysés avec, conformément à la définition de Dell (1973), une tendance à l'harmonisation légèrement plus importante dans les mots bi-morphémique. Il est à propos de noter cependant que « *that most of the correlation coefficients were statistically non-significant* » (Nguyen & Fagyal, 2008), allant même jusqu'à une corrélation négative (opposée à la tendance générale) dans certains cas étudiés.

Le dernier aspect abordé, et que nous tenons à mentionner ne serait-ce que brièvement puisqu'il n'a plus été réexaminé depuis, a concerné un impact éventuel de la fréquence lexicale sur les schémas

d'assimilation, supposément susceptible d'occasionner une réduction vocalique plus forte pour les mots « faciles », c'est-à-dire les plus fréquents, comme cela avait été précédemment démontré par Wright, (2003). La fréquence lexicale moyenne a ainsi été calculée pour chaque paire de mots du corpus, pour être ensuite mise en rapport avec le degré d'harmonie vocalique. Pourtant, les résultats n'ont pas été concluants. Comme le précisent les auteurs : « *in general, the correlation between lexical frequency and our acoustic measures of VH was very low* » (Nguyen & Fagyal, 2008). Cela pourrait s'expliquer par la fréquence moyenne globale sensiblement basse pour l'ensemble des contextes d'harmonie vocalique (qui, de surcroît, n'est probablement pas une mesure fiable dès lors que l'on compare la fréquence moyenne des paires et non pas les mots séparés). Ainsi, l'effet des mots les plus fréquents (susceptibles d'entraîner une harmonie plus importante) peut être contrecarré par leurs contreparties ayant une fréquence beaucoup plus basse (comme *été* ou *école* par rapport à *éther* ou *écot*).

En somme, les deux études expérimentales menées par Fagyal *et al.* (2003) et Nguyen & Fagyal (2008) ont entraîné une remise en question de la réalité empirique de l'harmonie vocalique en français, suscitant ainsi de nouvelles investigations et invitant à une réévaluation de la nature de ce phénomène dans la tradition phonologique française. Les résultats qu'ils ont apportés peuvent être résumés comme suit :

- i. Il existe en français une propension à assimiler les voyelles moyennes en syllabe initiale en fonction du degré d'antériorité et, dans une moindre mesure, en fonction du degré d'aperture de la voyelle finale de mot.
- ii. Cet effet est graduel et correspond à un *détail acoustique propre à la langue* plutôt qu'à un phénomène dont il serait approprié de dire qu'il se rapproche de la catégorialité phonologique.
- iii. L'harmonie vocalique affecte aussi bien la variété méridionale que le français standard avec, toutefois, un effet plus prononcé dans ce dernier.
- iv. Les mots bi-morphémiques semblent légèrement plus affectés que les mots monomorphémiques.
- v. L'effet de la fréquence lexicale sur le degré d'assimilation est difficile à étudier en raison de la dissymétrie au sein des paires de contextes d'HV et, plus généralement, en raison de la fréquence globalement faible de ces paires dans le lexique français.
- vi. Le terme d'harmonie vocalique est trompeur puisqu'il renvoie à un phénomène de nature différente à la « vraie » harmonie vocalique.
- vii. L'harmonie vocalique en français correspond à une réalité empirique complexe et hétérogène, qui peine à être proprement intégré dans des descriptions phonologiques traditionnelles. Or, pour qu'elle puisse être considérée comme un phénomène phonologique, une approche plus fine serait nécessaire.

En revanche, force est d'observer que le protocole expérimental utilisé comporte des limites dont certains se répercutent sur le déroulement de l'étude ainsi que sur ses conclusions. Nous en citerons deux qui méritent d'être retenues :

- i. Nonobstant les bénéfices qu'il apporte à une étude expérimentale rigoureuse (étant donné qu'il permet notamment d'examiner, en l'isolant, un phénomène précis, toutes choses égales par ailleurs) un contexte d'élocution aussi peu écologique que celui de la parole lue enregistrée en laboratoire, impacte sensiblement la nature acoustique d'un « *fine-grained, gradient phenomenon* » (Nguyen & Fagyal, 2008) que représente l'harmonie vocalique en français. L'effet du contexte d'élocution a conduit jusqu'à neutraliser certains contrastes, comme en ont fait montre les locuteurs qui ont apporté une « correction » de certains de leurs traits phonétiques régionaux.
- ii. La structure d'un corpus utilisant des paires de mots permet, certes, d'isoler l'effet d'harmonie vocalique dans des conditions de laboratoire, mais limite le domaine d'action de celle-ci à un échantillon composé de lexèmes rares, ne représentant qu'une part infime de contextes d'harmonie vocalique présents en parole continue, contextes pour lesquels il n'est pas possible de trouver des contreparties phonétiquement correspondantes.

Pour conclure, qu'il nous soit permis de citer la toute dernière phrase de l'étude de Nguyen & Fagyal (2008), que nous trouvons intéressante dans le contexte de la présente étude : « *only time can tell whether the Northern dialect, exhibiting the strongest assimilatory pattern at this point in time, will evolve towards more categorical VH-like assimilatory distinctions in the future* » (Nguyen & Fagyal, 2008 ; nous soulignons). Les deux sections suivantes, ainsi que l'ensemble du travail expérimental rapporté ici auront pour vocation d'y apporter un éclairage nouveau.

### **2.3. « Que nous apprennent les gros corpus sur l'harmonie vocalique en français ? »**

Observé en synchronie, l'écart en termes du degré d'harmonie vocalique dû à la variation diatopique a donc été interprété, par Nguyen & Fagyal (2008), comme témoignant d'une inclination, de nature diachronique, à favoriser, dans un français régional, une assimilation de voyelle à voyelle plus importante, sous l'emprise de la variété standard, où cette assimilation est ressortie comme plus forte. Cette affirmation est ainsi issue d'une extrapolation des faits empiriques qui, force est de l'admettre, doivent être particulièrement intéressants en tant que tels, pour autant qu'ils mettent en exergue la capacité des locuteurs de la variété régionale du français (en l'occurrence, celle du Sud) à imiter, d'une façon probablement inconsciente et sous l'effet d'un contexte situationnel formel, un trait phonétique du français standard fin et subtil que représente l'harmonie vocalique. Or, Nguyen & Fagyal (2008) sont allés un pas plus loin pour postuler que le français standard tendrait lui-même

à maximiser le degré d'assimilation pour évoluer vers une harmonie vocalique plus forte, voire *davantage catégorielle*.

Le travail expérimental de Turco *et al.* décliné en deux publications (2016a, 2016b), aborde ces différents questionnements d'une façon nouvelle, en se proposant d'étudier l'effet d'harmonie vocalique dans deux grands corpus de parole journalistique, ESTER (Galliano *et al.*, 2006), et spontanée, NCCFr (Torreira *et al.*, 2010). Tout en s'inscrivant, l'explicitent-ils, « dans la lignée des études de Fagyal *et al.* (2003) et Nguyen & Fagyal (2008) » Turco *et al.* (2016a) ont décidé de réinterpréter la définition d'harmonie vocalique, en incorporant, en qualité de contextes possibles, tout élément lexical contenant une  $V_2 \in \{i, e, \varepsilon, o, a, y, u, \tilde{a}\}$  en syllabe finale et une  $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$  en syllabe pénultième, qu'elle soit initiale ou non. Il convient de préciser que les quatre voyelles-cibles n'ont pas été analysées séparément, mais regroupées soit par deux (et dans ce cas codées en tant qu'archiphonèmes /E/, /O/) soit analysées toutes ensemble (à savoir en fonction du facteur de variabilité examiné). Aucune contrainte n'a été retenue quant à la longueur du mot, sa structure syllabique ni le type de consonne (ou séquence de consonnes) séparant la  $V_1$  et la  $V_2$ . Les  $V_2$  ont été classées en deux catégories selon leur degré d'aperture phonologique et, contrairement à Fagyal *et al.* (2003) et Nguyen & Fagyal (2008), sans tenir compte des contrastes d'antériorité. La première catégorie englobait donc, en les regroupant sous la spécification commune [+haut], les voyelles fermées et mi-fermées :  $V_2^{[+haut]} \in \{i, e, o, y, u\}$  et la seconde, les voyelles ouvertes et mi-ouvertes  $V_2^{[-haut]} \in \{\varepsilon, a, \tilde{a}\}$ . Si la décision d'exclure de l'examen les voyelles nasales /ɔ̃, ẽ/ a possiblement été motivée par la complexité, inhérente à l'étude des corrélats acoustiques de la nasalité (Amelot, 2004), les raisons de l'exclusion des /ɔ, ø, œ/ n'ont en revanche pas été exposées.

Il ne saurait être omis de remarquer que la définition de l'harmonie vocalique dans les deux séries d'études est ainsi sensiblement différente. Certains aspects sont directement liés aux exigences relatives au format du corpus. Celui-ci est bien contrôlé chez Fagyal *et al.* (2003) et Nguyen & Fagyal (2008), puisque destiné à isoler, autant que faire se peut, le phénomène d'harmonie vocalique. Le format est bien plus permissif et plus proche de la réalité d'harmonie vocalique en parole écologique chez Turco *et al.* (2016a, 2016b). Ces derniers ont dès lors abandonné le concept de *paires* d'harmonie vocalique, sans étudier non plus l'effet du degré d'antériorité sur les modèles d'assimilation.

Pour mieux illustrer les différences dans la structure de contextes d'harmonie vocalique tels que définis par, d'un côté, Fagyal *et al.* (2003) et Nguyen & Fagyal (2008), et, de l'autre, Turco *et al.* (2016a, 2016b), le **Tableau 1** compare le répertoire des  $V_1$  et  $V_2$  possibles, ainsi que la structure syllabique autorisée, dans les quatre études citées.



	V <sub>1</sub>	V <sub>2</sub>	structure syllabique autorisée	paramètres étudiés
Fagyal <i>et al.</i> (2003)	/o/-/e/	/e/-/ɛ/ /o/-/ɔ/ /ø/-/œ/ /i/-/a/	#(C <sub>1</sub> )(C <sub>2</sub> )V <sub>1</sub> C <sub>3</sub> V <sub>2</sub> (C <sub>4</sub> )#	F2
Nguyen & Fagyal, (2008)				F1, F2
Turco <i>et al.</i> (2016a)	/O/-/E/ (/O+E/)	/i e o y u/-/ɛ a ã/	#(...)(C <sub>1</sub> )(C <sub>2</sub> )V <sub>1</sub> (C <sub>3</sub> )[ɐ]C <sub>4</sub> V <sub>2</sub> (C <sub>5</sub> )#	F1
Turco <i>et al.</i> (2016b)	/O/-/E/	/i e o y u/-/ɛ a/		

**Tableau 1.** Structure des contextes d’harmonie vocalique utilisées dans les études de Fagyal *et al.* (2003), Nguyen & Fagyal (2008), Turco *et al.* (2016a, 2016b).

Plusieurs aspects nouveaux ont été étudiés par Turco *et al.* (2016a), comme l’influence du lieu d’articulation de la consonne intervocalique (C<sub>4</sub> dans le **Tableau 1**) sur le degré d’assimilation par harmonie vocalique, la présence d’un schwa intervocalique non prononcé (formes du type *aidez* [ɛd(ə)ʔe], en comparaison avec *maîtrise* [mɛʔʁiz]), ou encore la présence de la graphie « é, au, eau, ô » pouvant interférer avec l’influence ouvrante d’une V<sub>2</sub> basse, en favorisant une prononciation plus fermée d’une V<sub>1</sub> ainsi orthographiée. En effet, les résultats ont montré que le degré d’assimilation par harmonie vocalique était moins important lorsque la consonne intervocalique était une labiale, C<sub>4</sub> ∈ {p, b, m, n} par rapport à une linguale C<sub>4</sub> ∈ {t, k, d, g, f, s, ʃ, v, z, ʒ, l, n, ʁ}, quand un schwa sous-jacent non prononcé intervenait entre la V<sub>1</sub> et la V<sub>2</sub>, ainsi qu’en présence d’une graphie fermante, sans pour autant neutraliser complètement l’action assimilatrice de la voyelle finale. Pour ce qui est de la comparaison des deux corpus, contrairement aux suppositions de départ, le degré d’harmonie vocalique s’est révélé « globalement plus important en parole journalistique qu’en parole conversationnelle » (Turco *et al.* 2016a). Si ce résultat va à l’encontre des descriptions historiques qui attribuaient le phénomène d’harmonie vocalique à un style de parole peu soutenu (cf. Cecelewski, 2022 pour une revue de la littérature), il contredirait également l’affirmation de Nguyen & Fagyal (2008), qui entrevoyaient plutôt une dynamique évolutionnelle conduisant vers davantage d’assimilation, du moins dans le français standard. Néanmoins, il est important de signaler que Turco *et al.* (2016a) ont quantifié un degré d’harmonie vocalique global, sans faire la distinction entre /E/ et /O/, et encore moins entre /e, ε, o, œ/. Or, il se pourrait qu’un examen plus approfondi révèle des comportements différents des voyelles moyennes en fonction du style de parole. Aussi cette question reste-t-elle ouverte et mérite sans doute une investigation plus fine.

Dans la seconde étude, Turco *et al.* (2016b), repartent de la même base définitionnelle et reviennent à l’analyse du même corpus de contextes. Mais ce, à une exception près, puisqu’ils renoncent à inclure la nasale /ã/ parmi les V<sub>2</sub> possibles, pour étudier deux facteurs de variation de plus, c’est-à-dire la position prosodique et la durée de la V<sub>1</sub> (codée en variable binaire comme *brèves* et *longues*, en dessous et en dessus de 50 ms, cette valeur de seuil ayant précédemment servi à l’étude des mêmes corpus, cf. Audibert *et al.*, 2016). Les résultats ont montré que le degré d’assimilation était significativement plus important lorsque la V<sub>1</sub> se trouvait en position initiale dans un syntagme intonatif (*Intonational Phrase*), par rapport à la position à l’intérieur de celui-ci (*Word medial*), la

magnitude de l'effet coarticulaire étant égale à environ 7.5 Hz contre 42 Hz, respectivement dans ces deux positions. En revanche, aucun effet significatif de la durée sur le degré d'harmonie vocalique n'a été mis en évidence, ce qui n'est pas en désaccord avec les résultats de travaux antérieurs qui ont conclu à un effet négligeable du débit de parole (Mok, 2011 ; Grosvald, 2009) ainsi que de la quantité vocalique distinctive (Hertrich & Ackermann, 1995) sur la magnitude de l'effet de coarticulation transconsonantique (cf. Turco *et al.*, 2016b).

Quant à la question de l'adéquation terminologique d'*harmonie vocalique*, soulevée par Fagyal *et al.* (2003), la position de Turco *et al.* semble quelque peu ambiguë, pour autant que, si dans la première des deux études, (2016a), c'est le terme d'*harmonie vocalique* qui apparaît systématiquement, en continuité avec la tradition phonologique française (*HV* utilisé 52 fois, *coarticulation* 17 fois, *métaphonie* une fois), dans le titre de la seconde étude (2016b), les chercheurs ont déjà opté pour le terme de *French V-to-V coarticulation* (utilisé 51 fois dans le corps du texte), tandis que l'usage du terme d'*harmonie vocalique* s'est raréfié, sans pour autant disparaître complètement (utilisé neuf fois). Bien que la définition de la nature du phénomène ainsi que de son domaine d'action demeure le même dans les deux études celui-ci semble avoir changé de statut dans la seconde étude : « *vowel-to-vowel coarticulation in French (better known as vowel height harmony)* » (Turco *et al.* 2016b).

En mettant à profit l'accès aux gros corpus de parole continue pour réinterroger, sous un jour nouveau, la réalité empirique de l'harmonie vocalique en français, les deux études de Turco *et al.* (2016a, 2016b) ont fourni un certain nombre de nouveaux résultats que nous pouvons résumer de la façon suivante :

- i. L'harmonie vocalique en français, définie comme un transfert à distance de la spécification d'aperture de la  $V_2$  à la  $V_1$ , induit un degré d'assimilation des voyelles postérieures arrondies, /o, ɔ/, de 30 Hz plus important par rapport à la série antérieure non-arrondie /e, ε/.
- ii. Le degré d'harmonie vocalique varie en fonction du lieu d'articulation de la consonne intervocalique (de 20 Hz plus important pour les *labiales* par rapport aux *linguales*), se montre plus important (15 Hz environ) dans les formes où la graphie de la  $V_1$  ne l'empêche pas d'opérer et ressort comme plus élevé (20 Hz environ) dans les formes dans lesquelles la séquence des consonnes intervocaliques était lexicale (et ne résultant pas de la chute d'un schwa non prononcé).
- iii. L'effet d'harmonie vocalique apparaît également comme plus prononcé lorsque la  $V_1$  se trouve dans une syllabe située au sein d'un groupe intonatif, tandis qu'il s'avère décorrélé par rapport à la durée de la  $V_1$ .
- iv. Contrairement à la description, véhiculée depuis les premières études impressionnistes du début du XX<sup>e</sup> siècle, de l'harmonie vocalique comme étant un phénomène appartenant à un style de parole relâché et conversationnel, celle-ci s'avère plus forte en parole journalistique (supposée plus conservatrice).

Ce dernier résultat constitue sans doute la contribution la plus intéressante de Turco *et al.* (2016a, 2016b). Cet aspect conduit à poser la question de l'évolution de l'harmonie vocalique en français, non plus en extrapolant une variation synchronique observée entre deux styles de parole supposés plus ou moins innovants, mais alors dans une perspective diachronique plus conséquente.

## 2.4. Harmonie vocalique en perspective diachronique

En se proposant pour objectif de réconcilier, d'une part, la suggestion de Nguyen & Fagyal (2008) et, d'autre part, les résultats contraires à l'intuition et aux descriptions antérieures, obtenus par Turco *et al.* (2016a), Cecelewski *et al.* (à paraître), se sont proposés de mesurer le degré d'harmonie dans un corpus de parole journalistique composé d'archives audiovisuelles embrassant la période de 1940 à 1999, corpus ayant été initialement conçu pour une série d'études menées par Boula de Mareüil *et al.* (2008, 2012a, 2012b, 2017).

La définition de contexte d'harmonie vocalique adoptée a été celle de Turco *et al.* (2016a), supposée mieux correspondre à l'étude de gros corpus de parole continue. L'harmonie vocalique a été quantifiée en termes d'écart de la hauteur du premier formant de la  $V_1$  suivie d'une  $V_2$  phonologique fermée ou mi-fermée par rapport à une  $V_2$  ouverte ou mi ouverte.

Les résultats ont montré que l'harmonie vocalique était non seulement présente en parole journalistique dès les années 1940, ce qui va à l'encontre de l'ensemble des descriptions impressionnistes anciennes décrivant la parole soignée comme quasi exempte d'harmonie vocalique. En outre, les résultats indiquent que le degré le plus important d'harmonie vocalique correspond aux deux décennies les plus anciennes du corpus, à savoir les années 1940-50. Par la suite, le degré d'harmonie tend à progressivement décroître jusqu'aux années 1990. Qui plus est, la tendance à la réduction d'harmonie vocalique a été confirmée pour les deux séries des voyelles moyennes antérieures non-arrondies et postérieures arrondies, ce qui a corroboré la cohérence des résultats obtenus.

Toutefois, une illusion d'optique éventuelle, liée à l'évolution de la hauteur de la fréquence du fondamental entre 1940 et 1999 (cf. Boula de Mareüil *et al.*, 2008, 2012a, 2012b), illusion susceptible de fausser les conclusions, devait être écartée. En effet, la période 1940-50, qui correspondait à un degré remarquablement plus important d'harmonie vocalique par rapport aux décennies suivantes, est à la fois celle où la fréquence du fondamental est particulièrement haute, reflétant de ce fait une diction spécifique à cette période. Afin de vérifier si le degré d'harmonie vocalique plus élevé n'était pas entièrement dû au « style Gaumont-Pathé » (cf. Boula de Mareüil *et al.*, 2008, 2012a, 2012b), le degré d'harmonie vocalique a été mesuré en fonction de la hauteur du  $f_0$ , codée en variable binaire *haut* et *bas* en dessus et en dessous de la médiane des valeurs, séparément pour la série des voyelles moyennes antérieures non-arrondies et postérieures. Les résultats ont indiqué qu'un  $f_0$  plus haut entraînait un rehaussement du premier formant, mais qu'il n'impactait pas le degré d'harmonie

vocalique d'une manière qui pourrait imputer l'évolution à la baisse de l'harmonie vocalique dans le temps à la seule diction emphatique des speakers radiophoniques dans les années 1940-1950. Toutefois, le rapport entre le style de parole (en l'occurrence entre la voix forte, ou la voix projetée) et l'harmonie vocalique ne saurait être réduit au seul paramètre de la fréquence du fondamental et, comme l'estiment les auteurs (Cecelewski *et al.*, à paraître), cette question mérite un approfondissement expérimental plus ciblé et systématique.

Toujours est-il que ce résultat a permis de jeter une lumière nouvelle sur les conclusions de Turco *et al.* (2016a), en apportant un argument en faveur d'une innovation phonétique actuellement en cours, conforme à l'évolution diachronique mise en évidence dans la parole journalistique dans la seconde moitié du XX<sup>e</sup> siècle. L'écart entre le degré d'harmonie vocalique dans les corpus NCCFr et ESTER apparaît dès lors comme un reflet synchronique de l'évolution étalée sur un temps plus long et confirmée expérimentalement (pour autant qu'il soit justifié de considérer les archives des années 1940-50 d'une part, et, de l'autre, le corpus NCCFr, comme deux pôles opposés d'un changement phonétique en cours) tel qu'il transparaît dans les limites de l'observable de ces deux expérimentations.

## 2.5. Hypothèses et objectifs

L'ensemble des travaux expérimentaux susmentionnés a ainsi sans aucun doute élargi nos connaissances sur la nature du phénomène d'harmonie vocalique en français sous ses multiples facettes, aboutissant parfois à des résultats surprenants ou, tout du moins, inattendus. Un certain nombre de questions demeurent cependant ouvertes et appellent à réinterroger tant la nature acoustique que perceptive de ce phénomène complexe.

Dès lors, la vocation du présent travail de recherche est de compléter ces lacunes quant à nos connaissances actuelles sur l'harmonie vocalique et ce, en mobilisant un outillage méthodologique diversifié de façon à pouvoir y apporter un éclairage nouveau.

En premier lieu, nous tenterons de reproduire, en l'élargissant, le résultat obtenu par Turco *et al.* (2016a) pour vérifier si l'effet d'harmonie vocalique est systématiquement plus important en parole spontanée par rapport à une parole plus formelle, soutenue, voire lue. Notre première partie mobilisera l'accès à trois gros corpus, NCCFr (Torreira *et al.*, 2010), ESTER (Galliano *et al.*, 2006) et BREF (Lamel *et al.*, 1991) consistant, respectivement, en une parole conversationnelle, journalistique et lue, en vue d'éclairer la corrélation entre le degré d'harmonie vocalique et le style de parole supposé refléter l'évolution diachronique en cours. Notre approche se décline désormais en deux sous-hypothèses :

**H1a.** *L'harmonie vocalique occasionne un écart d'aperture de la  $V_1$  plus important en parole journalistique par rapport à la parole spontanée et plus important en parole lue par rapport à la parole journalistique.*

**H1b.** *Si tant est qu'elle corresponde à la réalité empirique, la corrélation entre le style de parole et le degré d'harmonie vocalique est observable pour chacune des quatre voyelles moyennes, /e, ε, o, ɔ/, impliquées dans l'harmonie vocalique.*

En second lieu, nous nous reposerons la question fondamentale : « *Is there vowel harmony in French?* », qui n'est autre que le titre de la première étude expérimentale ayant examiné ce phénomène il y a tout juste vingt ans (Fagyal *et al.*, 2003). En sachant par ailleurs que l'harmonie vocalique n'a eu de cesse d'évoluer, et que, selon Cecelewski *et al.* (à paraître) (mais également et, indirectement, Turco *et al.* (2016a)), sa présence dans la variété standard du français d'Île-de-France semble diminuer, nous estimons qu'il n'est pas illégitime de reposer cette question : *Is there still vowel harmony in French ?* : En effet, l'harmonie vocalique est-elle toujours présente en français ? Afin de tenter d'y répondre, une expérience de production sera mise en place, s'inscrivant dans la continuité des travaux de Fagyal *et al.* (2003) et Nguyen & Fagyal (2008), avec pour objectif de quantifier le degré d'harmonie vocalique dans un corpus contrôlé enregistré en laboratoire, corpus de vingt ans plus récent que les productions sonores examinées par les deux études susmentionnées. L'hypothèse que nous avançons est dorénavant la suivante :

**H2.** *L'harmonie vocalique est toujours présente en français, mais son faible degré ne la distingue plus d'une forme de coarticulation spécifique à la langue.*

De surcroît, l'expérience de production sera l'occasion d'apporter deux modifications importantes aux protocoles antérieurs pour répondre à deux questions de recherche subsidiaires que nous formulerons ainsi : L'harmonie vocalique des voyelles moyennes antérieures (centrales) arrondies /ø, œ/, postulée par certaines descriptions phonologiques anciennes (Straka, 1950 ; Malmberg, 1969, Selkirk, 1972 ; Tranel, 1987) mais qui n'a jamais été examinée expérimentalement jusqu'à présent, correspond-elle à une réalité empirique ? Et, si c'est le cas, est-elle de nature similaire à l'harmonie vocalique sur /e, ε, o, ɔ/ ? Notre hypothèse sera donc comme suit :

**H3.** *L'harmonie vocalique sur /œ/ existe bel et bien en français et est de nature similaire à celle mise en évidence pour /E/ et /O/.*

La définition d'harmonie vocalique différant d'une étude à une autre, celle-ci a systématiquement été quantifiée jusqu'à présent sans tenir compte du degré d'aperture sous-jacent de la V<sub>1</sub>. Par conséquent, nous ne savons pas s'il existe en français une harmonie vocalique « ouvrante » occasionnant un rabaissement des voyelles moyennes phonologiques mi-fermées, aussi bien qu'une harmonie vocalique « fermante » entraînant un rehaussement des voyelles moyennes mi-ouvertes. Aussi notre protocole d'expérimentation sera-t-il conçu de façon à pouvoir y apporter un élément de réponse, avec comme hypothèse l'assertion suivante :

*H4. Puisque sa nature est par essence coarticulatoire et non phonologique, il doit exister en français une harmonie vocalique bidirectionnelle provoquant la fermeture des voyelles moyennes mi-ouvertes /ɛ, ɔ, œ/ et l'ouverture des voyelles moyennes mi-fermées /e, o, ø/.*

De plus, la mise en place d'une expérience de perception nouvelle sera une opportunité dont nous espérons tirer profit pour élucider, d'une manière plus systématique que ne l'ont fait Cecelewski *et al.* (à paraître), le rapport éventuel existant entre le degré d'harmonie vocalique et le style de parole. Effectivement, notre expérience sera construite de manière à pouvoir réaliser une comparaison des schémas d'assimilation de voyelle à voyelle en *voix modale* et en *voix projetée* ou *forte*, cette dernière étant supposée reproduire plus ou moins fidèlement (sans toutefois faire appel à des acteurs professionnels) la voix projetée spécifique au style emphatique des speakers radiophoniques des années 1940-1950. En suivant les conclusions de l'étude diachronique de Cecelewski *et al.* (à paraître, cf. aussi Cecelewski, 2022), notre hypothèse revêtira la formulation suivante :

*H5. Le degré d'assimilation de voyelle à voyelle dans les contextes d'harmonie vocalique n'est pas significativement plus important en voix projetée par rapport à la voix modale.*

La vérification des cinq hypothèses ainsi formulées, portant pour chacune d'entre elles sur un aspect acoustique du phénomène étudié, sera l'objet de la première partie du présent travail. Or, il ne saurait échapper à notre attention l'aspect perceptif de l'harmonie vocalique en français, en particulier « la capacité d'auditeurs francophones à exploiter la coarticulation vocalique anticipatoire pour améliorer les performances de l'accès lexical » (Turco *et al.*, 2016a) qui demeure non entièrement éclairée. En conséquence, la dimension perceptive de l'harmonie vocalique, que nous n'avons pas discuté jusqu'ici, sera l'objet de la seconde partie de notre étude.

### **3. Harmonie vocalique dans les grands corpus**

#### **3.1. Méthode**

##### **3.1.1. Corpus**

Trois grands corpus fréquemment utilisés dans la communauté phonétique, alignés automatiquement et représentant trois styles distincts de parole en français standard, ont été analysés. Le premier corpus, ESTER (Galliano *et al.*, 2006), comprenait près de 50 heures de discours journalistiques, composés d'extraits d'émissions radiophoniques diffusées entre 1999 et 2004 sur différentes stations telles que France Inter, Radio France International, France Culture, Radio Classique, et, dans une moindre mesure, sur trois chaînes de radio africaines, à savoir Africa 1, Radio Congo et Radiotélévision du Maroc (TVME). Le deuxième corpus, NCCFr (Torreira *et al.*, 2010), était constitué de 36 heures de conversations spontanées en binômes portant sur des sujets de société,

enregistrées à Paris en 2007 lors d'interactions en face à face non préparées. Enfin, le troisième corpus, BREF (Lamel *et al.*, 1991), comptait environ cent heures d'enregistrements de locuteurs non professionnels lisant des articles tirés du journal *Le Monde*.

Les corpus ESTER et NCCFr, et dans une certaine mesure le corpus BREF, ont servi précédemment à un nombre considérable d'études en phonétique et leurs contributions à la connaissance de l'influence du style d'élocution sur la réalisation de divers corrélats acoustiques de la parole continue sont bien connues (Adda-Decker *et al.*, 2005 ; Gendrot & Adda-Decker, 2005 ; Adda-Decker *et al.*, 2008 ; Bürki *et al.*, 2011, Adda-Decker *et al.*, 2013 ; Audibert *et al.*, 2015 ; entre autres). Les apports cumulés de ces études ont notamment permis d'établir des valeurs formantiques de référence des voyelles françaises basées sur l'analyse de grandes quantités de données, d'examiner les effets d'hypo- et hyper-articulation affectant le système vocalique en fonction du contexte consonantique, de la valeur de  $f_0$ , de la durée du segment, de sa position morphologique et prosodique, mais surtout de mettre tous ces facteurs en rapport avec style de parole. Qui plus est, les corpus ESTER et NCCFr ont été précédemment utilisés dans deux études sur l'harmonie vocalique en français (Turco *et al.*, 2016a, 2016b), dont le protocole, modifié et élargi, sera partiellement repris ici.

Les trois corpus comprennent des transcriptions orthographiques manuelles, qui ont été phonétisées et alignées de façon automatique (Lamel *et al.*, 1991 ; Galliano *et al.*, 2006 ; Torreira *et al.*, 2010). Dans le but d'écarter d'éventuelles discordances dues à des transcriptions réalisées dans des conditions différentes, une inspection automatique de la correspondance graphème-phonème des  $V_1$  sélectionnées pour l'analyse a été effectuée, comparativement dans les trois corpus. En cas de divergences, la transcription des voyelles moyennes en syllabe prétonique a ainsi été alignée sur celle présente dans le corpus BREF.

Étant donné que le corpus diachronique INA utilisé dans notre étude sur l'évolution de l'harmonie (Cecelewski *et al.*, à paraître) se compose exclusivement de productions sonores de locuteurs masculins (cela est dû à la très faible présence de femmes dans les archives radio d'époque, cf. Boula de Mareüil *et al.*, 2008), et que le petit corpus enregistré en laboratoire (voir section 4.1.1.) conçu pour tester l'hypothèse (cf. **H5**) concernant un éventuel impact de la voix projeté sur le degré d'harmonie vocalique (en lien direct avec cette première étude), ne contient également que des productions de locuteurs masculins, nous avons décidé de restreindre l'ensemble des analyses du présent travail de recherche aux productions sonores de sujets masculins. Nous considérons que cela permettra d'exclure le facteur de variation potentiel que constitueraient des productions de sujets féminins dès lors qu'elles ne seraient présentes que dans une partie de nos données. Néanmoins, nous sommes conscients des limites qu'impose cette précaution méthodologique sur le potentiel de généralisation de nos résultats et tenons à souligner qu'une étude d'harmonie vocalique plus complète et exhaustive que la nôtre devrait inclure également les voix de femmes.

Ainsi, les contextes éligibles à l'analyse ont été sélectionnés parmi les productions de sujets masculins provenant des corpus BREF (40 locuteurs sur 90), ESTER (744 sur 2172) et NCCFr (24 sur

46). De plus, nous n'avons inclus dans notre analyse que les locuteurs francophones natifs originaires de la région Île-de-France. C'est pourquoi les productions sonores de dix locuteurs du corpus NCCFr dont le lieu de résidence principale était les régions de Bourgogne, de Bretagne, de Haute/Basse Normandie, du Limousin ou de Poitou-Charentes, ainsi que trois locuteurs francophones non natifs du corpus BREF, dont un locuteur arabophone, un lusophone et un germano-/russophone, n'ont pas été intégrés dans le corpus étudié ici. Le filtrage appliqué au corpus ESTER a, quant à lui, été limité à l'exclusion des chaînes de radio africaines (Africa 1, Radio Congo et TVME). En revanche, l'exclusion d'éventuels locuteurs non natifs ou ceux parlant une variété phonétique régionale n'a pu être effectuée sur les données du corpus ESTER provenant de radios françaises qui contiennent, entre autres, des apparitions de courte durée d'un grand nombre de participants, par exemple des appels téléphoniques d'auditeurs intervenant dans l'émission « Le téléphone sonne », pour lesquels nous ne disposons pas de métadonnées suffisantes.

Les contextes possibles d'harmonie vocalique ont été sélectionnés parmi les mots ayant une  $V_1 \in \{e, \varepsilon, \text{ɔ}, o\}$  en syllabe pénultième et, en syllabe finale, l'une des dix voyelles orales  $V_2 \in \{i, y, e, \varepsilon, \emptyset, \text{œ}, a, \text{ɔ}, o, u\}$ . Les  $V_1$  ont été codées selon leur degré d'antériorité en une série antérieure, /e, ε/ et une série postérieure, /o, ɔ/ et, selon leur degré d'aperture, en tant que mi-fermées /e, o/ et mi-ouvertes /ε, ɔ/. Parallèlement, une approche concurrente a été adoptée, à l'instar d'études antérieures (Turco *et al.*, 2016a, 2016b), qui ne distinguaient pas les deux degrés d'aperture des voyelles moyennes, classées alors en tant qu'archiphonèmes /E, O/. Aucune contrainte sur la structure syllabique du mot, sur sa longueur et sur le type de consonnes environnant la  $V_1$  et la  $V_2$  n'a été retenue. Ainsi, le mot le plus long de notre corpus, *anticonstitutionnelle*, était composé de sept syllabes, dont seule l'avant-dernière, [sjɔ], a été considérée comme une voyelle-cible d'une harmonie vocalique suivie, en l'occurrence, d'une voyelle-source classée comme *ouverte*. Contrairement aux études de Fagyal *et al.* (2003) et de Nguyen & Fagyal (2008), mais en continuité avec celles de Turco *et al.*, (2016a, 2016b), les  $V_1$  en syllabe fermée ont également été incluses comme des contextes possibles d'harmonie vocalique, e.g. *Belgique, augmente*. De même, nous avons accepté des contextes contenant une séquence consonantique non lexicale, résultant de la chute d'un schwa non prononcé, e.g. *amènera*. Le nombre de contextes soumis à l'analyse, répartis selon la classe de la  $V_1$  et celle de  $V_2$ , en fonction du corpus, a été indiqué dans le **Tableau 2**.

corpus	/e/		/ε/		/o/		/ɔ/		
	fermée	ouverte	fermée	ouverte	fermée	ouverte	fermée	ouverte	
BREF	3512	6402	1324	1768	1235	671	2000	4214	21126
ESTER	2241	2989	1461	1198	1220	360	1846	3233	14548
NCCFr	905	1664	538	921	985	176	833	1184	7206
<i>total</i>	6658	11055	3323	3887	3440	1207	4679	8631	42880

**Tableau 2.** Nombre de contextes d'harmonie vocalique analysés, en fonction de la  $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$  et du type de  $V_2$ , fermée ou ouverte, séparément pour les corpus BREF, ESTER et NCCFr.

Les valeurs des deux premiers pics formantiques étaient extraits respectivement à  $\frac{1}{3}$ ,  $\frac{1}{2}$  et  $\frac{2}{3}$  de la durée de la  $V_1$  à l'aide de l'algorithme Burg implémenté dans Praat (Boersma & Weenink, 2016),

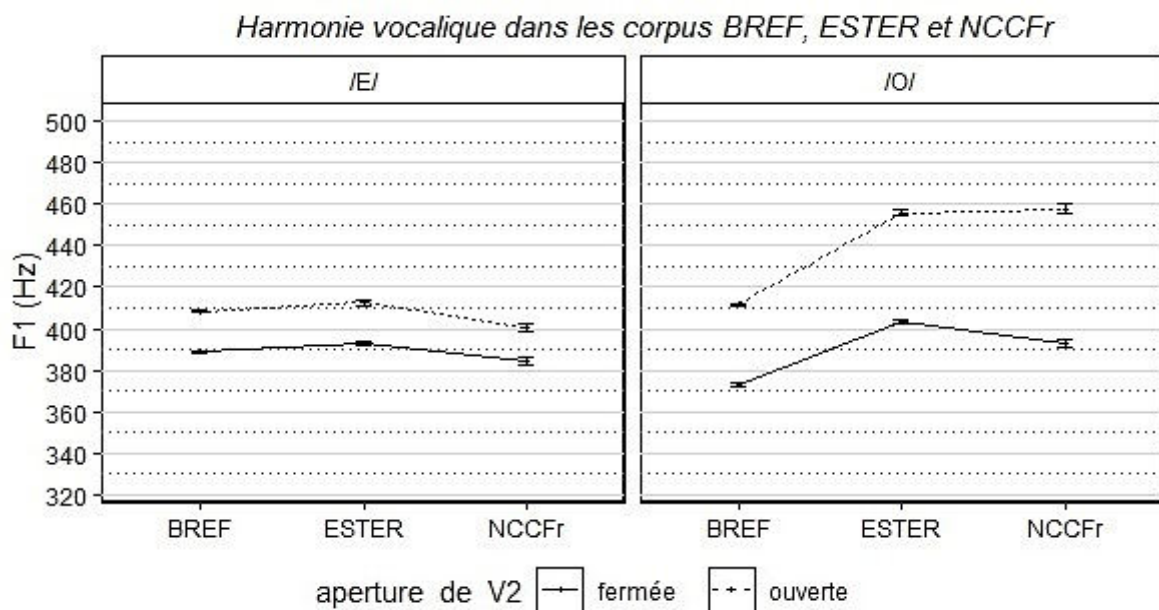


puis moyennées pour obtenir une valeur unique. Un script a été utilisé pour automatiser l'extraction de valeurs formantiques qui a été effectuée avec les paramètres suivants : préemphasis à 50 Hz, plage de détection inférieure à 4,8 kHz et fenêtre d'analyse de 25 ms. Les tests de validation statistique ont été conduits dans l'environnement R (R Development Core Team, 2010 ; version 2022.12.0+353).

## 3.2. Résultats

### 3.2.1. Harmonie vocalique sur /E, O/ dans les corpus BREF, ESTER et NCCFr

Dans un premier temps, nous avons mesuré le degré d'harmonie vocalique agissant sur les  $V_1 \in \{E, O\}$ , en fonction de l'aperture phonologique de la  $V_2$ , fermée ou mi-fermée d'une part et, de l'autre, ouverte ou mi-ouverte. Puisqu'il s'agit, à ce stade de l'analyse, d'appréhender les deux séries de voyelles moyennes sous une enseigne commune d'archiphonèmes /E, O/, la directionnalité d'harmonie vocalique n'est pas considérée. Aussi, les formes du type *était/rester* sont-elles comparées à *étude/prestige* et *auquell/local*, à *aussi/protagoniste*, respectivement pour les  $V_1$  de la série antérieure non-arrondie et postérieure arrondie. La **Figure 1** représente les valeurs moyennes de la hauteur du premier formant des /E, O/ en fonction du degré d'aperture de la voyelle finale, dans les corpus BREF, ESTER et NCCFr.



**Figure 1.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la  $V_1$  /E, O/, suivie d'une  $V_2$  fermée ou mi-fermée ou d'une  $V_2$  ouverte ou mi-ouverte, dans les corpus BREF, ESTER et NCCFr.

Afin d'étudier les relations entre la hauteur de F1 et les effets fixes d'*aperture de la  $V_2$*  et de *corpus*, nous avons ajusté un modèle de régression linéaire mixte (bibliothèque `lme4`, Bates *et al.*, 2014). En plus d'effets fixes, l'intercept pour *locuteur* a été intégré au modèle en tant qu'effet aléatoire, en vue de tenir compte de la variabilité inter-locuteur sur la hauteur de F1.

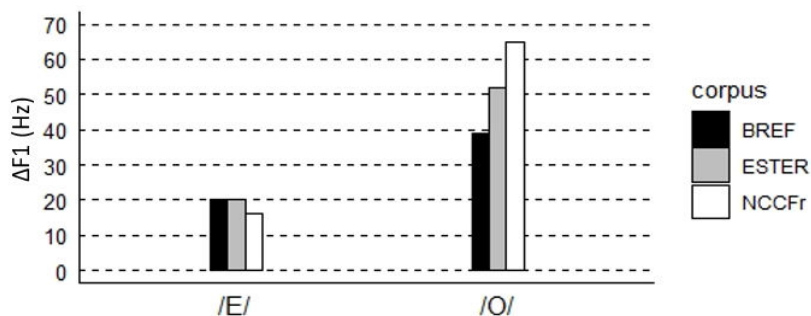
Le modèle effectué pour /E/ a indiqué un R<sup>2</sup> marginal de 0.12 et un R<sup>2</sup> conditionnel de 0.54, obtenus à l'aide de la fonction `model_performance()` (Lüdtke *et al.*, 2021). Le test a révélé un effet significatif d'*aperture de la V<sub>2</sub>* (p<.001), mais pas d'effet significatif de *corpus*. En revanche, nous avons observé une interaction significative entre *aperture de la V<sub>2</sub>* et *corpus:NCCFr*. En suivant la même procédure du choix de termes, nous avons ajusté un modèle de régression pour /O/, qui a indiqué un R<sup>2</sup> marginal de 0.06 et un R<sup>2</sup> conditionnel de 0.48. Nous avons constaté un effet significatif d'*aperture de la V<sub>2</sub>* (p<.001), ainsi qu'une interaction significative entre les deux effets fixes intégrés au modèle (p<.001 et p<.05, respectivement pour ESTER et NCCFr). En revanche, l'effet global de *corpus* n'a pas été significatif. Les coefficients des modèles ajustés pour /E/ et pour /O/, de même que les valeurs des tests de rapport de vraisemblance obtenues par des approximations de type *Satterthwaite* au moyen de la fonction `lmerTest` (selon la méthode suggérée par Turco *et al.*, 2016a) ont été indiqués dans le **Tableau 3**. Les résultats du test des comparaisons multiples effectué à l'aide de la fonction `emmeans` (Lenth *et al.*, 2017) ont, quant à eux, été détaillés dans l'**Annexe 1**.

V <sub>1</sub>	terme	coefficient	erreur-type	dl	t	p
/E/	(intercept)	411.92	13.03	271.52	31.59	<.001
	aperture_V2	20.85	1.07	24461.69	19.46	<.001
	ESTER	-8.92	13.29	276.85	-0.67	0.50
	NCCFr	-25.40	15.82	266.31	-1.60	0.10
	aperture_V2:ESTER	-1.85	1.72	24848.06	-1.07	0.28
	aperture_V2:NCCFr	-5.97	2.20	24455.67	-2.70	<.001
/O/	(intercept)	437.15	15.75	266.75	27.75	<.001
	aperture_V2	32.19	1.40	12881.45	22.85	<.001
	ESTER	-5.85	16.05	272.12	-0.36	0.71
	NCCFr	-38.03	18.99	266.04	-2.00	<.05
	aperture_V2:ESTER	9.20	2.09	13129.36	4.39	<.001
	aperture_V2:NCCFr	28.41	2.71	12879.52	10.47	<.001

**Tableau 3.** Résultats des effets fixes des modèle de régression linéaire mixte ajustés pour le F1 de /ε, œ/.

Ces résultats indiquent ainsi l'écart significatif dans la hauteur du premier formant de la voyelle subissant l'effet d'harmonie vocalique, induit par le degré d'aperture de la voyelle finale du mot. De surcroît, cet effet interagit avec le style de parole, encore que la dynamique de cet effet diffère nettement entre les deux voyelles analysées, ce qui aboutit à deux schémas apparemment contradictoires. En effet, il apparaît dans notre analyse que l'effet d'harmonie vocalique sur /E/ est légèrement plus modéré en parole spontanée par rapport à la parole plus contrôlée, qu'il s'agisse de la parole lue (corpus BREF) ou de la parole journalistique, partiellement lue (corpus ESTER). L'effet du style de parole sur l'harmonie vocalique sur /E/ converge dès lors avec les résultats énoncés par Turco *et al.*, (2016a). Néanmoins, le comportement de /O/ sous l'effet d'harmonie vocalique ne s'aligne pas sur celui de /E/. En effet, le premier subit un effet d'assimilation plus fort en parole journalistique que dans la parole lue, ainsi qu'en parole spontanée par rapport à la parole journalistique. Autrement dit, si tant est qu'une parole plus contrôlée apparaisse comme un domaine plus propice à l'harmonie vocalique, cet effet ne semble vrai que pour /E/, tandis que la variation de /O/ semble confluer avec les descriptions anciennes qui statuaient que « *vowel harmony occur naturally in spontaneous speech* »

(Tranel, 1987). Pour une visualisation plus commode, la **Figure 2** représente les  $\Delta F1$  moyens calculés pour les  $V_1 \in \{E, O\}$ , en fonction du style de parole.

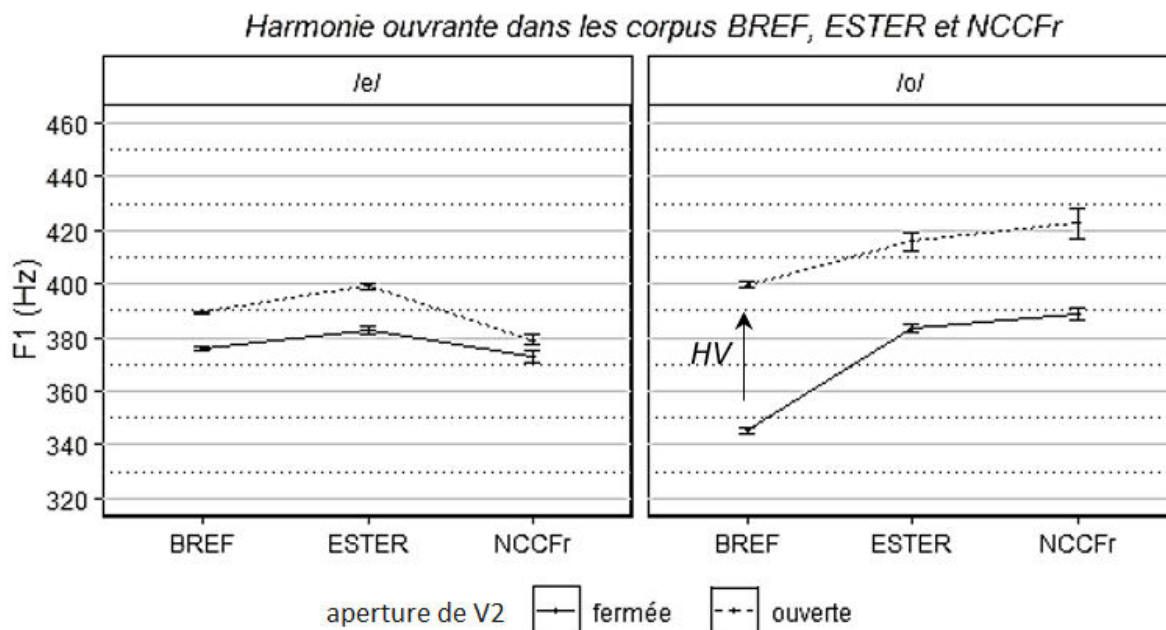


**Figure 2.**  $\Delta F1$  moyens (en Hertz) des  $V_1 \in \{E, O\}$ , dans les corpus BREF, ESTER et NCCFr.

Conséquemment, il n'est pas inexact de suggérer que l'effet d'harmonie vocalique se révèle plus nuancé que supposé et qu'il mérite un examen plus approfondi, en vue de mettre en évidence d'éventuelles différences en termes d'interaction entre le degré d'assimilation et le style de parole, en mettant en examen différentes catégories phonologiques des voyelles subissant l'harmonie vocalique.

### 3.2.2. Harmonie vocalique ouvrante sur /e, o/ dans les corpus BREF, ESTER et NCCFr

En conséquence, nous avons quantifié le degré d'harmonie vocalique *ouvrante*, exprimé en termes de rehaussement du premier formant de la  $V_1 \in \{e, o\}$ , lorsqu'elle était suivie par une  $V_2$  ouverte ou mi-ouverte, e.g. *égard*, *autant*. Cet écart a été comparé au cas où la  $V_1$  était suivie d'une  $V_2$  fermée ou mi-fermée, considéré ici comme un cas neutre, e.g. *prévu*, *aucune*. La **Figure 3** représente le degré d'harmonie vocalique ouvrante dans les trois corpus analysés.



**Figure 3.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la  $V_1 \in \{e, o\}$  phonologique, suivie d'une  $V_2$  fermée ou mi-fermée ou d'une  $V_2$  ouverte ou mi-ouverte, dans les corpus BREF, ESTER et NCCFr.

Pour étudier les relations entre les valeurs de F1 et les effets fixes d'*aperture de la V<sub>2</sub>* et de *corpus*, nous avons ajusté un modèle de régression linéaire mixte, selon le protocole détaillé dans la section précédente (cf. 3.2.1.). Ainsi, le modèle réalisé pour /e/ a indiqué un R<sup>2</sup> marginal de 0.04 et un R<sup>2</sup> conditionnel de 0.51, ainsi qu'un effet significatif d'*aperture de la V<sub>2</sub>* (p<.001) mais pas d'effet significatif de *corpus*. En addition, nous avons constaté une interaction significative entre *aperture de la V<sub>2</sub>* et *corpus* (pour NCCFr, mais pas ESTER). Le modèle de régression ajusté pour /o/ a indiqué un R<sup>2</sup> marginal de 0.06 et un R<sup>2</sup> conditionnel de 0.48, de même qu'un effet significatif d'*aperture de la V<sub>2</sub>* (p<.001). Une interaction significative entre les deux effets fixes intégrés au modèle a également été observée (p<.001 et p<.05, respectivement pour ESTER et NCCFr). En revanche, l'effet de *corpus* n'a pas été significatif. Les coefficients des modèles ajustés pour /e/ et pour /o/, ainsi que les valeurs des tests de rapport de vraisemblance ont été indiqués dans le **Tableau 4**. Dans l'**Annexe 1**, nous rapportons les résultats des tests de comparaisons multiples.

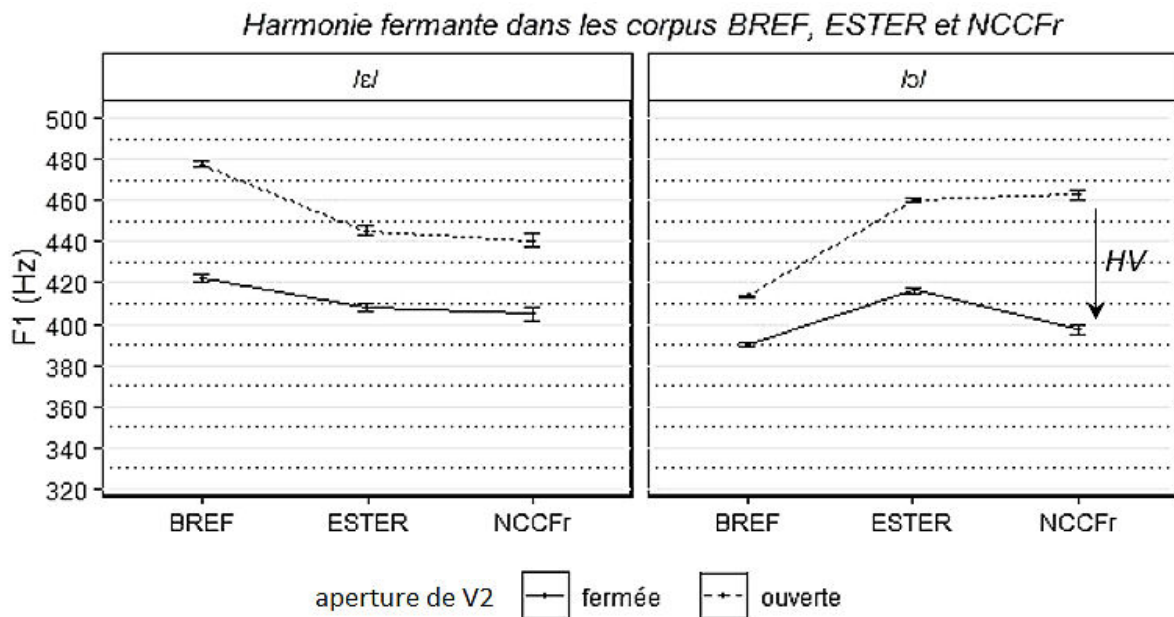
V <sub>1</sub>	terme	coefficient	erreur-type	dl	t	p
/e/	(intercept)	398.22	13.21	265.65	30.12	<.001
	aperture_V2	16.33	0.98	17263.22	16.60	<.001
	ESTER	-6.69	13.48	270.59	-0.49	0.62
	NCCFr	-22.78	16.08	261.06	-1.41	0.15
	aperture_V2:ESTER	-0.32	1.65	17632.36	-0.19	0.84
	aperture_V2:NCCFr	-11.73	2.13	17256.38	-5.50	<.001
/o/	(intercept)	389.31	13.79	161.40	28.22	<.001
	aperture_V2	41.80	2.32	4365.68	17.95	<.001
	ESTER	4.78	14.15	165.71	0.33	0.73
	NCCFr	3.65	16.19	155.80	0.22	0.82
	aperture_V2:ESTER	-9.38	3.73	4574.23	-2.51	<.05
	aperture_V2:NCCFr	-11.52	4.45	4349.98	-2.58	<.001

**Tableau 4.** Résultats des effets fixes des modèle de régression linéaire mixte ajustés pour le F1 de /e, o/.

Ainsi, l'analyse indique que l'harmonie vocalique ouvrante, c'est-à-dire agissant sur les V<sub>1</sub> phonologiques /e, o/ opère dans nos données, occasionnant un rehaussement significatif de F1 dans le cas où la V<sub>1</sub> était suivie d'une V<sub>2</sub> ouverte ou mi-ouverte. De plus, cet effet se montre sensible au style de parole, quoique cette interaction se révèle plus nuancée que supposée. En effet, le degré d'harmonie vocalique, aussi bien sur /e/ que sur /o/ est, certes, moins important en parole spontanée (corpus NCCFr) qu'en parole journalistique (corpus ESTER), en convergeant avec les résultats de l'étude de Turco *et al.* (2016a). Or, dans le cas de /o/, le degré d'harmonie est plus ténu, tant et si bien que les ΔF1 moyens calculés pour les deux corpus sont presque égaux (32 Hz dans ESTER et 34 Hz dans NCCFr). Quant à la parole lue, l'harmonie vocalique sur /e/ dans le corpus BREF atteint une magnitude presque égale à celle du corpus ESTER. En revanche, de la comparaison entre la parole lue et la parole journalistique, un schéma différent apparaît pour /o/, étant donné que le degré d'harmonie vocalique se révèle d'environ 15 Hz plus important dans le corpus BREF par rapport à ESTER.

### 3.2.3. Harmonie vocalique fermante sur /ε, ə/ dans les corpus BREF, ESTER et NCCFr

Nous avons ensuite quantifié le degré d'harmonie vocalique fermante, c'est-à-dire induisant une prononciation plus fermée des  $V_1 \in \{\epsilon, ə\}$ , exprimée par un abaissement du premier formant sous l'influence d'une  $V_2$  fermée ou mi-fermée, e.g. dans *irréversible*, *reconnu*, par rapport au cas où elle était suivie d'une  $V_2$  ouverte ou mi-ouverte, e.g. *permettre*, *totale*, que nous considérerons par la suite comme des cas neutres. La **Figure 4** illustre le degré d'harmonie vocalique fermante dans les trois corpus analysés.



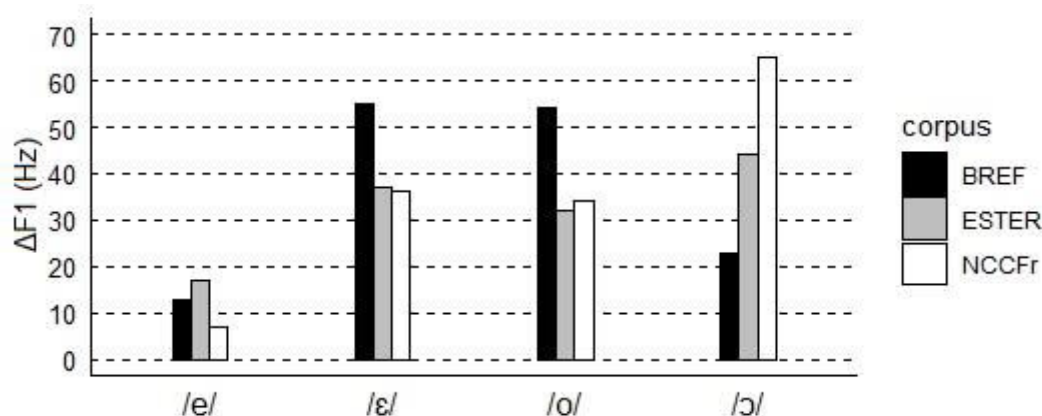
**Figure 4.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la  $V_1$  /ε, ə/ phonologique, suivie d'une  $V_2$  fermée ou mi-fermée ou d'une  $V_2$  ouverte ou mi-ouverte, dans les corpus BREF, ESTER et NCCFr.

En vue de valider ces résultats, nous avons ajusté un modèle de régression linéaire mixte incluant les mêmes effets fixes et les intercepts aléatoires que ceux définis ci-avant. Ainsi, le modèle réalisé pour /ε/ a indiqué un  $R^2$  marginal de 0.09 et un  $R^2$  conditionnel de 0.38, de même qu'un effet significatif d'*aperture\_V2* ( $p < .001$ ), mais pas celui de *corpus*. En outre, nous avons trouvé une interaction significative entre ces deux prédicteurs ( $p < .001$  pour les deux comparaisons). Le modèle réalisé pour /ə/ en suivant la même procédure a révélé un  $R^2$  marginal de 0.09 et un  $R^2$  conditionnel de 0.38. En addition, nous avons trouvé un effet significatif d'*aperture\_V2* ( $p < .001$ ) et de *corpus* ( $p < .05$  pour NCCFr, mais pas d'effet significatif pour ESTER), ainsi qu'une interaction significative entre ces deux termes ( $p < .001$  pour les deux comparaisons). Le **Tableau 5** résume les résultats obtenus pour les effets fixes des modèles ajustés pour les  $V_1 \in \{\epsilon, ə\}$ . Les résultats des tests de comparaisons multiples ont été détaillés dans l'**Annexe 1**.

V <sub>1</sub>	terme	coefficient	erreur-type	dl	t	p
/ε/	(intercept)	417.08	14.18	154.23	29.39	<.001
	aperture_V2	57.54	2.35	6885.48	24.43	<.001
	ESTER	-0.73	14.58	161.29	-0.05	0.96
	NCCFr	-13.36	16.99	158.34	-0.78	0.43
	aperture_V2:ESTER	-22.75	3.52	7165.62	-5.36	<.001
	aperture_V2:NCCFr	-14.11	4.19	6875.71	-3.65	<.001
/ɔ/	(intercept)	437.15	15.75	266.75	27.75	<.001
	aperture_V2	32.19	1.40	12881.45	22.85	<.001
	ESTER	-5.85	16.05	272.12	-0.36	0.71
	NCCFr	-38.03	18.99	266.04	-2.00	<.05
	aperture_V2:ESTER	29.20	2.09	13129.36	4.39	<.001
	aperture_V2:NCCFr	8.41	2.71	12879.52	10.47	<.001

**Tableau 5.** Résultats des effets fixes des modèles de régression linéaire mixte ajustés pour le F1 de /ε, ɔ/.

Ainsi, l'harmonie vocalique fermante provoque un abaissement significatif du premier formant des V<sub>1</sub> /ε, ɔ/ dans les trois corpus analysés. Ce qui plus est, elle apparaît interagir avec le style de parole, tant et si bien que son degré, dans le cas de la voyelle moyenne mi-ouverte antérieure, s'avère plus grand dans une parole lue par rapport à la parole journalistique ou spontanée, ce qui est en continuité avec les résultats de Turco *et al.* (2016a). En revanche, cette tendance s'inverse dans le cas de la voyelle moyenne mi-ouverte postérieure. Ainsi, le degré d'harmonie vocalique sur /ɔ/ se montre significativement plus important dans ESTER par rapport à BREF, et également plus important dans NCCFr par rapport à ESTER, en traçant une continuité d'un effet d'assimilation qui va s'atténuant au fur et à mesure que le style de parole devient de plus en plus contrôlé. En résumé, pour un repérage visuel plus commode, la **Figure 5** représente les ΔF1 moyens calculés pour les quatre voyelles moyennes analysées, V<sub>1</sub> ∈ {e, ε, o, ɔ}, en fonction du style de parole.

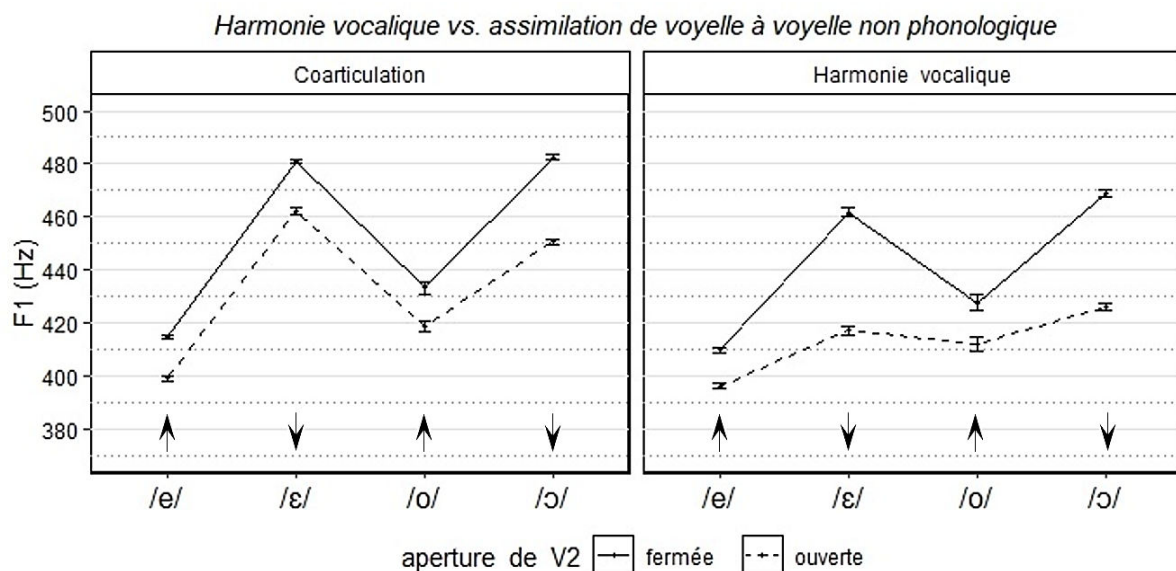


**Figure 5.** ΔF1 moyens (en Hertz) des V<sub>1</sub> ∈ {e, ε, o, ɔ} phonologiques suivies d'une V<sub>2</sub> fermée/mi-fermée ou ouverte/mi-ouverte, dans les corpus BREF, ESTER et NCCFr.

### 3.2.4. Harmonie vocalique versus assimilation de voyelle à voyelle dans des syllabes initiales prétoniques internes

Nous avons ensuite comparé le degré d'assimilation de voyelle à voyelle dans des contextes d'*harmonie vocalique* phonologique à celui agissant dans d'autres contextes, que l'on qualifiera dorénavant de *coarticulation*. La **Figure 6** représente les valeurs moyennes du premier formant, en

Hertz, des  $V_1 \in \{e, \varepsilon, o, \circ\}$ , en fonction de la spécification d'aperture de la voyelle suivante. Les flèches indiquent la direction d'assimilation : la pointe tournée vers le haut correspond à l'harmonie vocalique ou coarticulation ouvrante, c'est-à-dire occasionnant un rehaussement du premier formant des voyelles moyennes mi-fermées. Les flèches pointant vers le bas correspondent à l'harmonie vocalique/coarticulation fermante agissant sur les voyelles moyennes mi-ouvertes. Les  $V_1$  en syllabe pénultième ont ainsi été codée comme subissant l'harmonie vocalique, tandis que les  $V_1$  se trouvant en syllabe initiale ou en syllabe interne de mot, non pénultième, ont été codées comme étant sujettes à la coarticulation de voyelle à voyelle. En guise d'exemple, dans *nécessairement* [ne.se.sεκ.mã], la syllabe initiale [ne], ainsi que la syllabe interne [se] ont toutes les deux été considérées comme sujettes à une coarticulation de voyelle à voyelle non phonologique, induite par une  $V_2$  fermée [e] et une  $V_2$  ouverte [ε], respectivement. La syllabe pénultième [sεκ] a, quant à elle, été considérée comme un cas neutre d'une harmonie vocalique ouvrante, puisque suivie d'une  $V_2$  ouverte [ã]. Ainsi, nous ne distinguons pas de positions différentes de la  $V_1$  au sein du mot, de même que nous n'incluons pas dans notre analyse des contextes possibles d'assimilation de voyelle à voyelle agissant au travers de la frontière de mot. Néanmoins, pour sommaire et approximative que soit une telle comparaison, et indépendamment des limites méthodologiques dont elle n'est guère exempte, nous considérons qu'elle est susceptible de rendre compte de la nature d'harmonie vocalique, si tant est qu'elle laisse paraître un comportement distinct par rapport à la coarticulation non phonologique.



**Figure 6.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la  $V_1 \in \{e, \varepsilon, o, \circ\}$ , suivie d'une  $V_2$  fermée ou mi-fermée ou d'une  $V_2$  ouverte ou mi-ouverte, séparément pour les contextes d'harmonie vocalique (*Harmonie vocalique*) et dans une syllabe initiale de mot ou interne non pénultième (*Coarticulation*).

En vue de valider ces résultats, nous avons effectué des régressions linéaires mixtes, séparément pour les voyelles moyennes de la série antérieure non-arrondie et postérieure arrondie. Chaque modèle a inclus les effets fixes suivants :  $V_1$  (deux niveaux correspondant à la classe, 'antérieure' ou 'postérieure' de la  $V_1$ ),  $aperture\_V_2$  (deux niveaux pour les  $V_2$  fermées et ouvertes),  $coarticulation\_type$



(deux niveaux, ‘Harmonie vocalique’ et ‘Coarticulation’), ainsi que les termes d’interaction à deux et trois niveaux. De plus, l’intercept aléatoire a été inclus dans les deux modèles afin de tenir compte de la variabilité inter-locuteur. Ainsi, le modèle ajusté pour /e~o/ a indiqué un R<sup>2</sup> marginal de 0.02 et un R<sup>2</sup> conditionnel de 0.20, tandis que celui effectué pour /ε~ɔ/, un R<sup>2</sup> marginal de 0.17 et un R<sup>2</sup> conditionnel de 0.29. Outre plus, nous avons observé l’effet significatif de V<sub>1</sub>, d’*aperture\_V<sub>2</sub>* et de *coarticulation\_type* dans les deux modèles (p<.001). En revanche, les termes d’interaction ont laissé apparaître deux schéma distincts selon le degré d’aperture de la V<sub>1</sub>. En effet, aucun des quatre termes d’interaction intégrés au modèle ajusté pour /e~o/ ne s’est pas révélé significatif, alors que les mêmes termes inclus dans le modèle effectué pour /ε~ɔ/ ont tous atteint la significativité (p<.001). Les résultats détaillés obtenus pour les deux modèles ont été indiqués dans le **Tableau 6**.

V <sub>1</sub>	terme	coefficient	erreur-type	dl	t	p
/e~o/	(intercept)	414.84	0.72	22946.13	568.51	<.001
	V <sub>1</sub>	18.55	2.19	24665.22	8.46	<.001
	aperture_V <sub>2</sub>	-15.82	1.01	24831.09	-15.57	<.001
	coarticulation_type	-4.70	1.38	25105.59	-3.38	<.001
	V <sub>1</sub> : aperture_V <sub>2</sub>	0.24	2.84	24676.27	0.08	=.93
	V <sub>1</sub> : coarticulation_type	-0.90	4.05	25105.45	-0.22	=.82
	aperture_V <sub>2</sub> : coart.	-2.35	2.05	24649.04	1.14	=.25
	V <sub>1</sub> : aperture_V <sub>2</sub> : coart.	-2.56	5.09	25012.41	-0.50	=.61
/ε~ɔ/	(intercept)	482.49	0.96	29843.86	499.59	<.001
	V <sub>1</sub>	-1.34	1.32	30277.22	-1.22	<.001
	aperture_V <sub>2</sub>	-31.44	1.30	30277.02	-24.12	<.001
	coarticulation_type	-14.02	1.66	30390.48	-8.44	<.001
	V <sub>1</sub> : aperture_V <sub>2</sub>	12.56	1.98	30242.60	6.62	<.001
	V <sub>1</sub> : coarticulation_type	-5.14	2.72	23024.46	-1.88	<.001
	aperture_V <sub>2</sub> : coart.	-9.85	2.57	30000.87	-3.82	<.001
	V <sub>1</sub> : aperture_V <sub>2</sub> : coart.	-16.73	4.14	30196.37	4.04	<.001

**Tableau 6.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la V<sub>1</sub> /E, O/, suivie d’une V<sub>2</sub> fermée ou mi-fermée ou d’une V<sub>2</sub> ouverte ou mi-ouverte, en fonction de l’année de la naissance du locuteur : 1923-45, 1946-66, 1967-1973.

De la comparaison du degré d’assimilation de voyelle à voyelle dans des contextes d’harmonie vocalique, c’est-à-dire en syllabe pénultième, et celui dans des syllabes initiales ou internes non pénultièmes, il appert qu’il existe deux schémas nettement distincts régissant le comportement des voyelles moyennes dans ces positions. Ainsi les voyelles mi-fermées /e, o/ subissent-elles une assimilation d’aperture d’un degré comparable dans les deux positions. Exprimée en termes d’abaissement du premier formant dû à la présence d’une voyelle finale de mot, l’harmonie vocalique ne se distingue dès lors guère de la coarticulation de voyelle à voyelle transconsonantique mise en évidence dans d’autres positions au sein de mot. En guise d’illustration, le /e/ initial de *égalité* ou le /e/ interne de *libération* est prononcé plus ouvert en présence d’un /a/ dans la syllabe suivante avec un degré d’aperture comparable à celui de /e/ pénultième dans *libéral*. Or, ce n’est que le dernier qui représente un contexte d’harmonie vocalique. En revanche, les deux voyelles moyennes mi-ouvertes, lorsqu’elles se trouvent dans une syllabe pénultième, subissent une fermeture – exprimée en termes du rabaissement du premier formant – d’un degré significativement plus important que lorsqu’elles appartiennent à une syllabe initiale de mot ou interne non pénultième. Ainsi, le /ɔ/ de *colissimo* ou le



/ɛ/ de *cervical* subissent une assimilation significativement moins forte que les mêmes voyelles de *colis* ou de *servi*, ce qui démontre l'existence d'un schéma d'assimilation propre à l'harmonie vocalique, qui ne saurait donc être réduite à une instance de coarticulation.

### 3.2.5. Évolution d'harmonie vocalique : étude en temps apparent du corpus BREF

En dernier lieu, nous avons quantifié l'évolution du degré d'harmonie vocalique, dans le corpus BREF, *en temps apparent* (Labov, 1972). En effet, contrairement à ESTER et à NCCFr, le corpus BREF comporte des productions de locuteurs dont l'âge s'étend sur une plage de valeurs embrassant plus ou moins trois générations, nés entre 1923 et 1973, c'est-à-dire ayant entre 68 et 18 ans au moment des enregistrements. Étant donné que l'un des objectifs du présent travail de recherche est de mettre au défi la conclusion énoncée dans Cecelewski *et al.* (à paraître), selon laquelle le degré d'harmonie vocalique avait diminué depuis le milieu du XX<sup>e</sup> siècle, une étude exploratoire en perspective synchronique en temps apparent, soit à l'aide de sujets de générations différentes enregistrés au même moment, pourrait potentiellement apporter un éclairage nouveau sur l'évolution de ce phénomène dans le temps (cf. Rathcke & Stuart-Smith, 2015).

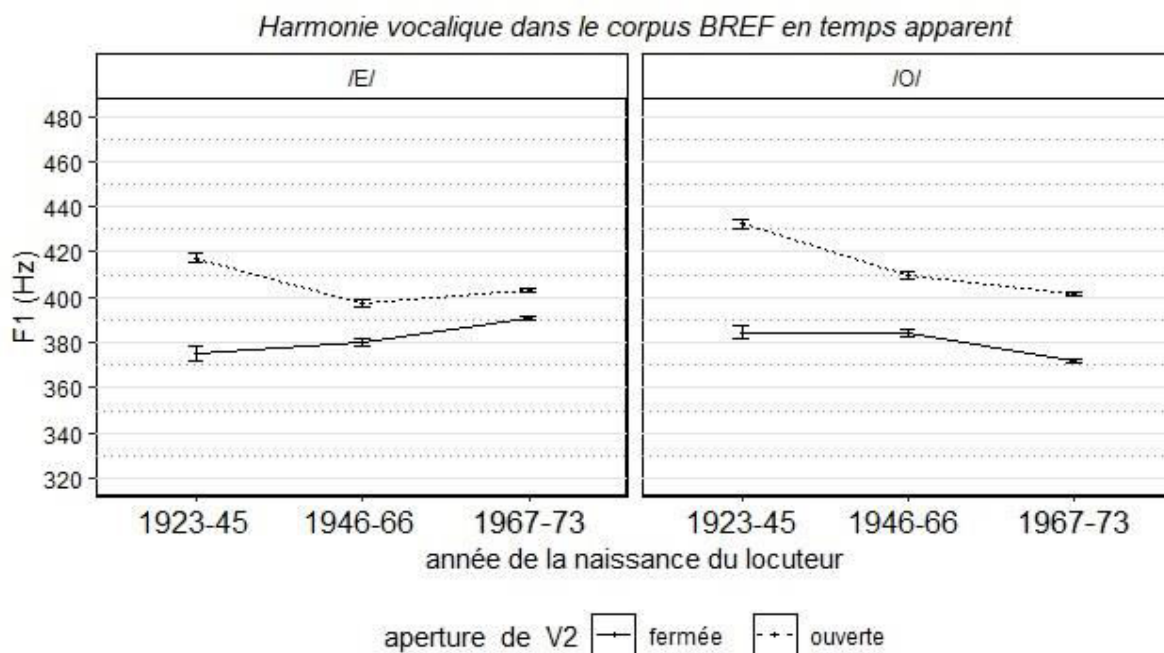
Toutefois, il ne s'agira ici que d'une étude pilote, contrainte par la nature des données dont nous disposons. Force est en effet d'observer que le corpus BREF n'a pas été conçu spécialement en vue d'une étude diachronique en temps apparent conséquente et systématique. Par conséquent, les résultats énoncés ci-après ne sauraient être qu'une indication et point de repère pour une éventuelle étude future, plus approfondie et exhaustive.

Dans le **Tableau 7**, nous avons indiqué la répartition des données analysées en fonction de l'année de naissance des locuteurs. Les trois tranches d'âge ont été délimitées de façon à équilibrer les données, en réconciliant, dans la mesure du possible, le nombre restreint des locuteurs les plus âgés et la nécessité de comparer des périodes éloignées dans le temps au point d'être susceptible de refléter une évolution du phénomène étudié.

tranche de temps	locuteurs (n)	/E/+V <sub>2</sub> fermée	ouverte (n)	/O/ + V <sub>2</sub> fermée	ouverte (n)	
1923-45	7	341	726	312	332	1711
1946-66	12	993	957	774	811	3535
1967-73	21	3502	6490	2149	3742	15883
	40	4836	8173	3235	4885	total

**Tableau 7.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la V<sub>1</sub> /E, O/, suivie d'une V<sub>2</sub> fermée ou mi-fermée ou d'une V<sub>2</sub> ouverte ou mi-ouverte, en fonction de l'année de la naissance du locuteur : 1923-45, 1946-66, 1967-1973.

Les écarts dans la hauteur de F1 d'une V<sub>1</sub> ∈ {E, O} dû à l'harmonie vocalique et en fonction de l'année de la naissance du locuteur ont été représentés dans la **Figure 7**. En raison d'un petit nombre d'occurrences dans la catégorie la moins représentée, 1923-45, les voyelles mi-ouvertes et mi-fermées ont été regroupées ensemble.



**Figure 7.** Valeurs moyennes de F1 (en Hertz) et erreur-type de la  $V_1 \in \{E, O\}$ , suivie d'une  $V_2$  fermée ou mi-fermée ou d'une  $V_2$  ouverte ou mi-ouverte, en fonction de l'année de la naissance du locuteur : 1923-45, 1946-66, 1967-73.

Afin d'étudier la relation entre le degré d'harmonie vocalique et la tranche d'âge des locuteurs, nous avons ajusté un modèle de régression linéaire mixte incluant les effets fixes d'*aperture\_V2*, *classe\_âge* et un terme d'interaction entre ces deux prédicteurs. En addition, l'intercept aléatoire pour *locuteur* a été intégré au modèle. Le modèle ajusté pour /E/ a ainsi indiqué un  $R^2$  marginal de 0.05 et un  $R^2$  conditionnel de 0.42. Nous avons observé un effet significatif d'*aperture\_V2* sur la hauteur de F1 ( $p < .001$ ). L'effet de *tranche\_d'âge*, quant à lui, n'a pas été significatif. Toutefois, nous avons pu conclure à la significativité du terme d'interaction entre *aperture\_V2* et *tranche\_d'âge* ( $p < .05$  pour 1946-66 et  $p < .01$  pour 1967-1973). Ensuite, nous avons effectué un modèle pour /O/, en suivant le même protocole du choix de termes. Ainsi réalisé, le modèle a montré un  $R^2$  marginal de 0.14 et un  $R^2$  conditionnel de 0.32. Un effet significatif a été observé pour le prédicteur *aperture\_V2* ( $p < .001$ ). En addition, nous avons observé un effet significatif du terme d'interaction entre *aperture\_V2* et *tranche\_d'âge* ( $p < .05$  pour 1967-73 mais pas significatif pour 1946-66) Cependant, l'effet de *tranche\_d'âge* ne s'est pas révélé significatif. Les coefficients des deux modèles, ainsi que les résultats des tests de rapport de vraisemblance, ont été indiqués dans le **Tableau 8**.

V <sub>1</sub>	terme	coefficient	erreur-type	dl	t	p
/E/	(intercept)	384.76	12.82	14.27	23.12	<.001
	aperture_V <sub>2</sub>	37.23	3.33	12993.14	21.03	<.001
	1946-66	-2.55	7.22	3877.61	-0.05	=.21
	1967-73	-4.31	8.90	1489.64	-0.23	=.06
	aperture_V <sub>2</sub> : 1946-66	-20.25	4.01	13001.31	-5.75	<.05
	aperture_V <sub>2</sub> : 1967-73	-20.74	3.48	12995.67	-4.84	<.01
/O/	(intercept)	391.16	16.32	12.11	27.11	<.001
	aperture_V <sub>2</sub>	42.32	2.56	8110.10	25.34	<.001
	1946-66	-3.83	6.27	5686.80	-0.33	=.09
	1967-73	-5.01	7.75	3508.44	-3.12	=.12
	aperture_V <sub>2</sub> : 1946-66	9.20	3.01	8109.24	4.39	=.03
	aperture_V <sub>2</sub> : 1967-73	18.41	2.69	8109.47	10.47	<.05

**Tableau 8.** Résultats des effets fixes des modèles de régression linéaire mixte ajustés pour le F1 de /ε, ɔ/ en fonction de l'année de la naissance du locuteur : 1923-45, 1946-66, 1967-73.

Ainsi, ces résultats suggèrent que le degré d'harmonie vocalique varie en fonction de la classe d'âge du locuteur, du moins laisse-t-il apparaître l'interaction significative entre *aperture\_V<sub>2</sub>* et les classes d'âge délimitées pour les besoins de l'expérience. Cette tendance est plus visible entre le groupe de sujets nés entre 1923-45 et le groupe né entre 1967-73. En revanche, les différences entre chacun de ces deux groupes et le groupe intermédiaire sont plus ténues (et non significatives pour l'effet d'harmonie vocalique sur /O/). Ainsi, il ne serait pas inexact d'estimer, sans pour autant que l'échantillon analysé ici ne nous autorise à l'affirmer avec certitude, que la réalisation d'harmonie vocalique, telle qu'exprimée au travers des trois générations de locuteurs enregistrés dans le corpus BREF, suit la tendance à réduire le degré d'harmonisation avec le temps.

En résumé, l'analyse des corpus BREF, ESTER et NCCFr a indiqué la présence d'un effet d'assimilation régressive de voyelle à voyelle (*harmonie vocalique*) affectant les V<sub>1</sub> ∈ {e, ε, o, ɔ} en syllabe pénultième sous l'influence du timbre (*fermé/mi-fermé* ou *ouvert/mi-ouvert*) de la V<sub>2</sub>.

De plus, le degré d'harmonie vocalique se montre sensible au style de parole (*lue, journalistique* et *spontanée*), quoique cet effet diffère en fonction de la voyelle analysée. Ainsi, l'harmonie vocalique sur /e, ε, o/ est légèrement plus forte en parole lue par rapport à la parole journalistique et plus forte en parole journalistique par rapport à la parole spontanée. En revanche, le degré d'harmonie vocalique sur /ɔ/ se révèle significativement plus important en parole spontanée par rapport à la parole journalistique et à la parole lue.

Une analyse de l'effet d'assimilation régressive de voyelle à voyelle des V<sub>1</sub> en syllabe pénultième, comparée au même type d'assimilation en syllabes initiales ou internes non pénultième, sous l'influence de la voyelle en syllabe adjacente, a conduit à distinguer deux schémas distincts, selon l'aperture phonologique de la V<sub>1</sub>. Ainsi, les /ɔ, ε/ subissent une assimilation plus forte dans les contextes d'harmonie vocalique par rapport à d'autres positions dans le mot, tandis que le degré d'assimilation des /o, e/ se révèle pratiquement égal sous l'effet d'harmonie vocalique et dans d'autres positions.

Enfin, une étude préliminaire du corpus BREF *en temps apparent* suggère que le degré d'harmonie vocalique est plus important dans le groupe de locuteurs nés entre 1923 et 1945 par rapport aux groupes de locuteurs nés dans des tranches de temps plus récentes.

## 4. Expérience de production

### 4.1. Méthode

#### 4.1.1. Corpus

La structure du corpus que nous avons conçu a été, dans une large mesure, inspirée de celui utilisé par Fagyal *et al.* (2003) et Nguyen & Fagyal (2008). 33 paires de mots dissyllabiques présentant un contexte possible d'harmonie vocalique ont été sélectionnées pour représenter un adjectif, un nom (commun, sauf l'unique mot propre *Rosine*), ou un verbe (conjugué ou à l'infinitif). Les contextes d'harmonie vocalique sont définis en tant que *paires (quasi) minimales* représentant une quantité relativement restreinte dans le lexique français, si ce n'est pour des mots très peu fréquents du type *Messire* (comme paire de *message*) ou *Tolède* (paire de *tollé*). 27 paires dans notre corpus ont été choisies parmi celles analysées par Nguyen & Fagyal (2008), de façon à correspondre à une fréquence d'usage possiblement élevée.

Dans chaque paire, la première syllabe, toujours ouverte, contient ainsi une  $V_1 \in \{e, \varepsilon, o, \circ\}$ . Contrairement à toutes les études expérimentales antérieures, c'est-à-dire celle de Nguyen & Fagyal (2008), Fagyal *et al.* (2003), Duniec & Crouzet (2013) et Turco *et al.* (2016a, 2016b), nous avons inclus six paires supplémentaires contenant une  $V_1 \in \{\emptyset, \text{œ}\}$  également en syllabe ouverte. En effet, les voyelles moyennes antérieures arrondies ont été décrites comme faisant partie des contextes d'harmonie vocalique dans plusieurs traités phonologiques (Malmberg, 1969 ; Morin, 1971 ; Casagrande, 1984 ; Tranel, 1987), sans avoir fait jusqu'alors l'objet d'une étude empirique.

L'attaque de la première syllabe, si elle est remplie, est constituée soit d'une seule obstruante  $C_1 \in \{p, k, d, s, \text{ʒ}, m, n, \text{ʁ}\}$ , soit d'un groupe consonantique obstruante + liquide parmi les suivants :  $C_1C_2 \in \{kl, fl, fr\}$ . Dans chaque paire, l'une des voyelles finales est fermée ou mi-fermée  $V_2 \in \{e, i, o, \emptyset\}$  et l'autre, ouverte ou mi-ouverte  $V_2 \in \{\varepsilon, a, \circ, \text{œ}\}$ .

En vertu des travaux menés par, entre autres, Öhman (1966), Recasens (1987) et Fowler & Brancazio (2000), il est admis que les consonnes intervocaliques ont un effet significatif sur le phénomène d'assimilation de voyelle à voyelle. Ceci a également été observé pour l'harmonie vocalique en français, notamment dans Turco *et al.* (2016a) et Cecelewski (2022). Partant, dans la présente étude, nous avons privilégié les contextes coronaux et labiaux comme ayant un impact relativement limité sur la nature de l'effet coarticulatoire de voyelle à voyelle. Nous avons accepté cependant quelques contextes dorsaux parmi les paires en /E/, principalement en raison d'une plus grande rareté de ces dernières dans le lexique français (en comparaison avec /E/ et /O/). Cela nous conduit à un compromis méthodologique, nécessaire pour les cas où la fréquence lexicale trop basse pourrait elle-même influencer sur le degré d'harmonie vocalique, en induisant de l'hyper-articulation. Le **Tableau 9** illustre la structure de notre corpus en termes du nombre total d'occurrences enregistrées dans les différents contextes mentionnés ci-dessus.

C <sub>3</sub> (C <sub>4</sub> ) \ V <sub>1</sub>	/E/	/O/	/œ/
<i>coronale</i>	672	868	228
<i>labiale</i>	502		
<i>dorsale</i>			336
<i>labiale+coronale</i>			112
<i>dorsale+coronale</i>	228		

**Tableau 9.** Nombre total d'occurrences en fonction de V<sub>1</sub> ∈ {E, O, œ} et de C<sub>3</sub>(C<sub>4</sub>) *coronale* ∈ {t, s, ʃ, z, l, n}, *labiale* ∈ {p, v}, *dorsale* /j/, *labiale+coronale* /p/, *dorsale+coronale* /g/.

En plus des contextes d'harmonie vocalique ainsi définis, nous avons inclus quatre paires de mots distracteurs présentant des pseudo-contextes d'harmonie, c'est-à-dire avec une V<sub>1</sub> /a/ en syllabe initiale suivie d'une V<sub>2</sub> haute ou basse, ainsi que dix mots monosyllabiques choisis de façon à inclure l'ensemble des voyelles orales du français standard, soit V ∈ {i, y, e, ε, ø, œ, a, o, ɔ, u}.

Les mots ainsi sélectionnés ont été insérés dans la même phrase-cadre que celle utilisée par Nguyen & Fagyal (2008) et avant par Fagyal *et al.* (2003) :

*Il retape \_\_\_\_\_ parfois, \_\_\_\_\_.*

Ainsi, le même mot-cible était produit deux fois, la première fois au sein d'un groupe rythmique et, la seconde fois, en fin de phrase avant la finale absolue. Cette phrase-cadre, mis à part sa formulation quelque peu artificielle, a le mérite d'enchaîner le mot-cible entre deux contextes labiaux identiques, susceptibles, d'une part, de favoriser une segmentation univoque et, de l'autre, de minimiser un effet coarticulatoire autre que celui qui faisait l'objet de l'étude-même.

#### 4.1.2. Sur la loi de position en français standard : stratégie de transcription

Dans leurs deux études complémentaires sur l'harmonie vocalique, Nguyen & Fagyal (2008) et Fagyal *et al.* (2003) ont décidé d'adopter systématiquement une transcription des voyelles-cibles et des voyelles-sources comme étant des voyelles moyennes mi-fermées. Leur transcription fait appel à la *loi de position* supposée régir de façon uniforme la distribution des voyelles moyennes dans la syllabe française, quelle que soit sa position dans le mot : « *VIs are systematically transcribed (...) as high-mid vowels. This is consistent with the loi de position (...), which we adopted as a rule of transcription* » (Nguyen & Fagyal, 2008). Ce choix s'explique par l'un des objectifs de l'étude qui a été de comparer la réalisation des voyelles moyennes dans des contextes d'harmonie vocalique dans deux variétés du français dont, d'une part, une variété du Nord, que l'on pourrait approcher du *français standard*, ou de la *variété de référence*, (« *Standard or General French* » pour citer le terme de Nguyen & Fagyal, 2008, cf. Di Cristo, 1998) sans que ce terme soit lui-même univoque, et, d'autre part, une variété du Sud, dans laquelle :

there is no contrastive distinction between /e/-/e/, /ø/-/œ/ and /o/-/ɔ/, and the distribution of the mid-high and mid-low variants is said to be entirely governed by a variant of the loi de position: the mid-high allophone occurs in open syllables and the mid-low allophone in closed syllables and whenever the next syllable contains a schwa (Nguyen & Fagyal, 2008).

Si tant est qu'elle fasse un accord, ne serait-ce que partiel et souvent fragile, entre les théoriciens, la loi de position, en ce qui concerne sa définition en *position accentuée*, c'est-à-dire en syllabe finale de mot (cf. Dauzat, 1930 ; Fouché, 1935 ; Delattre, 1951, 1959 ; Valdman, 1978), est bien moins nette pour définir le degré d'aperture en *position inaccentuée*. Dans cette première, la question qui prête à débat est avant tout de savoir s'il s'agirait d'une loi diachronique qui advient à la fin du XV<sup>e</sup> siècle et qui « n'a pas encore épuisé ses possibilités » (Fouché, 1935), ou plutôt d'une règle phonologique mise en place pour décrire une distribution synchronique, auquel cas, comme le remarque Morin (1998), elle serait tout de même « grossièrement inexacte ».

Concernant la *position inaccentuée* (initiale de mot), selon la définition de Straka (1981), la loi de position définie en fonction de la présence ou non d'une coda syllabique, imposerait autant une fermeture de /ɛ/ et /œ/ en syllabe ouverte, qu'une ouverture de /o/ dans la même position. En réalité, une définition simple en forme de règle (ou *loi*) de la distribution des timbres des voyelles moyennes semble limitée ne serait-ce que par le fait qu'il est « difficile d'isoler l'évolution phonétique des effets de l'analogie ou de l'influence de la graphie dans cette position » (Morin, 1998).

En ce qui concerne les contextes d'harmonie vocalique, outre l'hésitation des ouvrages de référence quant à la transcription phonologique dans de nombreux cas (e.g. *bêtise* [betiz]/[be-] ou *rêver* [ʁeɛv]/[re-] selon le CNRTL), une difficulté supplémentaire est que les voyelles moyennes antérieures arrondies en position inaccentuée sont rares en français, et encore plus en tant que paires harmoniques satisfaisant aux critères définis ci-avant (ainsi, l'unique paire avec un /ø/ phonologique a été trouvée pour être incluse dans notre corpus, c'est-à-dire *meuler – meulage*).

En conséquence, notre stratégie de transcription se déclinera en deux approches complémentaires. Dans un premier temps, nous adopterons une transcription des V<sub>1</sub>, en respectant les deux degrés d'aperture des voyelles moyennes, afin de distinguer les voyelles mi-fermées /e, o/ et les voyelles mi-ouvertes /ɛ, œ/ en syllabe initiale. En revanche, par manque de contextes d'harmonie vocalique contenant une V<sub>1</sub> /ø/, les voyelles moyennes antérieures arrondies ne feront pas l'objet de cette partie de l'analyse. La transcription sera conforme à celle adoptée par CNRTL sauf pour les mots du corpus pour lesquels deux transcriptions possibles étaient proposées. Auquel cas nous respecterons l'aperture de la voyelle présente dans la base dérivationnelle (*rêver* sera donc transcrit [ʁeɛv] car *rêve* [ʁɛv], cf. liste complète des transcriptions dans l'**Annexe 5**).

Pourtant, les divergences et incertitudes parmi les théoriciens sur le statut phonologique des voyelles moyennes en position prétonique impliquent que quelle que soit la stratégie de transcription adoptée, elle ne sera pas nécessairement dénuée d'un biais de catégorisation arbitraire. Enfin, comme le précisent à juste titre Nguyen & Fagyal (2008) « *The actual quality of V1 was of course an empirical issue that this acoustic study was precisely intended to address* ». C'est pourquoi, dans un second temps, nous adopterons une transcription des V<sub>1</sub> comme étant des archiphonèmes /E, O, Œ/, ce qui rendra nos résultats comparables, toutes propositions gardées, à ceux obtenus par Nguyen

& Fagyal (2008), en imposant toutefois d'abandonner, dans cette seconde partie, la question phonologique de la directionnalité du processus d'harmonie vocalique.

### 4.1.3. Locuteurs et enregistrements

Six sujets masculins ont participé à l'étude. Tous sont nés dans la région d'Île-de-France où ils résidaient encore au moment des enregistrements. Aucun d'entre eux ne présentait de traits de prononciation dialectaux discernables ni de troubles de la voix et de la parole. Leur niveau d'éducation et leur profil social étaient également comparables, étant tous soit diplômés d'études supérieures, soit des étudiants à l'Université Paris-Cité (locuteurs BA, BS, PA, TR) ou à l'Université Paris-Saclay (locuteur HU). À l'exception du locuteur BA, qui était étudiant en Sciences du Langage (parcours « Linguistique Théorique et Expérimentale »), les autres locuteurs n'avaient pas de connaissances particulières en phonétique et n'ont pas été informés de l'objet de l'étude. Le **Tableau 3.** résume les métadonnées des locuteurs.

code du locuteur	âge	lieu de résidence
AR	31	Paris 17 <sup>e</sup>
BA	22	Massy (Essonne)
GA	29	Malakoff (Hauts-de-Seine)
HU	22	Paris 13 <sup>e</sup>
PA	24	Meaux (Seine-et-Marne)
TR	21	Paris 8 <sup>e</sup>

**Tableau 3.** Code du locuteur, âge et lieu de résidence.

Les enregistrements ont été réalisés en chambre sourde et à l'aide d'un microphone de qualité. La production sonore a été enregistrée à la fréquence d'échantillonnage de 44,1 kHz (16 bits), et reconvertie ultérieurement à 22,05 kHz.

La liste des 84 phrases contenant les contextes d'HV, les mots distracteurs « pseudo-contextes d'HV » et les mots monosyllabiques, a été lue quatre fois par chaque locuteur. Lors de la première lecture, les locuteurs ont été invités à lire chaque phrase à un rythme normal, sans mettre l'accent sur un mot spécifique et sans marquer de pause entre les mots dans le premier groupe intonatif (jusqu'à « parfois »). Pour minimiser l'effet possible de l'habitude ou de l'ennui, la position d'une phrase qui apparaissait sur l'écran d'ordinateur était à chaque fois légèrement décalée par rapport à la phrase précédente. Pendant la deuxième lecture, les locuteurs se tenaient debout dans un coin de la chambre sourde pour donner l'illusion d'un espace légèrement plus grand. Ils ont reçu la consigne de parler fort, comme s'ils s'adressaient à une personne située dans l'angle opposé de la pièce, à une distance d'environ 2 à 2,5 mètres. Avant chacune des deux sessions d'enregistrements, les locuteurs ont été invités à maintenir autant que possible le même débit de parole, le même rythme et le même contour intonatif tout au long de leur lecture du corpus. La séquence de lectures en voix modale et en voix forte a été ensuite répétée après une pause (dont la durée variait pour chaque locuteur, allant de 30

minutes à environ 2 heures), de manière à obtenir quatre lectures au total, dont deux en voix modale et deux en voix forte.

#### 4.1.4. Mesures acoustiques

L'ensemble du corpus consistant en 2016 phrases porteuses ainsi enregistrées a été segmenté à l'aide du logiciel WebMAUS (Kisler *et al.*, 2017). La segmentation de chaque mot-cible a été ensuite inspectée visuellement et, si nécessaire, corrigée à la main. Un script a été utilisé pour l'extraction automatique des valeurs correspondant à la hauteur des deux premiers formants à  $\frac{1}{3}$ ,  $\frac{1}{2}$  et  $\frac{2}{3}$  de la durée de la  $V_1$ , qui ont été ensuite moyennés pour obtenir une valeur unique. L'algorithme Burg implémenté dans Praat (Boersma & Weenink, 2016) a été utilisé avec les paramètres suivants : préemphasis à 50 Hz, plage de détection inférieure à 4,8 kHz et fenêtre d'analyse de 25 ms. Les computations statistiques ont été conduites dans l'environnement R (R Development Core Team, 2010 ; version 2022.12.0+353).

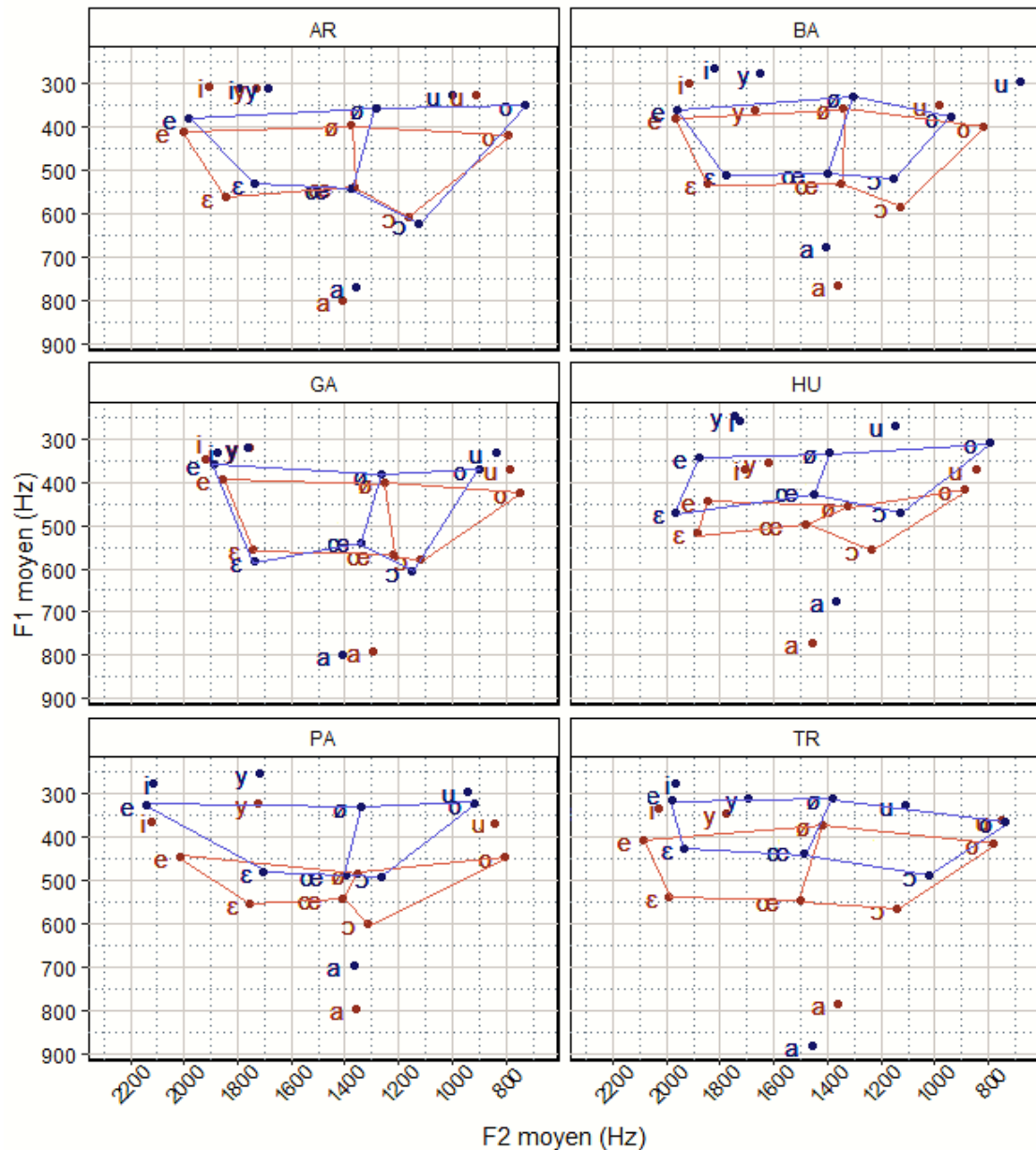
## 4.2. Résultats

### 4.2.1. Système des voyelles orales dans les productions des six locuteurs

Avant de commencer l'analyse du degré d'harmonie vocalique dans les données, nous représentons, dans la **Figure 8**, les valeurs moyennes des deux premiers formants des voyelles orales /i, y, e, ε, ø, œ, a, o, ɔ, u/ produites dans les mots monosyllabiques indiqués dans l'**Annexe 5**, séparément pour les six locuteurs enregistrés. La composition de notre corpus n'étant pas délibérément destinée à permettre une analyse exhaustive du système vocalique complet de nos participants, cette représentation ne dépasse pas un cadre exploratoire. Les points bleus dans l'espace F1 x F2 représentent les valeurs moyennes des formants des voyelles produites en voix modale. En brun ont été indiquées les mêmes valeurs mesurées sur les voyelles produites en voix forte, dans les mêmes contextes. Les lignes de même couleur relient les voyelles moyennes mi-fermées et mi-ouvertes /e, ε, ø, œ, o, ɔ/.

Ainsi, la **Figure 8** semble confirmer la présupposition initiale, nécessaire à l'analyse de l'harmonie vocalique, qui est que tous les six locuteurs distinguent les deux degrés d'aperture des voyelles moyennes. Les formes des trapèzes indiquant sommairement les patrons individuels diffèrent nettement d'un locuteur à un autre, en suggérant notamment une centralisation plus accentuée de /ε, ɔ/, avec une tendance parallèle à la maximisation du trait d'antériorité/postériorité de /e, o/ pour le locuteur PA, ainsi qu'une articulation relativement antérieure de /ε/ pour le locuteur HU. La distinction la plus prononcée entre, d'une part, les voyelles moyennes mi-fermées /e, ø, o/ et, d'une autre part, les voyelles mi-ouvertes /ε, œ, ɔ/, semble être réalisée par le locuteur GA, du moins en termes d'aperture et en position accentuée, ce qui laisse prédire un degré d'harmonie vocalique plus important.





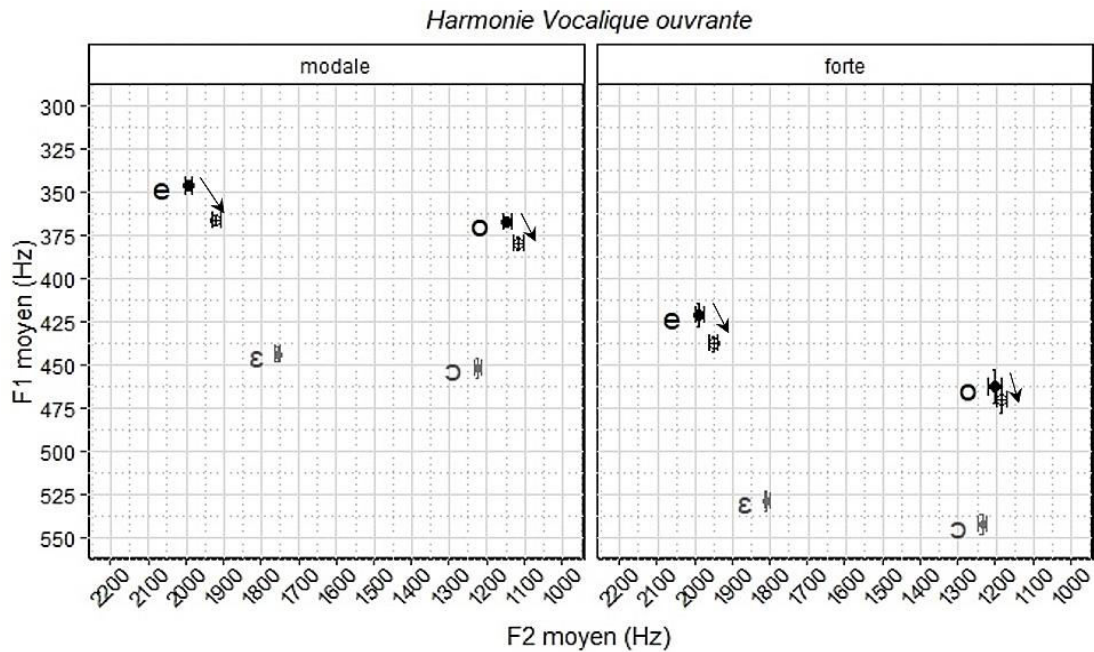
**Figure 8.** Formants moyens des 10 voyelles orales du français (en Hertz) prononcées par les locuteurs dans les mots monosyllabiques. En bleu : *voix modale* ; en brun : *voix forte*.

#### 4.2.2. Harmonie vocalique ouvrante sur /e, o/

Dans la **Figure 9**, nous illustrons l'effet d'harmonie vocalique sur les  $V_1$  phonologiques mi-fermées sous l'influence d'une  $V_2$  phonologique ouverte ou mi-ouverte, par exemple, *épaisse*, *rosette*, par rapport au cas neutre où la  $V_1$  est suivie d'une  $V_2$  phonologique fermée ou mi-fermée, par exemple, *épée*, *rosé*. Pour un repérage visuel plus commode et en guise d'illustration de la magnitude d'effet en question, le positionnement des  $V_2$  phonologiques mi-ouvertes / $\epsilon$ ,  $\text{ɔ}$ / a été inclus dans l'espace F1 x F2. Les flèches montrent la direction de l'effet assimilatoire. Contrairement au protocole d'étude de Nguyen & Fagyal (2008), nous n'avons pas examiné séparément l'effet des  $V_2$  ouvertes/fermées (paires en  $i-a$ ) par rapport aux mi-ouvertes/mi-fermées (paires en  $o-\text{ɔ}$ ,  $e-\epsilon$ ,  $\phi-\text{œ}$ ) en partant de la définition phonologique d'harmonie vocalique qui, elle, ne fait pas cette distinction.

Afin de tester si les différences de hauteurs de F1 et de F2 étaient significativement différentes pour les V<sub>1</sub> suivies d'une V<sub>2</sub> haute ou basse et, subsidiairement, si l'effet de la V<sub>2</sub> était significativement différent en fonction du style de parole, nous avons effectué une ANOVA à deux facteurs avec interaction en utilisant la formule suivante :  $\text{aov}(F\alpha \sim \text{aperture\_V2} * \text{voix})$ . L'analyse réalisée pour /e/ a indiqué un effet significatif d'*apertureV<sub>2</sub>* (F=16.398, p<.001, Df=1) et de *voix* (F=270.311, p<.001, Df=1) sur la hauteur de F1. En revanche, l'interaction entre ces deux prédicteurs n'a pas été significative (F=0.166, p=.06, Df=1). Un effet significatif d'*apertureV<sub>2</sub>* a également été mis en évidence pour le F2 (F=25.735, p<.001, Df=1), contrairement à l'effet de *voix* qui, lui, n'a pas été significatif (F=1.427, p=.02, Df=1), de même que l'interaction entre *apertureV<sub>2</sub>* et *voix* (F=2.615, p=.03, Df=1). L'analyse effectuée pour /o/ a indiqué que l'*apertureV<sub>2</sub>* n'avait d'effet significatif ni sur le F1 (F=2.745, p=.09, Df=1) ni sur F2 (F=2.655, p=.1, Df=1). En revanche, nous avons trouvé un effet significatif de la *voix*, pour F1 et F2, respectivement (F=190.486, p<.001, Df=1) et (F=18.206, p<.001, Df=1). Le terme d'interaction n'a pas non plus conduit à l'observation d'un effet significatif, tant pour F1 (F=0.124, p=.7, Df=1) que pour F2 (F=0.170, p=0.6, Df=1).

Ainsi, nos données suggèrent que, dans l'ensemble, l'effet d'harmonie vocalique fermante est négligeable, sauf pour le /e/ qui se révèle significativement plus ouvert et plus postérieur lorsque suivi d'une V<sub>2</sub> phonologique ouverte ou mi-ouverte. Or, force est de reconnaître qu'avec un  $\Delta F1$  moyen de 18 Hz et un  $\Delta F2$  moyen de 61 Hz, cet effet, quoique significatif, est très limité. Le fait que le degré d'harmonie vocalique mis en évidence pour le /o/ se soit révélé si peu important pourrait avoir comme cause plausible l'influence de la graphie « au », présente ici dans quatre sur six paires de mots contenant la V<sub>1</sub> postérieure arrondie mi-fermée phonologique. En effet, la tendance à réduire l'harmonie vocalique a été décrite pour les graphies « au, eau, ô » par Turco *et al.* (2016a), dans une parole journalistique et conversationnelle contemporaine, et par Cecelewski (2022), dans une parole journalistique de la seconde moitié du XX<sup>e</sup> siècle. Qui plus est, nous pouvons supposer que cet effet ne fait que s'accroître lorsqu'il s'agit d'une parole lue, précisément comme dans le cas du corpus analysé ici.



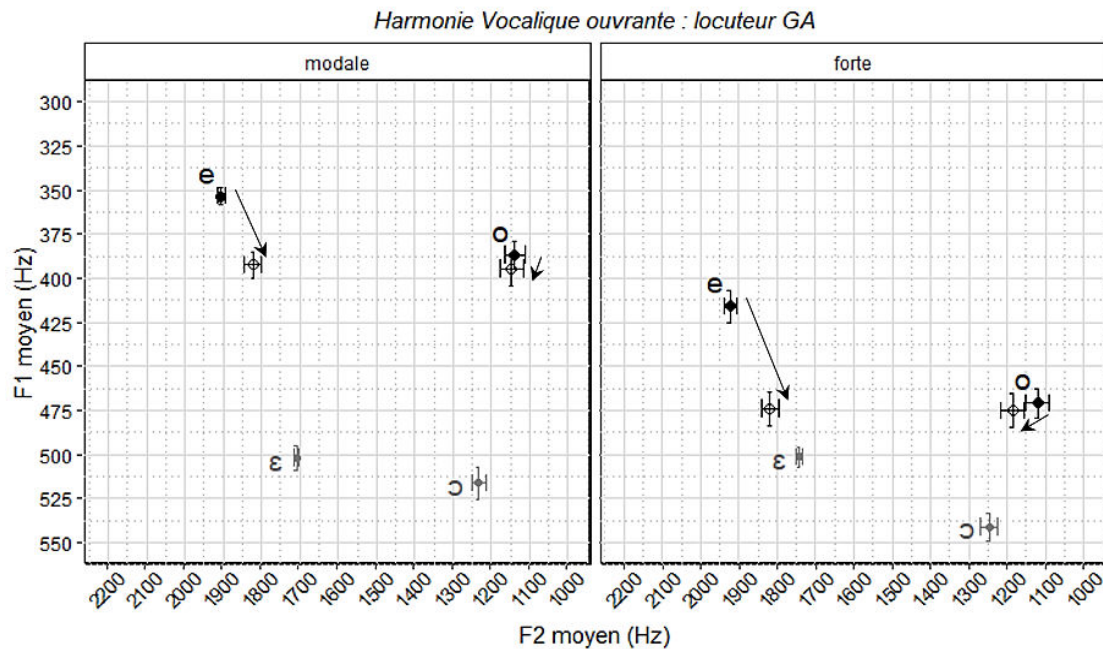
**Figure 9.** Valeurs moyennes de F1 et F2 (en Hertz) et erreur-type ; en noir :  $V_1 \in \{e, o\}$  phonologiques suivies d'une  $V_2$  haute (point plein) et d'une  $V_2$  basse (point vide) ; en gris :  $V_2 \in \{\epsilon, \omicron\}$  phonologiques ; tous les six locuteurs confondus.

Afin d'inspecter les différences individuelles entre locuteurs, une ANOVA à deux facteurs avec interaction a été réalisée séparément pour les six locuteurs du corpus. Les résultats statistiques détaillés sont consignés dans l'**Annexe 2**. En somme, nous n'avons pas trouvé d'effet significatif de l'aperture de la  $V_2$  sur la hauteur de F1 et de F2 dans le cas de /o/ chez aucun des six locuteurs, et ce, indépendamment du style de parole. De même, pour aucun des six locuteurs, le degré d'assimilation ne s'est montré sensible au style de parole. En revanche, un effet d'harmonie vocalique sur /e/ entraînant une valeur de F1 significativement plus élevée a été mis en évidence pour les locuteurs AR, BA, GA et HU, la tendance à l'ouverture de /e/ se révélant plus marquée en voix forte par rapport à la voix modale. Le locuteur GA a été le seul à avoir montré également une tendance significative à la postériorisation de /e/ sous l'effet d'harmonie vocalique. Dans le **Tableau 10**, les différences moyennes entre les valeurs des deux premiers formants pour les  $V_1$  /e/ et /o/ ont été présentées pour chaque locuteur, séparément pour la voix modale et la voix forte. Les valeurs en gras indiquent les différences significatives en Hertz.

HV	voix	$V_1$	formant	locuteur					
				AR	BA	GA	HU	PA	TR
ouvrante	modale	/e/	$\Delta F1$	+15	+27	<b>+39</b>	+12	+22	+15
			$\Delta F2$	-56	-36	<b>-83</b>	-41	-77	-105
	forte	/e/	$\Delta F1$	<b>+35</b>	<b>+33</b>	<b>+58</b>	<b>+39</b>	+4	+7
			$\Delta F2$	-55	-4	<b>-104</b>	-19	-28	-40
	modale	/o/	$\Delta F1$	+19	+3	+4	+11	+25	+7
			$\Delta F2$	-59	-52	-8	-20	-1	-12
forte	/o/	$\Delta F1$	+12	+22	+8	+14	+12	+15	
		$\Delta F2$	-5	-42	-64	-17	-41	-2	

**Tableau 10.** Différences de valeurs moyennes de F1 et de F2.  $\Delta F\alpha$  indique la différence de la valeur de formant  $\alpha$  lorsque la  $V_1$  était suivie d'une  $V_2$  phonologique ouverte ou mi-ouverte. Les différences significatives sont en gras ( $p < .05$ ).

Nous ne souhaitons pas encombrer la présente analyse par des visualisations séparées des patrons d'harmonie vocalique pour chacun des six locuteurs. Cependant, nous nous autoriserons tout de même à illustrer, dans la **Figure 10**, l'effet de l'harmonie ouvrante tel qu'il a été mis en évidence dans le cas du locuteur GA, dont l'articulation s'approche le plus, en contraste aux autres locuteurs, d'une image phonologique de l'harmonie vocalique en français. Nous pouvons ainsi observer une différence particulièrement prononcée entre une  $V_1$  /e/ suivie d'une  $V_2$  haute (points noirs pleins) par rapport au cas où elle était suivie d'une  $V_2$  phonologique basse (points noirs vides), en se rapprochant, en termes de degré d'aperture, d'un /ε/ mi-ouvert « canonique », en position forte en fin de mot (point gris). Remarquons également que l'effet ainsi décrit est sensiblement moins important dans le cas de /o/, contrairement à la tendance mise en évidence aussi bien par Nguyen & Fagyal (2008), mais également par Turco *et al.* (2016a), selon laquelle les voyelles moyennes postérieures tendraient à s'assimiler plus que les voyelles moyennes antérieures non-arrondies à l'aperture de la voyelle finale.

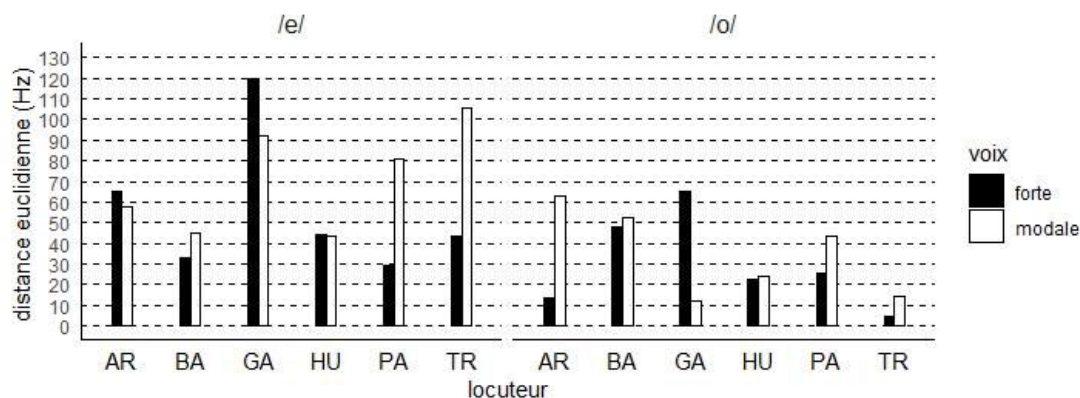


**Figure 10.** Valeurs moyennes de F1 et F2 (en Hertz) et erreur-type dans la production sonore du locuteur GA ; en noir :  $V_1 \in \{e, o\}$  phonologiques suivies d'une  $V_2$  haute (point plein) et d'une  $V_2$  basse (point vide) ; en gris :  $V_2 \in \{\epsilon, \circ\}$  phonologiques.

Pour conclure, afin de mieux appréhender le niveau d'assimilation des deux voyelles moyennes mi-fermées étudiées, la **Figure 11** illustre les distances euclidiennes moyennes indiquées en Hertz, qui séparent une voyelle initiale /e, o/ lorsqu'elle est suivie d'une voyelle finale haute ou basse, dans l'espace F1 x F2, pour chaque locuteur. Cette mesure, qui ne vise qu'à fournir une indication, a été utilisée pour évaluer l'écart entre deux segments vocaliques dans un espace articulatoire théorique.

Observons que le degré d'assimilation relatif à la voyelle antérieure /e/ s'avère particulièrement marqué chez les locuteurs PA et TR. Cette constatation contraste nettement avec l'hypothèse initiale suggérant que l'effet serait plus marqué en voix forte qu'en voix modale. Toutefois, il convient de

souligner que ces mêmes locuteurs ne manifestent qu'une assimilation limitée en ce qui concerne la voyelle moyenne postérieure.



**Figure 11.** Distances moyennes en Hertz entre les  $V_1 \in \{e, o\}$  suivies d'une  $V_2$  phonologique basse ou haute, calculées en tant que distances euclidiennes dans l'espace  $F1 \times F2$  selon la formule suivante :  $\sqrt{((F2_{V_1}-F2_{V_2})^2 + (F1_{V_1}-F1_{V_2})^2)}$ .

#### 4.2.3. Harmonie vocalique fermante sur /ε, ɔ/

Nous avons ensuite examiné l'effet d'harmonie vocalique fermante, c'est-à-dire agissant sur une  $V_1$  phonologique mi-ouverte /ε, ɔ/ sous l'influence d'une  $V_2$  fermée ou mi-fermée, comme dans *bêtise* et *noté* par rapport à *bêtasse* et *nota*, ces derniers étant considérés ici comme étant des cas « neutres ». Cette influence en termes du premier et du second formants est illustrée dans la **Figure 12**.

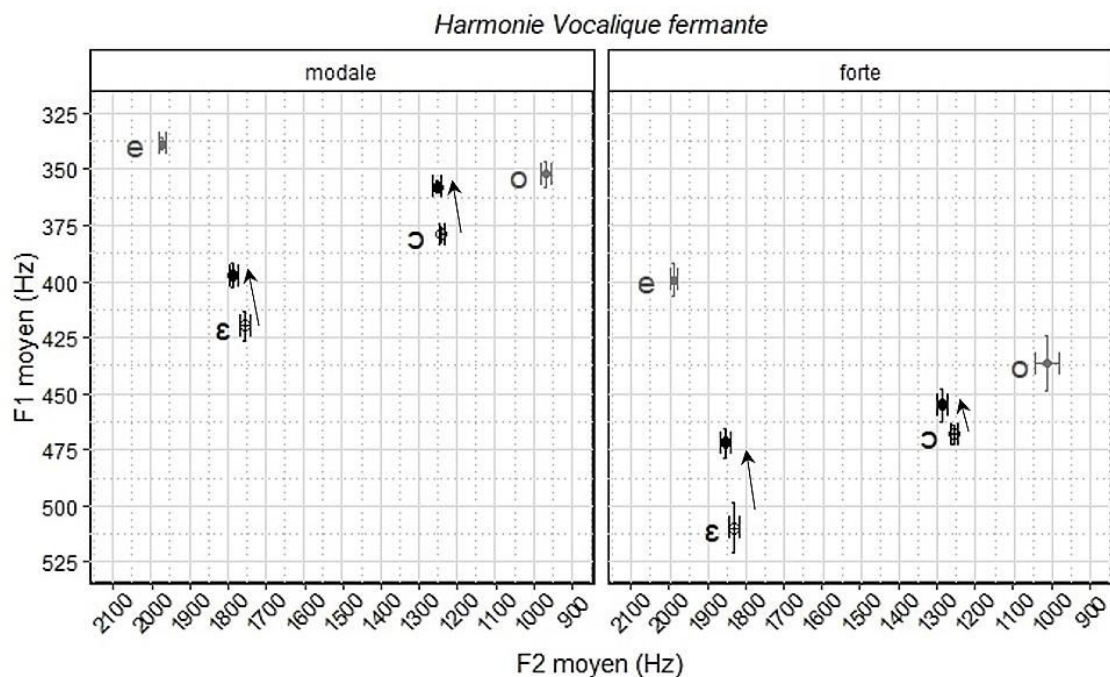
Il est à noter que notre choix méthodologique a été de limiter l'analyse de l'assimilation en termes de F2 aux paires de mots avec une  $V_2$  antérieure /e, i/, en excluant celles où la  $V_2$  était un /ø/, car cela aurait conduit à une neutralisation d'un effet antériorisant de /e, i/ par un effet agissant dans la direction opposée induit par /ø/. Par conséquent, seul l'effet de /e, i/ a été soumis à l'analyse.

Ainsi, une ANOVA à deux facteurs a été effectuée, qui a indiqué, dans le cas de /ε/, un effet significatif d'*aperture* $V_2$  ( $F=16.07$ ,  $p<.001$ ,  $DI=1$ ) et de *voix* ( $F=121.08$ ,  $p<.001$ ,  $DI=1$ ) sur la hauteur de F1, mais pas d'effet significatif du terme d'interaction entre ces deux variables ( $F=1.05$ ,  $p=.3$ ,  $DI=1$ ). Un effet significatif d'*aperture* $V_2$  ( $F=3.868$ ,  $p<.05$ ,  $DI=1$ ) et de *voix* ( $F=30.490$ ,  $p<.001$ ,  $DI=1$ ) a également été trouvé pour le second formant. Toutefois, l'interaction entre *aperture* $V_2$  et *voix* ne s'est pas révélée significative ( $F=0.071$ ,  $p=.79$ ,  $DI=1$ ). Dans le cas de /ɔ/, nous avons trouvé un effet significatif d'*aperture* $V_2$  ( $F=15.103$ ,  $p<.001$ ,  $DI=1$ ) et de *voix* ( $F=416.269$ ,  $p<.001$ ,  $DI=1$ ) sur la hauteur de F1, contrairement à l'interaction entre ces deux prédicteurs qui, elle, ne s'est pas montrée significative ( $F=0.834$ ,  $p=.36$ ,  $DI=1$ ). Un effet significatif a également été trouvé pour les deux variables, *aperture* $V_2$  ( $F=4.616$ ,  $p<.05$ ,  $DI=1$ ) et *voix* ( $F=5.416$ ,  $p<.05$ ,  $DI=1$ ), ces dernières révélant un impact sur la hauteur de F2. En revanche, le terme d'interaction n'a pas été significatif ( $F=1.076$ ,  $p=.29$ ,  $DI=1$ ).

Dans l'ensemble, nous pouvons ainsi conclure à la présence d'un effet d'harmonie vocalique fermante dans notre corpus, les voyelles moyennes mi-ouvertes montrant une tendance à la fermeture



sous l'influence d'une  $V_2$  phonologique haute ou mi-haute, ainsi qu'une tendance à l'antériorisation sous l'impact d'une  $V_2$  phonologique antérieure fermée ou mi-fermée. Toujours est-il que le degré d'harmonie vocalique exprimé en termes de différences de hauteurs de F1 et de F2 est de l'ordre de 25 Hz en moyenne pour le premier et, de 50 Hz pour le second formant. Ainsi, que ce soit en termes d'aperture ou d'antériorité, les  $V_1$  mi-ouvertes se trouvant dans un contexte d'harmonie vocalique restent relativement éloignées de leurs contreparties mi-fermées (indiquées en gris sur la **Figure 12**). De surcroît, cet effet ne se révèle pas sensible au style de parole. En effet, s'il est vrai que la voix forte induit un rehaussement du premier formant, indiquant, dans l'ensemble, une articulation plus ouverte, elle n'a pas d'impact sur le degré d'assimilation dû à l'harmonie vocalique.



**Figure 12.** Valeurs moyennes de F1 et F2 (en Hertz) et erreur-type ; en noir :  $V_1 \in \{\varepsilon, \text{ɔ}\}$  phonologiques suivies d'une  $V_2$  basse (point vide) et d'une  $V_2$  haute (point plein) ; en gris : les  $V_2 \in \{e, \text{o}\}$  phonologiques ; tous les six locuteurs confondus.

Examinons à présent l'effet d'harmonie vocalique sur la réalisation des  $V_1$  mi-ouvertes, séparément pour chaque locuteur. Dans l'**Annexe 3** ont été indiqués les résultats de tests statistiques effectués pour chacun des six locuteurs. Dans le **Tableau 11**, nous représentons les différences moyennes de la hauteur du premier et second formants entre une  $V_1$  suivie d'une  $V_2$  ouverte ou mi-ouverte et celle suivie d'une  $V_2$  fermée ou mi-fermée, en voix modale et en voix forte. Les valeurs significatives sont indiquées en gras. Nous pouvons observer que l'effet global d'harmonie vocalique s'estompe, du moins en termes de significativité statistique, dès lors que nous examinons les patrons individuels des locuteurs. Ainsi, les locuteurs AR, HU et TR ne montrent quasiment pas d'effet d'assimilation, autant pour /ε/ que pour /ɔ/. En revanche, l'harmonie vocalique sur /ε, ɔ/ est bel et bien présente dans les productions du locuteur GA, comme cela a également été mis en évidence pour les voyelles mi-fermées produites par le même locuteur. Remarquons tout de même que cet effet est

légèrement inférieur à l'effet global d'harmonie vocalique mis en évidence par Nguyen & Fagyal (2008). De plus, parmi les six locuteurs, une tendance à l'antériorisation, mais pas à la fermeture, a été trouvée dans les réalisations du locuteur PA, en voix modale et en voix forte, pour les deux voyelles moyennes mi-ouvertes en question.

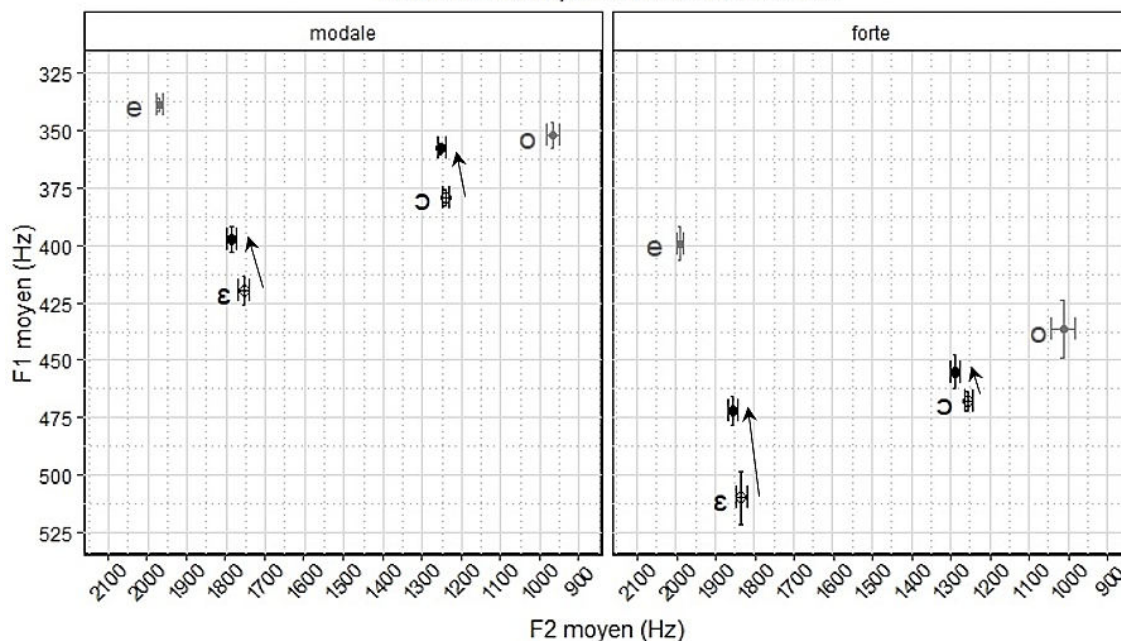
HV	voix	V <sub>1</sub>	formant	locuteur					
				AR	BA	GA	HU	PA	TR
fermante	modale	/ɛ/	ΔF1	-29	-39	<b>-53</b>	-34	-5	-9
			ΔF2	+17	-3	+81	+51	<b>+31</b>	+4
	forte		ΔF1	-22	<b>-30</b>	<b>-26</b>	-35	-13	-15
			ΔF2	+32	+32	<b>+181</b>	+38	<b>+28</b>	+24
	modale	/ɔ/	ΔF1	-10	-28	<b>-35</b>	-35	-21	-4
			ΔF2	+12	+6	+16	+4	<b>+58</b>	+22
			ΔF1	-17	-17	<b>-41</b>	-16	-31	-7
			ΔF2	+19	+15	+2	+4	<b>+112</b>	+27

**Tableau 11.** Différences de valeurs moyennes de F1 et de F2. ΔF $\alpha$  indique la différence de la valeur du formant  $\alpha$  lorsque la V<sub>1</sub> était suivie d'une V<sub>2</sub> phonologique fermée ou mi-fermée. En gras ont été indiquées les différences significatives ( $p < .05$ ).

Complétons notre analyse de l'harmonie vocalique fermante par un exemple d'effet assimilatoire individuel tel que présent dans la production sonore du locuteur GA. La **Figure 13** illustre cet effet pour les V<sub>1</sub> ∈ {ɛ, ɔ} en fonction de l'aperture de la V<sub>2</sub>. Comme cela a été visible pour l'harmonie vocalique ouvrante dans le cas du même locuteur, en particulier pour le /ɛ/, l'effet d'harmonie vocalique fermante se montre nettement plus prononcé, en termes de valeur moyenne absolue, en voix forte par rapport à la voix modale. Rappelons toutefois qu'aucune interaction significative n'a été trouvée entre *aperture*V<sub>2</sub> et *voix* pour l'ensemble des locuteurs, y compris le locuteur GA.

La **Figure 13** représente ainsi un *cas extrême* au sens d'un effet d'harmonie maximisé, chez le locuteur GA, relativement aux autres locuteurs du corpus. Ainsi, un /ɛ/ en syllabe initiale, lorsqu'il est suivi d'une V<sub>2</sub> fermée ou mi-fermée, se rapproche, en termes de son positionnement dans l'espace F1 x F2, d'un /ɛ/ accentué tel que produit par le même locuteur. Toujours est-il que l'effet d'harmonie vocalique sur /ɔ/ demeure, dans la production du locuteur GA, d'un degré similaire à celui des autres locuteurs.

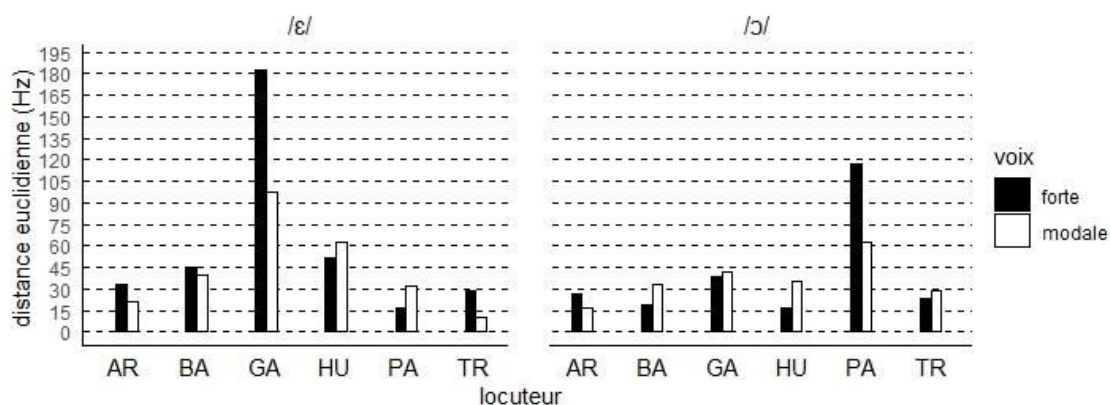
Harmonie Vocalique fermante : locuteur GA



**Figure 13.** Valeurs moyennes de F1 et F2 (en Hertz) et erreur-type dans la production sonore du locuteur GA ; en noir :  $V_1 \in \{\varepsilon, \text{ɔ}\}$  phonologiques suivies d'une  $V_2$  basse (point vide) et d'une  $V_2$  haute (point plein) ; en gris : les  $V_2 \in \{e, o\}$  phonologiques.

Pour conclure, la **Figure 14** illustre le degré d'assimilation exprimé en distances euclidiennes moyennes séparant une  $V_1 \in \{\varepsilon, \text{ɔ}\}$  lorsqu'elle était suivie d'une  $V_2$  basse ou haute, dans l'espace F1 x F2, en Hertz et pour chaque locuteur.

Il est à propos de remarquer que l'impact de l'harmonie vocalique est modéré et que l'influence du style de parole est difficilement généralisable chez les locuteurs AR, BA, GU et TR. Ces derniers contrastent avec les locuteurs GA et PA dont les tendances à l'harmonisation sont divergentes: le premier manifeste une forte propension à l'harmonie sur la voyelle / $\varepsilon$ / avec un effet peu marqué sur / $\text{ɔ}$ /, tandis que le second présente un schéma inverse, avec une harmonie plus marquée sur / $\text{ɔ}$ / et un effet moins prononcé sur / $\varepsilon$ /.



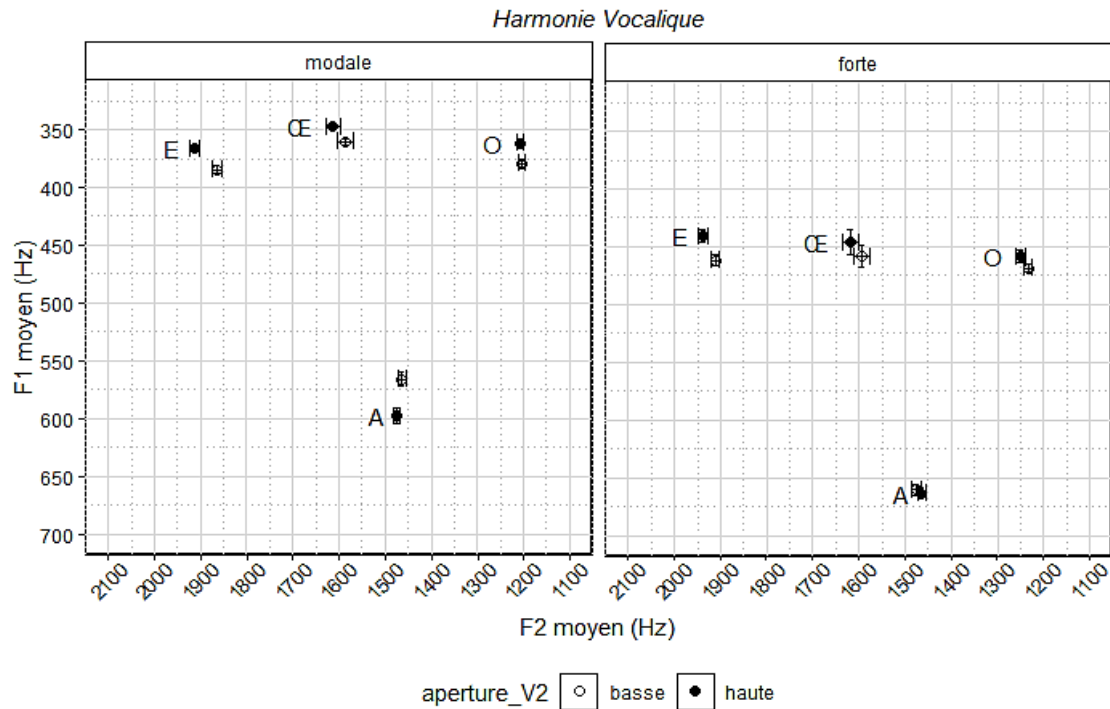
**Figure 14.** Distances moyennes en Hertz entre les  $V_1 \in \{\varepsilon, \text{ɔ}\}$  suivies d'une  $V_2$  phonologique basse ou haute, calculées en tant que distances euclidiennes dans l'espace F1 x F2 selon la formule suivante :  $\sqrt{((F2_{V_1}-F2_{V_2})^2 + (F1_{V_1}-F1_{V_2})^2)}$ .



#### 4.2.4. Harmonie vocalique sur /E, œ, O, A/

Dans cette seconde partie de l'analyse, notre choix méthodologique a donc été de considérer les voyelles moyennes dans les contextes phonologiques d'harmonie vocalique comme représentant des archiphonèmes /E, œ, O/, non spécifiés en termes de hauteur dans cette position. Ainsi, les V<sub>1</sub> dans *été* et *rêveuse*, *meuler* et *jeunisme*, *sauter* et *colline* seront considérés comme appartenant aux mêmes catégories, respectivement /E, œ, O/ suivis d'une V<sub>2</sub> phonologique haute, représentées par des points pleins dans la **Figure 15**. Leurs contreparties suivies d'une V<sub>2</sub> basse, respectivement *éther* et *rêveur*, *meulage* et *jeunesse*, *sauternes* et *collage*, ont été représentés par des points vides. De plus, suivant, d'une part, l'étude de Turco *et al.* (2016a) et, d'une autre part, celle de Cecelewski (2022) où des comparaisons similaires ont été effectuées, nous incluons la voyelle /A/ en qualité d'une V<sub>1</sub>-contrôle exclue des contextes d'harmonie vocalique dans la phonologie française. Les contextes en /A/ que nous analyserons ici sont détaillés dans l'**Annexe 5** et ont été choisis pour correspondre aux contextes d'harmonie vocalique telle que décrite pour les voyelles moyennes.

Enfin, dans cette partie de notre analyse, nous nous limiterons à l'influence de l'aperture de la V<sub>2</sub> sur le premier formant de la V<sub>1</sub>. En effet, selon notre présomption de départ et la généralisation phonologique adoptée, l'harmonie vocalique, dès lors que considérée comme un processus affectant les V<sub>1</sub> sous l'influence d'une V<sub>2</sub> haute ou basse, ne peut tenir compte, d'une manière globale cohérente, d'un effet corollaire d'antériorisation ou de postériorisation qui, lui, serait différent pour chaque paire des V<sub>2</sub>.



**Figure 15.** Valeurs moyennes de F1 et F2 (en Hertz) et erreur-type des  $V_1 \in \{E, O, \text{Œ}, A\}$  suivies d'une  $V_2$  haute (point plein) ou une  $V_2$  basse (point vide) ; tous les six locuteurs confondus.

Afin de tester l'influence de la  $V_2$  sur le F1 de la  $V_1$ , ainsi qu'un impact éventuel du style de parole sur le degré d'harmonisation, nous avons effectué une ANOVA à deux facteurs avec interaction selon la procédure mentionnée ci-avant. Les coefficients des modèles d'analyse de la variance, pour l'ensemble des locuteurs, sont indiqués dans le **Tableau 12**, tandis que les résultats statistiques des modèles réalisés pour chaque locuteur séparément sont détaillés dans l'**Annexe 4**.

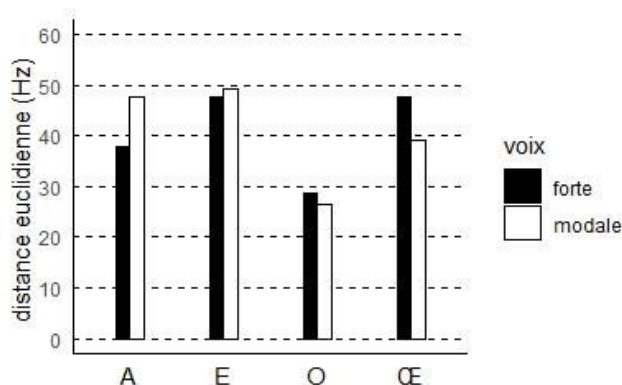
$V_1$	variation	SC	dl	CM	F	p
/E/	aperture_V2	128985	1	128985	21.466	<.001
	voix	2005779	1	2005779	333.800	<.001
	aperture_V2:voix	676	1	676	0.113	NS
	résiduelle	8358405	1391	6009		
/Œ/	aperture_V2	27617	1	27617	2.911	NS
	voix	1628009	1	1628009	171.580	<.001
	aperture_V2:voix	6	1	6	0.001	NS
	résiduelle	6338219	668	9488		
/O/	aperture_V2	93610	1	93610	15.20	<.001
	voix	3710067	1	3710067	602.53	<.001
	aperture_V2:voix	5355	1	5355	0.87	NS
	résiduelle	10646331	1729	6158		
/A/	aperture_V2	38908	1	38908	6.269	<.001
	voix	913771	1	913771	147.228	<.001
	aperture_V2:voix	28597	1	28597	4.608	NS
	résiduelle	3959763	638	6207		

**Tableau 12.** Résultats des tests d'ANOVA à deux facteurs avec interaction effectués pour les archiphonèmes /E, O, Œ, A/ ; formule utilisée :  $\text{aov}(F1\_V1 \sim \text{aperture\_V2} * \text{voix})$ .

Ainsi, l'analyse a mis en évidence un effet significatif de l'aperture de la  $V_2$  sur la hauteur du premier formant d'une  $V_1$ , dans le cas de /E/, /O/ et /A/. En revanche, la différence d'aperture trouvée pour la voyelle moyenne antérieure arrondie /Œ/ n'était pas significative. Le degré d'harmonie

vocalique ne s'est pas non plus révélé sensible au style de parole pour aucune des quatre voyelles examinées.

Pour terminer, et en vue de mieux visualiser le degré d'assimilation de chacune des quatre voyelles analysées, la **Figure 16** représente les distances euclidiennes moyennes séparant une  $V_1 \in \{E, O, \text{œ}, A\}$  lorsqu'elle était suivie d'une  $V_2$  basse ou haute, dans l'espace F1 x F2, en Hertz.



**Figure 16.** Distances moyennes en Hertz entre les  $V_1 \in \{E, O, \text{œ}, A\}$  suivies d'une  $V_2$  phonologique basse ou haute, calculées en tant que distances euclidiennes dans l'espace F1 x F2 selon la formule suivante :  $\sqrt{((F2_{V_1}-F2_{V_2})^2 + (F1_{V_1}-F1_{V_2})^2)}$ .

En résumé, l'analyse d'un corpus bien contrôlé enregistré en laboratoire a indiqué la présence d'un effet global modéré d'harmonie vocalique, qui a été mesurée en termes de rabaissement/rehaussement du premier et du second formant d'une  $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$  en syllabe pénultième sous l'influence d'une  $V_2$  en syllabe finale. Ainsi, les écarts moyens du premier et du second formant que nous avons observés sont relativement faibles et correspondent à un degré d'assimilation sensiblement inférieur à la distance acoustique séparant les voyelles moyennes phonologiques pleines mi-fermées et mi-ouvertes. En revanche, nous n'avons trouvé pas d'effet significatif d'harmonie vocalique sur /ø, œ/ (considérés ici sous le terme commun d'archiphonème /œ/).

De plus, aucun impact significatif de la *voix forte* (ou *projetée*) par rapport à la *voix modale* sur le degré d'harmonie vocalique n'a été trouvé.

Une variation importante inter-locuteur a également été mise en évidence. Lorsqu'analysées séparément, seules les productions sonores de deux locuteurs (GA et PA) présentent un effet significatif d'harmonie vocalique. Le locuteur GA produit ainsi une harmonie vocalique plus forte sur /e, ε/, et ne présente qu'un degré d'assimilation négligeable sur /o, ɔ/. En revanche, les productions sonores du locuteur PA présentent une harmonie vocalique plus importante sur /o, ɔ/, mais très peu d'harmonie vocalique sur /e, ε/. En contrepartie, les productions des quatre autres locuteurs du corpus (AR, BA, HU et TR) ne présentent généralement pas d'effet d'harmonie vocalique significatif, sauf certains contrastes isolés.

En outre, un effet d'assimilation de voyelle à voyelle a été mis en évidence pour /A/ (produit dans des pseudo-contextes d'harmonie vocalique). Qui plus est, la magnitude de cet effet est pratiquement égale, en termes de valeurs absolues, à celle de l'harmonie vocalique sur /E/ et légèrement supérieure à celle de l'harmonie vocalique sur /O/.

## 5. Expérience de perception

### 5.1. Coarticulation, métaphonie et harmonie vocalique dans l'optique de la phonétique auditive

#### 5.1.1. Coarticulation de voyelle à voyelle et la perception de la parole

Traditionnellement, les modèles de perception de la parole mettent l'accent sur le traitement perceptif limité aux propriétés acoustiques, non seulement associées à chaque contraste phonologique limité au domaine segmental, mais au détriment des propriétés plus fines. Cette perspective « atomiste » a eu pour conséquence d'avoir entériné pour longtemps une vision des indices de domaine temporel plus long comme étant dépourvus de signification perceptuelle : « *There is ample evidence that words are stored in memory in terms of sequences of segmental units, and that these segmental units are further represented in terms of the values of a set of binary features* » (Stevens, 2002). Or, il est désormais acquis que les auditeurs sont sensibles à des indices acoustiques des contrastes phonologiques distribués sur des intervalles temporels dépassant le domaine d'une syllabe, ce qui a remis en valeur ceux parmi les modèles d'accès lexical qui ont tenté d'en tenir compte en misant, par exemple, sur l'inclusion des effets coarticulatoires entre segments adjacents (Marslen-Wilson & Warren, 1994).

En effet, depuis la parution de l'étude spectrographique pionnière d'Öhman (1966) consacrée à la coarticulation dans les séquences VCV en anglais, suédois et russe, nombreuses ont été des tentatives d'inclure la composante perceptive dans la recherche expérimentale sur la nature de la coarticulation, la question primaire étant notamment de savoir si les auditeurs sont capables de percevoir les effets de la superposition partielle ou complète des indices acoustiques associés à différentes unités phonologiques (e.g. Lehiste & Shockey, 1972 ; Fowler, 1981). Une question différente étant de comprendre comment cette sensibilité aux différences d'un degré sub-phonémique affecte le traitement cognitif de la parole, un effort de recherche considérable a été déployé dès les années 1970 pour soumettre à des investigations diverses la façon dont s'opère le décodage d'un signal marqué par les effets de la coarticulation (Winitz *et al.*, 1971 ; Ostreicher & Sharf, 1976). Ces travaux, et bien d'autres, ont très tôt permis de faire apparaître que les effets de la coarticulation constituent une source d'information exploitée par l'auditeur dans le processus du traitement de la parole, en permettant notamment que soit anticipée l'apparition dans le signal d'un indice acoustique, ou encore que puisse être prédite, dans une certaine mesure, l'identité d'un segment supprimé du signal à partir d'un segment adjacent (cf. Nguyen, 2001 pour une revue exhaustive de la littérature).

Quoique les phénomènes de recouvrement temporel entre des unités adjacentes aient sans doute été le cas le plus largement interrogé, du moins dans un premier temps, les effets de coarticulation à travers un ou plusieurs segments ont presque aussitôt fait l'objet de l'attention des chercheurs, Martin & Bunnell (1982) étant parmi les premiers à s'être intéressés à la façon dont l'auditeur

s'emploie à tirer parti des effets acoustiques de la coarticulation de régressive de voyelle à voyelle. Ils ont montré en effet que la modification par *cross-splicing* (ou *transcollage*, selon le terme français employé par Nguyen, 2001) engendrait une rupture de correspondance phonétique, observable en termes du rallongement du temps de latence pour l'identification des stimuli *cross-splicés*. D'autres paradigmes expérimentaux ont été déployés depuis lors pour aboutir souvent à des résultats convergents. Ainsi, Hawkins & Slater (1994) ont trouvé que des effets coarticulatoires de voyelle à voyelle inclus dans des phrases synthétisées présentées dans du bruit augmentait l'intelligibilité du signal de 15 %. En partant de ce constat, Grosvald (2009), puis Grosvald & Corina (2012) ont trouvé que la perceptibilité des effets de coarticulation régressive de voyelle à voyelle, observés sur des stimuli enregistrés en laboratoire, diminue, certes, proportionnellement à la distance séparant la voyelle source (*i* et *a* dans l'étude en question), mais qu'ils demeurent perceptibles, pour environ un dixième des auditeurs, jusqu'à la distance de cinq segments séparant la  $V_1$  et la  $V_2$ . La conclusion globale de l'expérience semble résumer l'état des connaissances dans le domaine, en confirmant une fois de plus que « *anticipatory coarticulation might also be seen as aiding listeners as they process the incoming speech stream, because such coarticulation offers early information about the nature of upcoming segments* » (Grosvald & Corina, 2012)

### **5.1.2. Harmonie vocalique et traitement perceptif du signal**

L'harmonie vocalique phonologique, telle qu'elle a été décrite pour un nombre incalculable de systèmes vocaliques des langues parlées dans des aires géographiques les plus variées et appartenant à toutes les grandes familles linguistiques, et qui présente par ailleurs un éventail apparemment infini de configurations structurelles et une richesse typologique hors pair (cf. Krämer, 2003 ; Nevins, 2004 ; Archangeli & Pulleyblank, 2007 pour des revues de littérature sectorielles), tout comme des instances d'assimilation non phonologique dont elle n'est qu'une rémanence fossilisée selon l'affirmation bien connue d'Ohala (1994), a également fait l'objet de plusieurs études perceptives.

Si le cas d'harmonie vocalique phonologique est intéressant d'un point de vue de la perception et du traitement du signal, c'est parce que, comparé à l'assimilation de voyelle à voyelle non phonologique, il impose de reformuler la question de recherche principale qui n'est plus de savoir si les modifications de timbre qu'elle induit sont ou ne sont pas perçues par l'auditeur – ce qui, *a priori*, ne fait plus de doute dès lors que l'on a affaire à un phénomène opérant sur des contrastes phonologiques – mais plutôt quelle est son influence sur le traitement du signal ou, dit autrement, quel est l'avantage cognitif offert par les langues à harmonie vocalique par rapport aux langues non harmoniques. Suomi (1983) tente de répondre à cette question, en supposant que l'harmonie palatale en turc et en finnois est un moyen de faciliter la perception des contrastes de F2, car elle rend la spécification du trait d'aperture prévisible au-delà des positions prosodiquement saillantes (dans ce cas, les syllabes initiales). En s'appuyant sur la proposition de Suomi (1983), Kaun (1995) soutient que l'extension de la durée de réalisation d'un trait particulier (en l'étendant sur plusieurs segments)

équivalent, d'un point de vue de l'auditeur, à une plus grande probabilité d'identifier correctement cette valeur de trait. D'ailleurs, on retrouve, dans l'argumentation de Kaun (1995), un écho de l'argument de Steriade (1995), qui suggère que les processus de neutralisation positionnelle affectant les voyelles sous l'effet d'harmonie vocalique servent à fournir à l'auditeur des indices plus robustes pour le contraste des voyelles qui régissent cette dernière. Des études plus récentes, dont une série de quatre expériences de perception, ont été menées par Kimper (2017), avec l'utilisation de pseudo-mots à trois et quatre syllabes. Ils ont été construits de façon à représenter des structures harmoniques versus dysharmoniques variées, y compris des cas d'harmonie agissant à travers un segment vocalique dit *transparent*, donc violant le principe de stricte localité, et ont été présentés à des locuteurs anglophones. L'hypothèse générale étant que l'harmonie vocalique de hauteur est universellement avantageuse sur le plan perceptif, « *phonetically as well as articulatorily grounded* » (Kimper, 2017), ce sont des locuteurs parlant une langue sans harmonie vocalique qui ont été choisis comme participants à l'expérience. Les résultats ont été strictement favorables à l'hypothèse avancée, en apportant une forte preuve en faveur du gain de traitement qu'apportent les phénomènes d'harmonisation de nature diverse à l'auditeur, et ce, au-delà de mécanismes cognitifs propres à un système linguistique particulier. Néanmoins, au milieu de ces voix concordantes, l'étude de Tilsen (2007), basée sur le paradigme d'amorçage phonémique masqué, initialement conçue pour examiner dans quelle mesure la coarticulation de voyelle à voyelle est engendrée par des mécanismes cognitifs, par opposition aux contraintes physiques du mouvement des articulateurs, allait mettre en évidence des effets dissimilatoires significatifs, en suggérant l'existence d'un « *speechplanning mechanism restricting the extent of (...) vowel-to-vowel coarticulation* » (Tilsen, 2007).

### **5.1.3. Métaphonie et la reconnaissance de mots**

Pourtant, l'harmonie vocalique en français n'étant liée que par son nom – si ce n'est par le mécanisme d'assimilation prétendument universel dont serait issue toute harmonie des voyelles (Ohala, 1994) – à ses cousines lointaines turques, fenniques ou ougriennes, il serait délicat d'avancer à son propos des conclusions en lien direct avec les résultats susmentionnés. Or, s'il est un processus qui s'apparente le plus, au plan typologique, à l'harmonie vocalique telle qu'elle est présente en français, ce serait sans doute le phénomène de métaphonie décrite, pour ne citer que quelques exemples, pour certains parlars italo-romans, notamment dans plusieurs dialectes sardes centraux ou encore dans le dialecte italien de Piverone (Loporcaro, 2005).

Connue pour présenter un large éventail de phénomènes de métaphonie, l'aire dialectale couvrant une partie de l'Italie du Sud, dite aire de Lausberg, d'après le nom du premier dialectologue à lui avoir consacré son œuvre de recherche (cf. Lausberg, 1939, cité dans Greca, 2023), s'apparente, sous certains de ses aspects, à l'harmonie vocalique du français. Ainsi, les voyelles moyennes mi-ouvertes non finales de l'italien de l'aire de Lausberg s'assimilent en termes de hauteur, partiellement pour la série antérieure non arrondie, e.g. ['vecca] – ['veccu], et totalement pour la série postérieure

arrondie, e.g. ['mɔru] – ['murisi], sous l'influence des voyelles finales présentes dans les suffixes flexionnels. Mis à part la structure accentuelle différente, les dialectes italiens en question possédant, contrairement au français, un accent lexical, la métaphonie diffère de l'harmonie vocalique française par son lien inhérent aux alternances morphologiques, la voyelle-source entraînant à la fois une fermeture du timbre vocalique et un changement de catégorie morphologique.

Dans le cadre de sa thèse de doctorat récente consacrée à la métaphonie dans les dialectes italiens de l'aire de Lausberg, Greca (2023) s'est proposé d'examiner si les auditeurs exploitaient l'information acoustique contenue dans la voyelle du radical, et donc encourageant l'assimilation par métaphonie. Or, au lieu de puiser dans un cadre protocolaire utilisant une mesure du temps de latence, Greca (2003) a misé sur un paradigme inspiré de la tâche de *complétion* ou de *reconnaissance* de mot (*word completion task* ou *word recognition task*) ou avec une question à choix forcé. La question était, en effet, de savoir si, en absence du suffixe flexionnel, l'information acoustique sur le timbre vocalique résultant de la métaphonie, supposément stockée en mémoire du locuteur, suffirait pour reconnaître la catégorie morphosyntaxique du mot dont était extrait le radical tronqué. Si d'autres avant Greca (2023) avaient formulé des hypothèses similaires, en postulant par exemple, comme Clayards *et al.* (2021), que l'information sur le détail phonétique pouvait favoriser l'identification des formes d'un paradigme morphologique, ou encore, comme Manca *et al.*, (2019), qui ont utilisé des paradigmes comportementaux et neurophysiologiques, comme la réponse à la négativité de discordance (MMN) et la magnétoencéphalographie pour une série d'expériences s'appuyant sur l'hypothèse d'un gain de traitement rendu possible grâce à la métaphonie, la contribution de Greca (2023) reste intéressante justement par le fait d'avoir exploité le paradigme de complétion de mot – rare parmi les études sur la coarticulation (Nguyen, 2001) et d'un intérêt limité dans les études sur l'harmonie vocalique phonologique, la question de recherche se positionnant ailleurs (Kimper, 2007) – en tentant d'apporter un argument plus direct en faveur du rôle de la métaphonie dans le processus de reconnaissance de la parole.

Ainsi, 197 stimuli du type [bEi-] et [mOr-] issus soit de « bella »/« bellu » ou « morta »/« mortu », le premier mot de chaque paire possédant une voyelle moyenne 'neutre', que l'on classifierait de mi-ouverte, par contraste avec celle présente dans le second mot, produite mi-fermée ou fermée sous l'influence de la V<sub>2</sub>, ont été présentés à 170 participants, âgés de 19 à 80 ans, recrutés parmi les habitants de villages situés dans trois sous-régions, la variation diatopique étant l'une des variables étudiées, à qui on a demandé de choisir la catégorie morphologique (masculin/féminin ou singulier/pluriel) qu'ils percevaient d'après le stimulus tronqué. Le score des réponses correctes a oscillé entre 60 % et 77 % selon la région et la voyelle radicale /E, O/, avec des scores les plus bas, moins de 30 %, pour les formes en O-i et les plus élevés, environ 80 %, pour celles en O-u. Les résultats ont ainsi permis à Greca (2023) de conclure à la validation de l'hypothèse initiale, en affirmant que « *all three regions performed overall well in recognising stems* » (2023). Sans véritablement discuter la question du statut, phonologique ou phonétique, de la métaphonie, Greca

(2023) s'appuie sur cette expérience pour argumenter en faveur d'une relative stabilité du phénomène de métaphonie dans les dialectes en question, comme étant un indice d'un changement diachronique abouti, ou une forme de coarticulation fossilisée et spécifique à la langue.

#### **5.1.4. Deux expériences de perception sur l'harmonie vocalique en français**

Dans le sillage des travaux sur la perception des effets de coarticulation VCV, l'harmonie vocalique en français a fait, quant à elle, l'objet d'une étude perceptive menée par Nguyen *et al.* (2004). Ainsi, deux expériences complémentaires ont été conçues en vue d'apporter « *further empirical evidence for the role of long-domain acoustic properties in speech perception* ». Dans un premier temps, il s'agissait de déterminer si la modification de timbre, restreinte à un degré sub-phonémique, induite par l'harmonie vocalique était ou non perceptible à l'auditeur. Dans un second temps, la question était de savoir dans quelle mesure l'harmonie vocalique en français était mise à profit par l'auditeur pour accélérer l'accès lexical, en diminuant le temps d'évaluation et de classification de stimuli.

La première expérience avait pour objectif de déterminer si les effets d'assimilation  $V_2-V_1$  pouvaient être perçus par l'auditeur. Pour ce faire, une tâche dans le paradigme ABX, telle que définie par Majors (1998), a été mise en place, dans laquelle les participants écoutaient des séries de séquences (C)V<sub>1</sub> extraites des mots « épice/éplate » et « notice/nota » enregistrées en laboratoire dans l'expérience de production complémentaire (Fagyal *et al.*, 2003). Dans chaque série, l'un des stimuli était extrait d'un mot différent des deux autres. Or, contrairement à la conception standard de ce type de tâche, les deux stimuli 'identiques' provenaient, eux aussi, d'occurrences différentes du même mot : en d'autres termes, tous les stimuli étaient acoustiquement différents. L'hypothèse était alors que la variation sur la V<sub>1</sub> due à l'harmonie vocalique prévalait sur d'autres différences acoustiques, de telle sorte que les auditeurs seraient capables de distinguer le stimulus issu du mot lexical différent.

Les résultats obtenus ont permis de valider cette première hypothèse. En effet, l'effet d'harmonie vocalique s'est révélé perceptivement plus saillant que les variations acoustiques entre différentes occurrences du même mot, les scores de réponses correctes atteignant 55 % à 70 % en moyenne pour les quatre mots testés. Il est donc légitime d'avancer un pas plus loin pour demander si les variations acoustiques de la V<sub>1</sub> entraînés par l'harmonie vocalique, perceptibles donc pour les auditeurs, apportaient un gain de traitement et un accès lexical plus rapide ?

Pour y répondre, une seconde expérience, inspirée de l'étude de Martin & Bunnell (1982) sur la coarticulation régressive de voyelle à voyelle en anglais, a été mise en place. Son objectif était de déterminer si l'harmonie vocalique facilitait l'identification de la V<sub>2</sub>. Pour ce faire, les mêmes paires « épice/éplate » et « notice/nota » ont été présentées, enchâssées dans des phrases porteuses. Un groupe de stimuli était formé de phrases *cross-splitées*, dans lesquelles la première partie était tronquée après la syllabe initiale du mot-cible, [E-] ou [nO-]. La suite du stimulus, commençant par [-pis]/[-pat] ou [-tis]/[-ta], était *transcollée* d'une autre phrase porteuse. Ainsi créées, les versions *cross-splitées* et *non-*



*cross-splicées* étaient présentées aux participants qui devaient cliquer sur un bouton à chaque fois que la voyelle spécifiée à l'avance – [i] ou [a] – apparaissait dans le stimulus. L'hypothèse était que le temps de réaction serait plus important pour les stimuli croisés, si tant est qu'ils contenaient des indices acoustiques perçus comme inappropriés pour l'identification de la V2.

Il se trouve cependant que cette seconde expérience n'a confirmé l'hypothèse émise que partiellement, car le gain de traitement exprimé en termes de temps de latence n'a été trouvé que pour la paire « notice »/« nota », tandis que « *cross-splicing had no measurable impact on the response patterns for the épice-épate set of word pairs* » (Nguyen *et al.*, 2004). Or, selon les auteurs de l'étude, ce résultat trouverait une explication dans l'asymétrie mise en évidence dans l'expérience de production dont étaient issus les stimuli, étant donné que l'effet d'assimilation était globalement plus important pour les paires en O par rapport aux contextes en E, y compris dans les stimuli utilisés dans le test de perception. Un élément d'explication supplémentaire avancé par les chercheurs a consisté dans le fait de dire qu'une plus grande variation de F2 dans une bande de fréquence plus basse dans le cas des voyelles postérieures arrondies par rapport aux voyelles antérieures non-arrondies aurait été un indice acoustique plus saillant, entraînant une facilitation de traitement lexical mise en évidence par le test, pendant qu'une plage de variation plus ténue dans « épice »/« épate » se serait révélée insuffisante pour être exploitée cognitivement par l'auditeur (Nguyen *et al.* 2004).

Toujours est-il que ces résultats ont été interprétés comme satisfaisants au point d'affirmer que « *vowel harmony in French is perceptually relevant and may facilitate the identification of word-final vowels* » (Nguyen *et al.*, 2004) et, plus largement, comme suffisants pour être considérés comme un argument en faveur d'un modèle de traitement cognitif de la parole, admettant que l'information acoustique peut s'étendre au-delà du domaine d'au moins une syllabe. En contrepartie, il ne serait pas inadéquat de remarquer que, si l'information acoustique véhiculée par l'harmonie vocalique était effectivement mise à profit par l'auditeur, elle ne l'est que faiblement. Aussi le caractère graduel et variable de ce phénomène déjà acquis grâce aux études acoustiques précédentes se trouve-t-il confirmé dans le domaine de la perception.

## **5.2. Objectifs et hypothèses**

En s'appuyant d'une part, sur les résultats obtenus par Nguyen *et al.* (2004) et d'autre part, sur ceux de Greca (2023), nous avons réalisé une expérience de perception sur l'harmonie vocalique en français. Les deux expériences de Nguyen *et al.* (2004) s'étant positionnées dans le sillage des expériences perceptives sur la coarticulation de voyelle à voyelle, elles ont fourni des résultats cohérents avec les connaissances acquises auparavant dans ce domaine d'étude, sans pour autant aborder directement la question d'un rôle éventuel de l'harmonie vocalique dans la reconnaissance de mots. Les apports de Greca (2023), quant à eux, ont confirmé que la métaphonie peut jouer ce rôle, en accord avec l'inclusion de celle-ci parmi les règles morphophonologiques faisant partie intégrante du

système linguistique des dialectes examinés. La question de recherche que nous nous sommes proposé d'approfondir sera désormais de savoir si l'information acoustique sur l'harmonie vocalique en français est stockée en mémoire de l'auditeur au point de rendre possible la reconnaissance de mots ou bien, tout en étant un détail acoustique spécifique à la langue, elle ne peut être considérée que comme une instance de coarticulation.

Afin d'y apporter un élément de réponse, nous formulons l'hypothèse suivante qui guidera la conception du protocole expérimental et apportera d'autres réflexions par la suite :

**H6.** *L'harmonie vocalique en français est un phénomène de coarticulation fossilisé, mis à profit par l'auditeur pour faciliter l'accès au lexique.*

Ainsi avons-nous mis en place un protocole utilisant une tâche de *complétion de mot* ou de *reconnaissance de mot*, dont les résultats seront présentés ici en deux parties, suivant, dans un premier temps, la question subsidiaire, qui sera de savoir quels sont les corrélats acoustiques des stimuli qui ont contribué aux décisions des participants et qui ont eu un impact significatif sur la décision lexicale, sans prendre en considération la correction ou non de la réponse. Dans cette première partie, nous privilégierons une approche exploratoire plutôt qu'explicative, sans émettre d'hypothèses *a priori* sur l'importance relative des corrélats acoustiques susceptibles de conditionner la réponse du participant. Cela étant dit, puisque la structure même du test, ainsi que la définition phonologique traditionnelle de l'harmonie vocalique en français, présuppose un rôle de la hauteur du premier formant dans la perception par les participants des stimuli comme appartenant à un mot avec une  $V_2$  phonologique basse ou haute, nous retenons cette présupposition comme une hypothèse de travail, sans toutefois exclure un rôle potentiel d'autres paramètres, tels que le  $F_2$  de la  $V_1$  et sa durée.

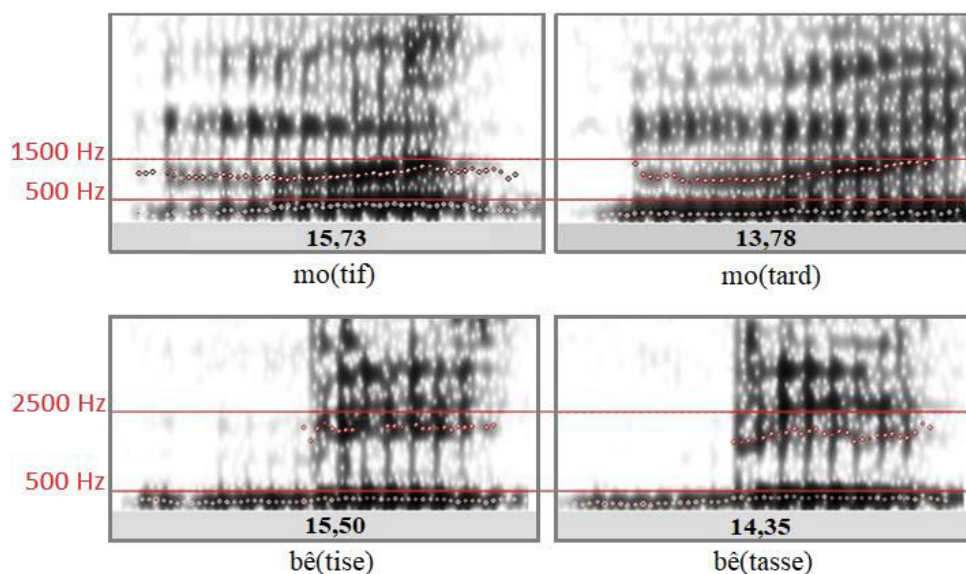
Dans un second temps, nous tenterons de mettre à l'examen l'hypothèse principale formulée ci-avant, en tâchant également de comprendre l'apport respectif des corrélats acoustiques différents à la façon dont l'harmonie vocalique est exploitée perceptivement par l'auditeur.

### **5.3. Méthode**

#### **5.3.1. Stimuli**

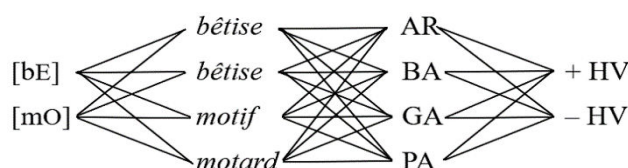
Les stimuli utilisés ont été sélectionnés parmi les paires de mots présentant un contexte d'harmonie vocalique issus du corpus enregistré en laboratoire. Les paires « bêtise »/« bêtasse » et « motif »/« motard » ont été retenues comme étant comparables en termes de structure syllabique  $C_1V_1C_2V_2C_3$ . Chacun des quatre mots contenait ainsi une  $V_1 \in /E, O/$  en syllabe ouverte, une  $C_1$  bilabiale /m, b/ dans l'attaque syllabique, une  $C_2$  intervocalique dentale /t/ et une  $V_2 \in /i, a/$  en syllabe fermée. Quatre occurrences de « bêtise », « bêtasse », « motif » et « motard » produits par quatre locuteurs (AR, BA, GA et PA) ont ensuite été sélectionnées de façon à présenter un spectre possiblement large du degré d'harmonie vocalique. Les locuteurs AR, BA, GA et PA ont été

sélectionnés comme ceux parmi les sept locuteurs du corpus, pour qui les différences des hauteurs des pics du premier et du second formant entre la voyelle-cible et la voyelle finale de mot étaient dus principalement à la variation de la V<sub>1</sub> et non de la V<sub>2</sub>, comme cela était le cas, dans une plus grande mesure, pour les trois locuteurs restants. En conséquence, pour chaque locuteur, les occurrences ont été séparées en deux groupes avec plus et moins d'assimilation de la V<sub>1</sub> en fonction de l'aperture de la V<sub>2</sub>. Les syllabes initiales ont été extraites suivant les frontières entre segments, telles que définies automatiquement à l'aide du logiciel WebMAUS. Une vérification visuelle des spectrogrammes des stimuli ainsi créés et un examen auditif de chaque extrait a permis de confirmer que la consonne intervocalique était bel et bien inaudible. Nous représentons ci-dessous les exemples de spectrogrammes correspondant à deux stimuli [mO] issus de « motard » et de « motif » ainsi que deux stimuli [bE] correspondant à « bêtise » et « bêtasse », obtenus à l'aide du logiciel Praat. Les tracés des deux premiers formants ont été surlignés pour une meilleure visibilité. L'on peut noter un espacement plus important entre F1 et F2 dans « motif », et une trajectoire montante d'un F2 anticipant le [i] final, contre une relative compacité entre F1 et F2 dans « motard », avec un pic de F2 atteint vers le milieu de la voyelle avant de décroître nettement devant le [a] final. Pour ce qui est de « bêtise »/« bêtasse », l'on notera que, du moins dans les extraits présentés ci-dessous, la différence entre les V<sub>1</sub> se fait principalement sur la hauteur du second formant, visiblement plus élevé dans « bê(tise) ».



**Figure 17.** Spectrogrammes des stimuli « mo(tard) », « mo(tif) », « bê(tise) » et « bê(tasse) » produits par le locuteur PA, générés à l'aide du logiciel Praat, durée totale (ms).

Au total, le corpus ainsi créé était composé de 64 stimuli, dont 32 syllabes [bE] et 32 syllabes [mO], répartis suivant la structure suivante :



Pour chacun des quatre locuteurs, il y avait donc quatre occurrences de « bêtise » et de « motif » et autant d'occurrences de « bêtasse » et de « motard », classées en deux groupes de deux occurrences pour chacun des quatre mots, présentant un degré plus élevé ou plus bas d'assimilation parmi les productions de chaque locuteur.

### **5.3.2. Tâche de complétion de mot**

L'expérience a été codée à partir de PsyToolkit (Stoet, 2010, 2019) afin d'être réalisée en ligne. Les participants ont commencé par remplir un questionnaire où ils ont renseigné leur âge et répondu à la question de savoir si oui ou non ils avaient des notions en phonétique, afin d'éviter un biais qui pourrait être induit par un déséquilibre de compétence entre un groupe de participants naïfs et un autre constitué de linguistes, prédisposés ne serait-ce qu'à deviner le sujet de l'étude qui ne devrait pas être *a priori* révélé aux participants.

Les stimuli étaient présentés aux auditeurs dans un ordre aléatoire, de façon à ce que chaque extrait soit joué deux fois. Les syllabes tronquées issues de l'un des quatre mots analysés, produits par l'un des quatre locuteurs étaient donc présentées au participant, sans imposer de contrainte sur les séquences de stimuli autorisées. Le temps de réponse, limité à 10 000 ms, a été mesuré sans être pour autant directement exploité en tant qu'une mesure peu fiable, dans le cas d'une expérience réalisée en ligne, sans contrôle expérimental du matériel utilisé.

Le protocole d'étude ainsi défini, deux limites méthodologiques méritent d'être signalées. Premièrement, l'harmonie vocalique étant un processus hautement variable et hétérogène, nos stimuli ne représentent qu'une sous-catégorie d'harmonie vocalique fermante, tandis que sa contrepartie ouvrante du type « rosace » ou « étal » est absente dans notre corpus. Nous n'étudierons pas non plus le modèle potentiellement distinct des formes avec la graphie « eau/ô », en les comparant à celles avec une graphie neutre et, enfin, nous n'examinerons pas l'impact – que nous savons par ailleurs non négligeable pour ce qui est des mesures acoustiques – du lieu d'articulation des consonnes intervocaliques. Tous ces aspects devront attendre de faire l'objet d'une expérience perceptive future plus étendue que l'expérience réalisée ici.

Enfin, une limite est induite par la nature lexicale des paires harmoniques en français, au sein desquelles un déséquilibre en termes de fréquence d'emploi fait la règle, et dont l'influence sur le stockage de l'information acoustique, mais également sur la décision lexicale de l'auditeur – ici non prise en compte parmi les facteurs étudiés – a sans doute son importance. La nature essentiellement morphophonologique du phénomène de métaphonie étudié par Greca (2023) avait, quant à elle, permis d'opérer sur des formes d'une fréquence d'usage comparable puisque provenant d'un même paradigme nominal. Autant il avait été possible ici de répliquer ce principe et de sélectionner les formes du type « noté »/« notait », autant nous avons estimé qu'une tâche de complétion de mot ainsi définie serait trop difficile, en sachant que la distinction *-é/-ait* en finale de mot est pour le moins incertaine en français parisien (Gendrot & Audibert, 2019).

## 5.4. Résultats

### 5.4.1. Participants

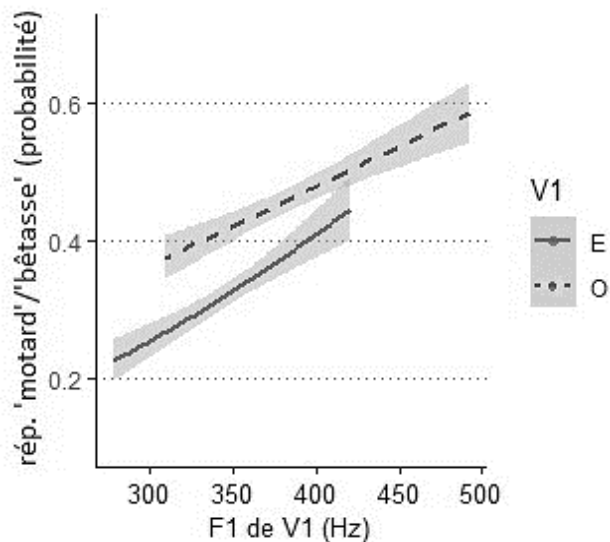
L'expérience a été réalisée auprès de 51 auditeurs francophones natifs (dont 45 sujets féminins), âgés de 20 à 26 ans avec une moyenne de 24 ans. Tous les locuteurs ayant indiqué être familiers avec les notions de base en phonétique, nous avons eu à notre disposition un groupe de répondants homogène qui, il convient de l'admettre, représentait un milieu spécialisé d'étudiants en linguistique, ce qui impose toutes les précautions quant à la généralisation possible des résultats de notre expérience. En échange, notons que l'âge des locuteurs dont les productions ont été exploitées (20 à 28 ans) et que celui des auditeurs ayant répondu au test coïncident, à deux années près, en formant un groupe sociolinguistique uniforme, témoignant de l'état de la langue à un moment bien précis de son évolution.

### 5.4.2. Effet des caractéristiques acoustiques des stimuli sur la décision des participants

Afin d'explorer l'effet des propriétés des stimuli sur les réponses obtenues, nous avons ajusté deux modèles de régression logistique, séparément pour les données correspondant aux stimuli [bE] et aux stimuli [mO], en utilisant la fonction `glm()` pour une distribution binomiale avec un lien logit (`family=binomial`). Les réponses obtenues ont été codées en variable dichotomique avec les niveaux suivants : 0 = 'motif'/'bêtasse', 1 = 'motard'/'bêtise'. Des modèles plus simples ont été comparés à des modèles plus complexes à l'aide du critère d'information bayésien (commande `BIC()` dans R) afin d'évaluer si l'ajout de chacun des prédicteurs améliorerait de manière significative le degré d'adéquation du modèle.

Ainsi, pour le groupe de stimuli [mO], le F1 de V<sub>1</sub>, le F2 de V<sub>1</sub> et la durée de V<sub>1</sub> ont été inclus comme prédicteurs dans le modèle. Le modèle obtenu explique 11 % de la variance des réponses au test, avec un R<sup>2</sup> de Nagelkerke de 0.14, et indique qu'un F1 plus élevé (B=0.004, p<.001), un F2 plus bas (B=-0.0006, p<.05) et une durée de V<sub>1</sub> plus importante (B=0.005, p<.05) augmentent la probabilité de répondre 'motard'.

Le modèle ajusté pour le groupe de stimuli [bE] a inclus l'unique prédicteur F1 de V<sub>1</sub>, la durée de V<sub>1</sub> ainsi que le F2 de V<sub>1</sub> ne s'étant pas montrés significatifs (B=-0.0005, p=0.27 et B=-0.005, p=.98). Le modèle ainsi formulé expliquait 10 % de la variance de la variable réponse et a indiqué un R<sup>2</sup> de Nagelkerke de 0.11, montrant qu'un F1 plus élevé (B=0.006, p<.001) augmentait la probabilité de répondre « bêtasse ». La **Figure 18** représente les prédictions des deux modèles pour la variable F1 des V<sub>1</sub> ∈ /O, E/. Notons que dans l'interprétation de ces résultats, il est nécessaire de tenir compte du faible pouvoir explicatif des deux modèles, ce qui est un indice indirect de la difficulté du test.

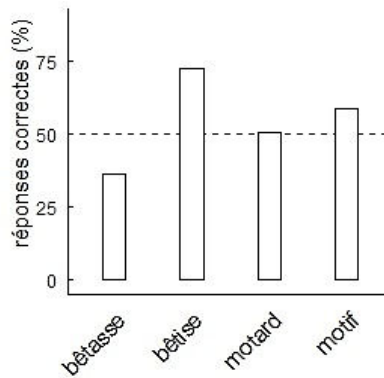


**Figure 18.** Probabilité de répondre 'motard'/'bêtasse' en fonction de F1 de V<sub>1</sub> pour les stimuli [mO] et [bE].

### 5.4.3. Étude de l'effet du degré d'harmonie vocalique sur la capacité à anticiper la fin du mot

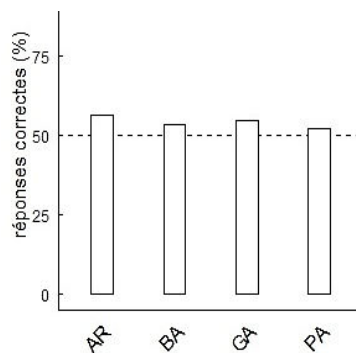
La **Figure 19** montre les pourcentages d'identification correcte pour chaque groupe de mots du corpus, tous locuteurs confondus. Sans dépasser un cadre d'interprétation approximative propre à ce type de calcul, on peut observer une disproportion de réponses correctes avec, d'un côté, les contextes d'harmonie vocalique fermante « bêtise » et « motif », pour lesquels le pourcentage de réponses correctes a dépassé le seuil de 50 % et, de l'autre côté, les contextes phonologiquement neutres, que l'on pourrait qualifier d'assimilation à distance, pour lesquels ce seuil n'a pas été atteint (« bêtasse ») ou a été atteint de justesse (« motard »). Notons cependant que les scores obtenus ne sont que légèrement plus bas que ceux trouvés par Greca (2023) dans une expérience similaire à la nôtre, avec une moyenne inférieure à 70 % d'identification correcte, tous contextes confondus.

Afin de valider les résultats obtenus dans les catégories « bêtise » (72,15 %) et « motif » (58,26 %), nous avons réalisé un test binomial qui a indiqué que la proportion de réponses correctes était significativement supérieure à 50 % ( $p < .001$  pour les deux comparaisons). En revanche, le pourcentage de réponses correctes pour la catégorie « motard » n'a pas atteint la significativité (50,4 %,  $p = .75$ ).



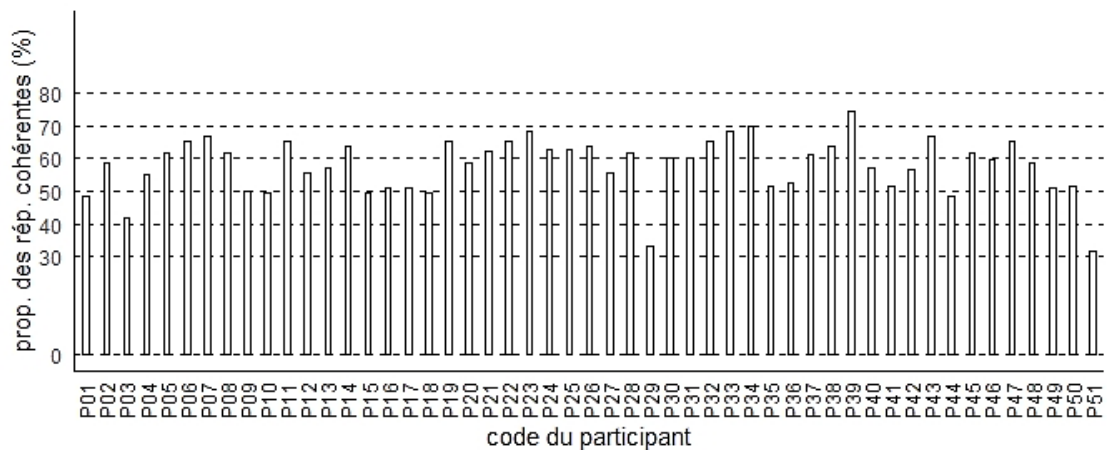
**Figure 19.** Proportion moyenne des réponses correctes pour chaque groupe de stimuli (%).

Dans la **Figure 20**, nous présentons les pourcentages de réponses correctes, tous les quatre mots confondus, séparément pour chaque locuteur dont les productions ont été utilisées comme stimuli. Un test binomial a été effectué pour vérifier si la proportion d'identification correcte était significativement supérieure au seuil du hasard (50 %). Ainsi, les tests ont indiqué que pour les locuteurs AR, BA et GA, la significativité a été atteinte, avec des valeurs de 56,4 % ( $p < .001$ ), 54,2 % ( $p < .01$ ) et 55,4 % ( $p < .001$ ) respectivement. En revanche, le pourcentage de réponses correctes pour les stimuli produits par le locuteur PA n'a pas été significativement supérieur au seuil du hasard, avec une valeur de 52,2 % ( $p = .07$ ). Malgré le caractère approximatif et rudimentaire de cette mesure, elle nous suggère cependant qu'il n'y a pas eu de biais non contrôlé en termes d'échantillonnage et que les productions des quatre locuteurs ont reçu une réaction perceptive comparable.



**Figure 20.** Proportion moyenne des réponses correctes pour chaque locuteur (%).

Afin d'estimer le degré de cohérence parmi les répondants lors du test, nous avons calculé les proportions moyennes de réponses identiques pour les deux écoutes des mêmes stimuli, représentées dans la **Figure 21**. Un score de 50 % signifie que le participant a donné la même réponse, que celle-ci soit correcte ou non, pour la moitié des paires de stimuli, soit 32 sur 64 paires. Pour les stimuli restants, ses réponses étaient différentes entre la première et la deuxième écoute. Ainsi, parmi les 51 répondants, 42 ont obtenu un taux de réponses cohérentes de 50 %, tandis que 23 ont dépassé le seuil de 60 %.



**Figure 21.** Proportion des réponses cohérentes pour les écoutes I et II du même stimuli ; par répondant (%).

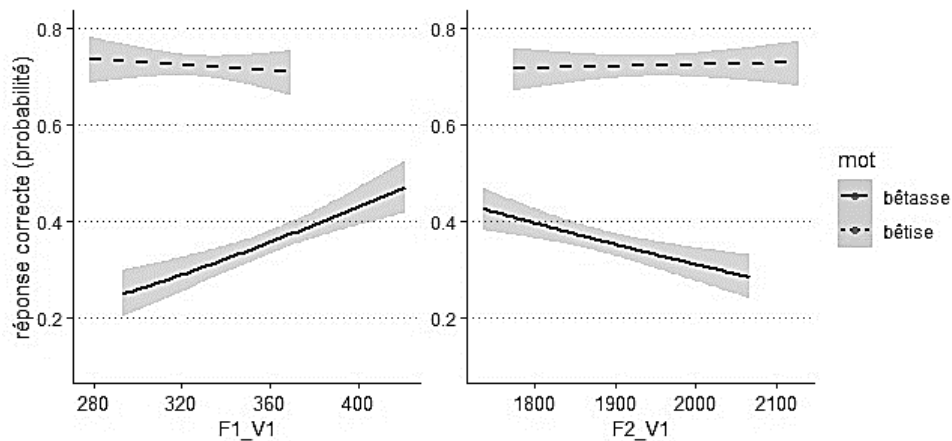
En vue d'examiner l'effet du degré d'harmonie vocalique sur la capacité des répondants à anticiper la fin du mot correcte d'après le stimulus perçu, nous avons ajusté un modèle de régression logistique selon le même protocole du choix de termes. Ainsi, les prédicteurs durée de V<sub>1</sub> et locuteur ont été exclus du modèle qui a comporté les termes F1 de V<sub>1</sub>, F2 de V<sub>1</sub> et l'aperture de V<sub>2</sub> (deux niveaux : 'basse'/'haute'). L'effet d'assimilation vocalique étant inversé, en termes de hauteur de F1 et de F2, pour les V<sub>1</sub> suivies de V<sub>2</sub> basses et hautes, nous avons également inclus dans le modèle les interactions entre ces prédicteurs. Les résultats du modèle ainsi formulé sont présentés dans le **Tableau 13**.

Prédicteur	B	St.Err.	z ratio	p
F1 de V <sub>1</sub>	0.003	0.001	3.826	<.001
aperture de V <sub>2</sub> (haute)	1.430	0.754	1.897	=.057
F2 de V <sub>1</sub>	-0.001	0.001	-6.549	<.001
F1 de V <sub>1</sub> :aperture de V <sub>2</sub> (haute)	-0.008	0.001	-5.283	<.001
F2 de V <sub>1</sub> :aperture de V <sub>2</sub> (haute)	0.001	0.001	7.668	<.001
$\chi^2 = 511.74$ p<0.0001	R <sup>2</sup> de Nagelkerke = 0.17		n = 6486	

**Tableau 13.** Résultats du modèle de régression logistique pour la variable 'statut réponse' (0 = incorrect, 1 = correct), formule : `glm(formula = status ~ F1_V1*aperture_V2 + F2_V1*aperture_V2, family = binomial)`.

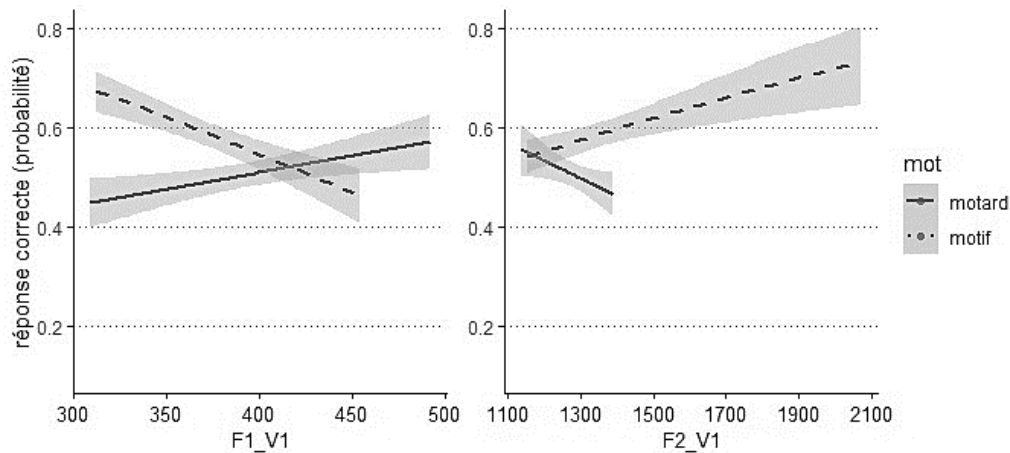
Dans la **Figure 22**, nous illustrons les prédictions du modèle de la probabilité d'identification correcte du mot correspondant au stimulus perçu en fonction de F1 et de F2 de V<sub>1</sub>, séparément pour « bêtise » et « bêtasse ». Ainsi, un F1 plus haut et un F2 plus bas augmentent le taux d'identification dans le cas des stimuli correspondant à « bêtasse », tandis que cette tendance s'inverse pour « bêtise », ce qui équivaut à une capacité de l'auditeur plus importante de reconnaître le mot lorsqu'il y a plus de coarticulation. Notons néanmoins une asymétrie entre les deux mots, l'effet de la hauteur de formants se montrant plus prononcé pour « bêtasse ».





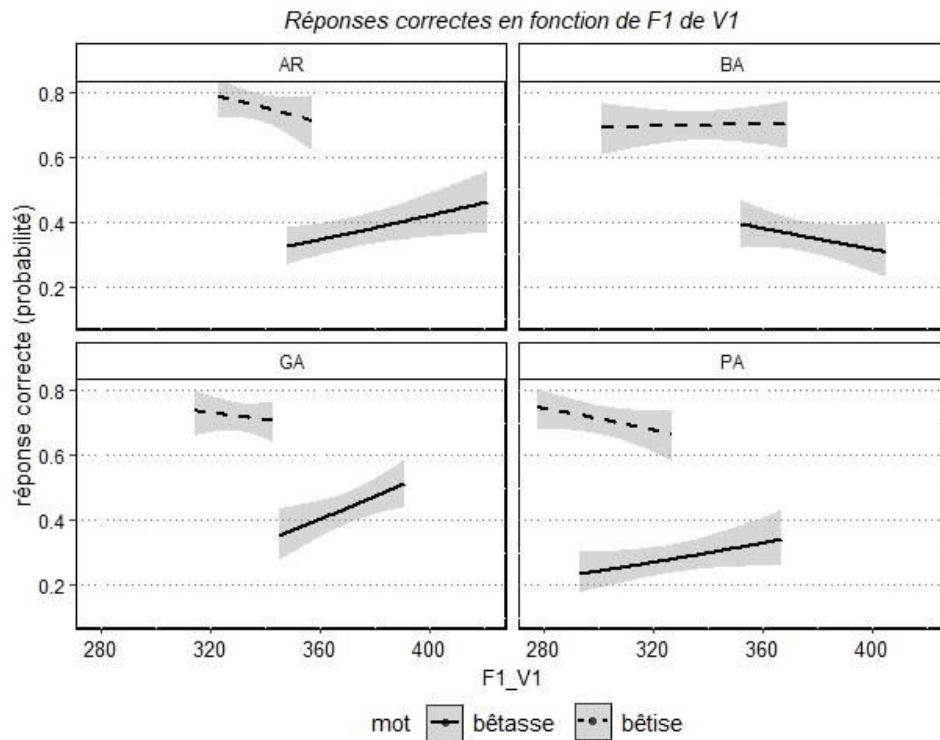
**Figure 22.** Probabilité d'identification correcte de la fin du mot pour les stimuli [bE] en fonction de F1 et de F2 de  $V_1$ .

La **Figure 23** représente l'effet des mêmes prédicteurs pour les stimuli [mO]. Observons l'interaction entre « motif » et « motard », le taux d'identification s'élevant au-dessus du seuil de 50 % pour un groupe d'occurrences avec une  $V_1$  plus assimilé à la  $V_2$ .



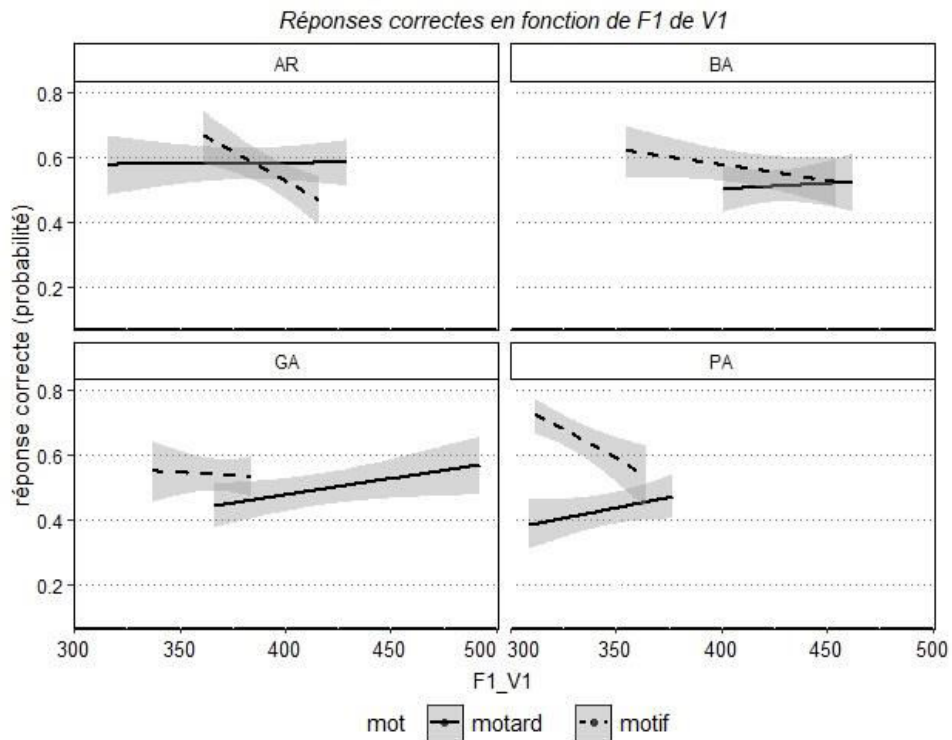
**Figure 23.** Probabilité d'identification correcte de la fin du mot pour les stimuli [mO] en fonction de F1 et de F2 de  $V_1$ .

Afin de visualiser d'éventuelles différences entre les stimuli produits par différents locuteurs, nous illustrons, dans la **Figure 24**, la corrélation entre la hauteur du premier formant et la probabilité de sélectionner la réponse correcte d'après l'écoute du stimuli [bE], séparément pour les quatre locuteurs du corpus. Ainsi pouvons-nous remarquer que les stimuli issus du mot « bêtise » présentant un contexte d'harmonie vocalique fermante se révèlent là encore plus faciles à reconnaître par rapport au contexte neutre, sans harmonie, que représente le mot « bêtasse ». Qui plus est, cette tendance est maintenue pour les stimuli provenant des productions sonores des quatre locuteurs.



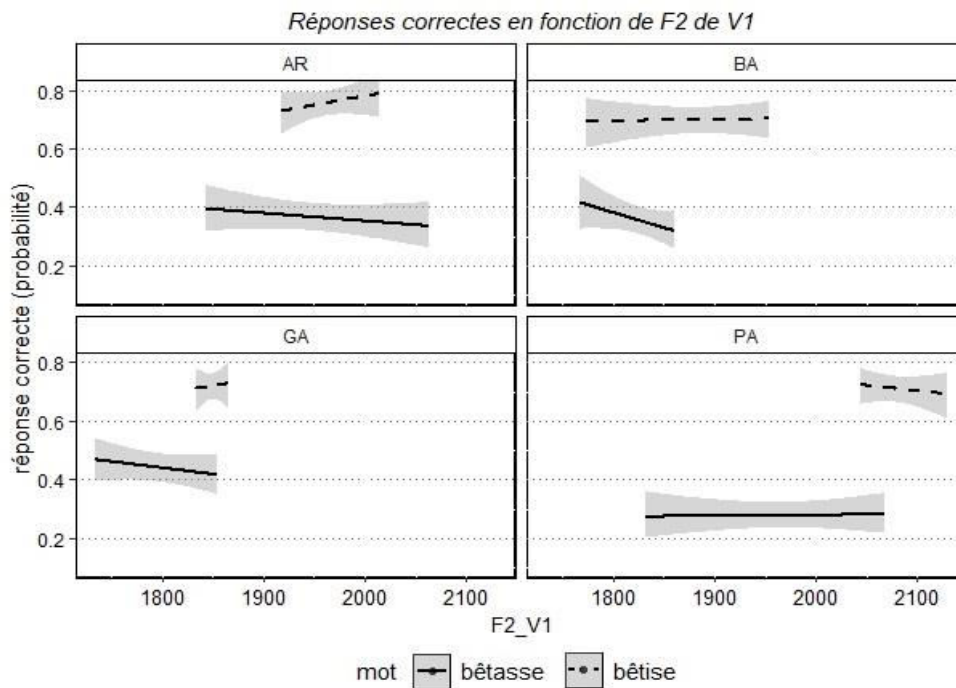
**Figure 24.** Probabilité d'identification correcte de la fin du mot pour les stimuli [bE] en fonction de F1 et de V<sub>1</sub>, séparément pour les stimuli produits par les locuteurs AR, BA, GA et PA.

La **Figure 25** représente la même corrélation pour les stimuli [mO], séparément pour les quatre locuteurs du corpus. Notons la variabilité acoustique des V<sub>1</sub> présentes dans les stimuli [mO] en fonction du locuteur. Les stimuli issus du contexte d'harmonie vocale « motif » sont généralement plus faciles à reconnaître, la probabilité d'identification correcte étant corrélée à la hauteur du premier formant. Ainsi, plus la V<sub>1</sub> est fermée (c'est-à-dire plus bas est le premier formant), plus les stimuli sont correctement identifiés comme provenant du mot ayant une V<sub>2</sub> fermée. Rappelons cependant que cette probabilité reste faible et ne dépasse le seuil de 50 % d'identification correcte que dans les cas des stimuli présentant un degré relativement plus important d'assimilation.



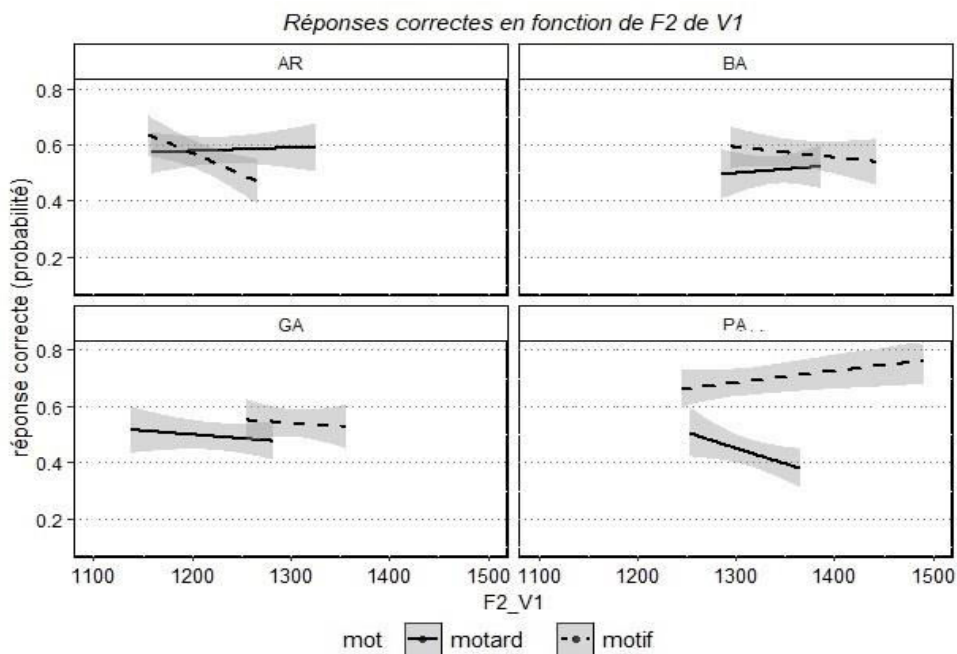
**Figure 25.** Probabilité d'identification correcte de la fin du mot pour les stimuli [bE] en fonction de F1 et de V<sub>1</sub>, séparément pour les stimuli produits par les locuteurs AR, BA, GA et PA.

La **Figure 26** représente les corrélations entre la hauteur du second formant de la V<sub>1</sub> présente dans les stimuli [bE] et la probabilité de choisir la réponse correcte (« bêtise » ou « bêtasse »), séparément pour les quatre locuteurs du corpus. Notons qu'en dépit des différences en termes de plages de valeurs de F2 présentes dans les stimuli, les tendances générales sont néanmoins maintenues. En d'autres termes, les contextes d'harmonie vocalique (« bêtise ») sont identifiés plus facilement que les contextes neutres (« bêtasse »). Cela amène à reconnaître que la hauteur de F2 joue sans doute un rôle dans la reconnaissance des mot par l'auditeur. En conséquence, l'harmonie vocalique, telle que définie traditionnellement, à savoir comme une assimilation *d'aperture* et non pas *d'antériorité*, n'est qu'un des facteurs qui entrent en jeu pour faciliter l'accès au lexique.



**Figure 26.** Probabilité d'identification correcte de la fin du mot pour les stimuli [bE] en fonction de F2 et de V<sub>1</sub>, séparément pour les stimuli produits par les locuteurs AR, BA, GA et PA.

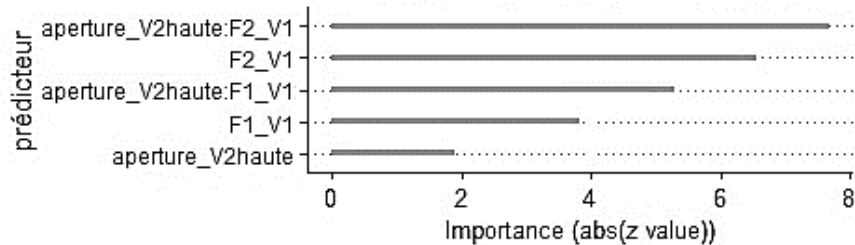
La **Figure 27** représente les corrélations entre la hauteur du second formant de la V<sub>1</sub> présente dans les stimuli [mO] et la probabilité de choisir la réponse correcte (« motif » ou « motard »), séparément pour locuteurs AR, BA, GA et PA.



**Figure 27.** Probabilité d'identification correcte de la fin du mot pour les stimuli [bE] en fonction de F2 et de V<sub>1</sub>, séparément pour les stimuli produits par les locuteurs AR, BA, GA et PA.

Pour tenter d'apporter un élément de réponse à la question de savoir quel est le rôle relatif du premier et du second formant, nous avons comparé l'importance relative de ces deux paramètres dans

la facilitation d'identifier correctement la fin du mot, la **Figure 28**, obtenue à l'aide de la commande `vip()` (Greenwell & Beehmke, 2020) représente la contribution relative des prédicteurs au modèle. Notons que tant le terme d'interaction F2 de  $V_1$ \*aperture de  $V_2$  que l'effet fixe de F2 de  $V_1$  indiquent une contribution plus importante au modèle par rapport à F1.



**Figure 28.** Contribution des prédicteurs à la significativité du modèle.

En résumé, un test de perception dans le paradigme de complétion de fin de mot a été mené afin de déterminer si l'information acoustique véhiculée par l'harmonie vocalique était mise à profit par l'auditeur pour faciliter l'accès au lexique. Les résultats indiquent que les auditeurs sont d'autant plus capables d'identifier correctement la fin du mot que la  $V_1$  présente dans le stimulus contient les caractéristiques acoustiques induites par l'assimilation par harmonie vocalique. Ainsi, les stimuli issus de contextes d'harmonie vocalique (« bêtise » et « motif ») sont reconnus plus facilement (avec des taux d'identification correcte dépassant le seuil de 50 %) que ceux provenant de contextes neutres (« bêtasse » et « motard »). Cela permet de confirmer le rôle joué par l'harmonie vocalique fermante dans la reconnaissance des mots. Notons que l'harmonie vocalique ouvrante (du type « bétail ») n'a pas été examinée dans cette étude pilote.

De plus, les résultats laissent transparaître la difficulté de la tâche, comme en témoignent les taux de réponses cohérentes qui dépassent faiblement le seuil de 50 %. Cependant, une analyse plus approfondie locuteur par locuteur montre que les tendances globales sont maintenues au sein des groupes de stimuli produits par chacun des quatre locuteurs : ainsi un F1 plus bas et un F2 plus haut augmentent la probabilité d'identifier correctement les contextes d'harmonie vocalique fermante (« bêtise », « motif »). De plus, les résultats suggèrent que le F2 pourrait avoir un rôle plus robuste dans la reconnaissance du mot par rapport à F1, ce qui invite à repenser la question de la pertinence de la définition traditionnelle d'harmonie vocalique en français en tant qu'une assimilation d'aperture plutôt qu'une assimilation de l'antériorité.

## 6. Discussion

### 6.1. Conclusions générales

#### *L'harmonie vocalique au travers des styles de parole*

Nous avons examiné le degré d'harmonie vocalique, appréhendée en termes de rehaussement ou de rabaissement du premier formant de la voyelle pénultième ( $V_1$ ) en fonction de l'aperture phonologique (ouverte, mi-ouverte ou fermée, mi-fermée) de la voyelle finale de mot ( $V_2$ ), dans trois grands corpus – NCCFr, ESTER et BREF – correspondant à trois styles de parole différents, à savoir la parole spontanée, journalistique partiellement scriptée et lue, respectivement.

Nos résultats montrent que les voyelles moyennes antérieures non-arrondies, /e, ε/, subissent une harmonie vocalique plus importante en parole lue par rapport à la parole spontanée. En outre, la même régularité est constatée dans le cas de la voyelle moyenne mi-ouverte postérieure arrondie /ɔ/. Ainsi le degré d'harmonie vocalique en parole journalistique est-il généralement proche de celui observé en parole spontanée (avec un  $\Delta F1$  inférieur à 5 Hz).

En revanche, le degré d'harmonie vocalique agissant sur la voyelle moyenne mi-fermée postérieure /o/ met en évidence une tendance inverse à celle des autres voyelles. Par conséquent, le degré d'harmonie vocalique sur /o/ est plus important en parole journalistique qu'en parole lue et bien plus encore en parole conversationnelle. Partant de ce constat, autant le comportement des /e, ε, ɔ/ confirme l'hypothèse initiale **HI**, autant la corrélation entre le style de parole et le degré d'harmonie vocalique sur /o/ va à l'encontre de celle-ci. Cela nous amène à conclure que l'influence du style de parole sur le degré d'harmonie vocalique n'est pas homogène en fonction de la voyelle analysée, laissant supposer une instabilité synchronique importante au sein du système des voyelles moyennes. Une autre approche consisterait à postuler que le terme d'*harmonie vocalique* englobe en réalité des phénomènes d'assimilation à distance de nature différente selon la voyelle qui en subit l'effet et dont le rôle et la dynamique au sein du vocalisme français ne se laissent pas réduire à une seule case terminologique.

#### *L'évolution récente du degré d'harmonie vocalique*

Dans un second temps, nous avons étudié le degré d'harmonie vocalique dans un corpus bien contrôlé de parole lue, enregistré en laboratoire, consistant en 33 paires quasi minimales présentant un contexte possible d'harmonie vocalique. L'analyse de ce corpus de trente ans plus récent que le corpus de parole lue BREF (dont il s'approche le plus en termes de style d'élocution) et de près de vingt ans plus actuel que le corpus de construction similaire de Nguyen & Fagyal (2008) a révélé un léger effet d'harmonie vocalique mesuré en termes d'écart de F1 et de F2 d'une voyelle pénultième selon l'aperture phonologique (ouverte/mi-ouverte ou fermée/mi-fermée) et le degré d'antériorité

(antérieure, postérieure) de la voyelle finale. Nos résultats indiquent que le degré d'harmonie vocalique est plus faible, dans notre corpus, par rapport aux corpus antérieurs du type similaire. En outre, quoique significatif pour l'ensemble du corpus, l'effet d'harmonie vocalique s'est montré non significatif pour la plupart des locuteurs analysés individuellement. Cela nous conduit à valider notre hypothèse **H2** et confirmer qu'une forme d'assimilation de voyelle à voyelle est toujours présente en français, mais qu'elle présente un degré plus faible dans les données plus récentes par rapport aux données antérieures fournissant un style d'élocution comparable.

### ***L'harmonie vocalique et l'assimilation de voyelle à voyelle dans d'autres positions au sein de mot***

Nous nous sommes également efforcé de comparer le degré d'assimilation de voyelle à voyelle dans des contextes d'harmonie vocalique phonologique à celui présent dans des contextes qui ne font pas partie des définitions traditionnelles d'harmonie vocalique.

Ainsi, l'analyse du degré d'harmonie vocalique sur /E, O, œ, A/ dans le corpus enregistré en laboratoire a conduit à constater un effet significatif d'harmonie vocalique (mesurée en termes de rehaussement ou de rabaissement du premier formant) sur /E, O, A/ mais pas sur /œ/. En conséquence, notre analyse a confirmé que le faible degré d'harmonie vocalique sur /E, O/ ne le distinguait guère d'une forme de coarticulation transconsonantique affectant une voyelle-contrôle (à savoir le /A/) qui ne fait pas partie de la définition phonologique d'harmonie vocalique. Qui plus est, la voyelle moyenne antérieure (ou centrale) arrondie /œ/, qui était parfois définie comme subissant l'harmonie vocalique (Straka, 1950), est sujette à un effet d'assimilation moins important que celui mis en évidence pour /A/ en syllabe pénultième sous l'influence de la voyelle finale. Cela nous conduit à invalider notre hypothèse **H3** et à affirmer que les données empiriques analysées ici suggèrent que la voyelle moyenne antérieure (centrale) arrondie /œ/ ne participe pas à l'harmonie vocalique au même titre que les autres voyelles moyennes du français.

Outre cela, nous nous sommes proposé d'étudier le degré d'assimilation transconsonantique régressive affectant les voyelles moyennes de la série antérieure /e, ε/ et postérieure /o, o/, dans des contextes d'harmonie vocalique (en syllabe pénultième), en comparaison avec le même type d'assimilation affectant d'autres positions dans le mot (en syllabe initiale ou interne non pénultième). Ainsi, l'analyse du corpus de parole lue (BREF) a mis en évidence la présence de deux schémas distincts selon la direction d'harmonie vocalique. Ainsi, le degré d'*harmonie vocalique ouvrante* agissant sur les voyelles moyennes mi-fermées /e, o/ ne permet pas de conclure à un effet d'assimilation plus élevé par rapport à celui observé dans d'autres positions dans le mot. En revanche, l'*harmonie vocalique fermante* affectant les voyelles moyennes mi-ouvertes /ε, o/ montre un effet significativement plus important lorsque celles-ci se trouvent en syllabe pénultième par rapport au même type d'assimilation présent dans les autres positions au sein de mot.

Accessoirement, cet aspect nous amène à souligner l'intérêt de distinguer une harmonie vocalique *fermante* et une harmonie vocalique *ouvrante*, qui possèdent chacune une dynamique propre

et méritent d'être considérées séparément. Ainsi, l'analyse de l'ensemble de nos données confirme la pertinence de distinguer une harmonie vocalique affectant les voyelles moyennes mi-fermées, prononcées plus ouvertes sous l'influence d'une V<sub>2</sub> ouverte ou mi-ouverte et celle occasionnant une fermeture des voyelles moyennes mi-ouvertes, ce qui permet de valider notre hypothèse **H4**.

#### ***L'harmonie vocalique dans la voix projetée***

De surcroît, l'analyse du corpus enregistré en laboratoire a également permis d'explorer l'influence éventuelle de la voix projetée sur le degré d'harmonie vocalique. L'analyse a indiqué que le degré d'harmonie vocalique n'est pas sensible à ce style de parole particulier, et ce chez aucun des six locuteurs examinés, ce qui confirme notre hypothèse **H5**. Ce résultat permet de corroborer notre hypothèse selon laquelle le degré d'harmonie vocalique particulièrement haut dans la parole journalistique des années 1940-1950 ne saurait être imputable à l'influence du style de parole emphatique spécifique aux speakers radiophoniques d'époque (style « Gaumont-Pathé », Boula de Mareüil *et al.*, 2008, 2012a, 2012b), et résulte d'une évolution diachronique propre à ce phénomène au cours de la seconde moitié du XX<sup>e</sup> siècle (Cecelewski *et al.*, à paraître).

#### ***Le rôle de l'harmonie vocalique dans le traitement cognitif de la parole***

Enfin, un test de perception a été mené en vue d'explorer dans quelle mesure l'harmonie vocalique était mise à profit par l'auditeur pour faciliter l'accès au lexique. Les résultats suggèrent que l'information acoustique sur l'harmonie vocalique est stockée dans la mémoire de l'auditeur. En effet, la capacité des répondants de reconnaître la fin du mot tronqué s'est révélée proportionnelle au degré d'harmonie vocalique (appréhendé en termes d'écart de F1 et de F2), présente dans le stimulus. Par conséquent, nous pouvons valider notre hypothèse **H6**.

## **6.2. Évolution d'harmonie vocalique entre 1940 et 2022**

Malgré d'éminents travaux théoriques consacrés à l'explication du changement phonétique (Ohala, 1974, 1981, entre autres ; Barnes, 2002 ; Blevins, 2004, 2008 ; cf. Yu, 2020 pour une revue de la littérature), les études expérimentales en phonétique diachronique sont rares et, comme l'ont remarqué récemment Sóskuthy & Stuart-Smith, « *typically focus on a small number of variables* » (Sóskuthy & Stuart-Smith, 2020). Or, il est un acquis que « toutes les parties de la langue sont soumises au changement » (Saussure, 1985 ; cité par Vaissière, 1996), ce qui constitue sans doute la difficulté principale et un « piège » tendu aux expérimentateurs, dans la mesure où « *looking at sound changes in a vacuum can produce misleading results* » (Sóskuthy & Stuart-Smith, 2020).

L'objectif de cette section sera d'illustrer la démarche méthodologique que nous avons suivie avec comme objectif premier d'apporter un éclairage expérimental sur l'évolution de l'harmonie vocalique en français, au travers de différents arguments empiriques que nous avons réunis et dont nous tenterons de donner un tableau récapitulatif.



Dans Cecelewski (2022), nous avons montré, au travers d'une analyse acoustique *en temps réel* d'un corpus d'archives radiophoniques INA, que le degré d'harmonie vocalique sur /o, ə/ a connu une décroissance progressive entre 1940 à 1997. Une analyse plus complète de ce même corpus, étendue sur /e, ε/, a été présentée dans Cecelewski *et al.* (à paraître), afin de mettre en évidence l'évolution à la baisse d'harmonie vocalique sur toutes les voyelles analysées. De plus, le changement en question a été modélisé en fonction de l'évolution de la hauteur moyenne de f0, ce qui a permis d'écarter l'hypothèse selon laquelle le changement du degré d'harmonie vocalique était corrélé à la « diminution du pitch moyen et de l'accent initial de mot » (Boula de Mareüil *et al.*, 2008), l'évolution la plus marquante s'étant produite entre les années 1940-50 et les décennies plus récentes.

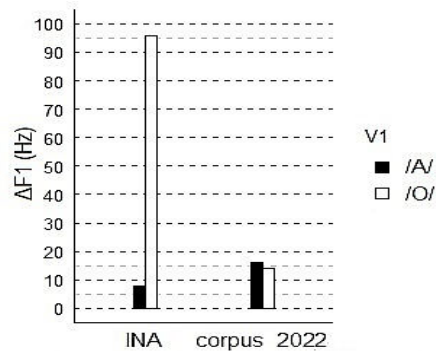
En effet, en 1908, la compagnie Pathé introduisit les films d'actualités, projetés avant le long métrage, ce sur fond d'un commentaire lu par un speaker professionnel s'adressant au public « avec une dynamique d'annonceurs de foire » (Boula de Mareüil *et al.*, 2017), le plus souvent sans amplification (Kermabon, 1994). Ainsi, le style « Gaumont-Pathé » s'était maintenu à la radio jusqu'à l'avènement du transistor. Or, le f0 n'était qu'un parmi d'autres corrélats acoustiques possibles du style emphatique qui caractérisait le style d'élocution de speakers radiophoniques des années 1940-50, entraînés à projeter la voix. C'est pourquoi nous avons conçu une expérience à base d'imitation (cf. par exemple Schlichting & Sullivan, 2013 ; Zetterholm, 2007), conçue dans l'esprit de *Historical Laboratory Phonology* (Ohala, 1993 ; Foulkes, 1997) :

If particular sound changes are posited to have a phonetic basis then one should be able to duplicate the conditions under which they occurred historically and find experimental subjects producing 'mini' sound changes that parallel them (Ohala, 1993).

Au lieu de faire appel à un professionnel de la parole, par exemple un comédien capable de reproduire une impression acoustique proche d'un style radiophonique ancien, nous avons opté pour un groupe de locuteurs non professionnels. En effet, comme aucune imitation n'est libre d'une empreinte individuelle du sujet, nous avons prédit une forte variabilité inter-locuteur, ce qui était censé multiplier la probabilité de détecter une corrélation entre le type de voix et le degré d'harmonie vocalique, ne serait-ce que chez une minorité des locuteurs, et que nous aurions pu manquer en travaillant avec un sujet unique. Les résultats détaillés ici ont permis de conclure qu'il n'y avait aucune incidence de la voix projetée sur le degré d'harmonie vocalique, et ce chez aucun des six locuteurs.

Qui plus est, le degré d'harmonie vocalique dans le corpus enregistré en laboratoire s'est révélé plus faible que dans les données antérieures, ce qui semblait aller dans le sens de notre hypothèse initiale sur la diminution du degré d'harmonie vocalique. Les données les plus récentes du corpus d'archives radiophoniques datant de 1999, nous disposons désormais d'un argument supplémentaire en faveur du changement actuellement en cours en français standard. Or, une comparaison des valeurs absolues extraites de données provenant de différentes sources matérielles, enregistrées à des époques différentes et dans des contextes d'élocutions différents, ne pourrait être une preuve robuste, en raison de multiples sources potentielles de variation non contrôlée. Cependant, dans le cas d'un paramètre

acoustique relatif, tel que le degré d'assimilation entre deux segments, une indication supplémentaire peut provenir de la comparaison de celui-ci avec un segment-contrôle, dans notre cas, le A. Dans la **Figure 29**, nous avons représenté le degré d'harmonie vocalique exprimé en termes d'écart de F1, en Hertz, d'une  $V_1 \in \{o, \text{ɔ}, a, \alpha\}$  (obtenu par la soustraction de F1 de la  $V_1$  suivie d'une  $V_2$  fermée/mi-fermée à celui de la  $V_1$  suivie d'une  $V_2$  ouverte/mi-ouverte), dans le corpus enregistré en laboratoire (*corpus 2022*) et dans le corpus d'archives radiophoniques INA (d'après les données présentées dans Cecelewski, 2022, qui ont inclus seule la série postérieure arrondie des voyelles moyennes).



**Figure 29.**  $\Delta F_1$ , en Hertz, d'une  $V_1 \in \{O, A\}$  suivie d'une  $V_2$  fermée/mi-fermée ou ouverte/mi-ouverte, dans le corpus enregistré en laboratoire (*corpus 2022*) et dans le corpus INA (d'après Cecelewski, 2022).

Ainsi, nous constatons que le degré d'harmonie vocalique sur /O/ a diminué de près de 80 Hz, tandis que le degré d'assimilation sur A est resté pratiquement inchangé, ce qui nous permet de confirmer qu'il s'agit bien de l'évolution de l'harmonie vocalique, et non pas d'une *dissimilation* plus générale affectant d'autres voyelles en syllabe pénultième.

Enfin, nous avons intégré dans notre étude l'esquisse d'une analyse *en temps apparent* (Labov, 1994 ; Sankoff & Blondeau, 2007), à savoir en comparant le degré d'harmonie vocalique dans les productions sonores de locuteurs appartenant à trois générations distinctes, enregistrés au même moment. Toutefois, étant donné la répartition inégale de sujets d'âge différent dans le corpus BREF, dont la vocation première n'était pas de constituer un portrait sonore de plusieurs générations de témoins, il ne s'agissait ici que d'une étude-pilote censée fournir un indice supplémentaire plutôt que le résultat robuste et définitif. Néanmoins, pour imparfait que pût être l'échantillon ainsi constitué, les résultats se sont montrés éclairants, dans la mesure où le degré d'harmonie vocalique le plus important a été réalisé par le groupe le plus âgé, nés entre 1923-45, et qu'une légère diminution a pu être observée dans les tranches d'âge suivantes (1946-66 et 1967-73). De surcroît, ce résultat gagne en importance explicative, lorsque comparé aux résultats d'une étude longitudinale récente consacrée à l'évolution du degré d'assimilation régressive de voyelle à voyelle, dans plusieurs langues, en fonction de l'âge du locuteur (D'Alessandro & Fougeron, 2021). Les résultats ont montré une tendance générale à *réduire* le degré d'assimilation – « *the amount of anticipatory VtoV coarticulation is found to decrease with age in our population* » (D'Alessandro & Fougeron, 2021) – ce qui a été discuté en termes de changements du contrôle moteur et de la décroissance de la vélocité des articulateurs au cours du vieillissement. Nos résultats ont pourtant indiqué une tendance opposée, ce

qui peut s'expliquer, d'une part, par l'absence de témoins âgés de plus de 50 ans dans le corpus BREF (enregistré en 1991). Or, c'est dans ce groupe d'âge que la tendance à la réduction du recouvrement articulatoire dû à la coproduction des segments de parole a été la plus forte. D'autre part, cela suggère que l'évolution du degré d'harmonie vocalique en français dans la seconde moitié du XX<sup>e</sup> siècle relève d'un changement phonétique qui résiste à la tendance physiologique à réduire le degré de coarticulation avec l'âge. Enfin, cela suggère que l'harmonie vocalique en français ne peut être réduite à une simple coarticulation mécanique, mais plutôt à un phénomène d'assimilation spécifique à la langue.

En somme, nous avons tenté d'apporter des preuves empiriques en faveur de la diminution du degré d'harmonie vocalique en français entre 1940 et 2022, à travers diverses expériences menées selon différentes approches méthodologiques. Notre objectif a été de mettre en exergue « l'importance d'examiner plusieurs sources de preuves même lorsque l'accent est mis sur une seule variable » (Sóskuthy & Stuart-Smith, 2020), résultant de notre préoccupation à construire un argumentaire expérimental solide, essentiel dans une étude diachronique.

### 6.3. Entre *harmonisation* phonétique et *harmonie* phonologique

Dès les premières études expérimentales, une réflexion émerge quant à la nature de l'harmonie vocalique en français, dont le caractère graduel, hautement variable et difficilement prédictible, remet en question son statut phonologique traditionnel (Fagyal *et al.*, 2003 ; Nguyen & Fagyal, 2008). Afin de répondre à la question de savoir si l'harmonie vocalique en français devrait ou non être considérée comme un processus phonologique, nous nous sommes proposé de revisiter l'histoire du terme *harmonie vocalique* tel qu'il apparaît, somme toute assez tardivement, dans la tradition terminologique française. Une réflexion terminographique sur le calibrage de ce *terme-intrus* à l'aune du contenu conceptuel auquel il réfère a été amorcée par Fagyal *et al.* (2003), qui ont constaté que les résultats de leur étude empirique sur le français peuvent bien correspondre à une acception large d'*harmonie vocalique* :

Therefore, according to the broadest definition of 'harmony', i.e. 'the way the articulation of one phonological unit is influenced (is 'in harmony' with) another unit in the same word or phrase' (Crystal 1997), the observed assimilation phenomena can be called vowel harmony (Fagyal *et al.*, 2003).

Or, des définitions moins restrictives mettent davantage en avant la dimension phonologique, ou systémique – à savoir grammaticale – de ce phénomène, ce qui rend plus délicate l'adoption de ce terme dans le cas du français :

(...) **grammatical process** known in certain languages under the name of vowel harmony (Jakobson *et al.*, 1952 ; nous soulignons) ;

In vowel harmony, the vowels in a domain, such as the word, systematically agree, or 'harmonize', in some **phonological property** (Walker, 2012 ; nous soulignons).

Quant à Fagyal *et al.* (2003), ils passent en revue différents termes possibles, dont celui de *umlaut*, et concluent en conseillant l'adoption de celui de *métaphonie*, à la fois plus familier à la tradition de la linguistique des langues romanes et, contrairement à *harmonie vocalique*, référant exclusivement à l'assimilation régressive :

(...) following the Romanist tradition that reserved the term metaphony for such anticipatory assimilatory changes at the word-level, we suggest that vowel assimilation observed in our data be called metaphony rather than vowel harmony. (Fagyal *et al.*, 2003)

Toutefois, il n'est pas tout à fait exact de dire que l'harmonie vocalique soit un terme inconnu à la linguistique romane. Pour preuve, Mateus (1987) utilise le terme *d'harmonie vocalique* pour décrire des phénomènes d'assimilation et de réduction vocalique en portugais, tandis que Hualde (1989) alterne l'emploi de ce même terme avec celui de *métaphonie* pour décrire un large éventail de phénomènes d'assimilation affectant plusieurs dialectes castillans :

A number of dialects spoken in northwestern Spain (...) present vowel-harmony processes triggered by the word final vowel or by stressed vowel. Harmonization that has its source in the word-final vowel is known as metaphony in the Romanist tradition. (Hualde, 1989)

Loporcaro prône même l'abandon du terme de *métaphonie* dont l'emploi serait, selon lui, la cause de l'exclusion artificielle de phénomènes d'assimilation présents dans les langues romanes par la tradition typologique de l'harmonie vocalique :

(...) vowel harmony in Eurasia is not a structural peculiarity of, say, Turkic or Uralic languages for it also occurs within Romance, as the product of regular sound change (Loporcaro 2005)

En outre, il n'est pas tout à fait exact de connoter uniformément le terme de *métaphonie* aux langues romanes et celui de *umlaut*, au domaine germanique, pour autant qu'il n'était guère rare, à une époque, d'utiliser le terme de « *Umlaut in der Romania* » (Schürr, 1936) pour référer à des phénomènes d'assimilation (ou dissimilation) des voyelles dans les langues romanes (cf. aussi Lausberg, 1939 ; Rensch, 1964). En effet, la ligne de démarcation entre les trois termes n'est pas toujours très nette. Enfin, il s'agit plutôt d'un problème terminologique à paliers, mobilisant des notions-type et méta-notions d'un sémantisme mouvant :

Similarly to Umlaut, in fact, metaphony is a kind of regressive vowel harmony, originating from trans-consonantal vowel-to-vowel coarticulation (Greca, 2023)

En réalité, les trois termes désignent des phénomènes d'une nature phonologique (ou « grammaticale ») et catégorielle, opérant sur un niveau phonémique et dont on peut rendre compte au moyen des règles de dérivation, contraintes, ou par le biais de toute autre approche phonologique symbolique. Ainsi, privilégier l'un plutôt que l'un autre ne saurait être une façon adaptée à décrire les « *infra-phonemic variations* » (Fagyal *et al.*, 2003) qui caractérisent l'harmonie vocalique en français.

Une question s'impose, celle de savoir dans quelles circonstances ce terme a commencé d'être utilisé dans la tradition terminologique française. Le phénomène d'assimilation de voyelle à voyelle affectant les voyelles moyennes en syllabe pénultième, tel que décrit par Rousselot & Laclotte (1903), n'appelle pas de notion spécifique jusqu'en 1916. C'est alors que Grammont introduit le terme

d'harmonisation. Or, il est important de garder à l'esprit qu'à l'époque, la linguistique historique et comparée française connaît déjà, ce au moins depuis le troisième quart du XIX<sup>e</sup> siècle, le terme d'*harmonie vocalique*, ou plus rarement, *harmonie des voyelles* (Adam, 1874). Le terme de *métaphonie*, d'origine française, fut, quant à lui, introduit vingt ans plus tard par Henry, dans sa *Grammaire comparée de l'anglais et de l'allemand*, qui reprit les éléments morphologiques de *umlaut* pour forger un calque terminologique par le truchement d'une traduction savante en grec. Ce même chercheur sera également à l'origine de la première traduction de son terme en anglais : *metaphony* verra le jour un an plus tard dans une version abrégée de sa *Grammaire* (1894).

Nous sommes ainsi en droit de supposer que Grammont connaissait ces deux termes. En revanche, il ne pouvait connaître celui de *coarticulation* (*Koartkulation*), introduit dans les années 1920 et popularisé grâce à la monographie de Menzerath & Lacerda consacrée à ce phénomène (1933), le terme étant encore à l'époque limité à désigner des instances de *coarticulation* strictement locale, sans mentionner de phénomènes distants. Toujours est-il que Grammont n'a pas eu recours au terme de *métaphonie*, pourtant désignant des phénomènes d'assimilation régressive, souvent d'aperture, affectant typiquement une voyelle non-finale sous l'influence de la voyelle finale du mot. Il n'a pas non plus utilisé celui d'*harmonie*, préférant forger un terme nouveau, spécifique aux phénomènes qu'il a observés en français. Qui plus est, nombre de linguistes emprunteront à Grammont le terme d'*harmonisation*. Il sera employé par Martinet (1945), Straka (1950), Fouché (1956) ou encore Malmberg (1969). La question du statut phonologique de l'*harmonisation* ne sera alors pas soulevée. Ce n'est en réalité qu'en 1973 que le terme d'*harmonisation* sera substitué par *harmonie* par Dell qui en donnera une interprétation sous forme de règle dérivationnelle, dans le sillage de la phonologie générativiste (Dell, 1973). D'autres suivront, tant et si bien que l'*harmonie* et non plus *harmonisation* vocalique s'inscrira peu à peu dans la tradition phonologique française (cf. Casagrande, 1984 ; Tranel, 1987 ; Durand & Lyche, 2004), jusqu'à ce que les premières études empiriques ne remettent ce statut en doute :

Our data also raise questions regarding symbolic phonological approaches to VH in French. The effects we observed are subtle and gradual, and they do not correspond to the traditional characterization of VH as a plus–minus switch in the height of the harmonizing vowel. (Nguyen & Fagyal, 2008)

Notre argument sera désormais que l'harmonie vocalique en français, non seulement, n'est pas un processus phonologique, mais également, qu'elle n'a été considérée comme tel que le temps d'une parenthèse terminologique relativement brève, tel un artefact symbolique issu d'une approche générativiste, et qui n'a pas tardé à être remis en cause dès que les premières données expérimentales ont été mises à disposition. S'il est un terme qui puisse convenir mieux que celui d'*harmonie* pour décrire le phénomène en question, nous proposerions que ce soit le premier terme en date qui a été employé depuis Grammont (1916), à savoir celui d'*harmonisation*, dont le mérite est qu'il n'implique pas la catégorialité phonologique. Ainsi, les résultats de Nguyen & Fagyal, de même que les données

plus récentes que nous avons présentées, nous conduisent à définir l'harmonie vocalique en français comme une *règle phonétique* plutôt que *phonologique* :

The output phonological representations can be acted on by phonetic rules, which interpret these features in time and space. Phonetic rules can thus, for example, assign a segment only a slight amount of some property, or assign an amount that changes over time during the segment (Keating, 1990)

Mais qu'en est-il de la nature exacte de ce phénomène ? Nos résultats suggèrent qu'il ne pourrait être réduit à une instance de coarticulation mécanique. D'un côté, l'évolution d'harmonie vocalique, depuis 1940 et jusqu'aux données les plus récentes, témoigne d'une dynamique propre qui serait difficile à défendre s'il s'agissait de la coarticulation, auquel cas il serait nécessaire d'argumenter en faveur d'une tendance à l'hyper-articulation, en cours dans le français standard. Or, le fait que le degré d'harmonie vocalique sur /e, ε, o/ soit supérieur en parole lue et journalistique par rapport à la parole spontanée ne correspond pas au comportement auquel on s'attendrait d'une forme de coarticulation en temps réel.

De plus, nos résultats indiquent que l'harmonie vocalique, et notamment l'harmonie vocalique fermante sur /ε, o/, se caractérise par un degré supérieur à celui de l'assimilation à distance en syllabe pénultième par rapport à d'autres positions au sein de mot, une variation qui ne saurait être expliquée par une influence purement « mécanique » du contexte phonétique déclencheur d'assimilation.

En outre, l'étude de Nguyen & Fagyal (2008) apporte un argument supplémentaire en faveur d'une coarticulation renforcée spécifique à la langue, notamment en mettant en évidence la variation du degré d'harmonie vocalique en fonction de la variété géographique, avec une tendance plus forte à l'harmonisation en français septentrional :

(...) vowel harmony patterns in our data also showed important differences with respect to V-to-V coarticulation, as it is portrayed in standard phonetic theory at least. These patterns were to some extent speaker-dependent, and appeared more systematic in Northern than Southern French. (Nguyen & Fagyal, 2008)

En dernier lieu, nous avons montré que l'harmonie vocalique était mise à profit par l'auditeur pour faciliter la reconnaissance des mots, ce qui témoigne d'une prégnance cognitive propre à ce phénomène.

En définitive, cela nous amène à conclure en proposant que l'harmonie vocalique en français correspond à ce que Ohala (1990) définit comme une *coarticulation fossilisée* ou *partiellement phonologisée* au sens où la magnitude de son effet dépasse un degré imputable exclusivement à la mécanique de la coproduction des sons de la parole :

There is a need to differentiate between “real-time” coarticulation and the fossilized or phonologized coarticulation due to sound change. The primary difference between the two is that only in the former is there a mechanical link between the variation noted and the contextual (conditioning) environment. Thus, in fossilized coarticulation one often encounters variation that is disproportionate with respect to the conditioning environment, i.e., exaggerated either in magnitude or temporal extent. Of course, clear cases of fossilized coarticulation are those where the conditioning environment has been lost through sound change. (Ohala, 1994)

Dans cette optique, une *coarticulation en temps réel* et une *coarticulation fossilisée* se distinguent par la présence ou non d'un contexte phonétique responsable d'une extension temporelle partielle d'un

trait phonétique sur des segments environnants ou, dans le cas d'une coarticulation à distance, au travers de segments. Or, dans le cas d'une forme de coarticulation transconsonantique ou assimilation – partielle ou totale – de voyelle à voyelle voire métaphonie ou harmonie vocalique qui, selon Ohala (1995), équivalent à autant de stades plus ou moins *fossilisés* d'un même phénomène étalés le long d'un spectre allant du *plus phonétique* au *plus phonologique*, la perte du contexte phonétique occasionnant l'assimilation de timbre vocalique est rare, ou du moins ne fait pas la règle. Toutefois, bien que ce dernier critère n'exclue pas à lui seul qu'un phénomène puisse être considéré comme étant une instance de *coarticulation fossilisée*, et que le contexte phonétique déclencheur soit toujours présent, cela nous amène à la question de savoir comment interpréter la diminution du degré d'harmonie vocalique dans le temps. En d'autres termes, dès lors que l'harmonie vocalique peut être appréhendée en termes de *coarticulation fossilisée*, il est justifiable d'estimer que son évolution, actuellement en cours, gagne à être examinée à l'aune de la dynamique évolutive du sous-système des voyelles moyennes, voire même du système vocalique français dans son ensemble.

#### **6.4. Harmonie vocalique et l'aménagement du système vocalique français**

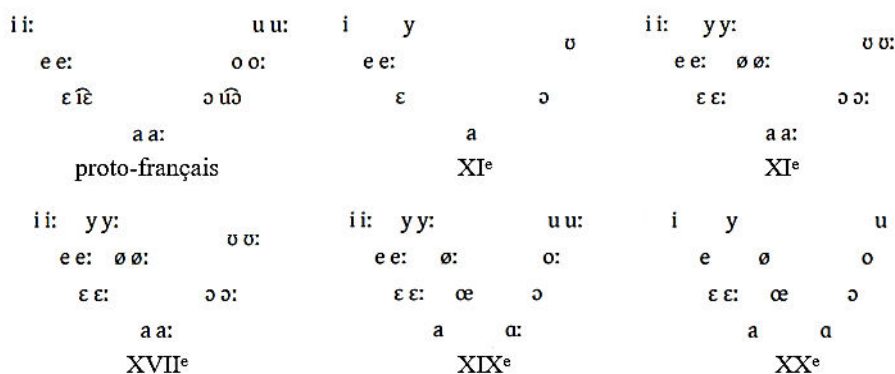
À commencer par l'ensemble des témoignages historiques et études impressionnistes anciennes (Cecelewski, 2022), en passant par les études empiriques contemporaines (voir sections 2.2. et 2.3.), jusqu'au présent travail de recherche, l'harmonie vocalique s'avère être un phénomène foncièrement variable, laissant apparaître une instabilité considérable qui conduit à des observations, sinon contradictoires, au moins fort hétérogènes qui peuvent aller jusqu'à empêcher de formuler des conclusions claires sur sa réalité acoustique effective.

Or, si la variabilité et l'instabilité sont autant de traits inhérents à l'identité de ce phénomène, et ce dès les premières mentions dans la littérature datant du début du XX<sup>e</sup> siècle (Rousselot & Laclotte, 1902), l'harmonie vocalique n'est qu'un « symptôme », sinon une retombée d'une instabilité plus générale qui, du moins selon l'avis de certains chercheurs, semble gouverner le vocalisme français dans sa totalité :

Le déséquilibre est flagrant, dans la phonologie du français, entre un système consonantique relativement équilibré, stabilisé pour l'essentiel depuis des siècles et démarqué de la structure latine qui demeure lisible dans ses grands traits, d'une part, et, d'autre part, un vocalisme construit autour d'oppositions précaires, à la fois élémentaires et fragiles. (Baraduc *et al.*, 1989)

Sans nous attarder sur cette affirmation sans doute trop simpliste et hâtive – en effet, si instabilité il y a, elle n'est certainement pas comparable pour l'ensemble du système –, force est tout de même de constater que l'histoire du vocalisme français est une histoire faite de mutations, de recouvrements, enfin de pertes de certains traits distinctifs (comme les longueurs vocaliques avant la fin du XIX<sup>e</sup> siècle, cf. Martinet, 1959) ou acquisitions de nouveaux types d'oppositions (comme l'apparition des timbres distinctifs des voyelles à double timbre, remontant à une époque difficile à déterminer avec certitude, cf. Pope, 1952). La figure ci-dessous, que nous reproduisons d'après Morin (2008), permet

de rendre compte sommairement des grandes tendances qui ont affecté le système vocalique français, limité ici aux voyelles orales toniques phonologiquement distinctives :



**Figure 30.** Mutations du système phonologique des voyelles orales toniques du français (d'après Morin, 2008)

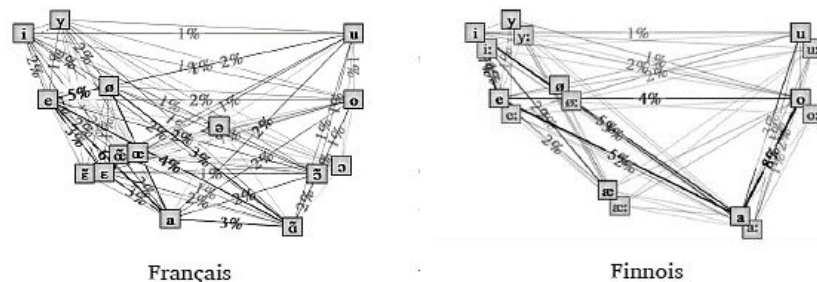
Ainsi, il suffit ne serait-ce que d'un survol rapide de l'histoire du vocalique français pour s'apercevoir qu'une relative stabilité des voyelles périphériques dans le temps contraste avec de nombreuses phases de changements et de transitions successives qui ont affecté les voyelles moyennes.

Une instabilité qui est, semble-t-il, une caractéristique *programmative* de sous-systèmes de voyelles moyennes ou *intermédiaires*. Un reflet indirect en est la difficulté de modéliser les larges systèmes vocaliques, à laquelle se sont heurtées différentes théories avancées, depuis Liljencrants & Lindblom (1972), pour prédire et comprendre l'architecture sonore des systèmes vocaliques à partir de contraintes auditives et articulatoires. Ainsi, les prédictions réalisées par la théorie de la dispersion adaptative (TD, Liljencrants & Lindblom, 1972) présentent un trop grand nombre de voyelles hautes non périphériques à la place de voyelles moyennes, plus ouvertes. Un problème qui se voit résolu par la théorie de la dispersion-focalisation (TDF, Schwartz *et al.*, 1997), qui met en compétition, avec le principe de distance perceptive entre voyelles, un principe de prégnance intra-voyelle : la focalisation, qui est la résultante d'une « convergence articulatoire produite autour d'un changement d'affiliation entre cavités » (Vallée *et al.*, 1999). Encore ici, tout comme dans des approches plus récentes mobilisant une composante du contrôle moteur (comme la Perception for Action Theory, PACT, Schwartz *et al.*, 2012), la force prédictive des modèles baisse sensiblement au-delà de cinq, six ou sept voyelles, en mettant parallèlement en exergue une certaine fragilité structurelle des larges systèmes vocaliques.

Une autre approche, davantage fonctionnaliste et abstraite, a puisé dans la notion de charge fonctionnelle, conçue pour qualifier le poids relatif des différentes unités phonologiques. Popularisée par Martinet (1933), dans le but de modéliser mathématiquement la résistance d'une opposition au changement phonétique (proportionnelle à sa charge fonctionnelle, selon son hypothèse attrayante mais infirmée depuis les premières études quantitatives de King, 1967 et Wang, 1967), la notion de charge fonctionnelle refait surface dans les années 2000 grâce à des études exploitant l'analyse de



grands corpus de parole continue (Surendran & Niyogi, 2003 ; Stokes & Surendran, 2005). Dans une approche basée sur la notion d'entropie telle que définie par Shannon (1948), qui permet notamment de décorrélérer le calcul de la charge fonctionnelle de la taille du système vocalique, Pellegrino *et al.* (2012) se sont ainsi proposé d'établir une typologie des systèmes vocaliques d'une dizaine de langues, dont le français. Dans la **Figure 31**, nous reproduisons les graphiques de la distribution de la charge fonctionnelle normalisée pour deux inventaires d'unités distinctives de taille comparable, français et finnois, d'après Pellegrino *et al.* (2012) :



**Figure 31.** Représentation des distributions de la charge fonctionnelle normalisée pour le français et le finnois (d'après Pellegrino *et al.*, 2012).

Les chercheurs soulignent ainsi que le système vocalique français se distingue des autres systèmes analysés par une répartition étonnamment progressive et équilibrée des charges fonctionnelles, ce qui serait notamment un reflet d'un tissu phonologique stable et économique, étayé par un ensemble des oppositions mises au pied d'égalité du point de vue de leurs contributions fonctionnelles respectives (Pellegrino *et al.*, 2012).

Une mise en rapport de l'approche fonctionnaliste de Pellegrino *et al.* (2012) d'une part et, d'une autre part, des théories de Liljencrants & Lindblom (1972) et Schwartz *et al.* (1997) – davantage ancrées dans la réalité phonétique (auditive et/ou articulaire) – laisse apparaître, dans le cas du vocalisme français, un système d'oppositions stables phonologiquement, puisque réparties de façon équilibrée dans la structure de la langue, mais faites d'une matière phonétique fragile et mouvante, du moins est-ce le cas des voyelles moyennes, occupant une région articulaire difficilement modélisable et auditivement peu saillant.

En conséquence, nous sommes convaincu que l'évolution d'harmonie vocalique, ainsi que sa variabilité synchronique que l'on observe en français à l'heure actuelle, mériterait d'être considérée dans un contexte plus large, qui est celui de la dynamique du système vocalique français dans son ensemble, ou du moins, celle du sous-système des voyelles moyennes. Or, comme en témoignent de nombreuses études, les voyelles moyennes seraient elles-mêmes sujettes à une série de changements de chaîne importants depuis le début du XX<sup>e</sup> siècle, changements qui attendent toujours d'être examinés par le biais d'une étude empirique approfondie.

## 6.5. Vers une perte des distinctions des voyelles à double timbre ?

Une tendance à la « confusion phonologique » (Hansen & Juillard, 2011) qui serait à l'œuvre en français standard (ou français parisien), est ainsi régulièrement dénoncée depuis déjà plusieurs décennies. Nombre de travaux puisant dans des méthodologies différentes (Martinet, 1945 ; Léon, 1973 ; Peretz, 1977 ; Peretz-Juillard, 1985 ; Landick, 1995 ; Storme, 2017 ; Gendrot & Audibert, 2019) ont noté des timbres intermédiaires là où on s'attendait à une distinction plus marquée entre timbre fermé et timbre ouvert. Partant, les quatre paires des voyelles dites à double timbre, /a, a/, /e, ε/, /ø, œ/ et /o, o/ ont été décrites comme encourageant une perte de distinctivité, chacune d'entre elles semblant posséder une dynamique et un parcours évolutif propres.

Ainsi la confusion entre /a/ et /a/ atteint un stade avancé dès le milieu du XX<sup>e</sup> siècle (Delattre, 1957). Le /a/ postérieur étymologique n'est alors plus présent qu'en parole soutenue et commence à acquérir une connotation « précieuse », faisant l'office d'un des traits caractéristiques de la prononciation « mondaine », parallèlement à l'émergence d'une tendance imitative, dans la jeune génération du sociolecte populaire, à remplacer /a/ en position accentuée par un [â] vélaire non étymologique (Mettas, 1970).

Concernant l'opposition des /ø, œ/, elle serait à peine stable parmi les locuteurs ayant fait des études supérieures, et complètement neutralisée dans les classes moins éduquées, ce dès les années 1980 (Landick, 1995).

L'opposition des /o, o/, quant à elle, quoiqu'également entamée, resterait la plus nette des quatre paires des voyelles à double timbre. Néanmoins, les résultats d'une enquête par panel menée par Hansen (2011) montrent qu'un changement vers la perte de cette opposition serait tout de même observable entre les années 1980 et 2010.

En revanche, un usage « assez normatif » de l'opposition /e, ε/ caractériserait le sociolecte des couches sociales populaires selon Léon (1973). Toujours est-il que des analyses empiriques plus récentes font état d'un recouplement presque total entre /e/ et /ε/, et confirment un rapprochement acoustique significatif entre ces deux sons en parole spontanée par rapport à la parole journalistique (Gendrot & Audibert, 2019).

D'autres tendances isolées, sources potentielles de confusion, sont également signalées, comme une surfréquence de timbres fermées /e, o, ø/ parmi les « jeunes socialement défavorisés en banlieue parisienne » (Coveney 2001, Jamin 1999), ou encore le phénomène d'antériorisation et de délabialisation de /o/, noté dans les travaux historiques des remarqueurs depuis le XVII<sup>e</sup> siècle comme véhiculant une connotation populaire (Armstrong & Low, 2008) et confirmé expérimentalement dans une étude plus récente de Boula de Mareüil *et al.* (2010).

Dans cette optique, la diminution du degré d'harmonie vocalique dans le temps pourrait s'expliquer par la confusion croissante des voyelles à double timbre, d'autant plus que les différences en termes de la dynamique de l'évolution du degré d'harmonie vocalique dans les cas de différentes

voyelles moyennes semblent correspondre, à grands traits, aux différentes dynamiques régissant ces confusions, telles que décrites dans la littérature, si tant est qu'elles correspondent à un rapprochement entre voyelles, comme proposé par Gendrot & Audibert (2019). La baisse d'harmonie vocalique ne serait alors qu'un corollaire mécanique d'un recouplement croissant entre voyelles entre elles, plutôt qu'un changement autonome.

Il n'empêche que cette question mérite un approfondissement expérimental. Il serait dès lors intéressant, dans le cadre d'une étude acoustique ciblée, de tenter de répondre à la question de savoir si la diminution du degré d'harmonie vocalique, en cours au moins depuis 1940 (Cecelewski *et al.*, à paraître) et jusqu'aux données les plus récentes, comme l'a indiqué l'analyse du corpus bien contrôlé de parole lue analysé ici, est entièrement due au processus de confusion entre les voyelles moyennes, à l'œuvre à l'heure actuelle en français standard, ou bien si le rapprochement acoustique n'est pas important au point d'entraîner une perte d'harmonie et que celle-ci demande de mobiliser d'autres facteurs explicatifs.

## 7. Conclusion

Nous nous sommes proposé ici d'examiner le phénomène d'assimilation de voyelle à voyelle affectant les voyelles moyennes en syllabe pénultième en fonction de l'aperture de la voyelle finale du mot, traditionnellement désigné par le terme d'*harmonie vocalique*. La présente étude s'inscrit, d'une part, dans la lignée des travaux de Fagyal *et al.* (2003), Nguyen *et al.* (2004), Nguyen & Fagyal (2008) et Turco *et al.* (2016a, 2016b) consacrés à des aspects acoustiques et perceptifs d'harmonie vocalique, et, d'une autre part, constitue le prolongement de nos travaux précédents sur l'évolution d'harmonie vocalique dans la seconde moitié du XX<sup>e</sup> siècle (Cecelewski, 2022 ; Cecelewski *et al.*, à paraître).

Ce travail s'est décliné en trois parties qui ont consisté, respectivement, en l'analyse du degré d'harmonie vocalique dans les grands corpus (BREF, ESTER et NCCFr), à l'analyse de l'harmonie vocalique dans un corpus plus récent, enregistré en laboratoire, et enfin à une expérience auditive consacrée à des aspects perceptifs d'harmonie vocalique en français.

Notre objectif ici était double : d'une part, nous avons voulu apporter des arguments empiriques supplémentaires afin d'étayer l'hypothèse de la diminution du degré d'harmonie vocalique dans le temps et, d'une autre part, nous nous sommes proposé de revisiter la question du statut phonologique de ce phénomène en français.

Dans un premier temps, nous avons montré, au moyen d'une analyse des corpus BREF, ESTER et NCCFr, que le degré d'harmonie vocalique était sensible au style de parole, mais que cette corrélation n'était pas uniforme pour toutes les voyelles affectées par l'assimilation, ce qui remet partiellement en cause les conclusions de l'étude de Turco *et al.* (2016a). Cela nous conduit à souligner l'intérêt d'une analyse plus approfondie du phénomène d'harmonie vocalique en français, notamment en distinguant une *harmonie vocalique fermante* – occasionnant une aperture des voyelles

moyennes mi-ouvertes /ɛ, ə/ –, ainsi qu’une harmonie vocalique ouvrante, agissant à son tour sur les voyelles moyennes mi-fermées /e, o/. Une autre conséquence, plus générale, en découle, qui consiste à postuler que la variation synchronique présente dans les contextes d’harmonie vocalique n’obéit pas à des tendances mécaniques propres à une forme de coarticulation à distance due exclusivement, par exemple, à l’inertie des articulateurs.

De plus, nous avons montré que les voyelles moyennes mi-ouvertes /ɛ, ə/ étaient sujettes à une assimilation plus importante dans des contextes d’harmonie vocalique par rapport à d’autres positions au sein de mot. Le degré d’assimilation des voyelles mi-fermées, en revanche, /e, o/, ne diffère pas selon la position de la syllabe dans le mot. Outre le schéma apparemment divergent dans le cas des deux séries de voyelles moyennes d’ouvertures différentes, ce résultat met en évidence une sorte de non-linéarité qui semble caractériser le comportement de l’harmonie vocalique en français, ce qui suggère qu’elle ne saurait être une résultante des « lois » phonétiques, et laisse supposer un certain degré de phonologisation.

Additionnellement, nous avons montré que l’assimilation de timbre induit par l’harmonie vocalique était, non seulement, perçue par l’auditeur ou, éventuellement, mise à profit pour accélérer le traitement cognitif du signal (selon l’étude perceptive de Nguyen *et al.*, 2004), mais que l’information acoustique véhiculée par l’harmonie vocalique pouvait être exploitée par l’auditeur dans la reconnaissance des mots. Les taux de réussite ne sont que faiblement inférieurs à ceux obtenus par Greca (2023) dans une expérience similaire consacrée à la métaphonie dans des dialectes italiens. Ce résultat corrobore notre hypothèse selon laquelle l’harmonie vocalique en français consiste en une forme de *coarticulation fossilisée*, c’est-à-dire *partiellement phonologisée*.

Dans un second temps, une analyse d’un corpus bien contrôlé enregistré en laboratoire nous a permis de confirmer que la *voix projetée* n’avait pas d’incidence sur le degré d’harmonie vocalique. Or, cela apporte un argument en faveur de l’hypothèse énoncée dans Cecelewski *et al.* (à paraître), (qui n’avait alors été testée que par le biais d’une mesure moins directe) selon laquelle le degré d’harmonie vocalique particulièrement important en parole journalistique des années 1940 et 1950 n’était pas induit pas le style de voix spécifique de l’époque (style « Gaumont-Pathé ») qui s’apparente le plus à de la voix projetée, que nous avons tenté de reproduire dans des conditions de laboratoire.

Deux arguments supplémentaires semblent confirmer la tendance à la baisse du degré d’harmonie vocalique en français. Tout d’abord, le degré d’assimilation, présent dans les données les plus récentes que nous avons analysées (à savoir dans le corpus bien contrôlé enregistré en laboratoire), s’est révélé plus bas que dans le corpus de construction similaire analysé vingt ans plus tôt par Fagyal *et al.* (2003), ainsi que dans les trois grands corpus, datant, eux aussi, en moyenne, de la même époque. La comparaison du degré d’harmonie vocalique sur *O* avec le degré d’assimilation dans des pseudo-contextes d’harmonie vocalique sur *A* révèle que l’évolution diachronique s’observe précisément dans les cas d’harmonie vocalique, mais pas dans d’autres types d’assimilation de voyelle à voyelle. Additionnellement, les résultats d’une expérience en *temps apparent* semblent dévoiler une

tendance diachronique similaire, constatée dans les productions de locuteurs appartenant à trois générations consécutives. Tous ces résultats mis ensemble nous amènent à souligner la nécessité de considérer les changements phonétiques dans un contexte plus large. Ainsi, un jeu de plusieurs facteurs explicatifs, examinés au gré d'une méthodologie variée, permet de mieux appréhender la nature d'une variation diachronique, tout en minimisant la probabilité d'aboutir à des conclusions fausses, induites par une perspective expérimentale trop restreinte.

Enfin, nous avons discuté l'évolution de l'harmonie vocalique dans le contexte plus large des évolutions, actuellement en cours, qui semblent affecter, selon un degré plus ou moins important, l'ensemble des voyelles moyennes du français, elles-mêmes singulièrement instables dans l'histoire du français, et plus généralement, dans d'autres systèmes linguistiques. Or, faute de données empiriques plus exhaustives destinées à documenter ces évolutions, nous postulons l'intérêt et la nécessité d'une étude diachronique plus approfondie des voyelles moyennes du français, qui serait susceptible d'apporter un éclairage nouveau sur le phénomène d'harmonie (ou harmonisation) vocalique.

En supplément, nous nous sommes efforcé d'apporter une contribution au débat sur l'adéquation du terme *harmonie vocalique*. Tandis que Fagyal *et al.* (2003) recommandent l'adoption du terme de *métaphonie* comme étant plus proche de la tradition terminologique de la linguistique romane, nous montrons qu'en réalité, les termes de *métaphonie*, de *umlaut*, voire d'*harmonie vocalique*, sont souvent usités de manière parfaitement interchangeable, et que la distinction entre, d'une part, le terme-parapluie d'harmonie vocalique et, d'une autre part, deux termes équivalents (*métaphonie*, *umlaut*) renvoyant à des traditions linguistiques différentes (*romane*, *germanique*), n'est guère respectée dans la littérature.

Qui plus est, nous montrons que le terme d'*harmonie vocalique* n'a en réalité connu qu'une carrière d'une durée limitée dans la tradition terminologique française, lorsqu'il est venu supplanter celui d'*harmonisation vocalique*. Or, le mérite de ce dernier terme, employé systématiquement pour désigner l'assimilation des voyelles moyennes en syllabe pénultième pendant plus de six décennies, avait été qu'il n'impliquait pas la *catégorialité phonologique*, désormais largement contestée, et mettait plutôt en avant le caractère partiel et graduel du phénomène en question. Parmi *métaphonie*, *umlaut*, *dilatation*, *harmonie* et *harmonisation*, c'est dès lors ce dernier terme qui nous paraît le mieux rendre la réalité empirique du phénomène qu'il désigne.

## 8. Bibliographie

- Abouda, L. & Skrovec, M. (2022). Micro-diachronie de l'oral. Présentation. *Langages*, 266(2), 9-24.
- Adam, L. (1874). *De l'harmonie vocalique dans les langues ouralo-altaïques*. Paris : Maisonneuve.
- Adda-Decker, M., de Mareüil, P.B., Adda, G. & Lamel, L. (2005). Investigating syllabic structures and their variation in spontaneous French. *Speech Communication*, 46(2), 119–139.
- Adda-Decker, M., Gendrot, C. & Nguyen, N. (2008). Contributions du traitement automatique de la parole à l'étude des voyelles orales du français. *Traitement Automatique Des Langues*, 49, 13-46.

- Adda-Decker, M., Gendrot, C., Snoeren, N. & Nguyen, N. (2013). *Apport du traitement automatique à l'étude des voyelles*. Hermes Science Publications.
- Amelot, A. (2004). *Étude aérodynamique, fibroscopique, acoustique et perceptive des voyelles nasales du français*. Linguistique. Thèse de doctorat, Université de la Sorbonne nouvelle - Paris III.
- Archangeli, D. & Pulleyblank, D. (2007). Harmony. In Paul de Lacy (ed.) *The Cambridge handbook of phonology*, 353-378. Cambridge : Cambridge University Press.
- Armstrong, N. & Low, J. (2008). C'est encœur plus jeuli, le Mareuc : some evidence for the spread of /ɔ/ - fronting in French. *Transactions of the Philological Society*, 106(2), 1-24.
- Audibert, N., Fougeron, C., Gendrot, C. & Adda-Decker, M. (2015). Duration-vs. style-dependent vowel variation: A multiparametric investigation. *18th International Congress of Phonetic Sciences (ICPhS'15)*.
- Audibert, N., Fougeron, C., Gendrot, C. & Adda-Decker, M. (2016). Duration- vs. style-dependent vowel variation: A multiparametric investigation. *18th International Congress of Phonetic Sciences (ICPhS 2016)*.
- Baraduc J., Bergounioux, G., Castellotti V., Dumont C. & Lansari M.-H. (1989). Le statut linguistique des voyelles moyennes. *Langage et société*, 49, 5-24.
- Barnes, J. A. (2002). *Positional neutralization: A phonologization approach to typological patterns*. PhD thesis, University of California, Berkeley, Berkeley.
- Bates, D., Maechler, M., Bolker, B. & Walker, S. (2014). lme4: Linear mixed-effects models using eigen and s4. R package version 1.1-31. <https://cran.r-project.org/package=lme4>.
- Blevins, J. (2004). *Evolutionary Phonology: The Emergence of Sound Patterns*. Cambridge University Press.
- Blevins, J. (2008). Phonetic explanation without compromise The evolution of Mussau syncope. *Diachronica*, 25(1), 1-19.
- Boersma, P. & Weenink, D. (2016). Praat: doing phonetics by computer (logiciel). <http://www.praat.org/>
- Boula de Mareüil, P., Rilliard, A. & Allauzen, A. (2008). Étude diachronique de l'accent initial au travers d'archives audio. *Actes des 27<sup>e</sup> Journées d'Étude sur la Parole JEP*, 1621-1624.
- Boula de Mareüil, P., Rilliard, A. & Allauzen, A. (2012a). A Diachronic Study of Initial Stress and other Prosodic Features in the French News Announcer Style : Corpus-based Measurements and Perceptual Experiments. *Language and Speech*, 55(2), 263-293.
- Boula de Mareüil, P., Rilliard, A. & Allauzen, A. (2012b). Variation diachronique dans la prosodie du style journalistique : Le cas de l'accent initial. *Revue française de linguistique appliquée*, 17(1), 97-111.
- Boula de Mareüil, P., Rilliard, A. & Allauzen, A. (2017). Perception de l'évolution diachronique de la prosodie dans le style journalistique. In : *Langue française mise en relief. Aspects grammaticaux et discursifs*, Presses Universitaires de Perpignan, pp. 187-198.
- Bürki, A., Fougeron, C., Gendrot, C. & Frauenfelder, U. H. (2011). Phonetic reduction versus phonological deletion of French schwa: Some methodological issues. *Journal of Phonetics*, 39(3), 279-288.
- Calabrese, A. (2011). Metaphony in Romance. In: *The Blackwell Companion to Phonology* (eds M. Oostendorp, C.J. Ewen, E. Hume and K. Rice).

- Casagrande, J. (1984). *The Sound System of French*. Georgetown University Press.
- Cecelewski, J. (2022). *Évolution de l'harmonie vocalique sur /O/ en français depuis 1940*. Mémoire de Master 1 Recherche (parcours Phonétique et phonologie), sous la direction de Cédric Gendrot, Paris, Université Sorbonne Nouvelle – Paris III, 80 p. en ligne <https://dumas.ccsd.cnrs.fr/MEM-UNIV-PARIS3/dumas-03978871v1> (lien valide le 8 mai 2023).
- Cecelewski, J., Gendrot, C., Adda-Decker, M. & Boula de Mareüil, P. (à paraître). A diachronic study of vowel harmony in French broadcast speech since 1940. *Proceedings of the 20th International Congress of Phonetic Sciences (ICPhS 2023)*, Prague, République Tchèque.
- Clayards, M., Gaskell, M.G. & Hawkins, S. (2021). Phonetic detail is used to predict a word's morphological composition. *Journal of Phonetics*, 87.
- CNRTL, Centre National de Ressources Textuelles et Lexicales, un outil de mutualisation de ressources linguistiques, en ligne : <https://www.cnrtl.fr> (lien valide le 8 mai 2023).
- D'Alessandro, D. & Fougeron, C. (2021). Changes in Anticipatory VtoV Coarticulation in French during Adulthood. *Languages*, 6(4), 181-195.
- Dauzat, A. (1930). *Histoire de la langue française*, Paris, Payot.
- Delattre, P. (1951). *Principes de phonétique française à l'usage des étudiants franco-américains*. Middlebury, Vermont.
- Delattre, P. (1957). La Question des deux "a" en français, *The French Review*, 31(2), 141-148.
- Delattre, P. (1959). Rapports entre la durée vocalique, le timbre et la structure syllabique en français. *French Review*, 32, 547-552.
- Di Cristo, A. (1998). Intonation in French. In D. Hirst, & A. Di Cristo, (eds.), *Intonation systems. A survey of twenty languages*, pp. 195-218. Cambridge, UK: Cambridge University Press.
- Duniec, A. & Crouzet, O. (2013). Dynamics of vowel-to-vowel assimilation in French. *PaPI2013 - Phonetics and Phonology in Iberia*.
- Durand, J. & Lyche, C. (2004). Structure et variation dans quelques systèmes vocaliques du français : l'enquête Phonologie du français contemporain (PFC). In A. Coveney, & C. Sanders, (eds.), *Variation et francophonie*, pp. 217-240. Paris : Harmattan.
- Fagyal, Z., Nguyen, N. & Boula de Mareüil, P. (2003). From dilation to coarticulation : Is there vowel harmony in French? *Studies in Linguistic Sciences*, 32, 1-21.
- Fouché, P. (1935). L'évolution phonétique du français du 16<sup>e</sup> siècle à nos jours. In A. Dauzat, (eds.), *Où en sont les études de français*, pp. 35-54, Bibliothèque du « français moderne », Paris.
- Fouché, P. (1956). *Traité de prononciation française*. C. Klincksieck.
- Fowler, C.A. (1981). Production and perception of coarticulation among stressed and unstressed vowels. *Journal of Speech and Hearing Research*, 24, 127–139.
- Fox, J. (2003). Effect displays in R for generalised linear models. *Journal of Statistical Software*, 8(15), 1-27.
- Galliano, S., Geoffrois, E., Gravier, G., Bonastre, J.-F., Mostefa, M. & Choukri, K. (2006). Corpus description of the ESTER Evaluation Campaign for the Rich Transcription of French Broadcast News. In *Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06)*, Genoa, Italy. European Language Resources Association (ELRA).

- Gendrot, C. & Audibert, N. (2019). La distinction /e/ vs /ɛ/ en français standard est-elle maintenue en finale de mot ? Étude sur des corpus de parole journalistique et de parole spontanée. *Langue française*, 203, 53-66.
- Georgia Z., Pycha, A. & Cohn, M. (2023). The perception of nasal coarticulatory variation in face-masked speech. *Journal of the Acoustical Society of America*, 153(2), 1084–1093.
- Grammont, M. (1916). *Traité pratique de prononciation française*. Delagrave.
- Greca, P. (2023). *An experimental analysis of metaphony and sound change in the dialects of the Lausberg area (Southern Italy)*. Thèse de doctorat, Ludwig-Maximilians-Universität, Munich.
- Greenwell, B.M. & Boehmke, B.C. (2020). Variable Importance Plots-An Introduction to the vip Package. *The R Journal*, 12(1), 343-366.
- Grosvald, M. & Corina, D. (2012). Perception of long-distance coarticulation: An event-related potential and behavioral study. *Applied Psycholinguistics*, 33, 55–82
- Grosvald, M. (2009). Interspeaker variation in the extent and perception of long-distance vowel-to-vowel coarticulation. *Journal of Phonetics*, 37(2), 173-188.
- Grosvald, M. (2010). Long-Distance Coarticulation in Spoken and Signed Language: An Overview. *Language and Linguistics Compass*, 4(6), 348–362.
- Hansen, A.B. & Juillard, C. (2011). La phonologie parisienne à trente ans d'intervalle – Les voyelles à double timbre. *Journal of French Language Studies*, 21, 313-359.
- Hawkins, S. & Slater, A. (1994). Spread of CV and V-to-V coarticulation in British English: Implications for the intelligibility of synthetic speech. In : *Proceedings of ICSLP 94*, volume 1, pp. 57-60.
- Henry, V. (1893). *Précis de grammaire comparée de l'anglais et de l'allemand rapportés à leur commune origine et rapprochés des langues classiques*, Paris.
- Henry, V. (1894). *A Short Comparative Grammar of English and German*, London.
- Hertrich, I. & Ackermann, H. (1995). Coarticulation in slow speech: durational and spectral analysis. *Language and Speech*, 38(2), 159-187.
- Hualde, J.I. (1989). Autosegmental and Metrical Spreading in the Vowel-Harmony Systems of Northwestern Spain. *Linguistics*, 27(5), 773-805.
- Jakobson, R. Fant, G. & Halle, M. (1952). Preliminaries to speech analysis. *Technical Report 13, Acoustics Laboratory*, Massachusetts Institute of Technology.
- Kaun, A. (1995). *The typology of rounding harmony: An Optimality Theoretic approach*. Thèse de doctorat, University of California, Los Angeles.
- Kaze, J. (1991). Metaphony and two models for the description of vowel systems. *Phonology*, 8(1), 163-170.
- Keating, P. A. (1990b). The window model of coarticulation: articulatory evidence. In Kingston, J., Beckman, M. E. (eds.) *Papers in Laboratory Phonology I: Between the Grammar and Physics of Speech*. 451–470. Cambridge: Cambridge University Press.
- Kermabon, J. (ed.) (1994). *Pathé: Premier Empire du Cinéma, Paris*. Centre Pompidou.
- Kimper, W. (2017). Not crazy after all these years? Perceptual grounding for long-distance vowel harmony. *Laboratory Phonology: Journal of the Association for Laboratory Phonology*, 8(1):19, 1-37.



- King, R.D. (1967). Functional Load and Sound Change. *Language*, 43(4), 831-852.
- Kisler, T., Reichel, U. & Schiel, F. (2017). Multilingual processing of speech via web services, *Computer Speech et Language*, 45, 326–347.
- Krämer, M. (2003). *Vowel Harmony and Correspondence Theory*. Berlin, New York: De Gruyter Mouton.
- Kühnert, B. & Nolan, F. (1999). *The origin of coarticulation. Coarticulation: theory, data and techniques*, (eds.) W. J. Hardcastle & N. Hewlett, pp. 7-30. Cambridge: Cambridge University Press.
- Labov, W. (1972). *Sociolinguistic Patterns*, Philadelphia, University of Pennsylvania Press.
- Lamel, L., Gauvain, J. L. & Eskénazi, M. (1991). BREF, a large vocabulary spoken corpus for French. In *2nd European Conference on Speech Communication and Technology (Eurospeech 1991)*.
- Landick, M. (1995). The Mid-Vowels in Figures: Hard Facts. *The French Review*, 69(1), 88-102.
- Landick, M. (2004). *Enquête sur la prononciation du français de référence : Les voyelles moyennes et l'harmonie vocalique*. Paris : Harmattan.
- Lausberg, H. (1939). *Die Mundarten Südlukaniens*. Halle an der Saale: Niemeyer.
- Lehiste, I., & Shockey, L. (1972). On the perception of coarticulation effects in English VCV syllables. *Journal of Speech & Hearing Research*, 15(3), 500–506.
- Lenth, R.V., Love, J. & Hervé, M. (2017). Package ‘emmeans: Estimated Marginal Means, aka Least-Squares Means’, [cran.r-project.org/web/packages/emmeans](https://cran.r-project.org/web/packages/emmeans).
- Léon, P.R. (1973). Modèle standard et système vocalique du français populaire de jeunes Parisiens. Dans: Rondeau, G. (dir.), *Contributions canadiennes à la linguistique appliquée*, pp. 55–79.
- Liljencrants, J. & Lindblom, B. (1972). Numerical Simulation of Vowel Quality Systems: The Role of Perceptual Contrast. *Language*, 48(4), 839-862.
- Lindblom, B.E.F. (1986). Phonetic Universals in Vowel Systems. *Experimental Phonology*, (eds). Ohala J.J., Academic Press, Orlando, Florida, 13-44.
- Lloret, M.-R. (2007). On the nature of vowel harmony : spreading with a purpose. *Proceedings of the XXXIII Incontro di Grammatica Generativa*. pp. 15-35.
- Loporcaro, M. (2005). Typological remarks on Sardinian. 1. Vowel harmony 2. Sardinian in a correlative typology of the Romance languages, *Sprachtypologie und Universalienforschung*, 58, 210-227.
- Lüdecke, D., Ben-Shachar, M. S., Patil, I., Waggoner, P. & Makowski, D., (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, 6(60), 3139, <https://CRAN.R-project.org/package=performance>.
- Majors, T. (1998). *Stress dependent harmony: Phonetics origin and phonological analysis*. Thèse de doctorat, University of Texas, Austin, TX.
- Malmberg, B. (1969). *Phonétique française*. Hermods.
- Manca, A. D., Di Russo, F., Sigona, F. & Grimaldi, M. (2019). Electrophysiological evidence of phonemotopic representations of vowels in the primary and secondary auditory cortex. *Cortex*, 121, 385-398.
- Marslen-Wilson, W. & Warren, P. (1994). Levels of perceptual representation and process in lexical access - words, phonemes, and features. *Psychological Review*, 101, 653-675.

- Martin, J. & Bunnell, H. (1982). Perception of anticipatory coarticulation effects in vowel-consonant-vowel sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 473-488.
- Martin, J. G. & Bunnell, H. T. (1982). Perception of anticipatory coarticulation effects in vowel-stop consonant-vowel sequences. *Journal of Experimental Psychology: Human Perception and Performance*, 8, 473-488.
- Martinet, A. (1933). Remarques sur le système phonologique du français. *Bulletin de la Société Linguistique de Paris*, 34, 192-202.
- Martinet, A. (1958). C'est jeu! le Mareuc ! *Romance philology*, 11, 345-355.
- Martinet, A. (1959). L'évolution contemporaine du système phonologique contemporain. *Free University Quarterly* 7(2), 1-16. Amsterdam: Université Libre.
- Mateus, M. H. M. (1987). L'harmonie vocalique en portugais. *Revue européenne des sciences sociales*, 25(77), 225-230.
- Menzerath, P., Lacerda, A. (1933). *Koartikulation, Steuerung und Lautabgrenzung: Eine experimentelle Untersuchung*. Berlin and Bonn: Ferd. Dümmler.
- Mettas, O. (1970). Étude sur le A dans deux sociolectes parisiens, *Revue Romane*, 5(1), 94-105.
- Meunier, C. & Ghio, A. (2018). Caractériser la distinctivité du système vocalique des locuteurs. *XXXII<sup>e</sup> Journées d'Études sur la Parole*, Aix-en-Provence, France. pp. 469-477.
- Mok, P. P. (2011). Effects of vowel duration and vowel quality on vowel-to-vowel coarticulation. *Language and Speech*, 54(4), 527-545.
- Morin, Y. C. (1971). *Computer Experiments in Generative Phonology : Low-level French Phonology. Natural Language Studies II*. Department of Computer and Communication Sciences, The University of Michigan, Ann Arbor.
- Morin, Y. C. (2008). Histoire des systèmes phonique et graphique du français. In G. Ernst, M. D. Gleßgen, C. Schmitt & W. Schweickard (eds.), *Romanische Sprachgeschichte / Histoire linguistique de la Romania - Ein internationales Handbuch zur Geschichte der romanischen Sprachen / Manuel international d'histoire linguistique de la Romania*, 3, pp. 2907-2926, Mouton de Gruyter.
- Morin, Y.-C. (1986). La loi de position ou de l'explication en phonologie historique. *Revue québécoise de linguistique*, 15(2), 199-231.
- Nevins, A. (2010). *Locality in vowel harmony*, Linguistic Inquiry Monographs, 55. Cambridge, MA: MIT Press.
- Nguyen, N. (2001). Rôle de la coarticulation dans la reconnaissance des mots. *Année psychologique*, 101(1), 125-154.
- Nguyen, N. & Fagyal, Z. (2008). Acoustic aspects of vowel harmony in French. *Journal of Phonetics*, 36, 1-27.
- Nguyen, N., Fagyal, Z. & Cole, J. (2004). Perceptual relevance of long-domain phonetic dependencies. *Journées d'Études Linguistiques (JEL)*, Nantes, France. pp. 173-178.
- Ohala, J. (1981). The listener as a source of sound change. In C. S. Masek, R. A. Hendrick, & M. F. Miller (eds.), *Papers from the Parasession on Language and Behavior*, pp. 178-203, Chicago Linguistics Society.
- Ohala, J. (1994). Towards a universal, phonetically-based, theory of vowel harmony. *ICSLP 3*, 491-494.

- Ohala, J. J. (1974). Experimental historical phonology. In J. M. Anderson & C. Jones (Eds.), *Historical linguistics II. Theory and description in phonology*. Amsterdam: North-Holland. pp. 353-389.
- Ohala, J. J. (1993). Coarticulation and Phonology. *Language and Speech*, 36(2-3), 155-170.
- Öhman, S. E. (1966). Coarticulation in VCV utterances : Spectrographic measurements. *The Journal of the Acoustical Society of America*, 39(1), 151-168.
- Ostreicher, H. J. & Sharf, D. J. (1976). Effects of coarticulation on the identification of deleted consonants and vowel sounds. *Journal of Phonetics*, 4, 285-301.
- Pellegrino, F., Marsico, E. & Coupé, Ch. (2012). La typologie des systèmes vocaliques revisitée sous l'angle de la charge fonctionnelle. *Proceedings of JEP-TALN-RECITAL 2012*, Grenoble, pp. 617-624.
- Peretz, C. (1977). *Les voyelles orales à Paris dans la dynamique des âges et de la société*. Thèse de 3ème cycle, Université de Paris V.
- Peretz-Juillard, C. (1985). Variétés vocaliques et différenciations sociales à Paris. Dans: Houdebine, A. (dir.), *La Phonologie de l'enfant français de six ans*. Variétés régionales. Hamburg: Helmut Buske Verlag, pp. 131-163.
- Pope, M. K. (1952). *From Latin to modern French with especial consideration of Anglo Norman*. Manchester: Manchester University Press.
- R Development Core Team. 2010. *R: A language and environment for statistical computing* [logiciel]. Version 2022.12.0+353 <https://www.r-project.org/>, Vienne, R Foundation for Statistical Computing.
- Rathcke, T. & Stuart-Smith, J. (2015). On the Tail of the Scottish Vowel Length Rule in Glasgow. *Language and Speech*, 58, 1-27.
- Recasens, D. (1987). An acoustic analysis of V-to-C and V-to-V coarticulatory effects in Catalan and Spanish VCV sequences. *Journal of Phonetics*, 15(4), 299-312.
- Rensch, K.-H. (1964). *Beiträge zur Kenntnis nordkalabrischer Mundarten*. Münster: Aschendorffsche Verlagsbuchhandlung.
- Rousselot, abbé P., Laclotte, F. (1902). *Précis de prononciation française*, Paris, Leipzig : Welter.
- Sankoff, G. & Blondeau, H. (2007). Language change across the lifespan: /t/ in Montreal French. *Language*, 83(3), 560-588.
- Schlichting, F. & Sullivan, K. (2013). The imitated voice-A problem for voice line-ups?. *International Journal of Speech Language and the Law*, 4, 148-165.
- Schürr, F. (1936). Umlaut und Diphthongierung in der Romania. *Romanische Forschungen*, 50, 275-316.
- Schwartz, J.L., Basirat, A., Ménard, L. & Sato, M. (2012). The Perception for Action Control Theory (PACT): a perceptuo-motor theory of speech perception. *Journal of NeuroLinguistics*, 25, 336-354.
- Schwartz, J.L., Boë, L.J., Vallée, N. & Abry, C. (1997). The dispersion-focalization theory of vowel systems. *Journal of Phonetics*, 25, 255-286.
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, 27, 379-423 & 623-656.
- Sóskuthy, M. & Stuart-Smith, J. (2020). Voice quality and coda /t/ in Glasgow English in the early 20th century. *Language Variation and Change*, 32(2), 133-157.
- Steriade, D. (1995). Positional neutralization. In: *NELS 24*. University of Massachusetts, Amherst.

- Stevens, K. (2002). Toward a model for lexical access based on acoustic landmarks and distinctive features. *Journal of the Acoustical Society of America*, 111, 1872-1891.
- Stoet, G. (2010). PsyToolkit - A software package for programming psychological experiments using Linux. *Behavior Research Methods*, 42(4), 1096-1104.
- Stoet, G. (2017). PsyToolkit: A novel web-based method for running online questionnaires and reaction-time experiments. *Teaching of Psychology*, 44(1), 24-31.
- Stokes, S. & Surendran, D. (2005). Articulatory complexity, ambient frequency and functional load as predictors of consonant development in children. *Journal of Speech and Hearing Research*, 48(3), 577-591.
- Storme, B. (2017). The loi de position and the acoustics of French mid vowels. *Glossa: a journal of general linguistics*, 2(1): 64, 1-25.
- Straka, G. (1950). Système des voyelles du français moderne. *Bulletin de la Faculté des lettres de Strasbourg*, 220-233.
- Straka, G. (1981). Sur la formation de la prononciation française d'aujourd'hui. *Travaux de linguistique et de littérature*, 19(1), 161-248.
- Suomi, K. (1983). Palatal harmony: A perceptually motivated phenomenon? *Nordic Journal of Linguistics*, 6, 1-35.
- Surendran, D. & Niyogi, P. (2003). Measuring the Usefulness (Functional Load) of Phonological Contrasts. *Technical Report TR-2003-12*. Department of Computer Science, University of Chicago.
- Tilsen, S. (2007). Vowel-to-vowel coarticulation and dissimilation in phonemic-response priming. *UC Berkeley PhonLab Annual Report*, 3(3), 416-458.
- Torreira, F., Adda-Decker, M. & Ernestus, M. (2010). The nijmegen corpus of casual French. *Speech Communication*, 52(3), 201-202.
- Tranel, B. (1987). *The Sounds of French : An Introduction*. Cambridge University Press.
- Turco, G., Fougeron, C. & Audibert, N. (2016a). Que nous apprennent les gros corpus sur l'harmonie vocalique en français ?. *Actes des 31<sup>e</sup> Journées d'Etudes sur la Parole JEP*, 1, 571-579.
- Turco, G., Fougeron, C. & Audibert, N. (2016b). The Effects of Prosody on French V-to-V Coarticulation : A Corpus-Based Study. *Proc. of Interspeech*, 998-1001.
- Ultan, R. (1973). Some reflexions on vowel harmony. *Working Papers on Language Universals*, 12, 37-67.
- Vaissière, J. (1996). From Latin to Modern French: on diachronic changes and synchronic Variations. *AIPUK, Arbetisberitche*, Institut für Phonetik und digitale Sprachverarbeitung, Universität Kiel, 61-74.
- Valdman, A. (1978). The 'loi de position' and the direction of phonological change in the French mid vowel System. In M. Suter (eds.), *Contemporary studies in Romance linguistics*, pp. 316-329, Washington, D . C ., Georgetown University Press.
- Vallée, N., Boë, L.-J. & Stefanuto, M. (1999). Typologies phonologiques et tendances universelles. Approche substantialiste, *Linx*, 11, 31-54.
- Walker, R. (2012). Vowel Harmony in Optimality Theory. *Language and Linguistics Compass*, 6(9), 575-592.
- Wang, W. (1967). The Measurement of Functional Load. *Phonetica*, 16, 36-54.

Winitz, H., Scheib, M. E. & Reeds, J. A. (1971). Identification of stops and vowels from the burst portion of /p,t,k/ isolated from conversational speech. *Journal of the Acoustical Society of America*, 51, 1309-1317.

Wright, R. (2003). Factors of lexical competition in vowel articulation. In Local, J., Ogden, R. & Temple, R., (eds.), *Papers in laboratory phonology, VI: Phonetic interpretation*, pp. 75-87. Cambridge, UK: Cambridge University Press.

Yu, A.C.L. (2020). The Phonetics of Sound Change. In *The Handbook of Historical Linguistics* (eds.) R.D. Janda, B.D. Joseph and B.S. Vance).

Zetterholm, E. (2007). Detection of Speaker Characteristics Using Voice Imitation. In C. Müller (eds.) *Speaker Classification, 192-205*, Springer.

## 9. Annexes

contraste	/e/				/o/			
	coeff.	err.-type	t	p	coeff.	err.-type	t	p
BREF V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-13.13	1.06	-12.40	<.001	-41.81	14.15	-17.95	<.001
ESTER V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-16.59	1.41	-11.78	<.001	-32.42	2.93	-11.08	<.001
NCCFr V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-6.24	2.08	-2.996	NS	-30.29	3.79	-7.98	<.001
BREF V <sub>2</sub> fermée : ESTER V <sub>2</sub> fermée	-6.54	1.36	-4.80	<.001	-4.78	14.15	-0.33	NS
BREF V <sub>2</sub> fermée : NCCFr V <sub>2</sub> fermée	-13.13	1.06	-12.40	NS	-3.65	16.20	-0.22	NS
ESTER V <sub>2</sub> fermée : NCCFr V <sub>2</sub> fermée	9.83	1.98	-1.96	<.001	1.13	9.06	0.125	NS
BREF V <sub>2</sub> ouverte : ESTER V <sub>2</sub> ouverte	-10.01	1.12	-8.96	<.001	4.60	14.33	0.32	NS
BREF V <sub>2</sub> ouverte : NCCFr V <sub>2</sub> ouverte	10.18	1.39	7.343	<.001	7.87	16.50	0.47	<.05
ESTER V <sub>2</sub> ouverte : NCCFr V <sub>2</sub> ouverte	20.19	1.54	13.09	<.001	3.27	9.83	0.33	<.001
	/ɛ/				/ɔ/			
BREF V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-57.54	2.34	-24.43	<.001	-32.19	1.41	-22.85	<.001
ESTER V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-37.29	2.62	-14.24	<.001	-41.40	1.55	-26.70	<.001
NCCFr V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-36.79	3.49	-10.59	<.001	-60.61	2.32	-26.13	<.001
BREF V <sub>2</sub> fermée : ESTER V <sub>2</sub> fermée	0.73	14.59	0.05	NS	5.85	16.06	0.36	NS
BREF V <sub>2</sub> fermée : NCCFr V <sub>2</sub> fermée	13.36	16.99	0.79	<.05	-22.57	18.98	-1.18	<.05
ESTER V <sub>2</sub> fermée : NCCFr V <sub>2</sub> fermée	12.63	9.94	1.27	<.05	32.19	11.07	2.90	<.05
BREF V <sub>2</sub> ouverte : ESTER V <sub>2</sub> ouverte	20.99	14.59	1.42	<.05	-3.36	16.04	-0.29	NS
BREF V <sub>2</sub> ouverte : NCCFr V <sub>2</sub> ouverte	34.11	16.89	2.19	NS	9.62	18.98	0.57	NS
ESTER V <sub>2</sub> ouverte : NCCFr V <sub>2</sub> ouverte	13.12	9.81	1.33	<.05	12.98	11.01	1.18	NS
	/E/				/O/			
BREF V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-20.85	1.07	-19.46	<.001	-42.62	1.17	-36.35	<.001
ESTER V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-19.00	1.35	-14.04	<.001	-50.70	1.31	-38.69	<.001
NCCFr V <sub>2</sub> fermée : V <sub>2</sub> ouverte	-14.88	1.93	-7.73	<.001	-59.86	1.85	-32.28	<.001
BREF V <sub>2</sub> fermée : ESTER V <sub>2</sub> fermée	8.93	13.29	0.67	NS	5.75	14.71	0.39	NS
BREF V <sub>2</sub> fermée : NCCFr V <sub>2</sub> fermée	25.40	15.83	1.60	<.05	28.58	17.58	1.62	NS
ESTER V <sub>2</sub> fermée : NCCFr V <sub>2</sub> fermée	16.48	9.35	1.76	NS	22.84	10.41	2.19	<.05
BREF V <sub>2</sub> ouverte : ESTER V <sub>2</sub> ouverte	10.78	13.29	0.81	NS	-2.33	14.71	-0.15	NS
BREF V <sub>2</sub> ouverte : NCCFr V <sub>2</sub> ouverte	31.37	15.80	1.94	NS	11.35	17.59	0.64	NS
ESTER V <sub>2</sub> ouverte : NCCFr V <sub>2</sub> ouverte	20.59	9.28	2.21	NS	13.68	10.42	1.31	NS

**Annexe 2.** Résultats des tests de comparaisons multiples pour les régressions linéaires multiples réalisées pour le F1 de la V<sub>1</sub> ∈ {e, o, ε, o, E, O} ; formule utilisée : lmer (F1~aperture\_V2\*corpus+(1|locuteur))

locuteur	V1	formant	variation	SC	dl	CM	F	p	
AR	/e/	F1	aperture_V2	18429	1	18429	16.768	<.001	
			voix	18429	1	18429	49.602	<.001	
			aperture_V2:voix résiduelle	3181	1	3181	2.894	NS	
		F2	aperture_V2	101671	1	101671	9.999	<.05	
			voix	264323	1	264323	25.994	<.001	
			aperture_V2:voix résiduelle	5	1	5	0	NS	
	/o/	F1	aperture_V2	5562	1	5562	4.075	<.05	
			voix	77351	1	77351	56.666	<.001	
		F2	aperture_V2						
			voix						
			aperture_V2:voix résiduelle						

			aperture_V2:voix	244	1	244	0.178	NS	
			résiduelle	114662	84	1365			
		F2	aperture_V2	23401	1	23401	0.923	NS	
			voix	11485	1	11485	0.571	NS	
			aperture_V2:voix	15945	1	115945	0.629	NS	
			résiduelle	2129319	84	25349			
BA	/e/	F1	aperture_V2	29343	1	29343	18.735	<.001	
			voix	27671	1	27671	17.668	<.001	
			aperture_V2:voix	264	1	264	0.169	NS	
			résiduelle	194211	124	1566			
		F2	aperture_V2	12660	1	12660	0.789	NS	
			voix	23193	1	23193	1.446	NS	
		/o/	F1	aperture_V2	3628	1	3628	0.226	NS
	voix			78303	1	78303	4.870	<.05	
	aperture_V2:voix			1960	1	1960	0.122	NS	
	résiduelle			1350694	84	16080			
	F2		aperture_V2	48969	1	48969	1.120	NS	
			voix	743360	1	743360	17.002	<.001	
			aperture_V2:voix	492	1	492	0.011	NS	
			résiduelle	3672718	84	43723			
GA	/e/	F1	aperture_V2	70410	1	70410	35.211	<.001	
			voix	162818	1	162818	81.422	<.001	
			aperture_V2:voix	2921	1	2921	1.461	NS	
			résiduelle	2433960	122	2000			
		F2	aperture_V2	276896	1	276896	25.900	<.001	
			voix	1944	1	1944	0.182	NS	
		/o/	F1	aperture_V2	339	1	339	0.249	NS
	voix			145237	1	145237	90.552	<.001	
	aperture_V2:voix			73	1	73	0.046	NS	
	résiduelle			131521	84	1604			
	F2		aperture_V2	29739	1	29739	1.682	NS	
			voix	1505	1	1505	0.085	NS	
			aperture_V2:voix	16874	1	16874	0.955	NS	
			résiduelle	1448793	84	17668			
HU	/e/	F1	aperture_V2	21814	1	21814	12.727	<.001	
			voix	382922	1	382922	223.409	<.001	
			aperture_V2:voix	5900	1	5900	3.442	NS	
			résiduelle	212535	124	1714			
		F2	aperture_V2	30227	1	30227	2.177	NS	
			voix	66749	1	66749	4.807	<.05	
		/o/	F1	aperture_V2	3905	1	3905	0.281	NS
	voix			1721980	124	13887			
	aperture_V2:voix			3605	1	3605	2.230	NS	
	résiduelle			358275	1	358275	221.647	<.001	
	F2		aperture_V2	35	1	35	0.022	NS	
			voix	135779	84	1616			
			aperture_V2	8211	1	8211	0.300	NS	
			voix	1986	1	1986	0.072	NS	
			aperture_V2:voix	53	1	53	0.002	NS	
			résiduelle	2301857	84	27403			
PA	/e/	F1	aperture_V2	2803	1	2803	1.545	NS	
			voix	524938	1	524928	289.379	<.001	
			aperture_V2:voix	5711	1	5711	3.148	NS	
			résiduelle	224934	124	1814			
		F2	aperture_V2	91271	1	91271	7.606	<.05	
			voix	4209	1	4209	0.351	NS	

			aperture_V2:voix	19159	1	19159	1.597	NS
			résiduelle	1488044	124	12000		
	/o/	F1	aperture_V2	10623	1	10623	4.487	<.05
			voix	367182	1	367182	155.092	<.001
			aperture_V2:voix	811	1	811	0.343	NS
			résiduelle	198871	84	2368		
		F2	aperture_V2	8782	1	8782	0.518	NS
			voix	94049	1	94049	5.551	<.05
			aperture_V2:voix	8802	1	8802	0.518	NS
			résiduelle	1423215	84	16943		
TR	/e/	F1	aperture_V2	504	1	504	0.330	NS
			voix	87258	1	87258	57.135	<.001
			aperture_V2:voix	4489	1	4489	2.939	NS
			résiduelle	189375	124	1527		
		F2	aperture_V2	170309	1	170309	5.075	<.05
			voix	12149	1	12149	0.362	NS
			aperture_V2:voix	33768	1	33768	1.006	NS
			résiduelle	4161483	124	33560		
	/o/	F1	aperture_V2	1406	1	1406	1.461	NS
			voix	64442	1	64442	64.546	<.001
			aperture_V2:voix	95	1	95	0.095	NS
			résiduelle	83865	84	998		
		F2	aperture_V2	453	1	453	0.050	NS
			voix	6	1	6	0.001	NS
			aperture_V2:voix	1199	1	1199	0.133	NS
			résiduelle	754809	84	8986		

**Annexe 2.** Résultats des tests d'ANOVA à deux facteurs avec interaction effectués pour les six locuteurs du corpus ; formule utilisée : `aov(Fα_V1~aperture_V2*voix)`.

locuteur	V1	formant	variation	SC	dl	CM	F	p
AR	/ε/	F1	aperture_V2	110681	1	110681	16.07	<.001
			voix	833835	1	833835	121.08	<.001
			aperture_V2:voix	7229	1	7229	1.05	NS
		résiduelle	3436446	499	6886			
		F2	aperture_V2	11167	1	11167	1.009	NS
			voix	232789	1	232789	21.045	<.0001
	aperture_V2:voix		974	1	974	0.088	NS	
	résiduelle	752197	68	11062				
	/o/	F1	aperture_V2	7454	1	7454	5.935	<.05
			voix	60840	1	60840	48.443	<.001
			aperture_V2:voix	394	1	394	0.314	NS
		résiduelle	195921	156	1256			
F2		aperture_V2	9972	1	9972	0.628	<.05	
		voix	72932	1	72932	4.590	NS	
	aperture_V2:voix	451	1	451	0.028	NS		
résiduelle	2478804	156	15890					
BA	/ε/	F1	aperture_V2	21661	1	21661	6.881	<.05
			voix	19767	1	19767	6.279	<.05
			aperture_V2:voix	302	1	302	0.096	NS
		résiduelle	214067	68	3148			
		F2	aperture_V2	5724	1	5724	0.487	NS
			voix	144991	1	144991	12.342	<.001
	aperture_V2:voix		3601	1	3601	0.306	NS	
	résiduelle	798856	68	11748				
	/o/	F1	aperture_V2	20369	1	20369	2.772	NS
			voix	216090	1	216090	29.404	<.001
			aperture_V2:voix	1307	1	1307	0.178	NS
		résiduelle	1146427	156	7349			
F2		aperture_V2	779	1	779	0.033	NS	

			voix	32710	1	32710	1.582	NS
			aperture_V2:voix	4793	1	4793	0.203	NS
			résiduelle	3670217	156	23527		
GA	/ɛ/	F1	aperture_V2	27436	1	27436	5.143	<.05
			voix	32715	1	32715	6.132	<.05
			aperture_V2:voix	3089	1	3089	0.579	NS
		F2	résiduelle	357434	67	5335		
			aperture_V2	297728	1	297728	20.145	<.001
			voix	5623	1	5623	0.380	NS
	/ɔ/	F1	aperture_V2:voix	43458	1	43458	2.941	NS
			résiduelle	990205	67	14779		
			aperture_V2	55387	1	55387	20.752	<.001
		F2	voix	205294	1	205294	76.918	<.001
			aperture_V2:voix	301	1	301	0.113	NS
			résiduelle	413697	155	2669		
HU	/ɛ/	F1	aperture_V2	2813	1	2813	0.265	NS
			voix	3401	1	3401	0.320	NS
			aperture_V2:voix	1635	1	1635	0.154	NS
		F2	résiduelle	1648156	155	10633		
			aperture_V2	22074	1	22074	8.276	<.01
			voix	252405	1	252045	94.633	<.001
	/ɔ/	F1	aperture_V2:voix	3	1	3	0.001	NS
			résiduelle	181369	68	2667		
			aperture_V2	801	1	801	0.052	NS
		F2	voix	385149	1	385149	24.984	<.001
			aperture_V2:voix	35601	1	35601	2.309	NS
			résiduelle	1048260	68	15416		
/ɔ/	F1	aperture_V2	29427	1	29427	5.640	<.05	
		voix	770796	1	770796	147.734	<.001	
		aperture_V2:voix	3681	1	3681	0.705	NS	
	F2	résiduelle	813921	156	5217			
		aperture_V2	627	1	627	0.015	NS	
		voix	1261	1	1261	0.028	NS	
PA	/ɛ/	F1	aperture_V2:voix	9	1	9	0.000	NS
			résiduelle	7139033	156	45763		
			aperture_V2	1677	1	1677	0.394	NS
		F2	voix	159236	1	159236	37.428	<.001
			aperture_V2:voix	299	1	299	0.070	NS
			résiduelle	289306	68	4254		
	/ɔ/	F1	aperture_V2	7147	1	7147	0.352	NS
			voix	47483	1	47483	2.355	NS
			aperture_V2:voix	2300	1	2300	0.113	NS
		F2	résiduelle	1382559	68	20332		
			aperture_V2	1002	1	1002	0.134	NS
			voix	628358	1	236358	90.933	<.001
/ɔ/	F1	aperture_V2:voix	26956	1	26956	3.592	NS	
		résiduelle	1170665	156	7504			
		aperture_V2	280303	1	280303	9.562	<.001	
	F2	voix	10693	1	10693	0.365	NS	
		aperture_V2:voix	27907	1	27907	0.952	NS	
		résiduelle	4573142	156	29306			
TR	/ɛ/	F1	aperture_V2	2836	1	2836	0.412	NS
			voix	191787	1	191787	27.828	<.001
			aperture_V2:voix	161	1	161	0.023	NS
		F2	résiduelle	468645	68	6892		
			aperture_V2	1872	1	1872	0.059	NS
			voix	32719	1	32719	1.166	NS
	/ɔ/	F1	aperture_V2:voix	3616	1	3616	0.113	NS
			résiduelle	2170366	68	31917		
			aperture_V2	1968	1	1968	1.951	NS



	voix	2765590	1	2765590	274.203	<.001
	aperture_V2:voix	73	1	73	0.073	NS
	résiduelle	157358	68	1009		
F2	aperture_V2	23867	1	23867	2.473	NS
	voix	9471	1	9571	0.981	NS
	aperture_V2:voix	315	1	315	0.033	NS
	résiduelle	1505343	156	9650		

**Annexe 3.** Résultats des tests d'ANOVA à deux facteurs avec interaction effectués pour les six locuteurs du corpus ; formule utilisée :  $\text{aov}(F\alpha\_V1 \sim \text{aperture\_V2} * \text{voix})$ .

locuteur	V1	variation	SC	dl	CM	F	p
AR	/E/	aperture_V2	10950	1	10950	4.67	<.05
		voix	63704	1	63704	27.19	<.001
		aperture_V2:voix	1942	1	1942	0.82	NS
		résiduelle	454598	194	2343		
	/O/	aperture_V2	13574	1	13574	10.440	<.001
		voix	132345	1	132345	101.795	<.0001
		aperture_V2:voix	19	1	19	0.014	NS
		résiduelle	317225	244	1300		
	/Œ/	aperture_V2	12173	1	12173	7.835	<.01
		voix	68748	1	68748	48.443	<.001
		aperture_V2:voix	556	1	556	0.358	NS
		résiduelle	142925	92	1554		
	/A/	aperture_V2	4301	1	4301	1.010	NS
		voix	67416	1	67416	15.825	<.001
		aperture_V2:voix	8878	1	8878	0.028	NS
		résiduelle	366384	86	4260		
BA	/E/	aperture_V2	42611	1	42611	15.518	<.001
		voix	47278	1	47278	17.218	<.001
		aperture_V2:voix	4	1	4	0.001	NS
		résiduelle	538195	196	2746		
	/O/	aperture_V2	32463	1	32463	3.112	NS
		voix	291673	1	291673	27.963	<.001
		aperture_V2:voix	26	1	26	0.306	NS
		résiduelle	2545107	244	10431		
	/Œ/	aperture_V2	5828	1	5828	2.772	NS
		voix	6801	1	6801	29.404	<.05
		aperture_V2:voix	165	1	165	0.178	NS
		résiduelle	149726	92	1627		
	/A/	aperture_V2	6693	1	6693	3.248	NS
		voix	32710	1	32710	130.157	<.001
		aperture_V2:voix	1769	1	1769	0.858	NS
		résiduelle	249360	121	2061		
GA	/E/	aperture_V2	81207	1	81207	17.864	<.001
		voix	186008	1	186008	40.919	<.001
		aperture_V2:voix	235	1	235	0.579	NS
		résiduelle	877327	193	4546		
	/O/	aperture_V2	41808	1	41808	17.940	<.001
		voix	349829	1	349829	150.115	<.001
		aperture_V2:voix	634	1	634	0.272	NS
		résiduelle	561629	241	2330		
	/Œ/	aperture_V2	9322	1	9322	5.229	<.05
		voix	153920	1	153920	76.918	<.001
		aperture_V2:voix	1820	1	1820	0.113	NS
		résiduelle	164004	92	1783		
	/A/	aperture_V2	1048	1	1048	0.306	NS
		voix	169	1	169	0.049	NS
		aperture_V2:voix	2089	1	2089	0.69	NS
		résiduelle	366661	104	3429		
HU	/E/	aperture_V2	38107	1	38107	16.564	<.001

		voix	634388	1	634388	275.752	<.001
		aperture_V2:voix	3709	1	3709	1.612	NS
		résiduelle	450913	196	2301		
	/O/	aperture_V2	26790	1	26790	6.791	<.01
		voix	1128152	1	1128152	285.975	<.001
		aperture_V2:voix	1557	1	1557	0.395	NS
		résiduelle	962565	244	3945		
	/Œ/	aperture_V2	7090	1	7090	5.640	<.05
		voix	327951	1	327951	314.334	<.001
		aperture_V2:voix	41	1	41	0.040	NS
		résiduelle	95986	92	1043		
	/A/	aperture_V2	13779	1	13779	0.015	NS
		voix	440659	1	440659	0.028	<.001
		aperture_V2:voix	110	1	110	0.000	NS
		résiduelle	414485	113	3668		
PA	/E/	aperture_V2	4	1	4	0.001	NS
		voix	670829	1	670829	211.235	<.001
		aperture_V2:voix	4181	1	4181	1.316	NS
		résiduelle	622447	196	3176		
	/O/	aperture_V2	762	1	762	0.352	NS
		voix	1051835	1	1051835	2.355	<.001
		aperture_V2:voix	12395	1	12395	0.113	NS
		résiduelle	1396777	244	5724		
	/Œ/	aperture_V2	9720	1	9720	0.134	<.05
		voix	455953	1	455953	90.933	<.001
		aperture_V2:voix	171	1	171	3.592	NS
		résiduelle	130005	92	1413		
	/A/	aperture_V2	18711	1	18711	9.562	NS
		voix	360727	1	360727	0.365	<.001
		aperture_V2:voix	3656	1	3656	0.952	NS
		résiduelle	546737	105	5207		
TR	/E/	aperture_V2	824	1	824	0.190	NS
		voix	249077	1	249077	57.561	<.001
		aperture_V2:voix	2700	1	2700	0.624	NS
		résiduelle	848126	196	4327		
	/O/	aperture_V2	1962	1	1962	1.834	NS
		voix	329504	1	329504	1.166	<.001
		aperture_V2:voix	374	1	374	0.113	NS
		résiduelle	260940	244	1069		
	/Œ/	aperture_V2	1176	1	1176	1.034	NS
		voix	2765590	1	2765590	96.134	<.001
		aperture_V2:voix	84	1	84	0.074	NS
		résiduelle	1-4648	92	1137		
	/A/	aperture_V2	99183	1	99183	2.473	<.001
		voix	68277	1	68277	0.981	<.05
		aperture_V2:voix	61300	1	61300	0.033	<.05
		résiduelle	1030700	89	11581		

**Annexe 4.** Résultats des tests d'ANOVA à deux facteurs avec interaction effectués pour les six locuteurs du corpus ; formule utilisée : `aov(F1_V1~aperture_V2*voix)`.

**CORPUS ENREGISTRÉ EN LABORATOIRE**

<b>V<sub>1</sub></b>	<b>V<sub>2</sub></b>	<b>contextes d’HV</b>		
<b>e/ɛ</b>	e/ɛ	<i>épée</i> /epɛ/	<i>épaisse</i> /epɛs/	
		<i>été</i> /etɛ/	<i>éter</i> /etɛːʁ/	
		<i>fléché</i> /fleʃɛ/	<i>fléchette</i> /fleʃɛt/	
		<i>rêvé</i> /ʁɛvɛ/	<i>réveil</i> /ʁɛvɛj/	
		<i>rainé</i> /ʁɛnɛ/	<i>rainette</i> /ʁɛnɛt/	
	i/a	<i>bêtise</i> /betiːz/	<i>bêtasse</i> /betas/	
		<i>éthyle</i> /etil/	<i>étal</i> /etal/	
		<i>réglisse</i> /ʁɛglis/	<i>réglage</i> /ʁɛglaːʒ/	
	ø/œ	<i>régleur</i> /ʁɛglœːʁ/	<i>régleuse</i> /ʁɛgløːz/	
		<i>rêveuse</i> /ʁɛvøːz/	<i>rêveur</i> /ʁɛvœːʁ/	
<b>o/ɔ</b>	ø/œ	<i>fraiseuse</i> /fʁɛzøːz/	<i>fraiseur</i> /fʁɛzœːʁ/	
		<i>épaule</i> /epoːl/	<i>époque</i> /ɛpɔk/	
	e/ɛ	<i>sonner</i> /sɔnɛ/	<i>sonnette</i> /sɔnɛt/	
		<i>sauté</i> /sote/	<i>sauternes</i> /sotɛʁn/	
		<i>noter</i> /nɔtɛ/	<i>notaire</i> /nɔtɛːʁ/	
		<i>potée</i> /pɔtɛ/	<i>poterne</i> /pɔtɛʁn/	
		<i>rosé</i> /ʁoze/	<i>rosette</i> /ʁozɛt/	
	i/a	<i>clocher</i> /kloʃɛ/	<i>clochette</i> /kloʃɛt/	
		<i>mollir</i> /mɔliːʁ/	<i>mollah</i> /mɔla/	
		<i>motif</i> /mɔtif/	<i>motard</i> /mɔtaːʁ/	
ø/œ	<i>notice</i> /nɔtis/	<i>nota</i> /nɔta/		
	<i>colline</i> /kɔlin/	<i>collage</i> /kɔlaːʒ/		
	<i>Rosine</i> /ʁozin/	<i>rosace</i> /ʁozas/		
	<i>sonneuse</i> /sɔnøːz/	<i>sonneur</i> /sɔnœːʁ/		
	<i>causeur</i> /kozœːʁ/	<i>causeuse</i> /kozøːz/		
<b>ø/œ</b>	o/ɔ	<i>auto</i> /oto/	<i>automne</i> /otɔn/	
		<i>mono</i> /mɔno/	<i>monocle</i> /mɔnɔkl/	
	e/ɛ	<i>cueillez</i> /kœje/	<i>cueillette</i> /kœjet/	
		<i>meuler</i> /møle/	<i>meulage</i> /mølaːʒ/	
	i/a	<i>peupler</i> /pœple/	<i>peuplade</i> /pœplad/	
		<i>jeunisme</i> /ʒœnism/	<i>jeunesse</i> /ʒœnes/	
		<i>cueillir</i> /kœjiːʁ/	<i>cueillage</i> /kœjaːʒ/	
		<i>cueilleur</i> /kœjøːʁ/	<i>cueilleuse</i> /kœjøːz/	
	<b>V<sub>1</sub></b>	<b>V<sub>2</sub></b>	<b>mots distracteurs (pseudo-contextes d’HV)</b>	
	<b>a</b>	e/ɛ	<i>pavé</i> /pavɛ/	<i>bavette</i> /bavɛt/
i/a		<i>affine</i> /afin/	<i>diaphane</i> /djafan/	
ø/œ		<i>fameux</i> /famø/	<i>faveur</i> /favœːʁ/	
o/ɔ		<i>capot</i> /kapɔ/	<i>capotent</i> /kapɔt/	
<b>mots monosyllabiques</b>				
<i>vif</i> /vif/	<i>vu</i> /vy/		<i>vous</i> /vu/	
<i>fée</i> /fɛ/	<i>peu</i> /pø/		<i>paume</i> /poːm/	
<i>peine</i> /pɛn/	<i>peuvent</i> /pœːv/		<i>pomme</i> /pɔm/	
		<i>paffe</i> /paf/		

**Annexe 5.** Liste complète de paires d’harmonie vocalique, mots distracteurs et mots monosyllabiques enregistrés en laboratoire.

