

## Que nous apprennent les gros corpus sur l'harmonie vocalique en français ?

Giuseppina TURCO<sup>1</sup>, Cécile FOUGERON<sup>1</sup>, Nicolas AUDIBERT<sup>1</sup>

(1) Université Sorbonne-Nouvelle Paris 3 - Laboratoire de Phonétique et Phonologie (LPP), UMR 7018, 19 rue des Bernardins, Paris, France

{giuseppina.turco; cecile.fougeron; nicolas.audibert}@univ-paris3.fr

### RESUME

---

Afin de mieux identifier le poids relatif des différents facteurs décrits dans la littérature comme influençant le phénomène d'harmonie vocalique (HV) en français, 33k mots extraits de deux corpus de parole continue et présentant un contexte d'HV possible  $V_1C(C)V_2$  ( $V_1 \in \{e, \varepsilon, o, \circ\}$ ) sont analysés. Le degré d'HV est mesuré en termes d'abaissement du F1 de  $V_1$  induit par la présence d'une  $V_2$  /+haut/ (fermée ou mi-fermée) par rapport à une  $V_2$  /-haut/ (ouverte ou mi-ouverte). Les résultats montrent une HV plus importante pour les voyelles moyennes postérieures que pour les antérieures, et plus faible lorsque l'orthographe favorise une prononciation mi-fermée de  $V_1$ . Comme attendu, l'HV est plus forte quand  $V_1$  est séparé de  $V_2$  par une consonne labiale vs. linguale ou par un cluster consonantique sous-jacent vs. un cluster résultant de la chute d'un schwa. En revanche, le style de parole (conversationnelle vs. journalistique) a un effet plus nuancé que celui attendu.

### ABSTRACT

---

#### What can we learn from big speech corpora about French vowel harmony?

This study sets to determine those highly debated - as yet poorly investigated - factors that affect French vowel harmony (VH). The analyses are based on 33k words extracted from two corpora

(journalistic vs. casual speech) with possible  $V_1C(C)V_2$  ( $V_1 \in \{e, \varepsilon, o, \circ\}$ ) harmonic contexts. VH is measured as the lowering of the first formant of the target  $V_1$  in relation to the  $V_2$  trigger /+high/ (i.e. high and mid-high) vs. /-high/ (i.e. low and mid-low)  $V_2$ . Results show that VH is stronger for mid-back target vowels than for mid-high vowels and when the realization of the target vowel is not driven by orthographic interference. As expected, VH is stronger with an intervening labial consonant (compared to a lingual consonant) and with consonant clusters without underlying schwa. Speech style is found to have a more subtle effect than the one described in the literature.

---

**MOTS-CLÉS :** harmonie vocalique, coarticulation, style de parole, gros corpus

**KEYWORDS:** vowel harmony, V-to-V coarticulation, speech style, corpus-based study

---

## 1 Introduction

Les variations de timbre des voyelles moyennes du français répondent à des influences diverses (variantes régionales ou individuelles, contraintes de distribution en fonction de la structure syllabique, analogie au sein des paradigmes, orthographe, position dans le mot, etc.). Ici, nous nous

intéressons à la modification de timbre que peuvent subir ces voyelles moyennes, lorsqu'elles sont en position non-finale, en fonction de l'aperture de la voyelle finale non-adjacente suivante. Par exemple, dans le paradigme *aime/aimait/aimer*, la première voyelle aura tendance à être réalisée mi-ouverte dans les deux premiers cas ([ɛm], [ɛmɛ]) mais mi-fermée dans le dernier [eme] sous l'influence de la voyelle finale. Décrit en phonologie comme un processus optionnel d'harmonie vocalique (HV) ou de métaphonie, avec transfert à distance de la spécification d'aperture de la source à la cible précédente, cette influence non-locale peut aussi être considérée comme un cas de coarticulation de voyelle à voyelle avec chevauchement des gestes articulatoires des deux voyelles co-produites (Ohala 1994, Fagyal et al., 2003, Nguyen et Fagyal 2008). Si la nature du processus et la question de savoir s'il résulte d'une planification ou d'une articulation anticipée de la voyelle finale est importante (cf. Tilsen 2007), nous n'aborderons cette question qu'indirectement dans cette étude. Notre objectif principal est de mieux comprendre ce phénomène, à la fois bien connu mais relativement mal défini dans la littérature et peu décrit empiriquement, en examinant son occurrence dans de gros corpus de parole continue en français. En d'autres termes, nous ne cherchons pas directement à savoir *pourquoi* ce phénomène apparaît mais *où* il apparaît (entre quelles voyelles), et puisqu'il est considéré comme optionnel, *quels facteurs* influencent son occurrence.

Cette étude s'inscrit donc dans la lignée des études de Fagyal, Nguyen, Boula de Mareüil (2003) et Nguyen & Fagyal (2008) qui, sur un corpus produit en laboratoire, ont étudié ce processus dans les productions de six locuteurs du français standard vs. méridional. Leur étude et leurs résultats montrent à quel point un état des lieux est nécessaire pour comprendre ce phénomène, auquel nous référons dans la suite en termes d'harmonie vocalique, qui est largement admis dans la phonologie du français, mais dont la définition n'est pas totalement claire. En effet, si l'HV est mentionnée dans la plupart des précis de prononciation ou descriptions phonologiques du français (Grammont 1926, Fouché 1959, Dell 1973, Tranel 1989, etc.) comme un processus anticipatoire, optionnel, affectant l'aperture de voyelles moyennes en position non-finale (ou non-accentuée) en fonction de l'aperture de la voyelle finale (accentuée) de la syllabe suivante (i.e. non adjacente), les conditions exactes de son application ne sont pas toujours décrites de la même façon (voir aussi la revue de Fagyal et al., 2003). Pour cela, l'examen de gros corpus de parole continue alignés automatiquement nous semble un paradigme approprié pour examiner plus systématiquement le phénomène dans des productions naturelles, incluant des locuteurs et des mots variés. Les tendances que nous arriverons à mettre à jour nous serviront dans un deuxième temps à construire un corpus plus contrôlé.

La première incertitude dans la définition des conditions d'application de l'HV en français concerne *la nature de la cible (V<sub>i</sub>)*, c'est à dire quelles voyelles moyennes peuvent subir l'harmonie dans un contexte /(#)V<sub>i</sub>C(C)V<sub>z</sub>#/. Chez Fouché (1959) par exemple, seules les voyelles moyennes antérieures non-arrondies /e, ɛ/ sont sujettes à l'HV (« l'harmonisation vocalique ne joue pas dans le cas de o inaccentué » p.77). Chez Tranel (1987), il est noté que les voyelles moyennes arrondies /œ/ et /ɔ/ sont moins sensibles à l'HV que la voyelle antérieure /ɛ/. Pour /œ/ et /ɔ/, l'HV n'opèrerait que lorsqu'elles sont respectivement suivies de leur contrepartie mi-fermée (comme dans *peureux* [pøRø] ou *auto* [oto]). Cette restriction n'est pas intégrée dans d'autres descriptions, comme celle de Walker (2001) par exemple. S'il relève que l'HV affecte principalement les voyelles /e, ɛ/, elle lui permet d'expliquer aussi l'alternance au sein des paires /o, ɔ/ et /ø, œ/ comme dans *œuvre/œuvrer* [œvR]/[øvRe], *pleut/pleuvoir* [plø]/[pløvwæR] ou *code/coder* [kød]/[kode], *gros/grossesse* [gRø]/[gRøses].

Cet effet partagé par les voyelles moyennes avant et arrière ressort également dans les données acoustiques de Nguyen & Fagyal (2008) où les F1 des paires /e, ɛ/ et /o, ɔ/ (les /ø, œ/ n'étant pas examinées) sont abaissés par une voyelle finale fermée ou mi-fermée. Il est à noter que les descriptions diffèrent également en ce qui concerne la directionnalité des modifications subies par la cible : pour certains l'HV en français est uniquement un processus 'fermant' par lequel les voyelles

cibles mi-ouvertes se ferment sous l'influence de la voyelle suivante (Fouché 1959, Tranel 1987, Casagrande 1984), alors que pour d'autres, le processus couvre aussi bien l'influence fermante que peuvent subir les mi-ouvertes que l'influence ouvrante subie par les mi-fermées (p.ex. Malmberg 1969). En lien avec cette définition de la cible potentielle de l'HV, certaines descriptions intègrent des facteurs pouvant interagir avec l'application de l'HV, comme par exemple la loi de position ou des contraintes de fidélité ou d'analogie avec la prononciation de la racine dans les dérivés morphologiques (*coder* [kɔde] < *code* [kɔd]). A celles-ci s'ajoutent des critères orthographiques favorisant telle ou telle prononciation : par exemple, les graphies *é* ou *au/eau* favoriseraient une prononciation mi-fermée ([e] et [o] respectivement) indépendamment de l'aperture de la voyelle suivante (p.ex. Tranel 1989).

La seconde incertitude concerne **la nature de la source (V<sub>2</sub>)**, c'est à dire la voyelle finale provoquant l'harmonie. Chez Fouché (1959), seules les voyelles /i, e, y/ ont le pouvoir de fermer la voyelle précédente (seulement /ɛ/ pour lui, ex. *aigre* /ɛgR/ - *aigri* /ɛgRi/). Chez Tranel (1989) la source est définie comme fermée ou mi-fermée mais les exemples qu'il donne n'incluent que des /i, e/ et /ø, o/ (ces dernières influençant que les cibles /œ, ɔ/). Dell (1985), sans préciser un inventaire de V<sub>2</sub>, propose des exemples de type *céder* /sede/ - *cédant* /seda~/, et semble inclure toutes les voyelles en fonction de leur spécification [+/-bas]. Les résultats de Nguyen & Fagyal (2008) montrent effectivement que les voyelles V<sub>2</sub> moyenne et non-moyenne ont le même effet : l'abaissement du F1 de V<sub>1</sub> est provoqué aussi bien par une V<sub>2</sub> fermée /i/ qu'une mi-fermée /e, ø, o/, et l'augmentation du F1 peut être provoquée aussi bien par une V<sub>2</sub> ouverte /a/ que par /ɛ, œ, ɔ/. Toutefois, ils notent que dans leur corpus, les V<sub>2</sub> non-moyennes (/i-a/) ont un effet systématique sur la fréquence du F1 de V<sub>1</sub> qui est partagé par les six locuteurs, que les V<sub>2</sub> /ø, œ/ ont un effet régulier notamment sur les V<sub>1</sub> postérieures (*porteur/porteuse*), mais que les V<sub>2</sub> /e- ε/ n'ont un effet que pour 3 des 6 locuteurs.

Une autre incertitude dans le conditionnement de l'HV concerne **la séquence intervenant entre la cible V<sub>1</sub> et la source V<sub>2</sub>**. La question de l'étendue des processus d'harmonie dans les langues a soulevé de nombreux débats dans la littérature, en relation avec la question de savoir si l'HV cible d'une manière séquentielle des segments linéairement ordonnés ou si elle se propage non-linéairement sur des segments partageant la même tier (p.ex. Gafos 1996). Dans la littérature phonétique, la question de l'empan coarticulatoire de voyelle à voyelle a été aussi largement débattue (p.ex. Fowler et Brancazio 2000). Pour l'HV en français, les descriptions sur ce point ne sont pas non plus uniformes. Pour Dell (1985), l'HV nécessite une frontière morphologique entre V<sub>1</sub> et V<sub>2</sub>, alors que Nguyen et Faygal (2008) montrent que l'HV opère au travers de celles-ci. Fouché (1959) note que la suite 'rr', un groupe consonantique ou la présence d'un schwa sous-jacent (prononcé ou non) entre V<sub>1</sub> et V<sub>2</sub> bloque l'HV. Dell (1973) au contraire propose une règle d'HV tardive qui s'appliquerait après la déletion du schwa et permettrait donc l'HV (p.ex. *aiderez* [edRe]).

Les nombreuses descriptions que nous avons étudiées convergent toutefois sur un point : l'HV est décrite comme un phénomène optionnel dont l'occurrence est **liée au style de parole**. Elle aurait tendance à être plus fréquente en parole conversationnelle (Fouché 1959), spontanée (Tranel 1987) et peu formelle (Walker 2001). Pour autant, cet aspect n'a jamais été vérifié empiriquement à notre connaissance. Nous examinerons donc le processus d'HV dans deux corpus incluant des styles de parole différents : journalistique, que l'on suppose plus soutenu, et conversationnel, plus relâché.

## 2 Méthodologie

### 2.1 Corpus et prétraitements

Deux corpus de grande taille correspondant à différents styles de parole en français standard ont été analysés. Le corpus ESTER (Galliano et al., 2009) comporte des interventions publiques partiellement pré-écrites, extraites de diverses émissions de radio et de TV (bulletins d'information et débats politiques ou de société). La majeure partie de ces données est produite par des locuteurs professionnels avec un degré de contrôle relativement important. Le corpus NCCFr (Torreira et al., 2010) est composé de conversations entre amis sur des sujets de société produites dans une interaction face-à-face informelle. Ces deux corpus ont été alignés automatiquement en mots et phonèmes à partir de transcriptions orthographiques ainsi que segmentés en locuteurs.

A partir d'une liste de mots produits dans ces corpus, nous avons identifié les mots présentant un contexte d'HV potentiel, c'est à dire ayant une  $V_2 \in \{i, e, \varepsilon, o, a, y, u, \tilde{a}\}$  en syllabe finale et en syllabe penult une  $V_1 \in \{e, \varepsilon, o, \text{ɔ}\}$ . Aucune contrainte sur la structure syllabique du mot, sa longueur ou sur la séquence de consonnes séparant  $V_1$  et  $V_2$  n'a été apportée. Ainsi, des formes comme *aidez* alignées sans schwa [ $V_1dRV_2$ ] sont incluses dans l'analyse. 33 325 mots (corpus NCCFr : 14k mots, ESTER : 19k mots) ont ainsi été sélectionnés.

Les voyelles moyennes  $V_1$  ont été classifiées selon leur lieu d'articulation (et arrondissement, variable *Antériorité* $V_1$ ) avec /e,  $\varepsilon$ / antérieures et /o,  $\text{ɔ}$ / postérieures. Ces voyelles ont été également codées (variable *Ortho* $V_1$ ) selon que leur graphie favorise ou non une prononciation mi-fermée : graphie *é* pour [e] et *au/eau* pour [o]. Les voyelles finales  $V_2$  ont été catégorisées en deux degrés d'aperture (variable *Aperture* $V_2$ ), [+haut] (fermée ou mi-fermée /i, y, e, u, o/) et [-haut] (mi-ouverte ou ouverte / $\varepsilon$ , a,  $\tilde{a}$ ). Etant donné la variabilité des prononciations des  $V_2/e/-\varepsilon/$ , nous sommes partis des étiquettes phonétiques attribuées par l'alignement forcé (sensés refléter la prononciation mais dépendants des variantes de prononciation prises en compte par le système) que nous avons harmonisées en combinant des informations orthographiques ( $V_2/e/$  pour les graphies *é/és/ée/ées/ez/er* ;  $V_2/\varepsilon/$  pour les graphies *è/ais/ait/aient* et les quelques cas de substantif en *er* comme *laser* ou *amer*) et la transcription phonologique de référence extraite de la base Lexique 3 (New et al., 2007). Les catégories vocaliques sont donc des classes phonémiques indépendantes de leur prononciation dans les corpus. Les voyelles / $\emptyset$ ,  $\text{œ}$ /, dont les relations graphie-phonie sont plus complexes et qui peuvent être confondues avec un schwa, ont été exclues pour cette première étude.

Concernant la séquence de consonne présente entre  $V_1$  et  $V_2$ , nous avons distingué plusieurs cas présentant suffisamment d'occurrences dans les différentes catégories définies plus haut. Dans les cas où  $V_1$  et  $V_2$  sont séparées par une seule consonne, nous avons codé (variable *TypeC*) si cette consonne était labiale (*économie*) ou linguale (*échappe*, *écho*). Dans les cas où  $V_1$  et  $V_2$  sont séparées par plusieurs consonnes, nous avons codé dans la variable *SchwaIntervocalique* si la suite de consonne était lexicale, ex. *maîtrise*, *esprit* ou si elle résultait de la chute d'un schwa (c'est à dire les cas où le système d'alignement considère que le schwa n'est pas produit), ex. *laissez* [lesRe], *souhaiterez* [swetRe].

### 2.2 Analyse statistique

Les relations entre les valeurs de F1 en Hertz et les effets fixes listés en section 3 ont été testées au moyen de modèles linéaires mixtes (bibliothèque 'lme4' de R, Bates et al., 2014). L'intercept pour

les locuteurs a également été intégré au modèle comme effet aléatoire. De plus, pour éviter un taux élevé d'erreur de type I, les pentes aléatoires par locuteur ont été incluses pour chaque effet fixe, correspondant à la variabilité inter-locuteur de l'effet de chaque facteur fixe sur F1. Les effets aléatoires incluent les pentes pour tous les effets principaux, mais pas pour leurs interactions car cela ne permet pas au modèle de converger. Les valeurs de p (tests de rapport de vraisemblance) sont obtenues par des approximations de type *Satterthwaite* à l'aide de la fonction 'lmerTest', qui permet d'obtenir des estimations plus fiables que les méthodes équivalentes pour les régressions linéaires. Les valeurs de  $R^2$  associées à chaque modèle ont été obtenues à l'aide de la fonction *r.squaredGLMM* intégrée dans la bibliothèque 'MuMIn'.

### 3 Résultats et discussion

Dans un premier temps, nous avons modélisé les valeurs de F1 en fonction des variables *ApertureV<sub>2</sub>* ([+/-haut]), *AntérioritéV<sub>1</sub>* (antérieure/postérieure), *OrthoV<sub>1</sub>* (oui/non) et *Corpus* (journalistique/conversationnel), selon les principes généraux de construction du modèle présentés ci-dessus. L'analyse indique un  $R^2$  marginal de 0.09 et un  $R^2$  conditionnel de 0.36, et met en évidence un effet significatif de chacun des quatre prédicteurs testés ( $p < 0.0001$  dans les quatre cas). Ainsi, le F1 de  $V_1$  est significativement plus élevé lorsque  $V_2$  est une voyelle basse, lorsque  $V_1$  appartient à la catégorie des voyelles postérieures (/o, ɔ/), lorsque la prononciation de  $V_1$  n'est pas influencée par des critères graphiques favorisant une prononciation mi-fermée, et lorsque le mot est produit en parole journalistique comparativement à la parole conversationnelle. Afin de déterminer quels facteurs interviennent dans l'HV, nous sommes tout particulièrement intéressés aux prédicteurs interagissant avec le facteur *ApertureV<sub>2</sub>*. L'HV étant mesurée comme l'abaissement du F1 de  $V_1$  induit par la présence d'une  $V_2$  /+haut/ (fermée ou mi-fermée) par rapport à une  $V_2$  /-haut/ (ouverte ou mi-ouverte), ces interactions nous permettent de comparer le degré d'HV associé aux différentes valeurs de ces prédicteurs.

Concernant le lieu d'articulation de  $V_1$ , nous observons une interaction significative entre *AntérioritéV<sub>1</sub>* et *ApertureV<sub>2</sub>* ( $p < 0.0001$ ). Ces interactions, ainsi que la quantification de l'HV pour les deux types de  $V_1$  considérés, sont illustrées par la Figure 1. Le degré d'HV, i.e. l'effet de  $V_2$  sur le F1 de  $V_1$ , est plus important pour les voyelles postérieures arrondies /o, ɔ/ que pour les voyelles antérieures /e, ε/. Ce résultat est en désaccord avec les descriptions classiques présentées dans la littérature (Fouché, 1959 ; Tranel, 1987). D'une part, nous observons que le timbre des voyelles postérieures moyennes est aussi affecté par l'aperture de la voyelle suivante, contra Fouché (1959) mais conformément aux résultats empiriques de Nguyen et Fagyal (2008). D'autre part, contra Tranel (1987) nous montrons que celles-ci sont plus sensibles à l'HV, avec une HV supérieure pour les voyelles moyennes postérieures vs. antérieures. Nguyen et Fagyal (2008) n'ont pas fait de comparaison directe entre  $V_1$  postérieure et antérieure, mais si pour les deux ils observent un effet de  $V_2$  sur  $V_1$ , ils relèvent un effet différent sur F2 : le F2 de /e-ε/ augmente alors que celui de /o-ɔ/ s'abaisse sous l'influence d'une  $V_2$  [+haute].

Notre analyse met également en évidence une interaction significative entre *ApertureV<sub>2</sub>* et *OrthoV<sub>1</sub>* ( $p < 0.0001$ ), montrant un degré d'HV plus important dans les formes où la graphie de  $V_1$  ne l'empêche pas d'opérer, c'est à dire quand la graphie favorise une prononciation mi-fermée. On remarque que les graphies *é/au/eau* favorisent effectivement une prononciation plus fermée de  $V_1$ , avec un F1 plus bas, indépendamment de la voyelle suivante.

Ce conditionnement orthographique de la prononciation de  $V_1$  ressort également de façon surprenante dans une interaction significative entre *OrthoV<sub>1</sub>* et *Corpus* ( $p < 0.0001$ ) montrant que la

différence de F1 entre une  $V_1$  marquée orthographiquement ou non est plus grande en parole conversationnelle par rapport à la parole journalistique. Ceci va à l'encontre de l'hypothèse selon laquelle les modulations des représentations phonologiques induites par l'orthographe sont plus fortes dans un style de parole plus formel (Taft & Hambly, 1985). Si on considère que les locuteurs professionnels du corpus ESTER lisent en partie des contenus scriptés, on pourrait s'attendre à une plus forte influence orthographique dans ce corpus et chez ces locuteurs ayant probablement une plus grande maîtrise de l'écrit que les jeunes étudiants du corpus NCCFr (cf. Olson 1996).

Concernant l'effet du style de parole sur l'HV, ici encore les résultats sont surprenants. Une interaction significative entre  $ApertureV_2$  et  $Corpus$  ( $p < 0.0001$ ) fait ressortir que l'effet de  $V_2$  sur le F1 de  $V_1$  est globalement plus important en parole journalistique qu'en parole conversationnelle. Ceci va à l'encontre des descriptions antérieures pour lesquelles l'HV serait plus marquée en parole peu formelle. Toutefois, nous remarquons également que ces deux facteurs sont impliqués dans une interaction à trois niveaux avec le facteur  $OrthoV_1$  ( $ApertureV_2 * OrthoV_1 * Corpus$ ,  $p < 0.0001$ ). Comme illustré Figure 2, l'effet du corpus sur l'HV interagit avec les contraintes orthographiques : si l'on ne considère que les voyelles  $V_1$  non marquées orthographiquement, l'HV ressort comme plus importante dans le corpus de parole conversationnelle par rapport au corpus journalistique.

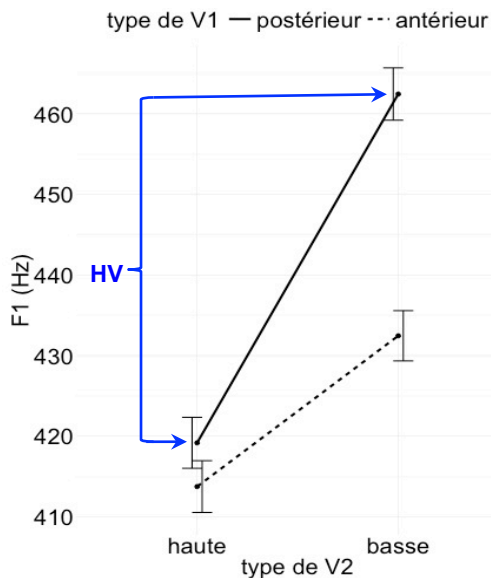


FIGURE 1 : F1 de  $V_1$  (ajustées par le modèle, moyenne et erreur type) en fonction de l'aperture de  $V_2$  et de l'antériorité de  $V_1$ . En bleu, la différence de F1 entre un contexte  $V_2$  haut vs. bas illustre la mesure du degré d'HV pour les  $V_1$  postérieures.

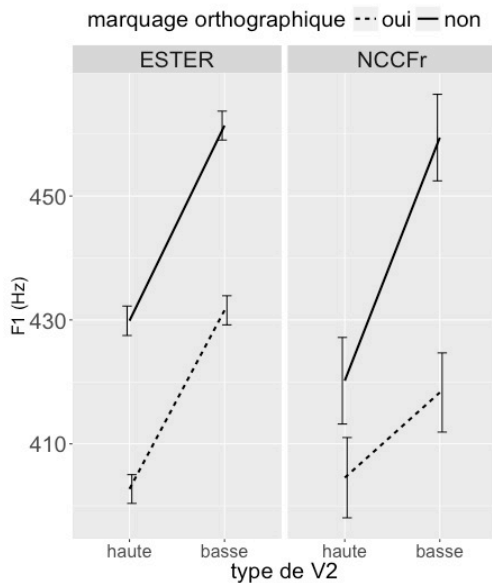


FIGURE 2 : F1 de  $V_1$  (ajustées par le modèle, moyenne et erreur type) en fonction de l'aperture de  $V_2$ , du corpus et des interférences orthographiques.

Afin de tester des hypothèses plus spécifiques sur les contextes susceptibles de favoriser ou non l'HV, nous avons construit deux modèles supplémentaires pour évaluer les liens entre l'HV et les éléments intervocaliques : à savoir (a) si l'articulation de la consonne, lorsqu'elle est seule, laisse la langue libre d'anticiper les mouvements linguaux entre les voyelles (consonne labiale) par rapport à une consonne nécessitant une articulation linguale, et (b) s'il existe entre les deux voyelles un autre

élément vocalique (un schwa) sous-jacent même s'il n'est pas prononcé. La variation relative à l'antériorité/postériorité de  $V_1$  a également été prise en considération dans ces modèles en incluant *Antériorité* $V_1$  comme facteur aléatoire, seules les voyelles  $V_1$  non marquées orthographiquement étant incluses dans l'analyse. Notons toutefois que pour ces deux modèles, nous obtenons des résultats très similaires en termes d'effets des facteurs et interactions pour les voyelles  $V_1$  marquées orthographiquement comme mi-fermées.

Nous avons tout d'abord testé l'influence du type de consonne intervocalique sur l'HV lorsque  $V_1$  et  $V_2$  ne sont séparés que par une consonne. Pour cela, nous avons modélisé les valeurs de F1 en fonction des variables *Aperture* $V_2$  et *TypeC* (labiale ou linguale). L'analyse indique un  $R^2$  marginal de 0.02 et un  $R^2$  conditionnel de 0.38, et ne montre pas d'effet significatif d'*Aperture* $V_2$  ( $p=0.18$ ) ni de *TypeC* ( $p=0.63$ ). En revanche, elle montre une interaction significative entre les deux facteurs ( $p<0.0001$ ). Comme le montre la Figure 3, l'HV est plus importante quand  $V_1$  et  $V_2$  sont séparées par une consonne labiale plutôt qu'une linguale : la différence de F1 entre  $V_2$  [-haute] et [+haute] est plus élevée dans le cas d'une consonne labiale. Ce résultat confirme sur de plus grands corpus les travaux bien connus sur les liens entre coarticulation voyelle-à-voyelle et résistance à la coarticulation (cf. par exemple Recasens, 1985 ou encore Fowler & Brancazio, 2000). Ces études montrent que la coarticulation voyelle-à-voyelle est favorisée au travers d'une consonne 'faiblement résistante' à la coarticulation telle qu'une labiale qui n'implique pas de constriction linguale.

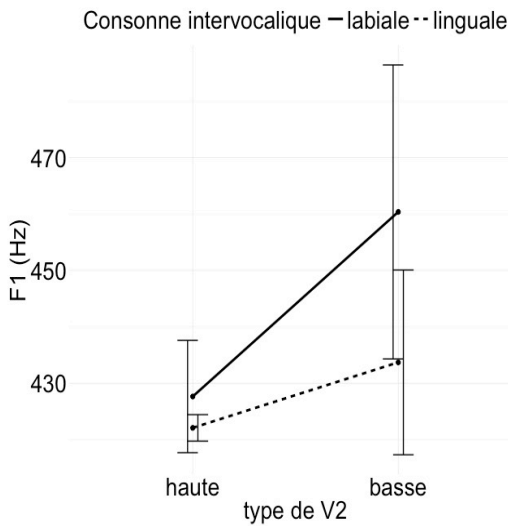


FIGURE 3 : F1 de  $V_1$  (ajustées par le modèle, moyenne et erreur type) en fonction l'aperture de  $V_2$  et du type de consonne intervocalique (labiale vs. linguale).

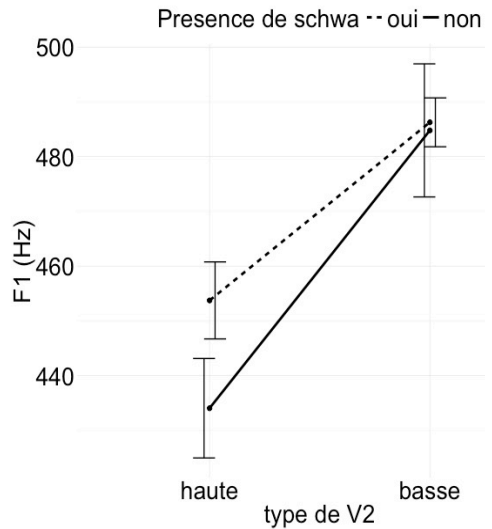


FIGURE 4 : F1 de  $V_1$  (ajustées par le modèle, moyenne et erreur type) en fonction l'aperture de  $V_2$  et de la présence ou non d'un schwa sous-jacent.

Enfin, notre dernière analyse permet de clarifier le rôle controversé d'un schwa intervocalique non-prononcé dans l'HV, en modélisant F1 en fonction des variables *Aperture* $V_2$  et *SchwaIntervocalique*. Pour cela, nous n'avons pris en considération que les cas dans lesquels  $V_1$  et  $V_2$  sont séparées par deux consonnes ou plus, soit les cas exclus dans le modèle précédent. L'analyse indique un  $R^2$  marginal de 0.06 et un  $R^2$  conditionnel de 0.34, et montre un effet significatif d'*Aperture* $V_2$  ( $p<0.0001$ ) mais pas d'effet significatif de *SchwaIntervocalique* ( $p=0.15$ ).

Elle montre également une interaction significative entre *ApertureV<sub>2</sub>* et *SchwaIntervocalique* ( $p < 0.05$ ), illustrée par la Figure 4. On remarque que les modulations de F1 en fonction de  $V_2$  sont plus importantes quand les deux voyelles ne sont pas séparées par un schwa sous-jacent. On remarque également que quand  $V_2$  est [-haute], le timbre de  $V_1$  semble mi-ouvert dans les deux conditions et que ce n'est quand  $V_2$  est [+haute] que l'abaissement du F1 de  $V_1$  semble freiné par la présence d'un schwa sous-jacent.

Plusieurs interprétations de ce résultat sont possibles. La première serait que la catégorie 'schwa présent' comprenne des productions dans lesquelles le schwa a vraiment été prononcé sans être détecté par l'alignement automatique, et donc ne seraient pas des contextes d'HV. La seconde serait que la présence d'une cible vocalique sous-jacente, même non réalisée, interfère dans cette interaction entre  $V_1$  et  $V_2$ , en étant opaque à l'harmonie. Pourtant, dans nos données l'HV, même si elle est réduite, ne semble pas absente dans les cas avec schwa sous-jacent (contra Fouché, 1959). Un examen plus détaillé des mots inclus dans cette catégorie sera nécessaire pour mieux comprendre cette interaction.

## 4 Conclusion

L'objectif de cette étude était de mieux décrire les contextes segmentaux, orthographiques et les effets de style, susceptibles d'après la littérature d'interagir avec l'HV en français. Les variations de timbre (déterminée par leur F1) des voyelles moyennes {e, ε, o, ɔ} ont été examinées en fonction de l'aperture de la voyelle finale suivante, ouverte/mi-ouverte vs. fermée/mi-fermée dans deux gros corpus de parole journalistique et conversationnelle. Le F1 des voyelles cibles varie en fonction des voyelles suivantes (plus bas devant  $V_2$  [+haute] et plus haut devant  $V_2$  [-haute]) et donc l'HV paraît clairement opérer dans nos données. Cet effet est modulé par le lieu d'articulation de  $V_1$  (plus fort pour /o-ɔ/) et peut être atténué par des interférences orthographiques et la séquence intervenant entre la source et la cible.

Nos prochains travaux testeront également dans quelle mesure le degré d'HV change en fonction de la prééminence prosodique, et plus spécifiquement si les voyelles cibles situées en position prosodique forte sont plus résistantes à l'HV. Ceci nous permettra d'approfondir notre compréhension de la relation entre coarticulation voyelle-à-voyelle et HV, et à un niveau plus général sur les liens entre manifestations phonétiques et processus phonologiques.

Enfin, au-delà de la nécessité déjà soulignée par Nguyen et al. (2004) d'établir la pertinence de l'HV en français sur le plan perceptif, les études que nous prévoyons de mener sur cette thématique exploreront la capacité d'auditeurs francophones à exploiter la coarticulation vocalique anticipatoire pour améliorer les performances de l'accès lexical (Tobin et al., 2010).

## Remerciements

Ce travail a été financé par le programme de recherche et d'innovation de l'Union Européenne Horizon 2020, à travers la bourse Marie Skłodowska-Curie n°662530 accordée au premier auteur et par le programme "Investissements d'Avenir" géré par l'Agence Nationale de la Recherche ANR-10-LABX-0083 (Labex EFL).



## Références

- BATES, D. MAECHLER, M., BOLKER, B. and WALKER, S. (2014). lme4: Linear mixed-effects models using eigen and s4 . R package version 1.1-7. <http://CRAN.Rproject.org/package=lme4>.
- DELL, F. (1972). *Les règles et les sons: Introduction à la phonologie générative*. Paris: Hermann
- FAGYAL, Z., NGUYEN, N., BOULA DE MAREÛIL, P. (2003). From dilation to coarticulation: is there vowel harmony in French? *Studies in Linguistic Sciences*, 32, 1-21.
- FOUCHÉ, P. (1959). *Traité de prononciation française*. Paris: Klincksieck
- FOWLER C. A. et BRANCAZIO, L. (2000) Coarticulation resistance of American English consonants and its effects on transconsonantal vowel-to-vowel coarticulation. *Language and Speech*, 43(1), 1-41.
- GAFOS, I. A. (1996). *The articulatory basis of locality in phonology*. (Doctoral dissertation), Johns Hopkins University [Published 1999, New York, NY: Garland].
- GALLIANO, S., GRAVIER, G., CHAUBARD, L. 2009. The ESTER 2 evaluation campaign for the rich transcription of French broadcasts. *Proceedings of Interspeech*, Brighton (UK), 2583–2586.
- GRAMMONT, M. (1926). *La prononciation française*. Delagrave.
- MALMBERG B. (1969) *Phonétique Française*, Malmô, Hermods
- NGUYEN, N., FAGYAL, Z., and COLE, J. (2004). Perceptual relevance of long-domain phonetic dependencies. *Proceedings of the IVth Linguistic Studies Workshop*, 173–178, Nantes, France.
- NGUYEN, N., FAGYAL, Z. (2008). Acoustic aspects of vowel harmony in French. *Journal of Phonetics*, 36, 1-27.
- NEW, M. BRYLSBAERT M., VERONIS J., PALLIER C. (2007) *The use of film subtitles to estimate word frequencies*, *Applied psycholinguistics*, 28(04), 661–677.
- OHALA, J. (1994). Towards a universal, phonetically-based theory of vowel harmony. *Proceedings of the ICSLP*, Yokohama. 10-14.11.-10-14.14
- OLSON, D. R. (1996) Towards a psychology of literacy: On the relations between speech and writing. *Cognition*, 60, 83–104.
- RECASENS, D. (1985). Coarticulatory patterns and degree of coarticulation resistance in Catalan CV sequences. *Language and Speech*, 28, 97-114.
- TAFT, M., HAMBLY, G. (1985). The influence of orthography on phonological representations in the lexicon. *Journal of Memory and Language*, 24(3), 320-335.
- TILSEN, S. (2007). Vowel-to-vowel coarticulation and dissimilation in phonemic-response priming. UK Berkeley Phonology Lab Annual report. 416-458
- TOBIN, S. J., CHO, P. W., JENNETT, P. M., Magnuson, J. S. (2010). Effects of anticipatory coarticulation on lexical access. *Proc. Mtgs. Cognitive Sci*, 2200-2205.
- TRANDEL, B. (1987). *The Sounds of French: An Introduction*. Cambridge: CUP.
- TORREIRA F., ADDA-DECKER M., ERNESTUS, M. (2010). The Nijmegen Corpus of Casual French. *Speech Communication*, 52, 201-221.
- WALKER, D. C. (2001). *French Sound Structure*. Calgary: University of Calgary Press.