

ARTICLE

# *Vot' artic' est formidab'*: une étude multifactorielle de la chute des liquides post-obstruantes finales de mot en français

Mathieu Avanzi 

Université de Neuchâtel

Email: [mathieu.avanzi@unine.ch](mailto:mathieu.avanzi@unine.ch)

(Received 25 August 2021; revised 06 January 2023; accepted 25 January 2023; first published online 16 March 2023)

## Résumé

Cet article présente une analyse multifactorielle du comportement des consonnes liquides post-obstruantes en position finale de mot en français. Plus précisément, nous nous intéressons au phénomène par lequel un <l> ou <r> final dans la transcription orthographique d'un mot n'est pas prononcé dans la parole continue, p. ex. comme dans les items « découvre » et « terrible », respectivement prononcés [dekuv] et [teʁvɛb] au lieu de [dekuvɛʁ] et [teʁvɛibl]. Au total, plus de 2500 items comportant un cluster obstruante+liquide en fin de mot ont été extraits d'un corpus de parole d'une durée de 13 heures, qui a été étiqueté à différents niveaux. Ce corpus comprend les productions de 120 locuteurs originaires de 3 pays francophones (Belgique, France et Suisse), enregistrés dans deux tâches différentes (lecture et conversation). Pour déterminer ce qui affecte la (non-)prononciation de <l> ou <r> dans ces contextes, une douzaine de prédicteurs de différentes natures sont testés dans un même modèle statistique. Les caractéristiques phonétiques telles que le lieu et le mode d'articulation, le contexte droit, le statut prosodique et le taux d'articulation; mais aussi les prédicteurs liés aux locuteurs (pays d'origine, sexe, âge, classe sociale) et au style de parole sont pris en compte. Des modèles mixtes linéaires généralisés révèlent que seulement la moitié d'entre eux ressortent comme jouant un rôle significatif sur la variable à l'étude. Les résultats sont discutés à la lumière d'études antérieures portant sur les aspects sociolinguistiques des variantes de prononciation en français.

**Mots-clés:** liquide finale post-obstruante; prononciation; variation; sociophonétique; corpus

## 1. INTRODUCTION

En français, les groupes consonantiques constitués d'une consonne obstruante (/p, b, t, d, k, g, f, v/) suivie d'une liquide (/l, ʁ/) en position finale de mot, comme dans les mots *découvre* [de.kuvɛʁ] ou *terrible* [te.ʁvɛibl], ont un statut phonologique particulier. Sur l'échelle de sonorité, les liquides sont en effet situées entre les

voyelles et les obstruantes. De fait, si une liquide est contenue dans la même syllabe que l'obstruante qu'elle suit, le principe de sonorité séquentielle (*Sonority Sequencing Principle*, v. Selkirk, 1984), qui stipule qu'une syllabe doit présenter une courbe croissante puis décroissante de sonorité, est violé (Dell, 1985; Côté, 2004). En pratique, il existe trois stratégies pour que la courbe de sonorité idéale soit conservée lors de la prononciation de mots français comme *découvre* ou *terrible* dans la parole continue. La première n'est applicable que si l'item cible est suivi d'un mot à initiale vocalique : le cluster de consonnes final est alors resyllabé avec la voyelle initiale du mot suivant ([te.ʁi.bl.e.pʁœv] > [te.ʁi.bl.e.pʁœv] pour *terrible épreuve*). La seconde stratégie consiste en la réalisation d'un schwa après la liquide, ce qui a pour conséquence, comme pour la première stratégie, que le groupe obstruante + liquide passe de la position de coda ([de.kuvʁ], [te.ʁi.bl]) à la position d'attaque ([de.ku.vʁə], [te.ʁi.blə]). La troisième stratégie consiste en l'élision de la consonne liquide, ce qui n'entraîne aucune restructuration syllabique et permet du même coup que le pic de sonorité ne remonte pas sur l'obstruante après avoir décré ([de.kuv], [te.ʁi.bl]). C'est à la découverte des facteurs qui déclenchent l'application de cette dernière stratégie par les locuteurs du français d'Europe que le présent article est consacré.

L'articulation d'un schwa dans ces contextes entraînant quasi-obligatoirement l'articulation de la liquide,<sup>1</sup> on peut facilement comprendre pourquoi dans les recherches antérieures, le comportement des liquides finales post-obstruantes (désormais LFPO) a surtout été abordé sous l'angle de l'application de la troisième stratégie. D'un point de vue sociolinguistique, ce n'est pas la prononciation de la liquide qui est un phénomène saillant, mais bien sa non-prononciation. Comme l'articulation d'une nasale alvéolaire (/n/) au lieu d'une nasale vélaire (/ŋ/) dans les clusters se terminant par *-ing* en anglais (Labov, 1966; Trudgill, 1974), phénomène auquel la chute des LFPO a été comparée (Armstrong, 2001; Boughton, 2015), le phénomène a en effet longtemps été perçu comme un marqueur propre au français 'populaire' (Guiraud, 1965 : 101) ou 'ordinaire' (Gadet, 1989 : 102),<sup>2</sup> surtout quand le mot qui suit commence par une voyelle (de Cornulier, 1978; Tranel, 1987). Partant, rien d'étonnant à ce qu'il constitue une variable de premier choix sur laquelle les auditeurs s'appuient pour juger du caractère plus ou moins naturel ou approprié d'extraits de parole, qu'ils soient produits par un synthétiseur (Brognaux et Drugman, 2014) ou par un humain (dans une tâche psycholinguistique d'identification lexicale p. ex., Peperkamp et Iturralde Zurita, 2019).<sup>3</sup>

<sup>1</sup>Laks (1977 : 113) donne des occurrences de schwas réalisés devant voyelles (/ɑ̃tʁəø/ pour *entre eux*), mais de tels phénomènes sont rares. P. ex. Milne (2014) évalue la production d'un schwa dans ce genre de contexte à 12,5% (contre 36,1% devant consonne et 30,3% devant pause). Chez Myers et Ranson (2014 : 1344), 19% des LFPO suivies d'une voyelle sont assorties d'un schwa (contre 49% devant consonne et 57% devant pause). Wu (2018 : 143) ne détaille pas le pourcentage de schwas prononcés alors que la liquide est éliée, mais écrit que « lorsque le /ʁ/ est absent, [elle n'observe] presque jamais d'insertion d'un schwa épenthétique ».

<sup>2</sup>Il n'en a pourtant pas toujours été de même à travers les siècles (v. Thurot, 1883; Martinon, 1913; Pustka, 2007 : 203 et Boughton, 2015 : 3 pour un aperçu historique sur la prononciation des LFPO).

<sup>3</sup>V. toutefois Léon et Tennant (1990) et Hansen (2015), qui livrent des résultats nuancés le caractère sociolinguistiquement marqué de cette absence en français parisien.

De façon générale, si les facteurs linguistiques et extralinguistiques qui motivent l'élosion des LFPO sont relativement bien étudiés, les mécanismes qui régissent leurs interactions demeurent encore largement méconnus à l'heure actuelle (§2).<sup>4</sup> Dans cet article, nous proposons de faire avancer nos connaissances en regard de ce problème, en reportant les résultats d'une analyse conduite à partir de l'analyse de plus de 2500 items extraits d'un corpus de 13 heures, annoté semi-automatiquement à différents niveaux linguistiques, et échantillonné de manière à tenir compte de la variation diatopique et diaphasique en français européen (§3). L'analyse des données est réalisée en deux temps. En ce qui concerne les mots grammaticaux, peu nombreux, les statistiques sont essentiellement descriptives. En ce qui concerne les mots lexicaux, des modèles linéaires généralisés à effets mixtes, qui permettent de considérer plusieurs variables dans un même modèle tout en tenant compte de la variation intra- et interlocuteur, ont été réalisés. Et il ressort que sur plus d'une dizaine de prédicteurs testés, seule la moitié d'entre eux s'avèrent jouer effectivement un rôle sur le comportement des LFPO (§4). Ces résultats sont discutés à la lumière d'études antérieures sur la variation sociophonétique en français (§5).

## 2. ÉTAT DE LA QUESTION

Si l'on considère l'ensemble des études sur corpus relatives aux LFPO en français conduites depuis le début des années 80, on peut estimer qu'en moyenne, un mot sur deux contenant une LFPO est prononcé sans sa liquide en parole continue :

Considérés en l'état, les chiffres reportés dans le Tableau 1 ne veulent toutefois pas dire grand-chose, tant il est vrai que le taux d'élosion des LFPO varie drastiquement selon les différents facteurs (extra)linguistiques que les uns et les autres prennent (ou non) en compte lors de leurs calculs. En nous basant sur les travaux susmentionnés et d'autres, nous en avons isolé six grands types.

### 2.1. Facteurs articulatoires

Sur le plan articulatoire, la (non-)réalisation d'une LFPO a été évaluée de façons différentes. D'aucuns se sont concentrés sur le mode d'articulation de la liquide (/l/ vs /ʎ/) : selon Armstrong (1993, 2001) et Pooley (1996), les clusters contenant un /ʎ/ seraient plus souvent simplifiés que les clusters comprenant un /l/; selon Pustka (2007, 2011) et Villeneuve (2010), ce trait n'aurait aucune incidence sur le comportement de la liquide. D'autres ont examiné le mode et le lieu d'articulation de l'obstruante. D'après les observations de Pooley (1996), Pustka (2007, 2011) et Villeneuve (2010), les consonnes fricatives (/f/ et /v/) entraînent plutôt la non-prononciation de la LFPO, alors que les occlusives ont plutôt tendance à en favoriser la prononciation. Finalement, un troisième groupe de chercheurs ont abordé la question en tenant compte des combinaisons possibles entre obstruantes et liquides. Ainsi, Pooley (1996), Hornsby (2006),

<sup>4</sup>En revanche, ce travail a été amorcé pour le schwa en position non-finale de mot (v. Bürki *et al.*, 2011) ainsi que pour la liaison (v. Durand *et al.*, 2011; Adda-Decker *et al.*, 2012; Eychenne et Laks, 2017).

**Tableau 1.** Inventaire des études portant sur l'élosion des LFPO en français, avec origine des locuteurs, nombre d'observations et pourcentage d'élosion.<sup>5</sup>

Référence	Origine des locuteurs	N	Élosion
Kemp et al. (1980 : 18)	Montréal	–	72%
Laks (Laks 1977 : 115)	Paris	–	76,7%
Pooley (1996 : 141)	Lille (Nord-Pas-de-Calais)	2.782	71%
Armstrong (1993, repris dans Armstrong 2001 : 67)	Dieuze (Lorraine)	1365	65,7%
Hambye (2005 : 341)	Wallonie	525	46,1%
Hornsby (2006 : 78)	Avion (Nord-Pas-de-Calais)	654	80%
Ranson (2010 : 1344)	Sud de la France	775	19%
Villeneuve (2010 : 137)	Vimeu (Picardie)	821	58%
Pustka (2007 : 193, repris dans Pustka 2011 : 26)	Guadeloupe, Aveyron (Occitanie) et Paris	4.314	41,8%
Myers et Ranson (2014 : 1349)	France septentrionale et méridionale	730	29,5%
Milne (2014 : 75, 79)	Québec et France septentrionale	15399	51,2%
Boughton (2015 : 141)	Rennes et Nancy (France septentrionale)	1494	56,9%
Wu (2018 : 141)	Île-de-France	32114	28,8%
Peperkamp et al. (2019)	Île-de-France	1205	58%
Brand et Ernestus (2021 : 256, 259)	Île-de-France	291	42,9%

Pustka (2007, 2011) de même que Myers et Ranson (2014) observent des différences selon les combinaisons : dans les clusters <gl> et <kr>, les liquides seraient plus souvent articulées que dans les clusters <fr> ou <pl>, notamment.

## 2.2. Facteurs phonotactiques

Un second facteur concerne la nature du contexte droit du cluster (pause, voyelle ou consonne) : Gadet (1989) stipule que la non-prononciation de la LFPO est plus fréquente quand le mot est suivi d'une consonne que quand il est suivi d'une pause, et quand il est suivi d'une pause que quand il est suivi d'une voyelle. Pooley (1996) et Hambye (2005) présentent des faits allant dans le même sens. Pourtant, Kemp et al. (1980), Chevrot et al. (2000), Hornsby (2006), Villeneuve (2010), Boughton (2013, 2015), Myers et Ranson (2014) mais aussi Milne (2014), Wu (2018) ainsi que Peperkamp et al. (2019), sont mitigés quant à cette

<sup>5</sup>Dans le tableau, les pourcentages comme les nombres sont parfois calculés par nous-même, c'est pourquoi nous avons indiqué les pages où figurent les informations que nous avons utilisées pour ce faire. L'étude de Chevrot *et al.* (2000) n'est pas mentionnée dans le tableau car elle implique un protocole expérimental avec des pseudo-mots.

assertion. Tous observent que si la non-prononciation de la liquide est plus fréquente devant consonne que devant pause ou voyelle, ils ne trouvent pas forcément de différence entre ces deux derniers contextes.

Dans le cas où la LFPO était suivie par un mot à initiale consonantique, certains chercheurs ont pensé à coder le type de consonne. Myers et Ranson (2014) rapportent que si la consonne initiale du mot suivant est similaire à l'obstruante en termes de lieux et modes, la non-prononciation de la liquide est favorisée. Wu (2018) observe que la liquide est davantage éliée si la consonne du mot suivant est une obstruante par rapport à une nasale ou une liquide.

### 2.3. Facteurs lexicaux

La catégorie grammaticale du mot hébergeant la LFPO a également été considérée comme un facteur susceptible d'avoir un rôle sur le comportement des LFPO. Selon Gadet (1989), la non-prononciation de /ʁ/ serait plus systématique dans les mots lexicaux (noms et verbes) que dans les mots grammaticaux (prépositions, cardinaux), alors que pour la liquide /l/, ce serait l'inverse : la non-prononciation serait plus fréquente dans les mots grammaticaux (pronoms) que dans les mots lexicaux (noms et verbes). L'importance des catégories grammaticales est également soulignée par Kemp et al. (1980) ainsi que Myers et Ranson (2014). Elle est toutefois contestée par Pooley (1996), qui remarque à la suite de l'analyse de ses données que la nature du mot (grammatical ou lexical) ne joue pas de rôle sur la prononciation ou la non-prononciation de LFPO (tout du moins s'il s'agit d'un /ʁ/).

Un autre facteur lexical touche à la fréquence du mot hébergeant la LFPO (Laks, 1977; Kemp et al., 1980; Armstrong, 2001; Brand et Ernestus, 2021). Selon une hypothèse communément admise, plus un mot est fréquent, plus il aurait tendance à être prononcé de façon « réduite » (v. notamment Wang, 1969; Bybee, 2002). Ainsi, Pooley (1996) remarque que les adverbes fréquents comme *peut-être* ou *par contre* sont plus souvent prononcés sans leur LFPO que des prépositions comme *entre* et *contre*, dont l'usage est moins fréquent. Dans la même veine, Ranson (2010), Myers et Ranson (2014) ou encore Villeneuve (2010) observent d'intéressantes corrélations positives entre la fréquence du mot et la non-prononciation des LFPO : plus un mot est fréquent, plus sa LFPO a tendance à ne pas être prononcée. Des observations que nuance toutefois Pustka (2007, 2011). L'auteure remarque en effet que des mots rares comme *astre*, *pègre*, *ténèbres* ou encore *oindre* sont rarement prononcés sans leur LFPO. Mais elle ne trouve toutefois pas de corrélations significatives entre le taux de non-prononciation des LFPO et la fréquence des mots contenant ce genre de cluster dans son corpus.

### 2.4. Facteurs prosodiques

Pooley (1996) et Pustka (2007, 2011) remarquent que les LFPO sont moins souvent prononcées quand elles sont en position accentuable (c'est-à-dire en position finale de groupe clitique, pour reprendre les termes de Garde, 1965). En pratique, cela veut dire que les mots intonés dans le même groupe accentuel que le mot qu'ils précèdent

voient leur LFPO ‘protégée’, alors que les LFPO des mots qui se trouvent en bord droit de groupe accentuel ont plus de chances de ne pas être prononcées. P. ex., selon cette hypothèse, il est plus probable que la LFPO du lexème *centre* dans le syntagme *le centre du pays* soit prononcée si le mot est phrasé dans un groupe accentuel différent que le mot qu’il suit (/l(ə)ˈɑ̃tʁ(ə) | dypɛˈi/) que s’il ne l’est pas (/l(ə)(ə) sɑ̃dypɛˈi/).

Dans l’étude de Brand et Ernestus (2021), deux critères comparables à ce dernier sont pris en compte : la position du mot dans l’énoncé (au début, au milieu ou à la fin) et le nombre de syllabes dans le mot. En se basant sur de tels critères, les chercheuses remarquent que lorsque le mot contenant une LFPO est au milieu d’un énoncé, la liquide a plus de chance de ne pas être prononcée que si le mot est dans une autre position. Elles observent également que la liquide est plus souvent prononcée sur les mots courts que sur les mots longs.

Dans la même veine, Myers et Ranson (2014) observent des effets significatifs de la position du mot dans l’énoncé, ce qui va dans le même sens que dans l’étude précédente. D’après leurs résultats, l’élision de la LFPO est favorisée dans un mot placé à l’intérieur d’un énoncé autre que dans une position prénominale et par rapport à un mot placé en position finale.

Enfin, d’autres travaux ont pris en compte la vitesse d’articulation, calculée comme la durée relative des syllabes du groupe dans lequel s’insère le mot avec une LFPO. Ce facteur s’est avéré significatif aussi bien dans l’étude de Milne (2014) que dans celles de Myers et Ranson (2014) et de Brand et Ernestus (2021). Les uns comme les autres montrent que plus la durée syllabique est courte (donc plus les locuteurs articulent rapidement), plus la tendance à ne pas prononcer la LFPO est grande.

## 2.5. Facteurs liés à la situation de parole

En ce qui concerne le rôle de la situation de parole (ou le degré de formalité du contexte de communication), le consensus entre les auteurs est plus grand. Laks (1977) et Armstrong (1993, 2001) observent que la chute des LFPO est moins fréquente dans les productions où l’enquêteur est présent (l’enquêteur joue le rôle d’interlocuteur) que dans les productions où l’interlocuteur est absent (la conversation se fait entre pairs uniquement). Sur la base de la comparaison de trois corpus alignés automatiquement,<sup>6</sup> Wu (2018) formule une conclusion similaire : plus la situation de parole est informelle, moins les LFPO sont réalisées. Enfin, Hambye (2005), Milne (2014) et Boughton (2015), remarquent tous que les locuteurs de leurs corpus prononcent davantage de LFPO lors d’une tâche de lecture (qui génère de la parole plus contrôlée) qu’en conversation (où la parole est moins contrôlée).

<sup>6</sup>ESTER, un corpus de journaux radio- et télédiffusés, est considéré comme le plus formel; NCCFr est un corpus d’entretiens spontanés, il est considéré comme le plus informel; ETAPE est un corpus de parole journalistique incluant des débats politiques et des interviews, son degré de formalité est considéré à mi-chemin entre les deux autres.

## 2.6. Facteurs relatifs au locuteur

Le dernier ensemble de variables ayant été discutées dans les études consacrées au comportement des LFPO en français sont relatives au profil socioéconomique des locuteurs. Laks (1977) a notamment constaté une corrélation marquée entre le profil socioéconomique des participants et leur propension à prononcer ou non les LFPO. D'après ses résultats, il ressort que plus les participants sont socialement 'défavorisés', plus ils ont tendance à ne pas prononcer les LFPO, et vice-versa (les taux d'élision allant de 30% pour les locuteurs en haut de l'échelle sociale à près de 100% pour les participants en bas de cette échelle). Des observations analogues ont été faites par Kemp et al. (1980) sur un corpus de français québécois.

Souvent, on observe que le profil socioéconomique interagit avec deux autres variables, à savoir le sexe et l'âge des locuteurs. Dans deux études faisant intervenir des groupes de locuteurs de différents âges, Pooley (1996) et Hornsby (2006) ont mis au jour des liens intéressants entre la non-prononciation des LFPO, l'âge, le sexe et le niveau socio-éducatif des locuteurs (les locuteurs de sexe masculin, moins éduqués et âgés de plus de 45 ans produiraient davantage de clusters sans LFPO que les autres). Villeneuve (2010) ne signale aucun effet de sexe dans son étude, mais remarque que les locuteurs âgés (55 ans et plus) réalisent moins de LFPO que les locuteurs plus jeunes (25–54 ans). Quant à Boughton (2015), elle observe que les locuteurs de classe ouvrière prononcent moins souvent les LFPO que les locuteurs de classe moyenne; que les locuteurs masculins prononcent moins souvent la LFPO que les femmes (ce qui coïncide avec ce que disent Pooley, 1996 et Hornsby, 2006). En revanche, l'auteure remarque que l'âge n'a aucun effet sur le comportement des LFPO, les jeunes locuteurs de son corpus (16–25 ans) présentant un taux de non-prononciation presque identique aux locuteurs plus âgés (40–60 ans).

Enfin, il convient de signaler que l'origine géographique du locuteur a été étudiée par certains auteurs. Milne (2014) a ainsi comparé le français canadien et le français standard européen, et a trouvé un taux de non-prononciation plus important dans la variété américaine que dans la variété européenne. Hambye (2005) trouve également des différences entre des locuteurs originaires de Tournai (ville proche de Lille, située sur la frontière franco-belge, et où le français qui y est parlé est généralement considéré comme plus proche de la variété de référence du français, Bardiaux, 2014) et des locuteurs de Gembloux et de Liège (deux variétés fortement marquées régionalement). D'après les comptages de Hambye (2005), il ressort en effet que les locuteurs de Liège et de Gembloux prononcent moins de LFPO que les locuteurs de Tournai. Pustka (2007, 2011) a comparé les productions de cinq groupes de locuteurs : des locuteurs de Paris, de Guadeloupe et du Midi de la France (département de l'Aveyron); des locuteurs originaires de Guadeloupe et d'Aveyron résidant à Paris. En pratique, elle constate que les locuteurs de Guadeloupe prononcent moins souvent les LFPO que les locuteurs de Paris, qui se comportent à leur tour différemment des immigrants aveyronnais, ces derniers présentant des pourcentages de non-prononciation légèrement supérieurs à ceux des locuteurs de l'Aveyron mais inférieurs à ceux des Parisiens. Quant à Boughton (2013), elle a comparé les productions des locuteurs de Rennes et de Nancy (deux villes à égale distance de

Paris, situées à l'Ouest et à l'Est de la France), et a constaté des différences entre les locuteurs de ces variétés, les locuteurs de Nancy étant plus enclins à prononcer les LFPO que ceux de Rennes. Enfin, Myers et Ranson (2014) comparent les productions de locutrices du sud de la France et du nord de la France : elles observent que dans les productions méridionales, le taux de non-prononciation des LFPO est plus bas que dans l'autre groupe, en raison de la propension plus importante à la prononciation du schwa dans cette région.

## 2.7. Synthèse et problématique

Les facteurs ayant un impact sur la (non-)prononciation des LFPO en français sont nombreux, mais le consensus qui se dégage quant à leur rôle respectif dans la communauté des chercheurs est mitigé. On a vu que des facteurs comme le mode d'articulation de la LFPO, la fréquence lexicale, la catégorie grammaticale, l'âge et le sexe des locuteurs étaient des critères qui ne faisaient pas l'unanimité. On a vu aussi que s'il y a accord sur le fait que le contexte droit (consonne, pause ou voyelle) permet d'expliquer le comportement de la LFPO, il est difficile de savoir dans quel sens ce facteur fonctionne, certains auteurs considérant le contexte préconsonantique comme défavorisant la prononciation de la LFPO au même titre que le contexte pré-pausal, d'autres ne trouvant pas de différences entre ces deux contextes. Le rôle du mode et du lieu d'articulation de l'obstruante, l'origine géographique des locuteurs et le statut prosodique du mot hébergeant le cluster prêtent moins à débat, mais ces variables n'ont pas été testées par la majorité des auteurs. Au final, seuls deux éléments semblent faire consensus : le style de parole (la parole préparée/lue étant plus encline à l'articulation de LFPO que la parole non-préparée/non-lue), le profil socioéconomique des locuteurs (les locuteurs présentant un profil social moins prestigieux socialement font chuter dans une plus grande proportion les LFPO que les locuteurs présentant un profil social plus prestigieux).

Comment expliquer l'existence de ces discordances ? Selon nous, l'une des principales raisons provient du fait que la majorité des auteurs considèrent les facteurs susceptibles d'influencer le comportement des LFPO de façon indépendante, alors qu'il est clair qu'ils doivent être considérés les uns par rapport aux autres, comme cela a été montré par ceux qui ont croisé différents facteurs dans leur analyse. À cet égard, Armstrong (2001) affirme que le comportement des LFPO est davantage influencé par le style de parole que par le sexe et les facteurs de niveau socio-économique. Hambye (2005) conclut que les intervenants de son corpus ne se comportent pas de la même façon en ce qui concerne leur niveau socio-éducatif et le style de parole, les écarts entre les différentes catégories socio-éducatives des locuteurs étant plus marqués en lecture qu'en conversation. Pustka (2007, 2011) trouve aussi une influence substantielle de la combinaison entre l'origine du locuteur et le contexte droit du cluster LFPO : les locuteurs de la Guadeloupe présentent un taux de non-prononciation plus élevé avant voyelle par rapport aux locuteurs originaires d'autres régions. Dans la même lignée, Villeneuve (2010) rapporte que dans la position prévoalique, la non-prononciation de la LFPO est plus fréquente dans les productions des locuteurs âgés que dans les productions des locuteurs jeunes.

Autre exemple : dans l'étude de Wu (2018), il y a une différence entre hommes et femmes dans les corpus de parole journalistique, mais pas en parole spontanée. Dans le travail de Milne (2014), il ressort que l'effet de vitesse d'articulation n'est pas aussi important en parole lue qu'en parole spontanée. Quant aux résultats de Boughton (2013), ils montrent que le contexte ne joue pas toujours le même rôle selon le profil socio-économique, l'âge et le sexe des locuteurs :

Middle-class speakers and females prefer the standard variant [LFPO prononcée] in all contexts, though we observe the greatest shift between groups in the behavior of the two classes in prevocalic/prepausal context, where deletion rates are 30% lower for middle-class speakers. This result is mirrored in the difference between males and females, where the gap in prevocalic/prepausal context nears 20%, almost double that for preconsonantal context (Boughton, 2013 : 13).

Dans un autre ordre d'idée, Pustka (Pustka 2011 : 28) s'interroge sur le fait que certains clusters obstruante + liquide, en plus d'être rares, n'apparaissent que dans certains mots diatopiquement distribués. Elle donne l'exemple de la séquence <gr>, qui n'apparaît que dans le mot *nègre*, un mot qui n'est prononcé que dans la bouche de locuteurs de Guadeloupe, où il n'a pas la même connotation péjorative qu'en Europe. Et de conclure : « La nature du groupe OL [obstruante + liquide] peut donc s'avérer être une variable parasite, derrière laquelle se cache une variation diatopique ou lexicale. »<sup>7</sup> Enfin, Myers et Ranson (2014) soulignent l'importance de croiser le facteur régional avec le facteur de fréquence lexicale, certains mots étant raccourcis plus souvent que d'autres dans une région mais pas dans l'autre.

Un second problème concerne le choix des méthodes pour évaluer l'effet de certains facteurs sur la variable à l'étude. Par exemple, Pustka (2007 : 208) remarque que « la liquide tombe plus souvent après une plosive (45%) qu'après une fricative (33%) », mais ne fournit pas d'analyse statistique pour s'assurer que cette différence n'est pas due au hasard, ou à l'influence d'un autre prédicteur qui n'aurait pas été contrôlé. Même imprudence chez Boughton (2013 : 128), quand elle observe que « Rennes participants consistently delete more liquids than Nancy speakers » en raison d'une « difference of over 20 percentage points between the two locations », mais sans avoir observé l'influence d'autres facteurs sur ces résultats. Dans certains cas toutefois, des analyses statistiques viennent appuyer les conclusions : on retrouve des tests d'indépendance ( $\chi^2$ ) chez Pooley (1996), chez Hambye (2005) et des tests d'analyse de variance (ANOVA) chez Armstrong (2001). Pourtant, ces tests ne permettent pas de prendre en compte plusieurs facteurs en même temps, et ne peuvent donc que confirmer ou infirmer l'existence de différences significatives quant à une variable isolée (ou entre des groupes de sous-corpus).

Villeneuve (2010) est la première à proposer une analyse multifactorielle du problème. L'approche est toutefois discutable, et ce pour au moins deux raisons : d'abord, parce que l'auteure ne tient pas compte des interactions entre les

<sup>7</sup>Signalons par ailleurs qu'en Guadeloupe, que le mot *nègre* se prononce justement sans sa consonne finale [neg], ce qui provoque d'autant plus de biais dans les résultats.

variables entrées dans le modèle (ni du fait que ces variables soient éventuellement corrélées); ensuite, parce que rien n'indique dans son modèle qu'elle a pris soin de tenir compte de la variation intra- et interlocuteur.<sup>8</sup> Dans ce contexte, il est possible que les effets observés soient dus au comportement de certains locuteurs, qui produiraient certains items plus souvent que d'autres – nonobstant la fréquence de ces items dans le corpus ou dans l'absolu; v. les remarques de Pustka (2011) au sujet du mot *nègre* reportées *supra*.

Dans les études récentes, des analyses statistiques plus complexes, à base de régressions, avec ou sans effets mixtes, ont été proposées. Le problème est que dans ces travaux, seuls quelques facteurs explicatifs sont testés dans un même modèle. Dans Milne (2014), les facteurs pris en compte comprennent la variation régionale (comparaison entre des locuteurs originaires du Québec et de la France septentrionale), le style de parole (lecture et spontanée), le contexte droit (devant pause, consonne ou voyelle) et la vitesse d'articulation. Des facteurs importants relatifs aux caractéristiques articulatoires des clusters, aux catégories de mots (grammaticaux ou lexicaux), à leur position dans la chaîne, sans parler des critères externes relatifs aux locuteurs (âge, sexe et profil socio-économique), sont laissés de côté. Wu (2018), qui travaille sur l'articulation de /ʁ/ seulement, ne considère que le style de parole, le sexe et le contexte droit (devant pause, consonne ou voyelle), en laissant de côté des paramètres pourtant tout aussi importants, à savoir la vitesse d'articulation et le nombre de syllabes dans le mot, de même que la catégorie grammaticale et la fréquence lexicale. Brand et Ernestus (2021) abordent le problème de la chute des LFPO dans le cadre d'une recherche sur la réduction phonétique.<sup>9</sup> Ce sont elles qui proposent le modèle le plus complet. Toutefois, comme c'est le cas de Wu, leur corpus n'est pas taillé pour tenir compte de facteurs variationnels qui touchent à l'âge, au niveau social ou à l'origine des locuteurs.

Au final, on l'aura compris : on ne sait toujours pas quels sont les mécanismes qui motivent la (non-)prononciation des LFPO en français à l'heure actuelle. Du travail reste donc à fournir pour aboutir à une meilleure compréhension du problème.

### 3. MÉTHODE

#### 3.1. Participants

Les matériaux que nous utilisons dans cette étude ont été extraits de la base Phonologie du Français Contemporain (PFC, Durand, Laks et Lyche, 2009). Plus précisément, le corpus de travail comprend les productions de locuteurs originaires de 15 régions de la francophonie d'Europe : 5 localités de la partie septentrionale de la France (Béthune, Brécey, Lyon, Paris et Ogéville), 5 localités

<sup>8</sup>Des critiques du même type peuvent être adressées à Ranson (2010) et Myers et Ranson (2014). Le problème repose en fait sur l'utilisation, dans ces trois études, de l'outil de statistique GoldVarb X (Sankoff et al., 2005). Outre le fait que cet outil n'est pas conçu pour traiter des variables numériques, il ne permet pas de faire des comparaisons deux par deux sans tenir compte des interactions entre les facteurs et de la structure interne des données (pour une critique complète de l'outil, v. Gradoville, 2019).

<sup>9</sup>C'est ce qui explique que les auteures aient choisi de travailler sur des mots où la LFPO suit une consonne qui est forcément sourde, et que tous les mots examinés dans le corpus sont des substantifs.

de Suisse romande (Fribourg, Genève, Martigny, Neuchâtel et Nyon) et 5 localités de Belgique (Bruxelles, Gembloux, Liège, Marche-en-Famenne et Tournai). La localisation de ces points d'enquête peut être visualisée sur la Figure 1.

Pour chacun des 15 points d'enquête, 4 hommes et 4 femmes ayant passé la plus grande partie de leur vie et étant bien ancrés dans lesdites localités ont été sélectionnés. Les 120 locuteurs présentent des âges et des profils socioéconomiques comparables entre eux et entre les régions. Le tableau 2 résume l'ensemble des informations y relatives.<sup>10</sup>

Chacun des 120 participants a été enregistré dans deux contextes de parole : dans une tâche de lecture et dans le cadre d'un entretien plus ou moins libre.<sup>11</sup>

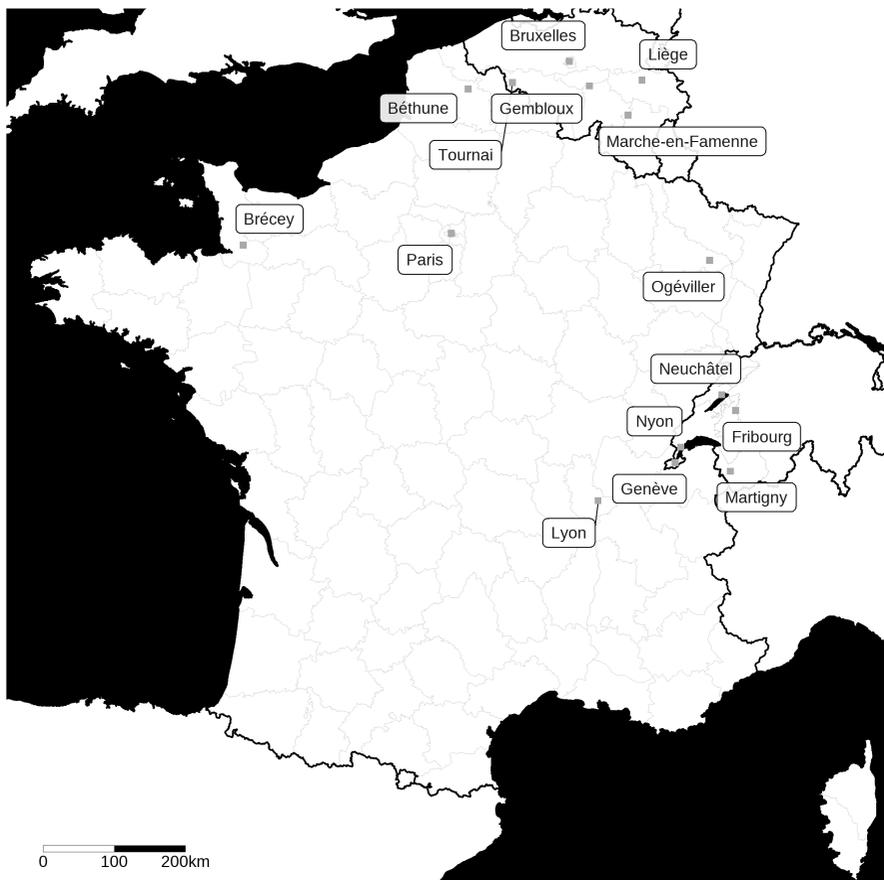
### **3.2. Annotations**

La totalité du texte lu (398 mots) ainsi que 3 minutes de conversation ont été annotés semi-automatiquement entre 2013 et 2016, dans le cadre d'un projet sur la variation prosodique régionale du français européen (Avanzi 2014). À un premier niveau, le corpus comprend une segmentation en phonèmes, syllabes et mots graphiques, réalisée à l'aide du plugin EasyAlign (Goldman, 2011) qui fonctionne sous Praat (Boersma 2001). Les alignements ont tous été vérifiés et corrigés manuellement par différents collaborateurs, qui se sont basés sur l'écoute des fichiers audio mais aussi sur le spectrogramme en vue de réajuster, ajouter ou supprimer des frontières de phonèmes, avec une attention spéciale aux phénomènes phonotactiques qui posent le plus problème aux aligneurs (schwas, liaison, élision, etc.). Sur la base du découpage en mots graphiques, une lemmatisation a été effectuée avec le logiciel Dismo (Christodoulides, Avanzi et Goldman, 2014), et cet étiquetage a également été révisé manuellement. Le corpus contient également une tire d'annotation où sont identifiées différentes informations relatives aux disfluences (Christodoulides et Avanzi, 2015), ainsi que les unités inter-pausales (Schwab et Avanzi, 2015).

---

<sup>10</sup>Nous avons classé les participants en deux catégories : ceux qui appartiennent plutôt à la classe ouvrière (CSP-) et ceux qui appartiennent plutôt aux classes moyenne et supérieure (CSP+). Pour ce faire, nous nous sommes basé sur les informations disponibles dans le corpus PFC, à savoir le niveau d'études des participants et sur le fait que leur profession implique plutôt une activité manuelle ou non (pour les témoins qui disposaient d'un emploi). Nous sommes conscient que les facteurs potentiellement déterminants du statut socio-économique des locuteurs sont plus nombreux que cela, mais nous avons suivi Chambers et Trudgill (1998) ainsi que Chambers (2009), en considérant que la profession est un bon indicateur de la classe sociale. Dans Avanzi (2014), on trouvera davantage de détails sur la sélection des points d'enquête (pour la Belgique et la Suisse, tous les points PFC ont été utilisés; pour la France septentrionale, l'idée était de disposer de points aussi éloignés les uns que les autres autour de Paris), mais aussi sur la nature des conversations utilisées (le corpus PFC présentant deux types de parole spontanée : l'une dirigée, l'autre plus libre; la distinction dans les faits n'étant pas toujours évidente dans ledit corpus, elle a été ignorée lors de la création de ce sous-corpus).

<sup>11</sup>Nous nous permettons de renvoyer aux travaux réalisés dans le cadre du projet PFC, qui explicitent en détail comment les témoins ont été choisis, et quel a été le protocole suivi pour le recueil des données.



**Figure 1.** Carte de la francophonie d'Europe, avec localisation des 15 points d'enquêtes. Les lignes épaisses signalent des frontières entre pays, les lignes plus fines des divisions internes aux provinces en Belgique, des départements en France et des cantons francophones en Suisse.

### 3.3. Sélection des items

Grâce à des scripts, nous avons pu extraire de ce corpus toutes les unités séparées par des blancs graphiques (*tokens*) se terminant par les graphèmes <t, d, p, b, g, k, c, f, v>, suivis de <r> ou <l> et de <e> (plus éventuellement de <s> et <nt>), soit 2801 items au total.<sup>12</sup> Pour travailler sur un corpus aussi homogène que possible, nous avons dû exclure un certain nombre de cas.

<sup>12</sup>280 mots (soit 10% du total) ont été choisis aléatoirement aussi bien dans le corpus de lecture que dans le corpus de conversation. Leur transcription phonétique (présence ou non d'un segment) a été vérifiée directement dans Praat par l'auteur de cet article. Le nombre de corrections apportées ne concernant pas plus de 8 items, l'annotation a été jugée fiable.

**Tableau 2.** Origine (pays et points d'enquête) des locuteurs, âge (en années, avec min., max., moyenne et écart-type) et nombre de locuteurs par profil socio-économique pour chaque groupe au moment de l'enregistrement

Pays	Localité d'origine	Âge (en années)		Profil socioéconomique	
		Min.-Max.	Moyenne (E.-T.)	CSP-	CSP+
Belgique	Bruxelles	27-65	44 (15)	0	8
	Gembloux	22-76	42 (21)	4	4
	Liège	21-76	48 (24)	3	5
	Marche	18-66	39 (16)	2	6
	Tournai	19-82	44 (26)	3	5
France	Béthune	21-89	46 (25)	4	4
	Brécey	19-80	47 (22)	4	4
	Lyon	21-74	42 (21)	4	4
	Ogéwiller	23-93	58 (24)	8	0
	Paris	24-86	50 (22)	0	8
Suisse	Fribourg	20-82	43 (24)	4	4
	Genève	21-61	41 (18)	2	6
	Martigny	22-80	49 (28)	4	4
	Neuchâtel	25-78	53 (24)	4	4
	Nyon	30-70	46 (17)	8	0
	Total	18-93	50 (29)	54	66

Nous avons d'abord écarté les mots prononcés à l'anglaise (*jingle*, N=1), les régionalismes (*encoupler*,<sup>13</sup> N=1) et les noms propres (*Londres*, *Alexandre*, N=21).<sup>14</sup> Nous avons également laissé de côté les items qui constituent le premier membre d'un mot composé (*quatre-vingts*, *quatre mille*, N=11) et les items suivis d'un mot commençant par la même liquide que la consonne initiale du mot précédent (avec <l>, par exemple *l'année prochaine*, N = 8; avec <r> *notre référence*, N=10), en raison des ambiguïtés d'interprétation qu'ils génèrent.<sup>15</sup> Nous avons aussi isolé tous les mots suivis d'un *ehh* d'hésitation (*prendre ehh*, *des êtres ehh*, N=59) ou associés à une disflueuse (*de notre #<sup>16</sup> liberté*, *montrer patte blanche ehh patte blanche*, N=97).

<sup>13</sup>En Suisse romande, le verbe *encoupler* est synonyme de *trébucher* (Thibault, 1997).

<sup>14</sup>Ces mots ont été exclus car nous ne disposons pas d'informations sur leur fréquence dans le corpus Lexique3, v. ci-dessous §3.4)

<sup>15</sup>L'ensemble des mots se terminant par <l> étaient assortis d'un schwa, tandis qu'un seul des mots terminés par <r> était suivi d'un schwa.

<sup>16</sup>Le symbole # indique une pause.

### 3.4. Variables considérées

Les annotations que comporte le corpus nous ont permis de juger du rôle de plus d'une douzaine de prédicteurs sur le comportement des LFPO des mots lexicaux de notre corpus, soit la quasi-totalité des paramètres généralement mentionnés dans la bibliographie sur le sujet (S§2). Un premier facteur concerne (i) le contexte droit de l'item hôte : est-il suivi d'une pause (N=633), d'un mot à initiale consonantique (N=1472) ou d'un mot à initiale vocalique (N=488) ? Les facteurs suivants concernent les caractéristiques phonétiques du groupe obstruante + liquide. Nous avons codé le (ii) la continuité de la consonne (occlusive, N=2370; fricative, N=223), (iii) le lieu d'articulation de cette consonne (labiodental, N=70; dorsal, N=77; bilabial, N=297; dental, N=2067) et (iv) la nature de la liquide (<l>, N=281; <r>, N=2312). Ensuite, nous avons calculé (v) la fréquence lexicale de l'item hôte dans le corpus Lexique3 (New et al., 2001), par million de mots.<sup>17</sup> Les mots les moins fréquents (*rectangle*, *admissible*, *cadastre*, etc.) reçoivent un score de 0,50 à 0,60, alors que le plus fréquent (le verbe *être*) obtient un score de 32236,5.<sup>18</sup> Dans un autre temps, nous avons calculé (vi) le nombre de syllabes dans l'item hôte (1 syllabe, N=1342; 2 syllabes, N=1201; 3 syllabes et plus, N=50<sup>19</sup>; moyenne=1,5, E.T.=0,5). Par ailleurs, puisque l'on sait que la vitesse à laquelle on parle influence la réalisation des segments, nous avons calculé la durée syllabique moyenne de l'unité inter-pause dans laquelle l'item hôte a été prononcé, obtenant ainsi (vii) la vitesse d'articulation locale pour ledit segment en ms (min.=73,4, max.=706,8, moyenne=219,7 et E.T.=61,8). De plus, nous avons codé (viii) le statut prosodique du mot, en différenciant les mots intonés dans le même groupe accentuel que le mot qu'ils précèdent (N=716) des mots intonés dans un groupe accentuel différent (N=1877), ce qui nous a permis de tenir compte du facteur intono-syntaxique (l'idée étant que si le mot appartient au même groupe accentuel que le mot qui le suit, il est plus 'lié' que s'il est phrasé dans une unité prosodique différente, Delais-Roussarie et al., 1996, Avanzi 2013). En ce qui concerne les informations relatives aux locuteurs, nous avons entré dans le modèle (ix) le pays d'où ils sont originaires (Belgique, N=855; France, N=871; Suisse, N=867), (x) leur âge (min.=18, max.=93, moyenne=45,9 et E.T.=21,2), (xi) leur genre (femme, N=1305; homme, N=1288) et (xii) leur profil socioéconomique (CSP-, N=1181; CSP+, N=1412). Enfin, nous avons codé la variable (xiii) style de parole (lecture, N=1780; conversation, N=813).

<sup>17</sup>Nous avons fait le choix d'un corpus indépendant car notre corpus (72797 mots, en ne comptant les mots du texte lu qu'une seule fois) est relativement 'petit' par rapport à Lexique3 (14,7 millions de mots). Cependant, à toutes fins utiles, nous avons calculé la fréquence lexicale des mots à l'intérieur de notre propre corpus, en ne comptabilisant le texte lu qu'une seule fois. Il ressort que les fréquences calculées sur la base de notre corpus sont hautement corrélées aux fréquences mesurées dans Lexique3 (avec  $\rho = 0,87$ ).

<sup>18</sup>Les fréquences lexicales ont été transformées de manière logarithmique.

<sup>19</sup>Le corpus ne comporte que 6 mots de quatre syllabes (*amphithéâtre(s)*, *épouvantable*, *inadmissible*, *incomparable* et *invraisemblable*) : nous les avons regroupés dans la même catégorie que les mots de 3 syllabes.

### 3.5. Statistiques

Les statistiques ont été conduites dans le logiciel R (version 4.2, R Development Core Team, 2013), à l'aide de la fonction *glmmTMB* de la librairie éponyme (Brooks et al., 2017). Les données ont été analysées à l'aide de modèles linéaires généralisés à effets mixtes (GLMM, v. McCulloch et Neuhaus, 2001), avec une fonction *logit* (transformation logistique et distribution binomiale) et une inflation nulle supposée constante dans l'ensemble de données (*ziformula=~1*).<sup>20</sup> La réalisation de la liquide (oui/non) a été entrée comme variable dépendante, les 13 prédicteurs présentés ci-dessus (§3.4) ont été entrés comme des variables indépendantes,<sup>21</sup> les items et les locuteurs comme des variables aléatoires. Les valeurs de *p* ont été obtenues au moyen de la fonction *drop1* de la librairie *stats*. Les tests *post-hoc* ont été effectués au moyen de la fonction *emmeans* de la librairie *emmeans* (Lenth, 2016).

## 4. RÉSULTATS

Sur les 2593 observables valides du corpus,<sup>22</sup> il est apparu que les mots grammaticaux étaient nettement moins nombreux que les mots lexicaux (N=238 contre N=2355, respectivement) et que la distribution de ces derniers en regard des variables présentées à la section 3.4. ne permettait pas l'utilisation de modèles statistiques tels que ceux décrits à la section 3.5. Les mots grammaticaux sont donc traités avec des outils de statistique descriptive uniquement (§4.1); les modèles statistiques inférentiels (présentés en 3.5) seront appliqués aux mots lexicaux dans un second temps (§4.2).

### 4.1. Mots grammaticaux

Sur les 238 mots grammaticaux du corpus, la moitié (N=120) sont prononcés lors de la lecture du texte. 116 impliquent le possessif *notre* (dans le SN *notre liberté*)<sup>23</sup> et 4 occurrences la préposition *entre* (dans l'expression *entre parenthèses*). Toutes les occurrences d'*entre* sont prononcées avec leur liquide assortie d'un schwa (/ã.tʁə.pa.ʁã.tɛz/). Sur ces 116 occurrences restantes de *notre*, la liquide est prononcée 102 fois (dont 82 occurrences comportant un schwa), ce qui veut dire que la liquide n'est élidée que dans 12,1% des cas (N=14).<sup>24</sup>

<sup>20</sup>En deux mots, l'intérêt de ces modèles est qu'ils permettent de tenir compte de la variation intra- et interlocuteur, du fait que certains items se répètent plus souvent que d'autres dans le corpus, mais aussi du déséquilibre dans la distribution des données du corpus.

<sup>21</sup>Contrairement à Brand et Ernestus (2021), mais tout comme Milne (2014) et Wu (2018), nous n'avons pas inclus le comportement du schwa (présent ou absent) dans nos modèles statistiques. Ceci s'explique par des raisons computationnelles. Comme on peut le voir sur la Figure 3, la distribution du schwa dans notre corpus est lacunaire (il ne précède quasi jamais un mot à initiale vocalique en l'occurrence).

<sup>22</sup>Les 120 locuteurs du corpus ont produit entre 14 et 39 items chacun, ce qui représente 21,6 mots par témoin en moyenne.

<sup>23</sup>On aurait attendu 120 occurrences, mais 4 d'entre elles ont été écartées car associées à une disfluece.

<sup>24</sup>Le profil socioéconomique des locuteurs n'explique pas cette distribution, puisque 7 locuteurs ne prononçant pas la LFPO présentent un profil CSP+. On n'observe aucun effet apparent de pays (6 Belges, 3 Français et 5 Suisses) ou de sexe (10 hommes vs 4 femmes).

En parole conversationnelle, la distribution des 118 items peut être résumée au moyen de la Figure 2. On peut voir que seulement quatre lexèmes sont concernés.<sup>25</sup> Trois d'entre eux (*notre*, *votre* et *entre*) ne sont jamais suivis d'une pause, et quand ils précèdent un mot à initiale consonantique, la liquide est soit majoritairement supprimée (*notre*, *votre*), soit majoritairement réalisée, auquel cas elle est assortie d'un schwa (23/27 occurrences d'*entre*). Suivie d'un mot à initiale vocalique, la liquide est systématiquement maintenue dans ces trois mots. Quant au cardinal *quatre*, on constate que la liquide est toujours prononcée devant voyelle (N=23), et majoritairement élidée devant pause (4/6 cas). Devant consonne, le /ʁ/ n'est guère prononcé (16/20 cas).

Au total, les faits examinés relatifs aux mots grammaticaux peuvent être résumés de la façon suivante. Dans le corpus de lecture, le mot *notre* est très majoritairement prononcé avec sa consonne liquide finale. Dans le corpus de conversation, pour les quatre mots répertoriés, la liquide est systématiquement prononcée devant voyelle; elle est majoritairement élidée devant consonne, sauf en ce qui concerne la préposition *entre*, majoritairement prononcée [ãtʁə]. Devant pause, la tendance est à la non-prononciation.

## 4.2. Mots lexicaux

### 4.2.1. Aperçu de la distribution des données

Avant de présenter les résultats des modèles statistiques examinant l'impact de différents prédicteurs sur la réalisation ou non d'une LFPO, il nous a paru intéressant de donner un aperçu de la distribution du comportement des 2355 LFPO de mots lexicaux dans les volets du corpus lecture et conversation du corpus.

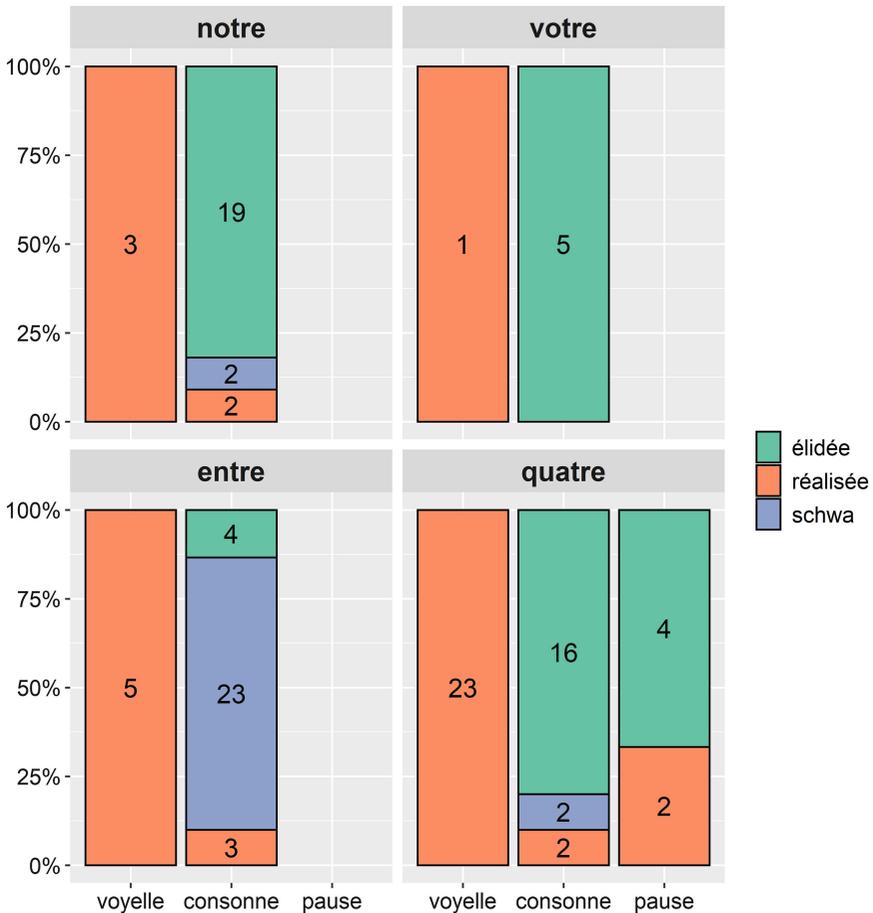
Sur la Figure 3, on peut voir que l'application des trois stratégies varie fortement d'un style à l'autre. En moyenne, 39% des LFPO ne sont pas prononcées en conversation, contre 10% en lecture. On observe également une différence importante quand on croise le style et le contexte, notamment en ce qui concerne les items suivant un mot à initiale consonantique : en lecture, 57,6% des liquides sont suivies d'un schwa, contre 16,3% seulement en conversation. On remarque également que dans l'un et l'autre style, les schwas ne sont jamais ou presque produits avant voyelle.

Compte tenu du fait que nous nous intéressons, dans cet article, à la présence ou l'absence d'une liquide indépendamment de la réalisation d'un schwa, les analyses qui suivent ont été conduites sur l'ensemble des clusters où aucun schwa n'avait été produit, soit sur 1687 unités lexicales.

### 4.2.2. Sélection du modèle statistique

Pour définir le meilleur modèle explicatif, nous avons dans un premier temps entré les 13 facteurs potentiellement pertinents présentés ci-dessus (§3.4) dans un seul et

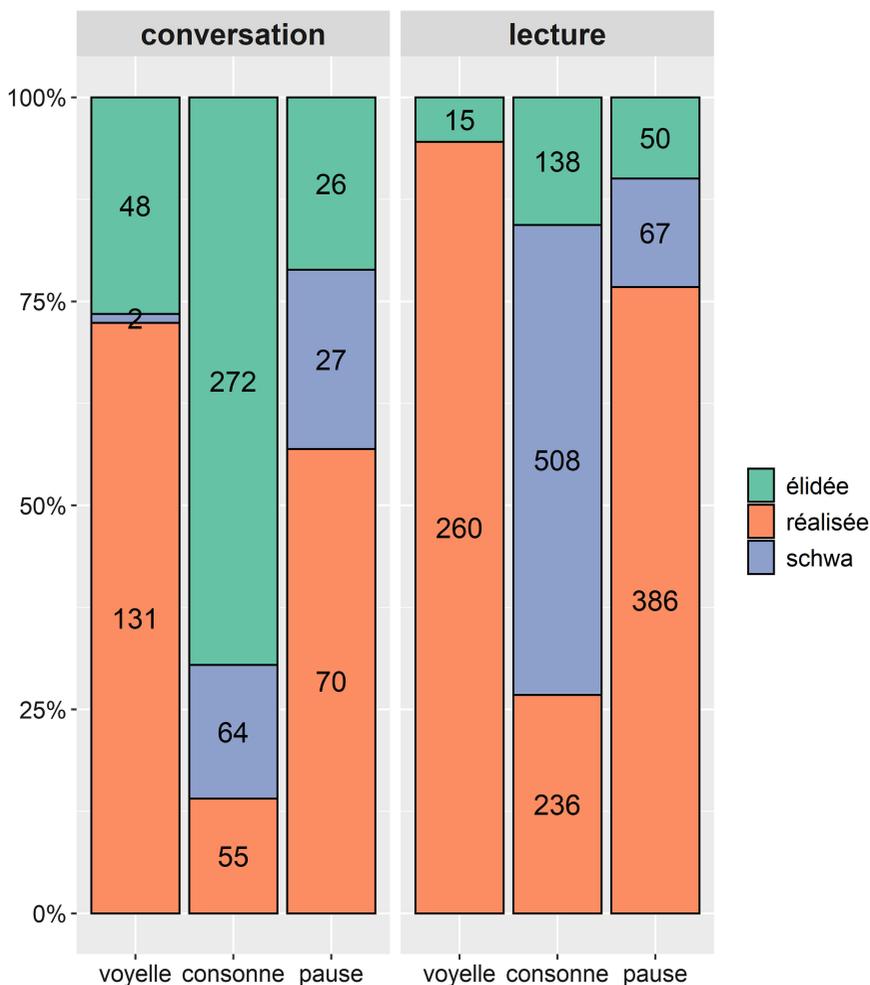
<sup>25</sup>Contextes de disfluences mis à part, la préposition *contre* n'apparaît que deux fois dans le corpus : une fois devant consonne (*contre les Genevois*) dans la bouche d'une locutrice genevoise, qui maintient le /ʁ/; une autre fois devant pause (*j'ai rien contre*) par un locuteur neuchâtelois, qui maintient le /ʁ/ et l'assortit d'un schwa posttonique ([kõtʁə]). On retrouve en revanche plusieurs occurrences de *par contre*, qui sont analysées comme des mots lexicaux, étant donné qu'elles jouissent de la fonction de marqueurs de discours.



**Figure 2.** Distribution (en %) des liquides pour chacun des quatre mots grammaticaux (*notre*, *votre*, *entre* et *quatre*) dans le corpus de conversation, en fonction de leur comportement (élidees, réalisées ou suivies d'un schwa) et du contexte droit (suivies d'un mot qui commence par une voyelle ou une consonne, suivies d'une pause).

même modèle, sans tenir compte des interactions possibles entre les prédicteurs. Nous avons alors testé le degré de multicolinéarité de ces 13 prédicteurs au moyen de la fonction *check\_collinearity* de la librairie *performance* de R (Lüdecke et al., 2021). Les valeurs de VIF (*Variable Inflation Factors*, « facteurs d'inflation variable ») étant toutes très faibles sauf pour le prédicteur lieu d'articulation de l'obstruante (72,14),<sup>26</sup> ce prédicteur a dû être écarté de

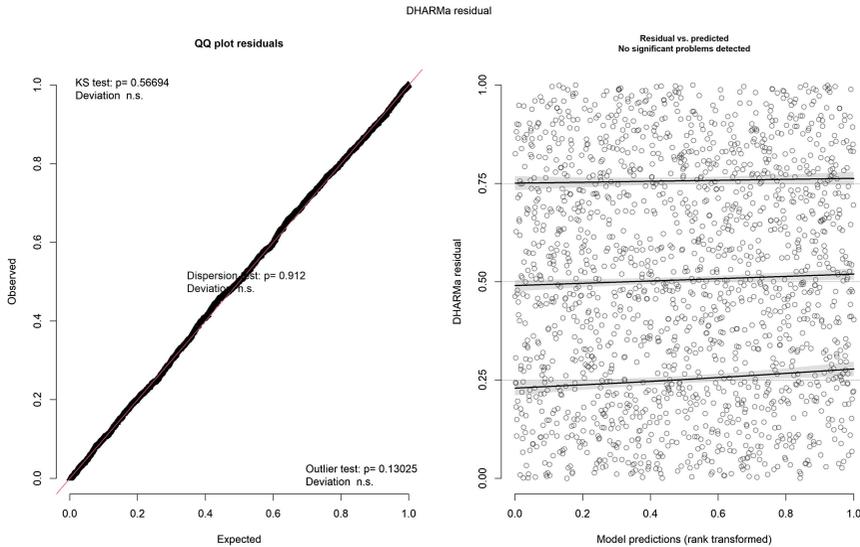
<sup>26</sup>Les valeurs oscillent entre 1 et 1,5. On considère que des valeurs entre 5 et 10 indiquent une corrélation modérée, et qu'au-delà de 10 la colinéarité est problématique (James et al., 2013).



**Figure 3.** Distribution (en %) des liquides dans le corpus de conversation (à gauche) et dans le corpus de lecture (à droite), en fonction de leur comportement (éliminée, réalisée ou suivie d'un schwa) et du contexte droit (suivies d'un mot qui commence par une voyelle ou une consonne, suivies d'une pause).

l'analyse.<sup>27</sup> Nous avons donc effectué un nouveau modèle avec les 12 prédicteurs restants, mais cette fois-ci en tenant compte d'une dizaine d'interactions impliquant la nature de la liquide et : (i) le contexte droit, (ii) la constituante de l'obstruante, (iii) le profil socioéconomique; (iv) le contexte droit et le style; le pays et : (v) la vitesse d'articulation, (vi) le style; l'âge et : (viii) la vitesse d'articulation, (ix) le pays, (x) le style; (xi) le style et la vitesse d'articulation. Nous avons enfin mis de côté tous les prédicteurs non-significatifs dans un

<sup>27</sup>Nous avons parallèlement vérifié si nous n'obtenions pas des résultats plus satisfaisants dans un modèle où le lieu d'articulation comportait trois niveaux (dorsal, N=76; labial, N=259 et coronal, N=1352), mais le modèle n'a pas convergé.



**Figure 4.** Distribution des résidus du modèle final, avec distribution attendue (*expected*) et distribution observée (*observed*). Les valeurs des différents tests statistiques sont affichées de part et d'autre du graphe de gauche (avec KS : Kolmogorov-Smirnov).

nouveau modèle en ne retenant que ceux qui présentaient des valeurs de  $p < 0,01$ . Nous avons alors examiné la structure des résidus du nouveau modèle ainsi mis au point. Afin de valider formellement la spécification de ce modèle final, nous avons effectué un test diagnostique des résidus à l'aide de la librairie DHARMA (Hartig, 2020). Cette procédure a permis de montrer que les résidus suivaient une distribution statistiquement uniforme (ce que confirme le test de Kolmogorov-Smirnov), comme on peut le lire sur la Figure 4.

#### 4.2.3. Résultats

Les résultats du modèle décrit ci-dessus sont présentés dans le Tableau 3.

En premier lieu, comme on aurait pu s'y attendre à la lecture de la Figure 3 plus haut et comme on le voit sur la Figure 5, on peut voir que le contexte droit a une influence sur le comportement de la LFPO : celle-ci est significativement moins souvent prononcée devant consonne (58,5%) que devant voyelle (13,8%, avec  $p < 0,001$ ) ou devant pause (14,3%, avec  $p < 0,001$ ). Le modèle ne fait pas ressortir pas de différence entre les contextes de pause et de voyelle.

Ensuite, on observe un effet de la vitesse d'articulation sur la variable dépendante. Comme on peut le voir sur la Figure 6, plus la durée syllabique augmente (donc que la vitesse d'articulation ralentit), plus les chances que la liquide soit prononcée sont élevées.

En ce qui concerne les variables extralinguistiques, on note un effet du pays d'origine des participants (Figure 7). Avec seulement 20,9% d'élision, les Suisses se démarquent significativement ( $p < 0,001$ ) des Belges (39% d'élision) et des Français (41%), qui ne se distinguent pas entre eux.

**Tableau 3.** Résumé du modèle linéaire généralisé à effets mixtes relatif au comportement de la liquide (prononciation vs élision) pour les items lexicaux du corpus de conversation

Effets fixes	$\beta$	z
intercept	2,12	4,71***
pays::France	0,23	0,57
pays::Suisse	-1,68	-3,96***
vitesse d'articulation	-0,019	-8,56***
style::lecture	-1,74	-4,46***
âge	-0,01	-0,15
profil socio.::CSP+	0,26	0,62
contexte droit::pause	-2,36	-9,20***
contexte droit::voyelle	-4,01	-13,50***
style::lecture*âge	0,031	3,38***
style::lecture*profil socio.::CSP+	0,031	3,38***
<b>Effets aléatoires</b>		<b>Ecart-Type</b>
locuteurs	intercept	1,47
items	intercept	0,87

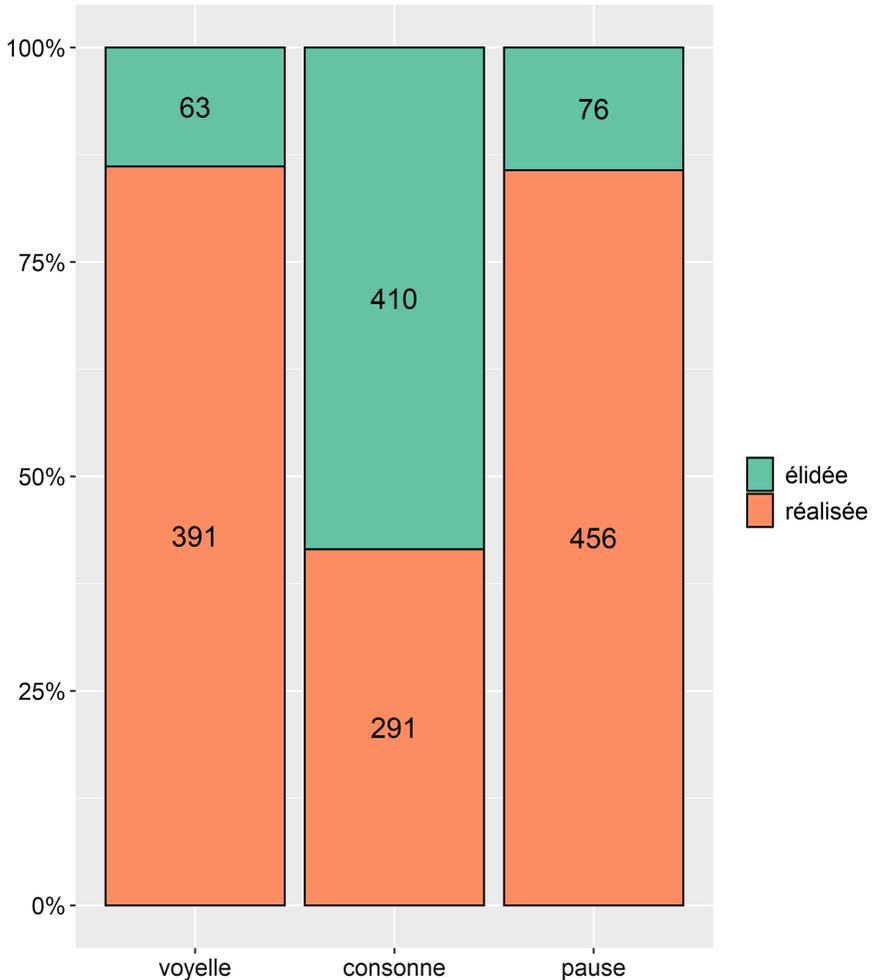
Note: \*\*\* indique que la probabilité ( $p$ ) est inférieure à 0,001.

On observe également un effet du style de parole sur le comportement de la variable dépendante : comme on l'a entraperçu sur la Figure 3 déjà commentée, en lecture, les liquides sont globalement plus souvent prononcées qu'en conversation (81,2% vs 42,5% d'élision). De façon intéressante, on observe toutefois que cet effet de style interagit avec l'âge des locuteurs. La Figure 8 montre que dans la tâche de lecture, plus les participants sont âgés, plus ils ont tendance à prononcer la liquide finale. En conversation à l'inverse, l'âge n'a strictement aucun impact sur le comportement de la variable à l'étude : le taux d'élision reste stable à travers les générations.

Dans le même ordre d'idée, le modèle fait ressortir l'existence d'une interaction significative entre le style et le profil socioéconomique des locuteurs. En conversation, avec respectivement 52,2% et 61,7% d'élision, les locuteurs CSP+ et CSP- ne se comportent pas différemment les uns des autres. En lecture en revanche, avec respectivement 12,5% et 25,1% d'élision, les CSP- produisent significativement moins de liquides que les CSP+ ( $p < 0,001$ ) (Figure 9).

## 5. DISCUSSION

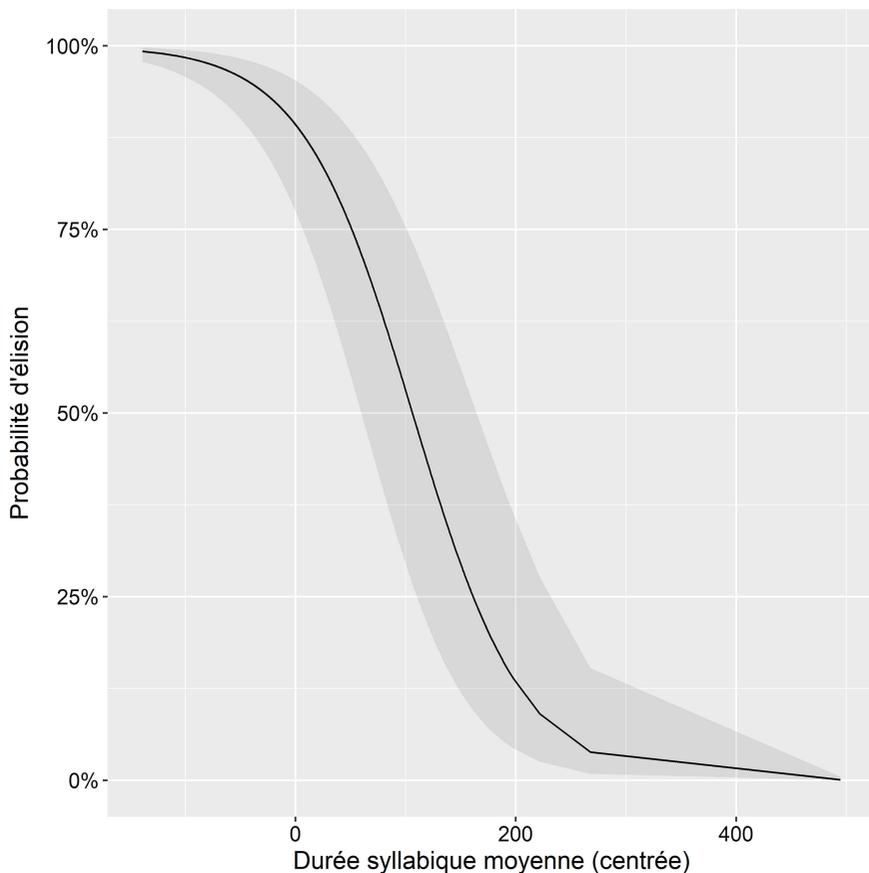
Sur la base d'un corpus comportant plus de 2500 unités analysées, nous avons pu dégager quelques grandes tendances permettant d'expliquer le comportement des liquides finales post-obstruantes en français.



**Figure 5.** Distribution (en %) des liquides dans le corpus, en fonction de leur comportement (élidee ou réalisée) et du contexte droit (suivies d'un mot qui commence par une voyelle ou une consonne, suivies d'une pause).

En ce qui concerne les mots grammaticaux, nos analyses sont restées descriptives, compte tenu de la rareté des occurrences dans le corpus analysé (238 items), si bien qu'il est difficile de tirer des conclusions fermes quant à leur comportement.

Pour les mots lexicaux en revanche, le corpus a permis dans un premier temps de comptabiliser la fréquence des différentes stratégies à disposition des locuteurs lorsqu'ils étaient confrontés à la violation du principe de sonorité séquentielle. Il est apparu que l'articulation d'un schwa ne se faisait jamais ou presque devant voyelle, et qu'elle semblait plus fréquente devant consonne en lecture qu'en conversation. L'articulation d'un schwa reposant sur des contraintes qui lui sont propres, nous avons préféré laisser de côté ce phénomène pour nous concentrer



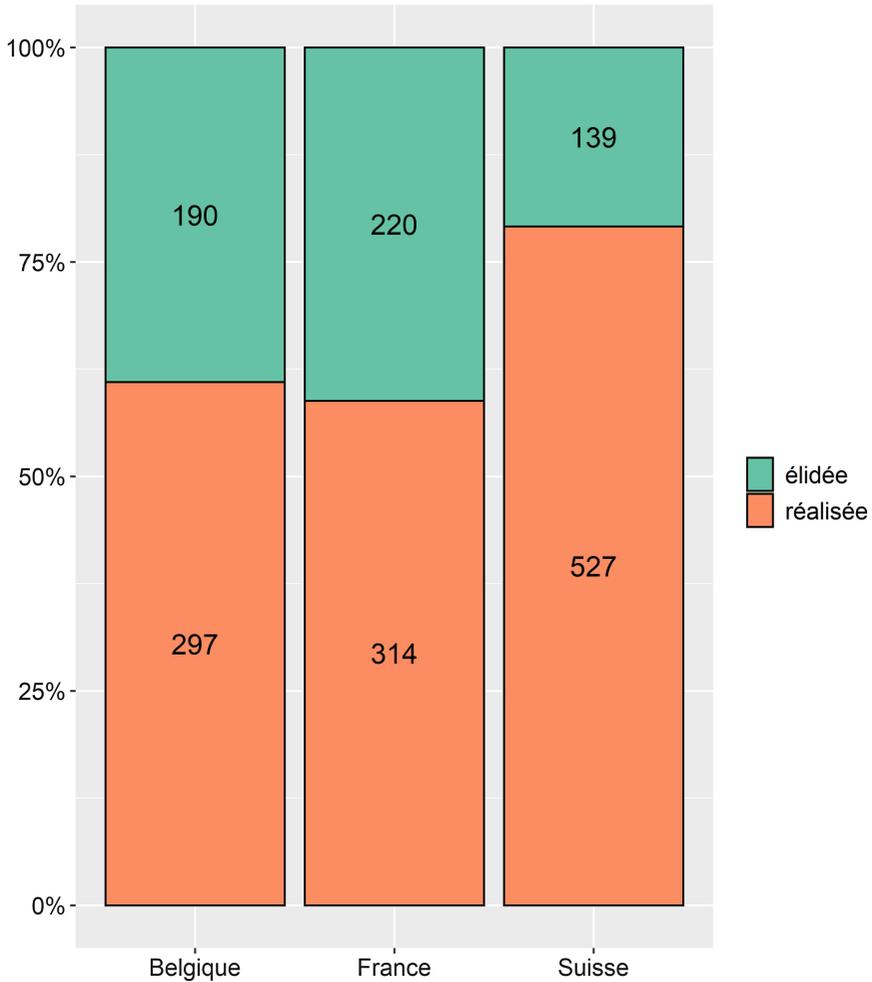
**Figure 6.** Probabilité d'élosion de la liquide dans le corpus en fonction de la vitesse d'articulation (en ms, les valeurs sont ici centrées); la zone grisée rend compte de l'écart-type.

sur les variables linguistiques et extralinguistiques qui peuvent expliquer la non-prononciation d'une liquide après une consonne obstruante en fin de mot.

Pour ce faire, nous avons conduit une analyse statistique au moyen de modèles généralisés à effets mixtes, qui permettaient de prendre en compte plus d'une douzaine de prédicteurs, ceux-là même qui ont été relevés dans la littérature. En pratique, nous avons examiné le rôle de prédicteurs relatifs aux propriétés phonétiques des LFPO (lieu et mode d'articulation des consonnes, types de cluster), à leur contexte (vitesse d'articulation, statut prosodique du mot hôte, contexte droit) et aux axes de la diavariation (diastatique, diatopique, diagénérationnelle, diagénique et diaphasique).

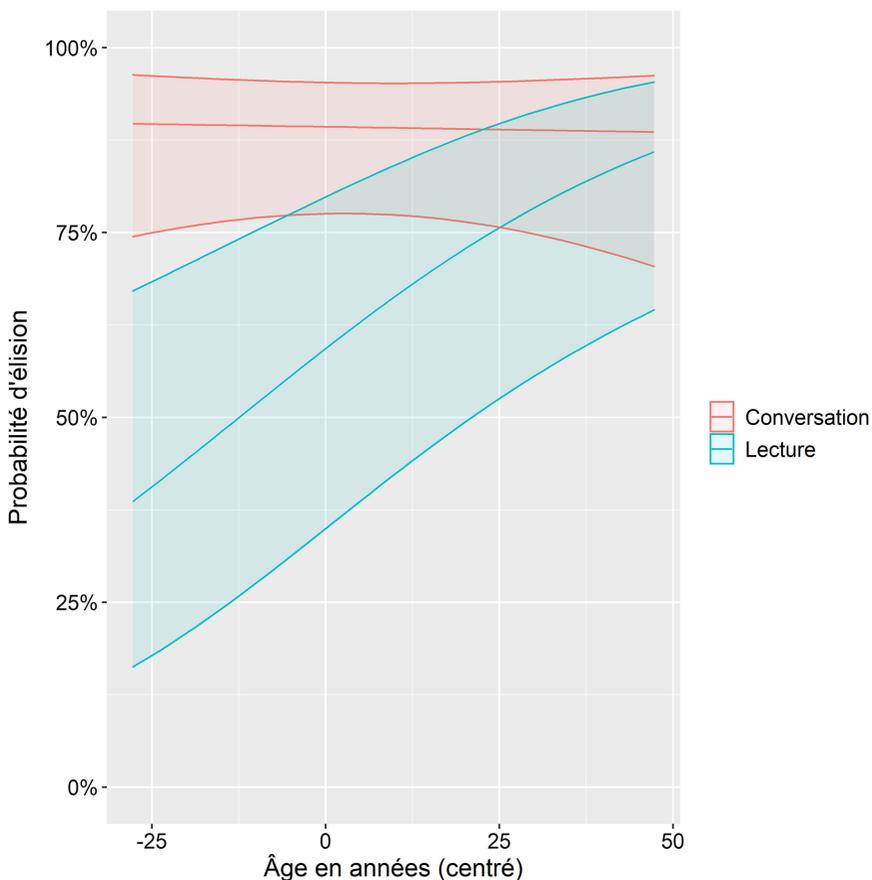
Du modèle effectué, il ressort que sur l'ensemble des prédicteurs analysés, la moitié d'entre eux se sont avérés jouer un rôle sur le comportement de la variable.

En ce qui concerne le contexte droit, on a vu qu'il n'y avait pas de différence entre les contextes vocalique et pausal, et que ce sont toujours les mots suivis par un mot à initiale consonnantique qui se différencient des mots suivis par une voyelle ou par



**Figure 7.** Distribution (en %) des liquides dans le corpus en fonction de leur réalisation (élidee ou réalisée) et du pays d'origine des participants.

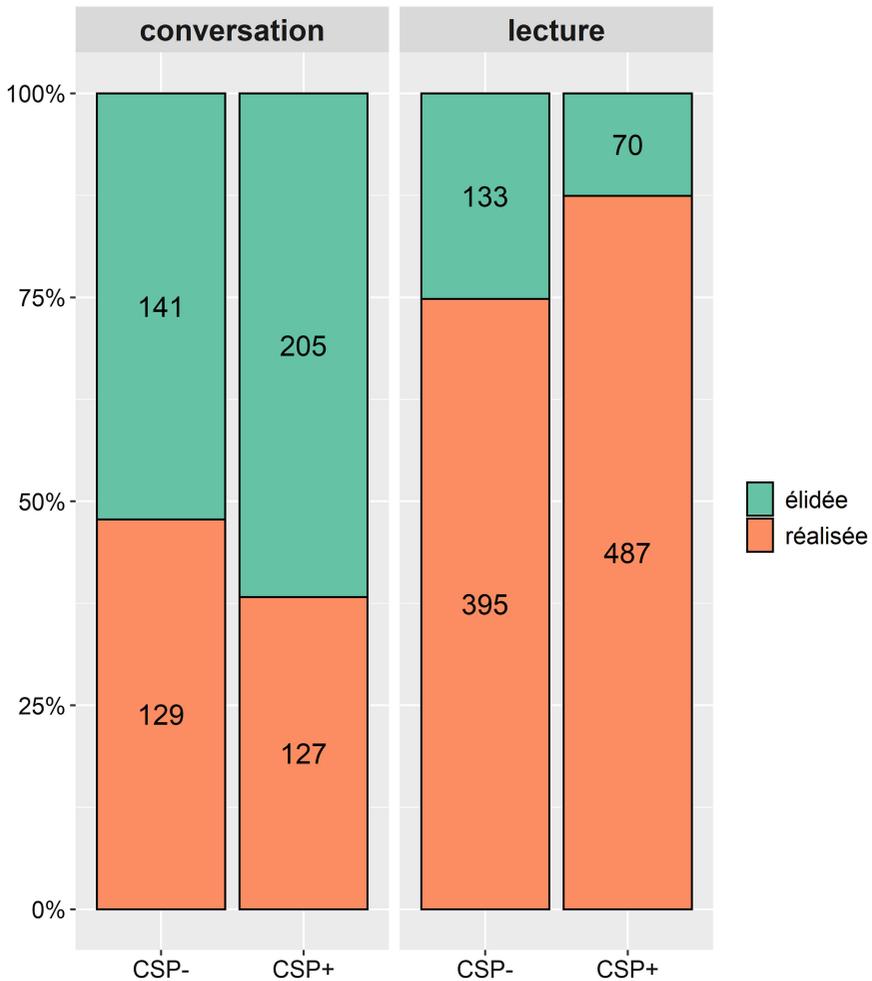
une pause. Ces résultats sont cohérents avec ce qui a été observé dans les recherches antérieures, la plupart des auteurs ayant conclu à l'importance d'une initiale consonantique par rapport à une initiale vocalique ou une pause dans le contexte droit de l'item hébergeant une LFPO (§2.2). En lecture comme en conversation, si la présence d'un mot à initiale vocalique favorise grandement l'articulation d'une liquide, on peut penser que c'est en raison de la possibilité de resyllabation qu'offre ce contexte (le syntagme *centre informatique* étant préférentiellement syllabé /sã.tʁẽ.fɔʁ.ma.tik/ que /sã.tʁẽ.ẽ.fɔʁ.ma.tik/. Quant aux contextes de pauses, ils s'avèrent coïncider avec les frontières droites des constituants les plus hauts dans la hiérarchie prosodique (constituants qu'on appelle 'syntagmes intonatifs', v. Nespor et Vogel, 1986; Delais-Roussarie, 1996).



**Figure 8.** Probabilité d'élosion de la liquide dans le corpus en fonction de l'âge des participants (en années, les valeurs sont ici centrées) et du style de parole (conversation ou lecture); les liserais autour des lignes rendent compte de l'écart-type.

Or, on sait que dans ces contextes, les segments sont plus souvent mieux articulés qu'ailleurs (Fougeron, 2001). Sur le plan perceptif, on comprend qu'il s'agit d'un contexte crucial, où l'absence d'une liquide se remarque davantage. Compte tenu du caractère profondément marqué socialement de l'élosion des LFPO, cela explique sans doute pourquoi les contextes pré-pausaux favorisent autant l'articulation des liquides que les voyelles.

Un second paramètre s'étant avéré significatif concerne la vitesse d'articulation. Là aussi, ces résultats ne sont pas surprenants, dans la mesure où il a été prouvé, dans toutes les études précédentes sur les LFPO qui en ont tenu compte, que plus les locuteurs parlent vite et moins ils ont tendance à articuler les LFPO. Ce résultat confirme donc ce qui a été montré dans les recherches récentes sur la LFPO (§2.4), mais aussi que le comportement des LFPO ne diffère pas fondamentalement des autres phénomènes phonotactiques phares de la



**Figure 9.** Distribution (en %) des liquides dans le corpus de conversation (à gauche) et dans le corpus de lecture (à droite), en fonction de leur comportement (élidee ou réalisée) et du profil socio-économique des participants (CSP- ou CSP+).

phonologie, comme le schwa (Bürki et al., 2011) et la liaison (Adda-Decker et al., 2012), tout aussi sensibles à ce paramètre.

Troisièmement, nos analyses ont fait ressortir un effet du pays d'origine des locuteurs. On savait que les locuteurs antillais et québécois avaient davantage tendance à simplifier les mots comportant une LFPO par rapport aux locuteurs de la partie septentrionale de la France (Pustka, 2007, 2011; Milne, 2014), en raison pour les uns de l'influence du contact avec le créole, pour les autres de la distance plus grande qu'ils entretiennent avec la norme hexagonale. On savait aussi que les locuteurs du sud de la France, parce qu'ils produisent davantage de schwas, faisaient beaucoup moins chuter les LFPO que ceux de la partie septentrionale de la France (Pustka, 2007, 2011; Myers et Ranson, 2014). Notre

étude permet de compléter nos connaissances du comportement des LFPO à l'échelle de la francophonie. Plus spécifiquement, il en ressort que les Suisses produisent davantage de LFPO que les Belges et les Français. On aurait pu penser que cette singularité des locuteurs suisses puisse être expliquée en raison d'un autre paramètre, la vitesse d'articulation, l'idée étant que les Suisses parlent plus lentement que les Français et les Belges (comme le montrent Schwab et Avanzi, 2015, qui ont travaillé sur le même corpus que nous avons utilisé dans cet article), et donc qu'ils articulent davantage de segments. L'absence d'interaction significative entre le pays et la vitesse d'articulation nous enjoint pourtant à exclure cette hypothèse. Une autre explication consisterait à dire que les Suisses se surveillent plus, par rapport aux Belges ou aux Français, quand ils parlent. Mais cela reste encore à prouver. Une comparaison avec les productions de locuteurs originaires de France voisine (Savoie, Franche-Comté) permettrait sans doute également d'y voir plus clair, et de confirmer que la tendance à articuler davantage les liquides post-obstruante est un marqueur phonétique propre au français de Suisse romande.

Enfin, on a vu que s'il existait des différences entre la lecture et la conversation, celles-ci dépendaient également du profil socioéconomique des locuteurs. À nos yeux, le résultat n'est pas si surprenant, car les locuteurs que nous avons catégorisés comme des CSP- ont a priori une moins bonne connaissance des normes de la lecture à voix haute par rapport aux locuteurs CSP+. Les CSP+ étant plus habitués à la pratique de la langue écrite, ils ont produit des enregistrements plus soignés et mieux articulés que les autres. Quant à savoir pourquoi cette différence ne se retrouve pas en conversation, c'est plus problématique en regard de ce qui avait été observé dans les travaux antérieurs. Nous proposerons deux explications, pas forcément incompatibles l'une avec l'autre. La première explication est que l'écart entre les CSP- et les CSP+ dans notre corpus n'est pas assez marqué, en d'autres termes que les différences sociales entre les locuteurs ne sont pas assez nettes. La seconde réside dans le fait que l'élimination des liquides n'est peut-être pas un marqueur sociolinguistique de classe aussi puissant que l'on n'aurait pu le croire en conversation, du moins dans les situations d'entretiens du corpus PFC, ceux-ci élicitant une parole relativement finalement peut-être plus contrôlée ? Il serait utile, pour trancher en faveur de cette seconde interprétation, de tester le comportement d'autres marqueurs sociolinguistiques, morpho-phoniques (liaisons, chutes de <l> dans les pronoms sujets *il* et *ils*, absence du *ne* de négation, etc., v. Léon et Tennant, 1990, *inter alia*) dans le corpus.

Enfin, on a vu que les différences que l'on observe entre la conversation et la lecture ne sont pas les mêmes selon l'âge des locuteurs. Pour mémoire, il est apparu qu'en lecture, plus les participants étaient âgés, plus ils tendaient à articuler leurs liquides, alors qu'en conversation, aucun effet d'âge ne se faisait sentir sur la (non-)prononciation de la liquide. Ici encore, on ne peut pas invoquer la vitesse d'articulation pour expliquer cette différence. D'une part, on n'observe pas d'interaction entre le style et la vitesse d'articulation dans notre modèle. D'autre part, on sait que les francophones âgés articulent moins vite que les plus jeunes, mais que le style de parole n'a pas d'influence sur ce paramètre :

en d'autres termes, la différence entre les participants âgés et les participants plus jeunes est stable entre la lecture et la conversation (Schwab et Avanzi, 2015). On serait donc tenté de voir ici un changement en cours dans les habitudes de lecture des francophones. De nos jours, les jeunes liraient moins soigneusement que les plus âgés.

Quant aux autres prédicteurs souvent invoqués dans la bibliographie sur le sujet, bon nombre ne sont pas ressortis comme pertinents pour décrire la variable à l'étude. Le fait d'avoir exclu de l'analyse les mots grammaticaux, qui se placent majoritairement dans le même groupe accentuel que le mot qu'ils précèdent, a sans doute permis de faire ressortir le caractère non pertinent de ce paramètre (ce qui n'est guère étonnant finalement, les frontières de groupes accentuels ne bloquant pas forcément les phénomènes de sandhis comme la liaison ou l'enchaînement, v. Avanzi et al., 2011). Pour les autres paramètres, les différences entre notre méthode et celles que d'autres ont suivies peuvent expliquer les différences. Dans notre étude, nous avons travaillé à partir de données alignées et vérifiées manuellement, ce qui n'est pas forcément le cas de nombreux chercheurs ayant publié avant les années 2010 (la technologie n'étant pas forcément aussi accessible qu'aujourd'hui). Or, quand on connaît la difficulté à identifier avec précision la présence ou non d'un segment dans la chaîne parlée, ce détail a son importance. Par ailleurs, l'utilisation de modèles statistiques élaborés, permettant de prendre en compte un grand nombre de paramètres dans un seul et même modèle, en ayant pris soin de tester auparavant le degré de multicollinéarité des facteurs, permet d'expliquer également les différences entre les résultats de recherches antérieures et les nôtres. Une explication, au sujet de l'absence de significativité de la fréquence lexicale, réside dans le format des modèles que nous avons utilisés. Ces modèles qui contiennent des effets 'mixtes' permettent de tenir compte aussi bien de la variation intra- et interlocuteurs que de la distribution des items dans le corpus. De fait, si ce paramètre de fréquence lexicale a pu ressortir comme significatif dans d'autres études, c'est parce que lesdites études ne tenaient pas compte du fait que certains locuteurs produisent plus de mots que d'autres, et que certains mots fréquents sont aussi plus souvent élidés que d'autres (v. Burki et al., 2011 pour des conclusions similaires en regard du schwa).

Enfin, pour terminer, signalons que cette étude comporte un certain nombre de limitations. Malgré la taille de notre corpus et le soin apporté à la sélection des items analysés, celui-ci n'a pas permis de faire de statistiques sur les mots grammaticaux. Or la non-prononciation des liquides finales post-obstruantes semble plus fréquente sur ce genre de mot : l'examen de ce genre de données doit être réalisé pour compléter les résultats présentés ici. Par ailleurs, le nombre de locuteurs (8 par point d'enquête) ne nous a pas permis de comparer plus finement les effets de région, certains auteurs ayant trouvé des différences entre l'est et l'ouest de la partie septentrionale de la France (Boughton, 2013) ou à l'intérieur de la Wallonie (Hambye, 2005). Un autre problème est que nous avons abordé la question du comportement comme une variable binaire (élision/articulation), alors qu'il est clair que le problème doit également tenir compte des phénomènes de réduction à l'œuvre dans la parole continue. Comme le montrent Burki et al. (2011) sur le schwa, et plus récemment Brand & Ernestus

(2021) sur les liquides, quand un segment est présent, il n'est pas toujours réalisé de la même façon, certaines formes étant plus raccourcies que d'autres. Reste également la question du schwa, stratégie alternative qui semble davantage utilisée en lecture qu'en conversation. L'analyse de ce paramètre, au moyen des mêmes prédicteurs pris en compte dans ce travail, devra être entreprise.

## 5. Conclusion

En conclusion, nous avons présenté dans cet article une analyse multifactorielle de la non-prononciation des liquides post-obstruantes finales de mot en français, un phénomène par lequel un <l> ou <r> final dans la transcription orthographique d'un mot tombe dans le flux de la parole, comme dans les items *découvre* et *terrible*, prononcés /dekuv/ et /teʁib/ et non /dekuvʁ/ et /teʁibl/. Plus de 2500 items contenant un cluster obstruante+liquide en fin de mot ont été extraits d'un corpus de parole de 13 heures, étiqueté à différents niveaux et comprenant des productions de 120 locuteurs originaires de 3 pays francophones (Belgique, France et Suisse) et enregistrés en deux tâches différentes (lecture et conversation). Les résultats obtenus ont permis de porter un regard nouveau sur les paramètres qui motivent l'élision des liquides post-obstruantes finales de mots. Nous avons pu voir que parmi les variables phonotactiques, lexicales, prosodiques, sociolinguistiques et stylistiques, le contexte droit (mot suivi d'une consonne, d'une voyelle ou d'une pause), la vitesse d'articulation, l'origine géographique et le profil socioéconomique, tout comme le style de parole (lecture vs conversation) permettaient d'expliquer en grande partie les phénomènes d'élision de LFPO en français. L'étude a laissé de côté un certain nombre de faits, comme le caractère binaire ou graduel des phénomènes de facilité de prononciation ou la présence d'un schwa. Ces facteurs devront être pris en compte dans le futur.

**Remerciements.** Nous sommes très reconnaissants envers Sandrine Brognaux, George Christodoulides, François Delafontaine, Jean-Philippe Goldman et Sandra Schwab pour leur aide précieuse dans la transcription, le codage et l'extraction des données. Nous tenons également à remercier chaleureusement les trois évaluateurs de la revue pour leurs commentaires constructifs sur une version antérieure de cet article. Leur expertise a été d'une grande aide pour améliorer la qualité de notre travail.

**Competing interests.** The author declares none.

**Conflits d'intérêts.** L'auteurs n'en déclare aucun.

## Références

- Adda-Decker, M., Delais-Roussarie, E. Fougeron, C., Gendrot, C. et Lamel L. (2012). La liaison dans la parole spontanée familière : une étude sur grand corpus. *Revue française de linguistique appliquée*, 16: 113–128.
- Armstrong, N. (1993). A study of phonological variation in French secondary school pupils. Thèse de doctorat, Université de Newcastle.
- Armstrong, N. (2001). *Social and stylistic variation in spoken French: a comparative approach*. Amsterdam/ New York : John Benjamins.
- Avanzi, M. (2013). Note de recherche sur l'accentuation et le phrasé à la lumière des corpus de français. *Travaux neuchâtelois de linguistique*, 59, 5–24.

- Avanzi, M.** (2014). A Corpus-Based Approach to French Regional Prosodic Variation. *Nouveaux cahiers de linguistique française*, 31 : 309–332.
- Avanzi, M., Bordal, G. et Obin, N.** (2011). Typological variations in the realization of French Accentual Phrase. In : *Proceedings ICPHS, Hong Kong*, 268–271.
- Bardiaux, A.** (2014). La prosodie de quelques variétés de français en Belgique : analyse perceptive et acoustique. Thèse de doctorat, Université catholique de Louvain.
- Boersma, P.** (2001). Praat, a system for doing phonetics by computer. *Glott International*, 5 (9/10): 341–345.
- Boughton, Z.** (2013). Locating variation in French: Geolinguistic patterns, Levelling and the ‘French Exception’. In : M. C. Jones et D. Hornsby (eds), *Language and Social Structure in Urban France*. Oxford : Legenda, pp. 119–132.
- Boughton, Z.** (2015). Social class, cluster simplification and following context: Sociolinguistic variation in word-final post-obstruent liquid deletion in French. *Journal of French Language Studies*, 25 : 1–21.
- Brand, S. et Ernestus, M.** (2021). Reduction of word-final obstruent-liquid schwa clusters in Parisian French. *Corpus Linguistics and Linguistic Theory*, 17 (1) : 249–285.
- Brooks, M. E., Kristensen, K., Benthem, K. J. van, Magnusson, A., Berg, C. W., Nielsen, A., et al.** (2017). Modeling Zero-Inflated count data with glmmTMB, *bioRxiv*, 132753.
- Brognaux, S. et Drugman, Th.** (2014). Variations phonétiques : impact de la situation de communication. *Nouveaux cahiers de linguistique française*, 31 : 63–76.
- Bürki, A., Ernestus, M., Gendrot, C., Fougeron, C. et Frauenfelder, U.** (2011). What affects the presence versus absence of schwa and its duration: a corpus analysis of French connected speech. *Journal of Acoustical Society of America*, 130 : 3980–3991.
- Bybee, J.** (2002). Word frequency and context of use in the lexical diffusion of phonetically conditioned sound change. *Language Variation and Change*, 14 : 261–290.
- Chambers, J. K.** (2009). *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford, Wiley-Blackwell.
- Chambers, J. K. et Trudgill, P.** (1998). *Dialectology*. Cambridge, Cambridge University Press.
- Chevrot, J.-P., Beaud, L. et Varga, R.** (2000). L'apprentissage des unités phonologiques variables : l'exemple du /R/ post-consonantique final en français. *LINX*, 42 : 89–109.
- Christodoulides, G. et Avanzi, M.** (2015). Automatic Detection and Annotation of Disfluencies in Spoken French Corpora. In : *Proceedings Interspeech*, Dresden, pp. 1849–1853.
- Christodoulides, G., Avanzi, M. et Goldman, J.-P.** (2014). DisMo: A Morphosyntactic, Disfluency and Multi-Word Unit Annotator. An Evaluation on a Corpus of French Spontaneous and Read Speech. In : *Proceedings LREC*, Rejkjavic, pp3902–3907.
- Côté, M.-H.** (2004). Consonant cluster simplification in Québec French. *Probus*, 16 : 151–201.
- de Cornulier, B.** (1978). Syllabe et suite de phonèmes en phonologie du français. In : B. de Cornulier et F. Dell (éds), *Études de phonologie française*, Paris, Éditions du CNRS, pp. 31–70.
- Delais-Roussarie, E.** (1996). Phonological Phrasing and Accentuation in French. In : M. Nespor et N. Smith (éds), *Dam Phonology*, La Haye, Holland Academic Graphics, pp. 1–38.
- Dell, F.** (1985). Consonant clusters and phonological syllables in French. *Lingua*, 95 : 5–26.
- Durand, J., Laks, B., Calderone, B. et Tchobanov, A.** (2011). Que savons-nous de la liaison aujourd'hui ? *Langue Française*, 169: 103–135.
- Durand, J., Laks, B. et Lyche, Ch.** (éds). (2009). *Phonologie, variation et accents du français*. Paris, Hermes.
- Eychenne, J. et B. Laks** (2017). La liaison en français contemporain : normes, usages, acquisitions. *Journal of French Language Studies*, 27 : 1–12.
- Fougeron, C.** (2001). Articulatory properties of initial segments in several prosodic constituents in French. *Journal of Phonetics* 29 : 109–135.
- Gadet, F.** (1989). *Le français ordinaire*. Paris, Armand Colin.
- Garde, P.** (1965). *L'accent*. Paris, PUF.
- Goldman, J.-P.** (2011). EasyAlign: an Automatic Phonetic Alignment Tool under Praat. *Proceedings Interspeech*, Firenze, pp. 3233–326.
- Gradoville, M.** (2019). The role of individual variation in variationist corpus-based studies of priming. *Italian Journal of Linguistics*, 31 (1) : 93–124.
- Guiraud, P.** (1965). *Le français populaire*. Paris, PUF.
- Hambye, P.** (2005). La prononciation du français en Belgique. Variation, normes et identités. Thèse de doctorat, Université catholique de Louvain.

- Hansen, A. B.** (2015). Sensibilité et insensibilité devant la variation phonétique : une étude perceptive sur le français de la région parisienne. *Langage et société*, **61** : 45–65.
- Hartig, Fl.** (2020). DHARMA: Residual diagnostics for hierarchical (multi-level/mixed) regression models. <https://CRAN.R-project.org/package=DHARMA>.
- Hornsby, D.** (2006). *Redefining Regional French: Koinéization and Dialect Levelling in Northern France*. Oxford, Legenda.
- James, G., Witten, D., Hastie, T. et Tibshirani, R.** (eds). (2013). *An introduction to statistical learning: with applications in R*. New York : Springer.
- Kemp, W., Pupier, P. et Yaeger, M.** (1980). A linguistic and social description of final consonant cluster simplification in Montreal French. In : R. Shuy et A. Shnukal (éds), *Language Use and the Uses of Language*. Washington, DC, Georgetown University Press, pp. 12–40.
- Labov, W.** (1966). *The Social Stratification of English in New York City*. Washington DC, Center for Applied Linguistics.
- Laks, B.** (1977). Contribution empirique à l'analyse socio-différentielle de la chute des /r/ dans les groupes consonantiques finals. *Langue Française*, **34** : 109–125.
- Lenth, R.** (2016). Least-squares means: The R package lsmeans. *Journal of Statistical Software*, **69** : 1–33.
- Léon, P. et Tennant, J.** (1990). 'Bad French' and nice guys: A morphophonetic study. *The French Review*, **63** (5) : 763–778.
- Lüdecke D., Ben-Shachar M., Patil I., Waggoner P. et Makowski, D.** (2021). performance: An R Package for Assessment, Comparison and Testing of Statistical Models. *Journal of Open Source Software*, **6**(60), 3139.
- Martinon, P.** (1913). *Comment prononce-t-on le français ?* Paris, Larousse.
- McCulloch, C. E. et Neuhaus, J. M.** (2001). *Generalized linear mixed models*. Wiley Online Library.
- Milne, P.** (2014). The variable pronunciations of word-final consonant clusters in a force aligned corpus of spoken French. Thèse de doctorat, Université d'Ottawa.
- Myers, E. L. et Ranson, D. L.** (2014). L'élision variable des /r/ et /l/ postconsonantiques finals en français méridional et septentrional : L'effet de l'âge, du lieu d'origine et des facteurs linguistiques. In : *Actes CMLF*, Berlin, pp. 1345–1364.
- Nespor, M. et Vogel, I.** (1986). *Prosodic Phonology*. Dordrecht, Foris.
- New, B., Pallier, C., Ferrand, L. et Matos, R.** (2001). Une base de données lexicales du français contemporain sur internet : LEXIQUE. *L'année psychologique*, **101** : 447–462.
- Peperkamp, S., Hegde, M. et Carbajal, M. J.** (2019). Liquid deletion in French child-directed speech. In : *Proceedings Interspeech*, Graz, pp. 3574–3578.
- Peperkamp, S. et Iturralde Zurita, A. M.** (2019). Compensation for French liquid deletion during auditory sentence processing. In : *Proceedings Interspeech*, Graz, pp. 1951–1955.
- Pooley, T.** (1996). *Chimi: The Urban Vernaculars of Northern France*. Clevedon, Multilingual Matters.
- Pustka, E.** (2007). *Phonologie et variétés en contact. Aveyronnais et Guadeloupéens à Paris*. Tübingen, Narr.
- Pustka, E.** (2011). Le conditionnement lexical de l'élision des liquides en contexte post-consonantique final. *Langue Française*, **169** : 19–38.
- R Development Core Team** (2013). *R: A language and environment for statistical computing*. Vienna, R Foundation for Statistical Computing.
- Ranson, D. L.** (2010). L'élision variable des /r/ et /l/ finals postconsonantiques en français méridional : l'interaction des facteurs lexicaux et diastatiques. In : *Actes CMLF*, La Nouvelle-Orléans, pp. 1341–1356.
- Sankoff, D., Tagliamonte, S. A. et Smith, E.** (2005). *Goldvarb X: A variable rule application for Macintosh and Windows*. Department of Linguistics, University of Toronto.
- Schwab, S. et Avanzi, M.** (2015). Regional Variation and Articulation Rate in French. *Journal of Phonetics*, **48** : 96–105.
- Selkirk, E.** (1984). *Phonology and Syntax: The relation between Sound and Structure*. Cambridge, MIT Press.
- Thibault, A.** (1997). *Dictionnaire suisse romand*. Carouge, Éditions Zoé.
- Thurot, C.** (1883). *De la prononciation française depuis le commencement du XVIIe siècle : d'après les témoignages des grammairiens*. Genève, Slatkine reprints.
- Tranel, B.** (1987). *The sounds of French: An introduction*. Cambridge, Cambridge University Press.
- Trudgill, P.** (1974). *The Social Differentiation of English in Norwich*. Cambridge, Cambridge University Press.

- Villeneuve, A.-J.** (2010). Word-final cluster simplification in Vimeu French: A preliminary analysis. *University of Pennsylvania Working Papers in Linguistics*, **15** (2) : 133–144.
- Wang, S. S. Y.** (1969). Competing changes as cause of residue. *Language*, **45** : 9–25.
- Wu, Y.** (2018). Étude de la réduction segmentale en français parlé à travers différents styles : apports des grands corpus et du traitement automatique de la parole à l'étude du schwa, du /ʁ/ et des réductions à segments multiples. Thèse de doctorat, Université Paris Sorbonne Nouvelle.
- Zuur, A., Ieno, E. et Elphick, Ch.** (2010). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, **1**(1), 3–14.

---

**Cite this article:** Avanzi M (2023). *Vot' artic' est formidab'*: une étude multifactorielle de la chute des liquides post-obstruantes finales de mot en français. *Journal of French Language Studies* **33**, 137–167. <https://doi.org/10.1017/S0959269523000029>