



HAL
open science

Caractérisation acoustique des réalisations approximantes du /v/ intervocalique en français spontané

Suyuan Dong, Nicolas Audibert

► To cite this version:

Suyuan Dong, Nicolas Audibert. Caractérisation acoustique des réalisations approximantes du /v/ intervocalique en français spontané. 35èmes Journées d'Études sur la Parole (JEP 2024) 31ème Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2024) 26ème Rencontre des Étudiants Chercheurs en Informatique pour le Traitement Automatique des Langues (RECITAL 2024), Jul 2024, Toulouse, France. pp.13-22. hal-04623055

HAL Id: hal-04623055

<https://inria.hal.science/hal-04623055v1>

Submitted on 1 Jul 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Caractérisation acoustique des réalisations approximantes du /v/ intervocalique en français spontané

Suyuan Dong¹ Nicolas Audibert¹

(1) Laboratoire de Phonétique et Phonologie (CNRS & Sorbonne Nouvelle),

4 rue des Irlandais, 75005 Paris, France

suyuan.dong@sorbonne-nouvelle.fr, nicolas.audibert@sorbonne-nouvelle.fr

RESUME

Les fricatives /v/ ont tendance à se réaliser comme une variante affaiblie en français spontané. Nous sommes appuyés sur les données du corpus NCCFr et avons évalué 5504 occurrences de /v/ intervocalique produites par 10 hommes et 10 femmes, à partir de l'inspection des spectrogrammes. Un ensemble de mesures acoustiques dynamiques ont été relevées sur chaque exemplaire, et comparées entre les deux types de réalisation (fricatives voisées et approximantes) par des modèles GAM pour prendre en compte leur évolution temporelle. Les résultats montrent une prépondérance des réalisations approximantes, également observables en positions accentuées, et indiquent que ces deux types de réalisation divergent tant au niveau temporel que spectral, particulièrement dans les hautes fréquences. La manifestation de cet affaiblissement varie également en fonction des locuteurs. Ces observations permettent de suggérer que le /v/ intervocalique serait sujet à un processus d'affaiblissement en cours, au-delà d'une simple réduction segmentale.

ABSTRACT

Acoustic characterization of approximant realizations of intervocalic /v/ in spontaneous French

The fricative /v/ tends to manifest as a weakened variant in spontaneous French. This study relies on data from the NCCFr corpus and evaluated 5504 occurrences of intervocalic /v/ produced by 10 males and 10 females, through the inspection of spectrograms. A set of dynamic acoustic measurements was taken on each exemplar and compared between the two types of realization (voiced fricative and approximant) using GAM models to account for their temporal evolution. The results show a predominance of approximant realizations, also observable in accentuated position, and indicate that these two types of realization diverge in both temporal and spectral dimension, particularly in the high frequencies. The manifestation of this weakening also varies according to speaker differences. These observations suggest that the intervocalic /v/ might be subject to an ongoing weakening process, beyond a simple segmental reduction.

MOTS-CLÉS : fricative ; approximante ; labiodentale ; français spontané ; grand corpus ; acoustique

KEYWORDS : fricative; approximant; labiodental; spontaneous French; large corpora; acoustic

1 Introduction

La variabilité phonétique est omniprésente dans la langue orale. La lénition désigne le processus phonologique par lequel un segment devient moins similaire à sa réalisation initiale ([Trask, 2006, p.190](#)). Sous une perspective synchronique, la lénition se manifeste par des alternances sonores au sein d'une langue dans une période donnée, transformant un son en un allophone « plus faible » ([ibid., p.201](#); [Kirchner, 2004, p.313](#)). Diachroniquement, elle peut entraîner une convergence progressive entre deux phonèmes, jusqu'à une perte de contraste phonologique ([Trask, 2000, p.216](#)). Selon [Trask \(ibid., p.191\)](#) et [Crystal \(2008, p.274\)](#), la lénition peut engendrer des changements sonores tels que Occlusive > Fricative > Approximante > Zéro. La lénition est aussi décrite comme un phénomène positionnel, souvent observé en position faible, telles qu'en intervocalique ou coda ([De Carvalho et al., 2008, pp.131-172](#); [Jatteau et al., 2019](#); [Lancien et al., 2023](#)).

Bien que les descriptions acoustiques pour les fricatives labiodentales fassent l'objet de nombreuses études, les recherches se concentrant sur leurs réalisations affaiblies sont moindres. Cependant, notre analyse préliminaire du corpus NCCFr (*The Nijmegen Corpus of Casual French*, [Torreira et al., 2010](#)) révèle que ce phénomène est couramment observé dans le français spontané. Par conséquent, cette étude cherche à examiner spécifiquement les /v/ intervocaliques, et à caractériser acoustiquement leurs réalisations affaiblies en tant qu'approximantes, afin de déterminer si ce phénomène résulte d'une simple réduction segmentale en parole spontanée ou s'il pourrait s'agir d'un potentiel phénomène de lénition en cours en français spontané parmi la jeune génération.

Du point de vue articulaire, les approximantes se distinguent des fricatives et des voyelles par le degré de fermeture entre des articulateurs ([Hewlett & Beck, 2006, pp.37-39](#)), ainsi que par le niveau d'implication du bruit de friction ([Crystal, 2008, p.32](#); [Trask, 2006, p.30](#)). Les fricatives sont définies comme des consonnes produites par un rétrécissement étroit et incomplet entre deux articulateurs, générant un bruit de friction audible (e.g., [Kent & Read, 2001, p.38](#); [Vaissière, 2020, p.70](#)). À l'inverse, les approximantes sont produites lorsque le conduit vocal est nettement rétréci, sans pour autant créer une constriction suffisamment étroite pour générer un flux d'air turbulent ([Ladefoged & Johnson, 2010, p.15](#); [Kent & Read, 2001, p.177](#)). Cette constriction modérée les distingue des voyelles, résultant en une stabilité moindre et une énergie acoustique plus faible. Sur le plan aérodynamique, les fricatives voisées combinent deux sources d'énergie : le voisement et le bruit de friction. Cette dualité génère un conflit aérodynamique : la génération du bruit turbulent nécessite une vitesse d'air élevée, tandis que les vibrations des plis vocaux tendent à ralentir le passage du flux d'air. En conséquence, les fricatives sonores ont tendance à perdre leur bruit de friction et ainsi se réaliser comme des approximantes ([Johnson, 2012, p.156](#)).

À partir de ces caractéristiques articulaires et aérodynamiques, les fricatives sont généralement décrites selon quatre attributs acoustiques : les propriétés spectrales et l'amplitude du bruit de friction, la durée, ainsi que les propriétés spectrales des transitions formantiques liées aux voyelles adjacentes ([Reetz & Jongman, 2009, pp.227-230](#)). Diverses méthodes ont été proposées pour mesurer et différencier les fricatives, incluant, sans s'y limiter, le pic de fréquence, les moments spectraux, la transition formantique, l'amplitude intégrale et dynamique, ainsi que la durée (e.g., [Jongman et al., 2000](#); [Maniwa et al., 2009](#); [Al-Tamimi & Khatib, 2015](#)). Récemment, la quantification de l'énergie dans les hautes fréquences (>7kHz) pour décrire les fricatives a suscité un regain d'attention ([Shadle et al., 2023a](#); [Kharlamov et al., 2023](#)).

Les descriptions acoustiques des approximantes sont pour leur part moins nombreuses. Malgré leurs similitudes avec les voyelles, elles présentent une structure formantique moins intense et moins

stable, accompagnée d'une durée plus courte que celles des voyelles ([Reetz & Jongman, 2009, p.225](#)). En comparaison avec les fricatives voisées, compte tenu de l'absence de turbulence, les approximantes devraient présenter une structure formantique plus marquée et stable, un signal sonore moins bruité, une amplitude supérieure mais avec moins d'énergie dans les hautes fréquences, ainsi qu'un voisement plus constant que les fricatives voisées. De plus, si ces réalisations en tant qu'approximantes correspondent simplement à une réduction due à l'hypoarticulation et à la rapidité de la parole spontanée, elles devraient présenter une durée réduite et être rares en position accentuée ([Adda-Decker & Snoeren, 2011](#) ; [Ernestus & Warner, 2011](#) ; [Lindblom, 1990](#)).

2 Méthodologie

2.1 Corpus

Le corpus NCCFr ([Torreira et al., 2010](#)) a été élaboré pour fournir des enregistrements de parole informelle et spontanée en français, adaptés aux investigations scientifiques. Cette base de données rassemble plus de 36 heures d'enregistrements, produits par 46 jeunes locuteurs natifs du français (22 femmes et 24 hommes, âgés principalement de 18 à 27 ans). Les participants, socio-géographiquement homogènes, ont été enregistrés en binômes d'amis pour capturer leurs conversations spontanées. Les enregistrements ont été recueillis avec un enregistreur stéréo à semi-conducteur Edirol R-09, équipé de microphones unidirectionnels Samson QV et d'un préamplificateur, avec une fréquence d'échantillonnage de 48kHz. Les données collectées ont ensuite été transcrites et annotées manuellement en suivant les directives développées par LIMSI, puis segmentées par alignement automatique forcé répété dans le même système. Pour cette étude, nous avons sélectionné un sous-ensemble de 20 locuteurs avec un total de 5504 occurrences du phonème /v/ intervocalique. Ces réalisations se répartissent de manière relativement équilibrée au sein des locuteurs, comprenant 2970 /v/ produits par 10 femmes et 2534 par 10 hommes.

2.2 Annotation des données

Nous avons commencé par la classification des différentes réalisations du /v/ intervocalique à l'aide de scripts Praat. Ces scripts permettent d'extraire automatiquement des informations du son cible, telles que la durée, les voyelles adjacentes et les informations lexicales. Les différents types de réalisation du /v/ ont été manuellement catégorisés à partir de l'observation du spectrogramme dans les bandes de fréquence de 0-5kHz (structure formantique) et 0-15kHz (bruit de friction), résultant en 9 catégories : fricative voisée, approximante, élision, dévoisement, voix craquée, voix soufflée, superposition, signal faible et autres. La catégorie « autres » regroupe les sons non-exploitable tels que les rires, les erreurs de transcription, les cas d'élision d'une des voyelles adjacentes, ou encore la saturation du signal due à la voix criée. Cette catégorisation s'est basée sur des indices acoustiques visuellement et auditivement identifiés, principalement la structure formantique, la transition formantique, le bruit de friction, la barre de voisement, la périodicité et la réduction d'amplitude. Ces indices ont été notés durant la catégorisation. Des critères complémentaires ont aussi été pris en compte, incluant les voyelles bruitées ou dévoisées et différentes catégories de positions prosodiques fortes : la « position accentuée » correspondant à la prééminence à la dernière syllabe des groupes rythmiques (hésitations exclues), l'« insistance » où le locuteur accentue un mot pour un objectif pragmatique (focus), et la « syllabe accentuée » qui englobe les deux catégories précédentes, mais

incluant cette fois-ci les cas d'hésitation. Lors de cette étape, nous avons affiné manuellement la position des frontières des sons cibles. Toutes les catégorisations et segmentations ont été revérifiées.

2.3 Analyses acoustiques

Étant donné que l'objectif principal de cette étude est d'identifier le potentiel contraste acoustique entre les réalisations modales et affaiblies du /v/ intervocalique et de les caractériser, nous avons sélectionné 30 mesures proposées dans la littérature, et les avons extraites à l'aide d'un autre script Praat. Ces mesures incluent la durée du /v/, les moments spectraux, l'intensité prise au milieu des /v/ pour illustrer la forme spectrale générale, l'amplitude dynamique dans différentes gammes de fréquences, le HNR, les formants F1 à F4 convertis en Bark, ainsi que les mesures spécifiques aux non-sibilantes dans les hautes fréquences (<7kHz) proposées par [Shadle et al. \(2023a\)](#). Il est à préciser que ces dernières suivent le paramétrage proposé par les auteurs, mais sont implémentées sur un spectre DFT au lieu de la méthode *multitaper*, ce qui selon [Kharlamov et al. \(2023\)](#) n'impacte pas la classification des fricatives dans l'analyse de la parole continue.

Pour rendre compte de la dynamique de ces mesures, nous avons retenu onze points de mesure équidistants répartis sur la durée totale des réalisations de /v/, complétés par cinq points de mesures sur la seconde moitié de la voyelle précédente (V1), et cinq points sur la première moitié de la voyelle suivante (V2). Les paramètres ont été adaptés pour tenir compte des différences entre hommes et femmes pour les mesures formantiques (méthode de Burg avec une fréquence maximale respectivement de 5kHz et 5.5kHz pour la détection de 5 formants). Les mesures des moments spectraux s'appuient sur les paramètres utilisés par [Al-Tamimi & Khattab \(2015\)](#). La forme du spectre de 0 à 14kHz avec une résolution de 20Hz a également été extraite au milieu du /v/.

2.4 Analyses statistiques

Nous avons d'abord comparé la distribution des différentes réalisations du /v/ et des indices acoustiques relevés. Par la suite, nous avons concentré notre analyse sur les réalisations en tant que fricatives voisées ou approximantes. L'effet du type de réalisation et du sexe des locuteurs sur la durée (transformée en log) a été évalué à l'aide d'un modèle linéaire mixte (fonction `lmer` du package `R lme4` ([Bates et al., 2015](#))) avec le locuteur comme ordonnée à l'origine (*intercept*) aléatoire.

La trajectoire des mesures acoustiques sur 21 points entre le milieu de la voyelle précédente V1 et le milieu de la suivante V2 a été modélisée par des modèles additifs généralisés (GAM), séparément pour hommes et femmes, incluant comme prédicteurs le type de réalisation, la durée du /v/, ainsi que celles des V1 et V2. Les différences entre locuteurs ont été considérées comme lissage aléatoire. Suite à une première analyse qui ne révèle pas d'interaction significative entre le contexte vocalique et le type de réalisation, le contexte a été intégré comme facteur aléatoire dans ces modèles. La forme spectrale au milieu du /v/ a également été modélisée par GAM séparément pour hommes et femmes avec les mêmes prédicteurs et facteurs aléatoires, cette fois avec une modélisation en fonction de la fréquence et non du temps.

3 Résultats

3.1 Distribution des types de réalisation : effet du locuteur et du sexe

Parmi les 5504 occurrences du /v/ intervocalique examinées, les réalisations comme fricatives voisées (39.9%, n=2197) et comme approximantes (28.8%, n=1583) sont les plus fréquentes. La Figure 1 présente la distribution de chaque type de réalisation en fonction des locuteurs (10F+10H). En outre, 6.6% (n=365) des /v/ correspondent à des superpositions de parole, en raison de la nature interactive des dialogues. Quant aux réalisations en fonction du sexe, les résultats suggèrent que les locuteurs masculins ont produit plus d'approximantes (33.4%, n=846) que les locutrices (24.8%, n=737). Les élisions présentées ici correspondent aux cas où aucune trace visuelle ni audible n'est présente pour indiquer la réalisation du /v/, en complément des cas déjà pris en compte dans l'annotation initiale du corpus.

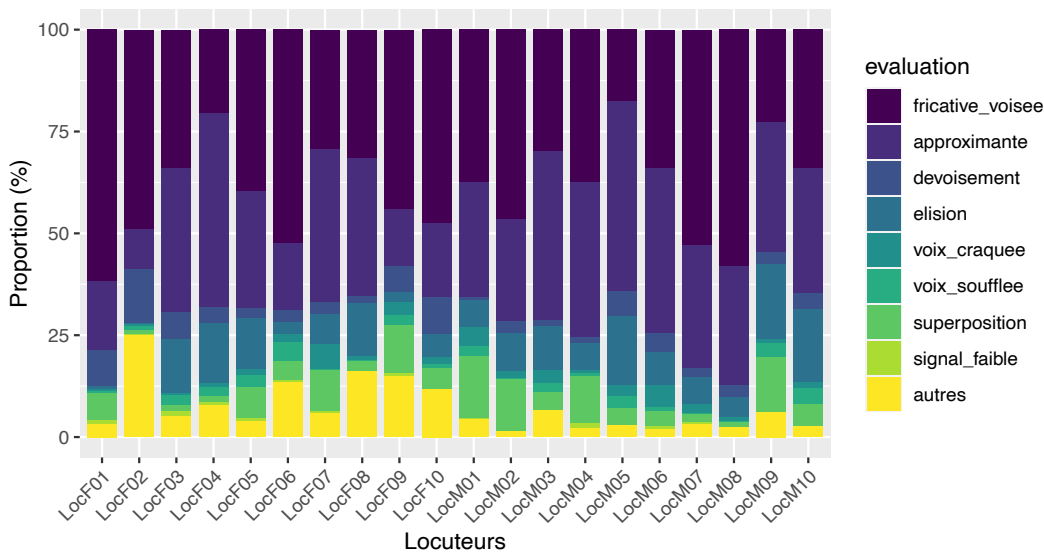


FIGURE 1 : Répartition générale des différents types de réalisations par locuteur/trice. Le sexe des participants est indiqué par « LocF » pour les femmes, et « LocM » pour les hommes.

Les indices acoustiques directement observables dans les spectrogrammes et le signal sonore sont globalement conformes aux descriptions de la littérature. Le bruit de friction et la structure formantique constituent les principaux indicateurs pour identifier visuellement les deux types de réalisation (fricatives voisées et approximantes). Une réduction d'amplitude relative aux voyelles adjacentes est plus fréquemment observée chez les fricatives voisées. Quant aux positions prosodiques, 21.8% (1200/5504) des /v/ ont été réalisés en position forte. Parmi ces réalisations, nous avons observé une présence relativement importante d'approximantes, avec 37% (311/839) en position accentuée et 11% (20/186) en focus, contre respectivement 57% (482/839) et 81% (151/186) de fricatives voisées. Une variabilité individuelle a également été constatée dans la production des approximantes en position accentuée, avec par exemple LocM09 à 71% contre seulement 2% pour LocF02. En moyenne sur l'ensemble des 20 locuteurs, 38% ($\sigma=18\%$) des /v/ en position accentuée sont réalisés comme approximantes.

3.2 Résultats d'analyses acoustiques

La comparaison des durées (Figure 2, gauche) suggère que les approximantes tendent à être plus courtes que les fricatives voisées pour les deux sexes, mais un recouvrement important est présent. Afin d'éviter un éventuel impact des différences individuelles de débit de parole, nous avons normalisé la durée par locuteur (Figure 2, droite). Une différence significative de durée est bien observée en fonction du type de réalisations ($p < 2.2e-16$), mais sans effet significatif du sexe des locuteurs ($p = 0.52$), ni d'interaction entre le sexe et le type de réalisation ($p = 0.71$).

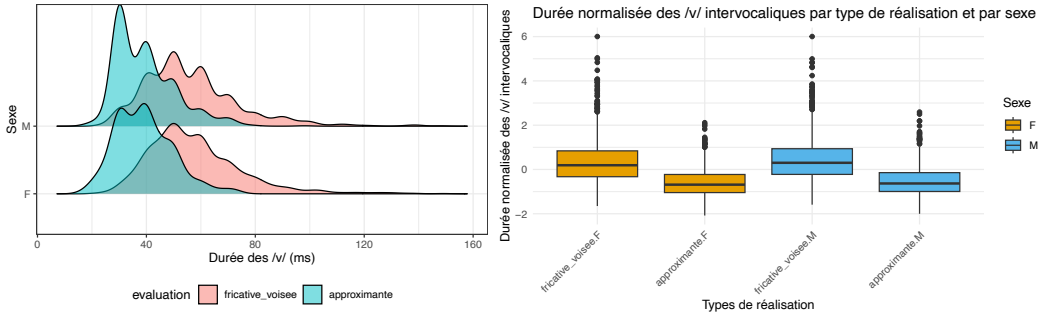


FIGURE 2 : Distribution des durées des sons cibles en fonction du sexe et du type de réalisations, avec l'échelle d'origine (gauche) et après normalisation par locuteur (droite).

Pour chacun des 60 modèles GAM correspondant aux mesures acoustiques (30 mesures * 2 sexes), les valeurs prédites au cours du temps ont été extraites pour les deux types de réalisation avec les intervalles crédibles associés. La Figure 3 illustre l'écart maximum entre fricatives voisées et approximantes prédit par les modèles pour chacune de ces mesures, après normalisation en z-scores. Parmi elles, nous pouvons distinguer les huit mesures sur la gauche de la figure qui discriminent le mieux les deux types de réalisation, avec une divergence maximum supérieure à un écart-type pour les deux sexes. Ces mesures incluent l'intensité relative (IntRel) dans les moyennes et hautes fréquences ($11-13\text{kHz} > 5-8\text{kHz} > 4-5.6\text{kHz}$), conformément aux observations sur la forme générale des spectres au milieu des /v/ non détaillées ici. Cette divergence entre fricatives et approximantes est également capturée par le centre de gravité spectral (CoG), la dispersion spectrale (SD), ainsi que les trois mesures (HiLevD, AmpRange, LevelH) proposées par [Shadle et al. \(2023a\)](#). Parmi ces trois dernières mesures, HiLevD et AmpRange sont suggérées par les auteurs comme les meilleurs indicateurs d'augmentation du bruit de turbulence pour les non-sibilants.

Le CoG et la SD font partie des mesures les plus couramment utilisées pour caractériser les fricatives. Un CoG élevé désigne une articulation plus antérieure, tandis qu'un SD élevé indique une dispersion spectrale plus large (e.g., [Forrest et al., 1988](#); [Al-Tamimi & Khattab, 2015](#)). Les résultats semblent correspondre à ce qui est attendu : les fricatives voisées présentent un spectre plus diffus et probablement un lieu d'articulation plus avancé que les approximantes ; ces dernières présentent une distribution spectrale plus similaires aux voyelles environnantes. Cependant, le CoG et SD peuvent être impactés par le degré de voisement, ainsi que par l'énergie dans les fréquences 11-13kHz et 5-8kHz comme le suggèrent les corrélations avec l'énergie relative dans ces bandes de fréquence (H : $\rho(\text{CoG}, 11-13\text{k}) = 0.67$, $\rho(\text{SD}, 11-13\text{k}) = 0.76$, $\rho(\text{CoG}, 5-8\text{k}) = 0.65$, $\rho(\text{SD}, 5-8\text{k}) = 0.66$; F : $\rho(\text{CoG}, 11-13\text{k}) = 0.74$, $\rho(\text{SD}, 11-13\text{k}) = 0.77$, $\rho(\text{CoG}, 5-8\text{k}) = 0.69$, $\rho(\text{SD}, 5-8\text{k}) = 0.62$). Parmi ces huit mesures, les corrélations en valeur absolue sont comprises entre 0.37 ($\rho(\text{CoG}, \text{AmpRange})$) et 0.8 ($\rho(\text{CoG}, \text{SD})$).

Le changement spectral dans les hautes fréquences est estimé par HiLevD qui mesure la différence d'intensité entre trois bandes de fréquence du spectre lissé : LevM (3-7kHz), LevH(7-11k) et LevHH(11-15k) (Shadle *et al.*, 2023a). Cette mesure de pente spectrale en hautes fréquences est suggérée comme l'indicateur le plus sensible de l'augmentation des énergies dans les hautes fréquences au milieu des fricatives. Nos résultats illustrent son efficacité pour distinguer les réalisations de /v/, marquées par une diminution notable pour les fricatives voisées et une augmentation pour les approximantes. Conformément aux observations initiales des auteurs sur les fricatives sourdes, la valeur minimale de cette mesure est très proche du milieu des fricatives /v/ (non illustré ici).

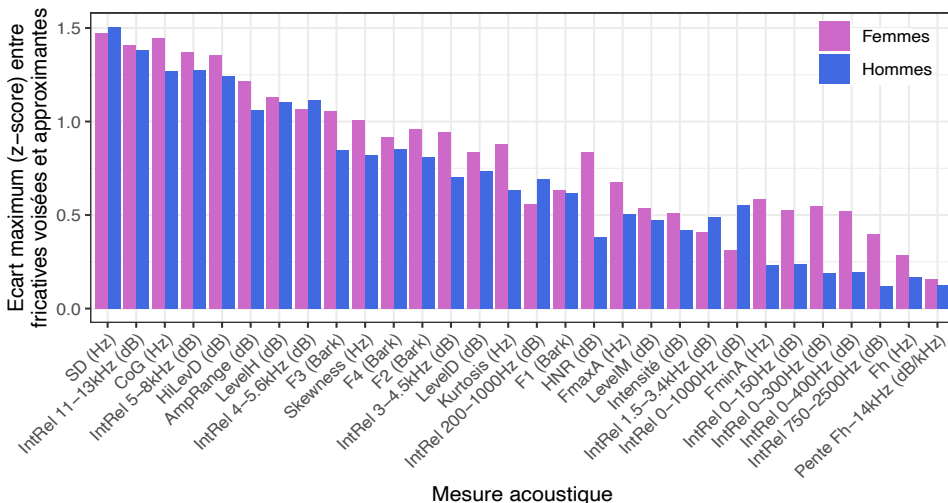


FIGURE 3 : Distance maximum modélisée par les GAM entre fricatives voisées et approximantes pour les 30 mesures acoustiques (normalisées en z-scores) en fonction du sexe des locuteurs.

La mesure AmpRange, quant à elle, est définie comme la différence d'amplitude (dB) entre le maximum en hautes fréquences (FmaxA : 2-13kHz) et le minimum en basses fréquences (FminA : 1-7kHz). Selon les auteurs, cette mesure correspond en partie à la réduction de la taille de constriction au fil du temps, dans le sens où plus la constriction est importante, plus la valeur de cette mesure est importante. Nos résultats indiquent bien une réduction de FminA et une augmentation de FmaxA qui correspondraient à une constriction plus importante pour les fricatives voisées, mais l'amplitude relative des variations sur ces mesures a pour conséquence des valeurs non conformes aux prédictions pour la mesure AmpRange.

4 Discussion

Cette étude a impliqué l'analyse acoustique de 30 paramètres, étendus jusqu'à 14kHz, pour étudier les réalisations des /v/ intervocaliques. Les résultats pour la plupart de ces mesures suggèrent que les /v/ intervocaliques ont tendance à se réaliser comme une variante affaiblie distincte des fricatives voisées. Malgré un recouvrement partiel entre gammes de fréquences capturée par les huit mesures les plus discriminantes, elles ne sont pas directement corrélées, à l'exception de relations modérées entre CoG et SD, aussi entre ces mesures et l'intensité relative dans les bandes 5-8kHz et 11-13kHz.

Cette distinction entre fricatives voisées et approximantes se manifeste d'abord au niveau spectral. Les principaux résultats acoustiques sont consistants avec les observations issues des études précédentes, indiquant un bruit de friction diffus notamment dans les hautes fréquences pour les fricatives voisées par rapport aux approximantes ([Shadle et al., 2023a](#); [Kharlamov et al., 2023](#)). Pour la mesure AmpRange, en revanche, l'inconsistance entre nos résultats et ceux de l'étude de [Shadle et al. \(2023a\)](#) pourrait être due au voisement présent dans nos données, aux différences de contexte (exclusivement intervocalique dans notre cas), voire au style de parole (spontanée dans nos données, parole lue et mots isolés dans celles de [Shadle et al.](#)).

En outre, nos modèles GAM (non illustrés ici) indiquent que la divergence est maximale en un point temporel très proche du milieu du /v/ pour toutes les mesures à l'exception de l'énergie relative sur les bandes de fréquence 5-8kHz et 11-13kHz (divergence maximale légèrement plus tardive, entre 58,5% et 63% de la durée totale de la consonne), et de la mesure Fh qui est toutefois peu discriminante entre types de réalisations. Pour la majorité des mesures acoustiques prises en compte, nous pourrions donc nous appuyer sur des mesures prises au milieu de la consonne afin de caractériser les réalisations fricatives ou approximantes dans nos travaux ultérieurs. Ce constat ouvre la voie à une extension à plus grande échelle de ces analyses, avec une application envisageable à des données pour lesquelles seul un alignement forcé automatique est disponible.

En outre, les données analysées proviennent de la parole spontanée caractérisée par des conversations informelles et relâchées dans un environnement calme. Ce style de parole tend à favoriser une production réduite, typiquement marquée par des sons plus courts et moins articulés ([Adda-Decker & Snoeren, 2011](#) ; [Ernestus & Warner, 2011](#)). Bien que nos résultats indiquent des durées significativement réduites pour les réalisations approximantes, le chevauchement important des distributions des durées suggère que les approximantes ne sont pas systématiquement plus courtes que les fricatives voisées. Tandis que les positions prosodiques accentuées impliquent souvent un effort articulatoire plus élevé ([Cho, 2016](#)), elles donnent lieu dans les données analysées à des réalisations approximantes relativement fréquentes et même majoritaires chez certains locuteurs. Ces observations suggèrent que les réalisations approximantes pourraient ne pas constituer de simples réductions segmentales, mais être plutôt une réalisation allophonique liée à la lénition. Dans cette perspective et compte tenu du fait que les données du corpus NCCFr ont été enregistrées en 2007, il serait intéressant de les comparer à des productions spontanées de locuteurs d'Ile de France de la même tranche d'âge en 2024 afin d'évaluer l'évolution de ce phénomène.

Nos résultats suggèrent également un effet du sexe sur la distinction acoustique entre ces deux types de réalisation. Par exemple, la distinction entre fricatives et approximantes est également identifiable dans les basses fréquences (0-400Hz) chez les femmes, sans pour autant être le cas chez les hommes, alors que la fréquence fondamentale est supposée incluse dans cette plage de fréquence dans les deux cas. L'observation qualitative des données suggère aussi une variabilité individuelle au-delà des importantes différences de distribution des types de réalisation, ainsi qu'une convergence au cours du temps au sein des binômes qui sera approfondie dans des recherches ultérieures.

Remerciements

Ce travail a été soutenu par le Partenariat Hubert Curien (PHC) Van Gogh n°49298RM et par le Laboratoire d'Excellence (LabEx) Empirical Foundations of Linguistics (EFL) n°ANR-10-LABX-0083. Il contribue à l'IdEx Université de Paris (ANR-18-IDEX-0001).

Références

- ADDA-DECKER, M. & SNOEREN, N. D. (2011). Quantifying temporal speech reduction in French using forced speech alignment. *Journal of Phonetics*, 39(3), 261-270.
- BATES D, MÄCHLER M, BOLKER B & WAKLER S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67(1), 1–48. DOI : [10.18637/jss.v067.i01](https://doi.org/10.18637/jss.v067.i01).
- AL-TAMIMI J. & KHATTAB G. (2015). Acoustic cue weighting in the singleton vs geminate contrast in Lebanese Arabic: The case of fricative consonants. *The Journal of the Acoustical Society of America*, 138(1), 344-360. DOI : [10.1121/1.4922514](https://doi.org/10.1121/1.4922514).
- CHO, T. (2016). Prosodic boundary strengthening in the phonetics–prosody interface. *Language and Linguistics Compass*, 10(3), 120-141.
- CRYSTAL D. (2008). *A dictionary of linguistics and phonetics* (6^e éd). Blackwell Pub.
- DE CARVALHO J. B., SEGERAL P., SCHEER T., ÉdS. (2008). *Lenition & fortition*. Mouton de Gruyter.
- ERNESTUS, M. & WARNER, N. (2011). An introduction to reduced pronunciation variants. *Journal of Phonetics*, 39(SI), 253-260.
- HEWLETT N. & BECK J. M. (2006). *An Introduction to the Science of Phonetics*. Routledge.
- FORREST K., WEISMER G., MILENKOVIC P. & DOUGALL R. N. (1988). Statistical Analysis of word-initial voiceless obstruents: Preliminary data. *Journal of the Acoustical Society of America*, 84, 115-123. DOI : [10.1121/1.396977](https://doi.org/10.1121/1.396977).
- JATTEAU A., VASILESCU I., LAMEL L., ADDA-DECKER M. & AUDIBERT N. (2019). “Gra[f]e!” Word-Final Devoicing of Obstruents in Standard French: An Acoustic Study Based on Large Corpora. *Interspeech 2019*, 1726-1730. DOI : [10.21437/Interspeech.2019-2329](https://doi.org/10.21437/Interspeech.2019-2329).
- JOHNSON K. (2012). *Acoustic and Auditory Phonetics*. Wiley-Blackwell.
- JONGMAN A., WAYLAND R. & WONG S. (2000). Acoustic characteristics of English fricatives. *The Journal of the Acoustical Society of America*, 108(3), 1252-1263. DOI : [10.1121/1.1288413](https://doi.org/10.1121/1.1288413).
- KENT R. D. & READ C. (2001). *Acoustic Analysis of Speech (2nd edition)*. Cengage Learning.
- KHARLAMOV V., BRENNER D. & TUCKER B. V. (2023). Examining the effect of high-frequency information on the classification of conversationally produced English fricatives). *The Journal of the Acoustical Society of America*, 154(3), 1896-1902. DOI : [10.1121/10.0021067](https://doi.org/10.1121/10.0021067).
- KIRCHNER R. (2004). Consonant lenition. In HAYES B., KIRCHNER R. & STERIADE D., ÉdS., *Phonetically Based Phonology* (1^{re} éd.), p. 313-345. Cambridge University Press. DOI : [10.1017/CBO9780511486401.010](https://doi.org/10.1017/CBO9780511486401.010).
- LADEFOGED P. & JOHNSON K. (2010). *A Course in Phonetics* (6^e éd). Wadsworth.
- LANCIEN M., HUTIN M., STUART-SMITH J., ADDA-DECKER M. & VASILESCU I. (2023). /R/ Lenition in Quebec French: Evidence from the Distribution of 9 Allophones in Large Corpora. *20th International Congress of Phonetic Sciences (ICPhS)*. Prague, Czech Republic. ISBN : [9788090811423](https://doi.org/9788090811423).
- LINDBLOM, B. (1990). Explaining phonetic variation: A sketch of the H&H theory. In *Speech production and speech modelling* (pp. 403-439). Dordrecht: Springer Netherlands.
- MANIWA K., JONGMAN A. & WADE T. (2009). Acoustic characteristics of clearly spoken English fricatives. *The Journal of the Acoustical Society of America*, 125(6), 3962-3973. DOI : [10.1121/1.2990715](https://doi.org/10.1121/1.2990715).
- REETZ H. & JONGMAN A. (2009). *Phonetics Transcription, Production, Acoustics, and Perception*. Wiley-Blackwell.

- SHADLE C. H., CHEN W. R., KOENIG L. L. & PRESTON J. L. (2023a). Refining and extending measures for fricative spectra, with special attention to the high-frequency range. *The Journal of the Acoustical Society of America*, 154(3), 1932-1944. DOI : [10.1121/10.0021075](https://doi.org/10.1121/10.0021075).
- TORREIRA F., ADDA-DECKER M. & ERNESTUS M. (2010). The Nijmegen Corpus of Casual French. *Speech Communication*, 52(3), 201-212. DOI : [10.1016/j.specom.2009.10.004](https://doi.org/10.1016/j.specom.2009.10.004).
- TRASK R. L. (2000). *The Dictionary of Historical and Comparative Linguistics*. Edinburgh University Press.
- TRASK R. L. (2006). *A dictionary of phonetics and phonology* (Reprinted). Routledge.
- VAISSIÈRE J. (2020). *La Phonétique* (4^e éd). Que sais-je ? Presses Universitaires de France.