

Les voyelles fermées tendues, relâchées et allongées du français québécois : la contribution d'indices statiques/dynamiques et absolus/normalisés à la détermination de leur identité acoustique

Caroline Sigouin, Vincent Arnaud

Département des arts et lettres, Université du Québec à Chicoutimi, 555, boulevard de l'Université,
Chicoutimi (Québec), G7H 2B1, CANADA

caroline.sigouin@uqac.ca, vincent.arnaud@uqac.ca

RÉSUMÉ

L'objectif de cette contribution est de déterminer les indices acoustiques discriminant le plus efficacement les variantes tendues, relâchées et allongées des voyelles fermées du français québécois. 1350 occurrences, extraites de la parole de 30 locuteurs des villes de Rouyn-Noranda, Saguenay et Québec, ont été analysées. Leur durée a été relevée, puis la fréquence fondamentale et la fréquence centrale des trois premiers formants (F_1 , F_2 , F_3) ont été estimées à 25-50-75 % de la durée. Les valeurs fréquentielles en Hertz et en scores-Z ont été considérées. Les résultats d'analyses discriminantes témoignent de la contribution de chaque indice acoustique, de la dynamique spectrale et de la normalisation à la répartition des occurrences au sein de leur classe vocalique d'appartenance. Alors que F_1 et F_2 , en Hertz, relevés au centre des voyelles permettent d'obtenir une classification correcte à 58 %, d'autres combinaisons de prédicteurs parviennent à des taux s'élevant jusqu'à 97 %.

ABSTRACT

Quebec French tense, lax and lengthened close vowels: The contribution of static/dynamic and raw/normalized cues to the determination of their identity

This study aims to estimate the contribution of various acoustic cues to the determination of tense, lax and lengthened variants of Quebec French close vowels identity. To do so, we analyzed the speech of 30 speakers from Rouyn-Noranda, Saguenay and Quebec cities who produced 1350 tokens of the variants under study. Their duration has been measured then the fundamental frequency and the central frequency of the first three formants (F_1 , F_2 , F_3) have been estimated at 25, 50 and 75% of this duration. Raw Hertz frequencies and Z-Scores transformed ones have been exploited. Discriminant analyses results show that each acoustic cue contributes in a manner to close vowels variants distinction. While mid-point Hertz values of F_1 and F_2 lead to an overall correct classification rate of 58%, other combinations of predictors may achieve as much as 97% of correct classification.

MOTS-CLÉS : voyelles fermées, français québécois, dynamique spectrale, normalisation.

KEYWORDS: close vowels, Quebec French, spectral change, normalization.

1 Introduction

1.1 Variantes phonétiques des voyelles fermées en français québécois

En français québécois (désormais FQ), les voyelles fermées /i/, /y/, /u/, dont les formes canoniques sont brèves et tendues, sont fréquemment soumises à différents phénomènes contextuels, notamment le relâchement, l’allongement, la diphtongaison, le dévoisement, la réduction et la syncope. Comme le remarque McLaughlin (1986), en observant la distribution de ces diverses réalisations, on constate qu’en syllabe non accentuée, différents phénomènes sont susceptibles de se produire (voire de se combiner) et que leur application est toujours facultative. Au contraire, les variantes présentes sous l’accent se distribuent catégoriquement dans trois contextes complémentaires : les variantes brèves tendues apparaissent en syllabe ouverte (abrégé V#), les variantes relâchées, en syllabe fermée par une consonne non allongeante (VK), et les variantes allongées (potentiellement diphtonguées), en syllabe fermée par une consonne allongeante (VR).¹ Par exemple, les mots *boue*, *boude* et *bouge* sont respectivement prononcés [bu], [bød] et [bu:ɜ] ou [b^uuɜ] en FQ. C’est à ces trois derniers types de variantes (sous accent) que cette contribution est consacrée.

1.2 Caractéristiques acoustiques : un bref état de la question

Si par définition les variantes allongées présentent des durées plus importantes que les tendues et les relâchées, le rapport de durée entre ces deux derniers types de réalisations, lui, ne fait pas l’unanimité. En effet, dans certains travaux ayant abordé la question du relâchement des voyelles fermées en FQ (par exemple Dumas, 1974), aucune différence de durée n’a été relevée entre les variantes tendues et relâchées. Dans d’autres études, les variantes tendues sont rapportées comme plus longues que les relâchées (Gendron, 1966 ; Martin, 2002 ; Poliquin, 2006). Arnaud et coll. (2011) ont pourtant observé le rapport de durée inverse. Cela dit, cette discordance peut être imputée aux différences méthodologiques entre ces études. En effet, il paraît peu probable qu’une analyse impressionniste comme celle de Dumas (1974) soit à même de révéler de subtiles différences de durée entre voyelles brèves. Par ailleurs, les relâchées seraient plus longues que les tendues lorsque suivies d’une consonne voisée (comme c’est le cas dans Arnaud et coll. 2011), mais plus courtes devant une non-voisée – ce qui est systématiquement le cas dans Poliquin (2006) et majoritairement le cas dans Martin (2002). Les variantes tendues et relâchées ne se distingueraient donc pas intrinsèquement par leur durée (Sigouin, 2013).

En ce qui a trait aux caractéristiques spectrales des voyelles fermées du FQ, on sait depuis Gendron (1966) que les variantes relâchées ont un F_1 plus élevé que celui des tendues (sans pour autant atteindre celui des voyelles mi-fermées homorganiques) et que les relâchés apparaissent comme légèrement centralisés (F_2 plus bas pour /i/ et /y/, plus élevé pour /u/) par rapport à leurs contreparties tendues. Arnaud et coll. (2011) ont de plus remarqué que dans un espace F_1 / F_2 , les variantes relâchées se centralisaient en cours d’émission alors que les tendues se déplaçaient en périphérie ; du moins lorsque ces segments étaient insérés dans des logatomes de forme /bV/ et /bVb/. Leur étude révèle également la pertinence de la prise en compte de la F_0 , de F_1 , de F_2 , de F_3 et de leur évolution en cours d’émission dans la répartition des occurrences dans leur classe vocalique attendue.

¹ Nous ne tiendrons pas compte ici des mots, somme toute marginaux, faisant exception à cette distribution, comme *cool* [ku:l] ou *quiz* [kwiz] (Côté, 2012 : 242).

En ce qui concerne les variantes allongées, leur timbre est sujet à débat et leurs caractéristiques spectrales ont été peu étudiées, généralement sous un angle ne permettant que difficilement la comparaison avec les variantes brèves. Sous l'accent, les voyelles fermées allongées sont-elles tendues (Gendron, 1966 ; Paradis, 1985), relâchées devant /ʁ/ mais tendues ailleurs (Côté, 2012), toujours diphtonguées (Dumas, 1974), parfois diphtonguées (Yaeger, 1979 ; Leblanc, 2012 ; Côté, 2012) ou ne présentent-elles qu'une variation de timbre (Dionne, 1971 ; Santerre et Millo, 1978) ? Au plan acoustique, certains chercheurs (Gendron, 1966 ; Santerre, 1971 ; Poliquin, 2006) n'ont considéré qu'un point de mesure par occurrence (ou une moyenne de plusieurs points), permettant difficilement de rendre compte d'une éventuelle diphtongaison. D'autres (Yaeger, 1979 ; Paradis, 1985) ont opté pour deux points de mesure uniquement lorsque les voyelles étaient jugées diphtonguées (et très peu de voyelles fermées – voire aucune – ont effectivement été jugées diphtonguées). Pour leur part, MacKenzie et Sankoff (2010) n'ont pris en compte qu'un point de mesure pour les voyelles brèves, mais deux pour les allongées, rendant hasardeuse toute comparaison entre ces deux classes vocaliques.

1.3 Objectifs

Les objectifs de cette étude sont a) de proposer une analyse acoustique unifiée des variantes tendues, relâchées et allongées des voyelles fermées du français québécois, b) d'estimer la contribution de la durée, de la F_0 , de F_1 , F_2 , F_3 , et de l'évolution temporelle de ces composantes spectrales à la classification d'occurrences de voyelles fermées dans leur classe phonétique attendue d'après leur distribution en syllabe accentuée, telle que proposée dans la littérature : (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK) ou (uR), et c) d'évaluer l'impact de l'application d'une procédure de normalisation aux valeurs fréquentielles sur cette même classification.

2 Méthodologie

2.1 Échantillon et corpus

La parole de 30 locuteurs du FQ, autant d'hommes que de femmes, étudiants ou diplômés universitaires, âgés de 20 à 31 ans et provenant à parts égales des villes de Saguenay, Québec et Rouyn-Noranda a été analysée. Les enregistrements de Saguenay et de Québec ont été réalisés en chambre anéchoïque ; ceux de Rouyn-Noranda, au domicile des locuteurs. Une partie de la tâche qui leur était demandée consistait à lire à haute voix, aussi naturellement que possible, une série de phrases porteuses pentasyllabiques et sémantiquement significatives contenant un mot cible en position finale. À la suite de la production de chaque phrase, le mot cible était répété de manière isolée après que l'enquêtrice eut simulé le fait qu'elle n'avait pas compris le mot final. Ce protocole permet 1) d'éviter que les locuteurs hésitent quant au sens des mots cibles, 2) de soustraire ces mots à l'influence de la phrase porteuse tout en prévenant un « effet de liste » et 3) de contrôler le nombre d'occurrences de chaque phonème et son contexte. Et comme les locuteurs naïfs du FQ ne semblent pas porter de jugement négatif à l'endroit du relâchement (Dumas, 1987 : 95) et que la diphtongaison des voyelles fermées passe le plus souvent inaperçue (Santerre et coll., 1985 : 51), il est peu probable que le degré de formalité induit par ces conditions expérimentales ait conduit les locuteurs à adopter une prononciation inhabituelle en ce qui concerne les voyelles fermées.

Des 523 mots cibles répétés isolément qui ont été produits une fois par chaque locuteur, nous en avons retenu 45. Les autres mots, qui contenaient l'une ou l'autre des voyelles orales du FQ et qui

sont destinés à d'autres études, agissaient comme distracteurs. Un total de 1350 occurrences lexicales constitue donc notre corpus. Il s'agit, sauf exception, de noms et de formes verbales monosyllabiques dont le noyau est la voyelle /i/, /y/ ou /u/. Ils se répartissent également entre les classes vocaliques (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK) et (uR), telles que définies par le contexte. Les voyelles en contexte (VR) sont suivies de /ʁ/, /v/, /z/ ou /ʒ/ ; chacune de ces consonnes apparaissant au moins une fois à la suite de chaque voyelle. Notons enfin que seules 7 occurrences ont été rejetées (hésitations lors de leur production ou difficultés d'analyse acoustique). Au final, c'est donc 1343 occurrences vocaliques qui ont été analysées.

2.2 Analyse acoustique

Les voyelles ont été segmentées et étiquetées à l'aide du logiciel PRAAT. Leur durée a ensuite été relevée. Puis les paramètres de détection automatique de la F_0 (par autocorrélation) et de la fréquence centrale des pics formantiques (algorithme de Burg) ont été adaptés manuellement pour chaque occurrence après examen minutieux de la représentation spectrographique. Lorsque la transition entre la voyelle et la consonne était très graduelle (notamment lorsque la voyelle était suivie de /ʁ/), nous avons posé la frontière au milieu de la transition, là où, à l'audition, l'influence mutuelle des deux phones était minimale.

Étant donné la possible diphtongaison des segments en contexte (VR), nous avons choisi de tenir compte de l'évolution des composantes spectrales en cours d'émission. Ainsi, la F_0 , le F_1 , le F_2 et le F_3 de chaque occurrence ont été relevés à 25 %, à 50 % et à 75 % de leur durée. La localisation de ces trois points de mesure n'est pas guidée par des considérations théoriques impliquant les notions de noyau et d'appendice vocalique. Ce choix méthodologique est plutôt inspiré des nombreuses études consacrées à la théorie du *Vowel Inherent Spectral Change* (VISIC) en anglais américain (pour un survol, voir Morrison et Assmann, 2013). La prise en compte d'un point de mesure dans la portion initiale de la voyelle et d'un autre dans la portion finale correspond notamment à l'hypothèse de la *dual target*, formulée par Nearey et Assmann (1986). La position exacte de ces deux points a été établie de façon à ce qu'ils soient « sufficiently extreme to include a substantial portion of observed formant frequency changes, while at the same time avoiding certain irregularities in source characteristics associated with absolute endpoints of the stimuli » (Nearey et Assmann, 1986 : 1298). Le point de mesure central permet quant à lui d'obtenir des résultats comparables aux études antérieures ne s'étant intéressées qu'à l'« état stable » des voyelles ou encore d'affiner la paramétrisation de la dynamique spectrale (lorsque les trois points de mesure sont pris en compte).

2.3 Normalisation

Comme notre échantillon comportait à la fois des hommes et des femmes, nous avons choisi de normaliser la F_0 , le F_1 , le F_2 et le F_3 . Pour ce faire, nous avons adapté la formule des scores-Z proposée par Lobanov (1971), qui consiste à exprimer les valeurs d'une composante fréquentielle par leurs distances, mesurées en écarts-types, à la fréquence moyenne de cette composante pour chaque locuteur. Cette opération a pour effet de superposer et d'égaliser les espaces vocaliques des différents locuteurs (et ce, même si dans le cas présent, une partie seulement du système vocalique est représentée). La formule utilisée est la suivante :

$$(N_x)_i = \frac{(F_x)_i - \overline{(F_x)}_i}{(s_x)_i} \quad \text{où :}$$

- $(N_x)_i$ correspond à la fréquence normalisée (au score-Z) du x^e formant (ou de la F_0) d'une occurrence donnée d'un locuteur i ;
- $(F_x)_i$ correspond à la fréquence en Hz du x^e formant (ou de la F_0) de cette occurrence ;
- $\overline{(F_x)}_i$ correspond à la moyenne du x^e formant (ou de la F_0) du locuteur i ;
- $(s_x)_i$ correspond à l'écart-type du x^e formant (ou de la F_0) du locuteur i .

2.4 Analyses statistiques

Nous avons procédé à une série d'analyses discriminantes quadratiques. Afin d'éviter une classification trop optimiste induite par un surajustement des fonctions discriminantes, nous avons utilisé une procédure de validation croisée. Celle-ci repose sur une partition de l'échantillon en un sous-échantillon d'entraînement et un sous-échantillon de test. Plus précisément, le discriminant était calculé sur les données provenant de la moitié des locuteurs de notre corpus (sexe et origine géographique le plus balancés possible) puis appliqué sur les données de l'autre moitié de notre échantillon. La variable dépendante était la classe vocalique. Diverses combinaisons de prédicteurs (F_0 , F_1 , F_2 , F_3 – normalisés ou non, relevés en 1, 2 ou 3 points de mesure – avec ou sans durée) ont ainsi été soumises à des analyses discriminantes.

3 Résultats et analyse

Comme on peut l'observer sur la FIGURE 1, à 50 % de leur durée, les voyelles /i/, /y/ et /u/ se distinguent sans surprise par leur F_2 (plus élevé pour les antérieures) et leur F_3 (plus faible pour les arrondies). Les variantes relâchées se différencient pour leur part des tendues et des allongées par leur F_1 plus élevé. De façon moins marquée, les allongées diffèrent également des tendues par un F_1 plus élevé en moyenne (voir aussi la FIGURE 2).

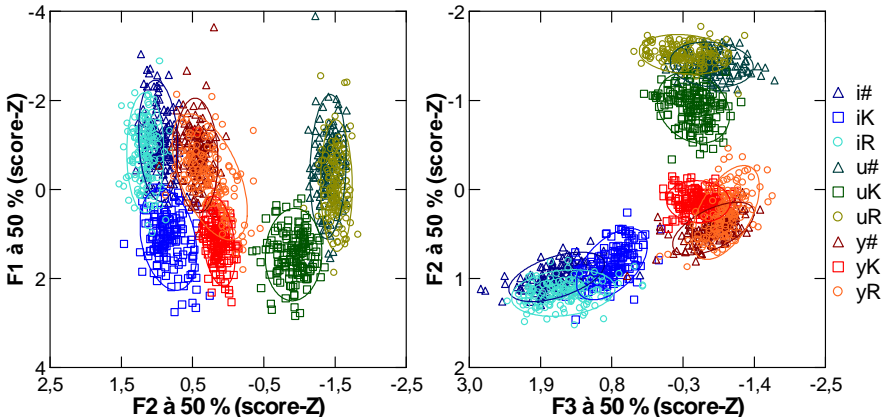


FIGURE 1 – Diagrammes bifonctiques F_1 / F_2 et F_2 / F_3 normalisés présentant les 1343 occurrences analysées à 50 % de leur durée regroupées en fonction des classes vocaliques (les ellipses de dispersion englobent 86,5 % des occurrences d'une même classe vocalique, soit $\pm 2 s$)

La FIGURE 2 présente quant à elle les trajectoires moyennes (de 25 % à 75 % de leur durée) des classes vocaliques dans un diagramme F_1 / F_2 . On constate que même lorsqu'elles sont extraites d'environnements consonantiques variés, la centralisation des relâchées en cours d'émission notée par Arnaud et coll. (2011) se maintient, à tout le moins dans le cas de /i/ et de /u/. Les tendues et les allongées, au contraire, tendent vers la périphérie du trapèze vocalique. Les allongées présentent les trajectoires les plus importantes. D'après MacKenzie et Sankoff (2010), les trajectoires des voyelles fermées allongées sont toutefois moins importantes que celles des autres voyelles allongées généralement considérées comme diphtonguées. Des études ultérieures pourraient permettre de déterminer dans quelle mesure la trajectoire des voyelles fermées est corrélée à la perception de diphtongaison.

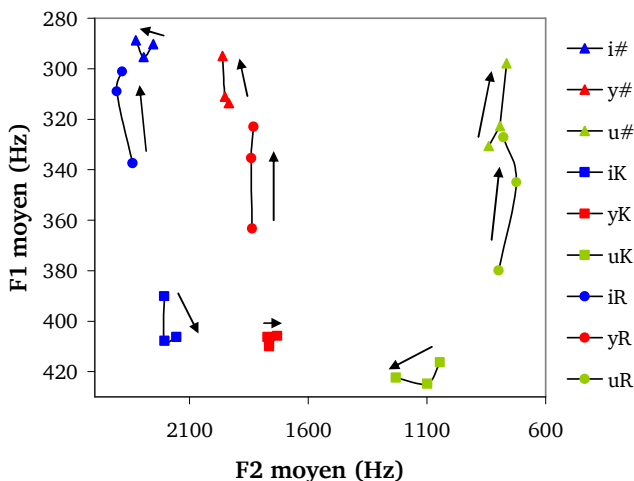


FIGURE 2 – Diagramme biformantique F_1 / F_2 présentant l'évolution moyenne des différentes classes vocaliques, tous locuteurs confondus

La TABLE 1 présente les taux de classification correcte obtenus lorsque différentes combinaisons de prédicteurs sont employées dans les analyses discriminantes.

		50 %		25, 75 %		25, 50, 75 %	
		sans durée	avec durée	sans durée	avec durée	sans durée	avec durée
F_1, F_2	Hz	58	82	68	87	73	85
	Score-Z	76	94	82	96	86	96
F_0, F_1, F_2	Hz	68	88	76	87	74	85
	Score-Z	81	95	87	89	89	95
F_1, F_2, F_3	Hz	66	90	76	93	79	92
	Score-Z	79	96	86	96	88	97
F_0, F_1, F_2, F_3	Hz	72	93	80	91	79	88
	Score-Z	84	96	89	96	92	96

TABLE 1 – Taux de classification correcte globaux (en %), résultats des analyses discriminantes visant à classifier les occurrences dans leur classe vocalique (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK), ou (uR) à partir de divers indices acoustiques relevés en différents points

On remarque tout d'abord que les données normalisées en scores-Z différencient plus efficacement les classes que les données brutes en Hertz, et ce, quels que soient les indices acoustiques et le nombre de points de mesure considérés. Conformément aux résultats d'Adank et coll. (2004), les valeurs relatives des composantes spectrales (ici exprimées en scores-Z) semblent donc un meilleur indicateur de l'identité des voyelles que leurs valeurs absolues en Hertz. Concernant la prise en compte de la dynamique spectrale, on peut constater que le relevé des indices acoustiques en deux points de mesure plutôt qu'en un seul permet d'améliorer la classification dans plusieurs cas. Cependant, lorsque la durée est également incluse dans les prédicteurs, l'impact positif de la dynamique sur la classification est moindre, voire disparaît selon les indices acoustiques considérés et selon que les données sont normalisées ou non. La prise en compte de trois points de mesure plutôt que deux améliore parfois la classification, mais jamais lorsque les données sont exprimées en Hertz et que la durée est incluse dans les prédicteurs, ni lorsque la F_0 est considérée.

La comparaison des matrices de confusion de chacune des analyses discriminantes dont le taux de classification correcte global est rapporté dans la TABLE 1 nous permet de déceler les classes vocaliques les mieux distinguées lorsqu'un indice acoustique donné est ajouté à F_1 et F_2 parmi les prédicteurs. Ainsi, même si les brèves (tendues et relâchées) se distinguent globalement des allongées sur le plan de la durée, l'inclusion de ce prédicteur permet plus particulièrement de mieux différencier les tendues et les allongées ; les relâchées étant déjà relativement bien séparées des allongées par leur F_1 . La prise en compte de F_3 réduit quant à elle la confusion entre les /i/ et les /y/. Étonnamment, bien que la F_0 moyenne des /i/ et celle des /y/ ne soient pas différentes (172 Hz tous contextes et locuteurs confondus, tant pour /i/ que pour /y/), la prise en compte de cet indice diminue également la confusion entre ces voyelles. En fait, comme on peut l'observer sur la FIGURE 3, la F_0 permet de séparer les /i/ produits par des hommes et les /y/ produits par des femmes, qui, autrement, présentent des valeurs de F_1 et de F_2 similaires.

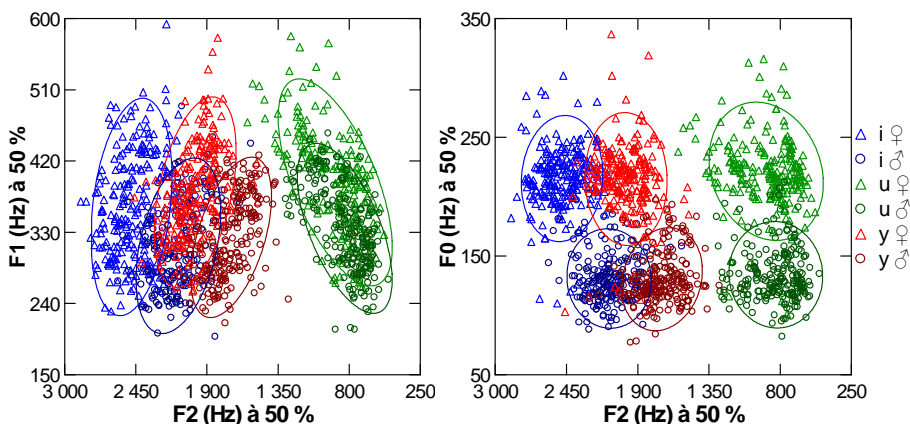


FIGURE 3 – Diagrammes de dispersion F_1 / F_2 et F_0 / F_2 présentant les 1343 occurrences de notre corpus à 50 % de leur durée regroupées d'après leur classe phonémique et le sexe du locuteur (les ellipses de dispersion englobent 86,5 % des occurrences d'un même groupe, soit ± 2 s)

En somme, les indices acoustiques pris en considération (F_0 , F_1 , F_2 , F_3 et durée) participent tous à la différenciation des voyelles fermées tendues, relâchées et allongées du FQ. La normalisation des valeurs fréquentielles et la prise en compte de la dynamique spectrale permettent, dans bien des cas, une différenciation encore plus grande. Cela dit, diverses combinaisons de prédicteurs

permettent de distinguer plus efficacement que d'autres les classes vocaliques de notre corpus. Ainsi, avec un ensemble très économique de prédicteurs (F_1 et F_2 exprimés en scores-Z, à 50 % de la durée des occurrences, et durée vocalique) on obtient un taux de classification correcte de 94 %. Mais il est possible d'atteindre un taux similaire (93 %) sans recourir à la normalisation, en prenant également en compte la F_0 et F_3 ; ou encore en ajoutant seulement F_3 mais en prenant en compte deux points de mesure. Le taux de classification correcte le plus élevé (97 %) est quant à lui atteint lorsque la durée ainsi que F_1 , F_2 , et F_3 relevés en trois points de mesure et normalisés sont fournis comme prédicteurs. Des solutions plus économiques parviennent cependant à des résultats similaires (96 %). C'est notamment le cas lorsque les prédicteurs sont la durée ainsi que F_1 et F_2 relevés en deux points de mesure et normalisés.

4 Discussion et conclusion

Les résultats de la présente contribution doivent être considérés à la lumière des données sur lesquelles ils reposent et du cadre expérimental dans lequel cette étude s'inscrit. Au plan méthodologique, la catégorisation utilisée lors des analyses discriminantes repose sur la classe vocalique *attendue*, la prise en compte du caractère diphtongué / monophthongué des segments lors de cette tâche de classification automatique aurait-elle modifié le poids des différents prédicteurs considérés ? Par ailleurs, s'il a été question, dans l'interprétation de ces résultats, du *français québécois*, une généralisation des caractéristiques acoustiques identifiées à l'ensemble du FQ est sans nul doute prématurée. En outre, une analyse préliminaire de la variation régionale au sein de notre corpus laisse entrevoir certaines particularités à Saguenay. Les voyelles fermées allongées y seraient plus tendues et les relâchées devant consonne voisée, plus longues qu'à Rouyn-Noranda et Québec (sans pour autant que ces différences n'affectent les tendances globales observées dans la présente contribution). Rappelons également que notre corpus a été récolté à la fois dans des conditions très formelles et lors d'une tâche de lecture oralisée. Nos résultats devraient donc être mis à l'épreuve sur un échantillon plus vaste et diversifié, et dans une autre modalité de parole, telle que la parole spontanée. Soulignons par ailleurs que les indices acoustiques discriminants que nous avons mis au jour ne sont peut-être pas ceux sur lesquels repose effectivement la perception des voyelles à l'étude. Entre autres, il est sans doute peu probable que le système perceptif calcule des scores-Z.

Cela étant dit, nos résultats indiquent qu'au-delà de F_1 et F_2 au centre de l'état stable, la prise en compte de la dynamique spectrale constitue une plus-value, y compris pour la description acoustique de variantes brèves réputées ne pas présenter de variations de timbre. En outre, la dynamique spectrale qui distingue ces variantes leur semble propre ; plus précisément, elle ne semble pas strictement liée aux effets acoustiques de la coarticulation ; ce qui ne signifie pas que le contexte, entendu dans un sens plus large, n'a pas d'impact sur la dynamique spectrale. En effet, nous sommes ici en présence de variantes *contextuelles* qui diffèrent non seulement par leur timbre et leur durée, mais également par la structure syllabique dans laquelle elles apparaissent et la nature allongeante ou non de la consonne subséquente. Il est donc difficile d'attribuer à l'une ou l'autre ou à plusieurs de ces caractéristiques la dynamique spectrale observée.

Remerciements

Ce projet a reçu l'appui financier du *Fonds de recherche québécois – Société et culture* ainsi que du *Conseil de recherche en sciences humaines* du Canada.

Références

- ADANK, P., R. SMITS et R. VAN HOUT (2004). A comparison of vowel normalization procedures for language variation research, *JASA*, 116(5): 3099-3107.
- ARNAUD, V., C. SIGOUIN et J.-P. ROY (2011). Acoustic description of Quebec French high vowels: First results. *In Proceedings of the 17th ICPhS*, Hong Kong, 244-247.
- CÔTÉ, M.-H. (2012). Laurentian French (Quebec). Extra vowels, missing schwas and surprising liaison consonants. *In* Gess, R., C. Lyche et T. Meisenburg (dir.). *Phonological Variation in French. Illustrations from three continents*, John Benjamins Publishing Company, 235-274.
- DUMAS, D. (1974). Durée vocalique et diphtongaison en français québécois. *Cahier de linguistique*, (4):13-55.
- (1987), *Nos façons de parler : les prononciations en français québécois*, PUQ.
- GENDRON, J.-D. (1966). *Tendances phonétiques du français parlé au Canada*, PUL.
- LOBANOV, B. M. (1971). Classification of Russian vowels spoken by different speakers, *JASA*, 49(2B):606-608.
- MACKENZIE, L. et G. SANKOFF (2010). A quantitative analysis of diphthongization in Montreal French. *UWPL: Selected Papers from NWAV 37*, 15(2):91-100.
- MARTIN, P. (2002). Le système vocalique du français du Québec. De l'acoustique à la phonologie. *La linguistique*, 38(2):71-88.
- MCLAUGHLIN, A. (1986). Une (autre) analyse de la distribution des variantes des voyelles hautes en français montréalais, *Revue québécoise de linguistique théorique et appliquée*, 5(4): 21-60.
- MORRISON, G. S. et P. F. ASSMANN (2013). *Vowel Inherent Spectral Change*, Springer.
- NEAREY, T. M. et P. F. ASSMANN (1986). Modeling the role of inherent spectral change in vowel identification. *JASA*, 80(5):1297-1308.
- PARADIS, C. (1985). *An Acoustic Study of Variation and Change in the Vowel System of Chicoutimi and Jonquière (Québec)*, thèse de doctorat, University of Pennsylvania.
- POLIQVIN, G. C. (2006). *Canadian French Vowel Harmony*, thèse de doctorat, Harvard University.
- SANTERRE, L. et J. MILLO (1978). Diphtongization in Montreal French. *In* Sankoff, D. (dir.). *Linguistic Variation: Models and Methods*, Academic Press, 173-184.
- SANTERRE, L., S.-P. DUFOUR et S. MCDUFF (1985). La perception de la diphtongaison : son importance dans les grands corpus, *Revue québécoise de linguistique théorique et appliquée*, 4(4): 33-53.
- SIGOUIN, C. (2013). *Caractéristiques acoustiques des voyelles fermées tendues, relâchées et allongées en français québécois*, mémoire de maîtrise, Université du Québec à Chicoutimi.
- YAEGER, M. (1979). *Context-determined Variation in Montreal French Vowels*, thèse de doctorat, University of Pennsylvania.