

Table des matières

Résumé	iii
Abstract.....	v
Table des matières	vii
Liste des tableaux	ix
Liste des figures.....	xi
Conventions et abréviations	xiii
Remerciements	xv
Introduction	1
Chapitre 1 : Problématique et objectifs.....	3
1.1 Variantes phonétiques des voyelles fermées en français québécois	3
1.1.1 Dévoisement.....	3
1.1.2 Réduction et syncope	3
1.1.3 Relâchement.....	4
1.1.4 Allongement.....	4
1.1.5 Diphtongaison.....	5
1.2 Choix des variantes à l'étude	6
1.3 État de la question	7
1.3.1 Durée	7
1.3.2 Caractéristiques spectrales.....	8
1.4 Objectifs	11
Chapitre 2 : Cadre expérimental	13
2.1 Descriptions pionnières du spectre des voyelles	13
2.2 Indices acoustiques supplémentaires	16
2.3 Normalisation vocalique	21
2.3.1 Échelles psycho-acoustiques.....	22
2.3.2 Objets et espaces psycho-acoustiques	23
2.3.3 Opérations extrinsèques.....	24
2.4 Dynamique spectrale	25
2.4.1 <i>Undershoot compensation theory</i>	25
2.4.2 <i>Dynamic cospecification theory</i>	27
2.4.3 <i>Vowel-inherent spectral change (VISC)</i>	27
2.5 Cadre expérimental adopté.....	30
Chapitre 3 : Méthodologie	33
3.1 Échantillon et corpus.....	33
3.2 Analyse acoustique.....	35

3.3 Normalisation	37
3.4 Dynamique spectrale.....	37
3.5 Analyses statistiques.....	39
Chapitre 4 : Résultats et analyse	43
4.1 Durée	43
4.2 Caractéristiques spectrales	51
4.3 Analyses discriminantes.....	62
4.4 Synthèse récapitulative	68
Chapitre 5 : Discussion conclusive.....	71
Bibliographie.....	77
Annexe 1 : Tableau de l'alphabet phonétique international	89
Annexe 2 : Sigouin (à paraître).....	91
Annexe 3 : Liste des phrases et des mots cibles du corpus.....	121
Annexe 4 : Script PRAAT vérifiant que les frontières sont bien posées à un passage par zéro	123
Annexe 5 : Script PRAAT relevant les mesures à chaque occurrence	131
Annexe 6 : Script Perl calculant les scores-Z.....	135
Annexe 7 : Valeurs moyennes et écarts-types de la F_0, de F_1, F_2 et F_3 (en Hertz) calculées à partir des moyennes individuelles.....	143
Annexe 8 : Résultats des ANOVAs portant sur les caractéristiques spectrales	149

Liste des tableaux

Tableau 1 : Durée moyenne et écart-type (en ms) des différentes classes vocaliques selon le sexe et la ville des locuteurs.....	43
Tableau 2 : Évolution des caractéristiques spectrales des différentes classes vocaliques.....	59
Tableau 3 : Taux de classification correcte globaux (en %), résultats des analyses discriminantes visant à classer 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	63
Tableau 4 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_1 et F_2 (en Hertz) relevés à 50 % de la durée	64
Tableau 5 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_1 , F_2 et F_3 (en Hertz) à 50 % de la durée	64
Tableau 6 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_0 , F_1 et F_2 (en Hertz) à 50 % de la durée	65
Tableau 7 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_1 , F_2 et F_3 (en Hertz) à 50 % de la durée et la durée	66
Tableau 8 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_1 et F_2 (en Hertz) relevés à 25 % et 75 % de la durée	66
Tableau 9 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_1 et F_2 (en scores-Z) relevés à 50 % de la durée	67
Tableau 10 : Taux de classification correcte globaux (en %), résultats des analyses discriminantes visant à classer les occurrences dans leur classe vocalique (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK), ou (uR) à partir des modèles du VISC	68

Liste des figures

Figure 1 : Coupe spectrale d'un [ɑ]. Adaptée de Bevier (1900).....	14
Figure 2 : Spectrogrammes à bande large de voyelles isolées. Tirée de Delattre (1948).....	14
Figure 3 : Diagramme de dispersion F_1 / F_2 de 1520 occurrences des dix monophthongues de l'anglais américain, produites en contexte [hVd] par 33 hommes, 28 femmes et 15 enfants. D'après les données de Peterson et Barney (1952).	16
Figure 4 : Position des quatre premiers formants de voyelles naturelles (indiqués par les lignes) et du second formant des voyelles synthétiques correspondantes (représenté par un rectangle). Tirée de Carlson et coll. (1970).	18
Figure 5 : Exemple de spectre d'un [i] synthétique présentant une forte pente spectrale (en rouge) et d'une version modifiée de ce spectre où la pente a été atténuée (en vert). Adaptée de Kale (2006).	19
Figure 6 : Diagramme bi-formantique F_2 / F_1 présentant les voyelles de l'anglais américain d'après les données de Peterson et Barney (1952). Les lignes courbes indiquent les ratios entre F_2 et F_1 des différentes classes vocaliques. Adaptée de Miller (1989).	23
Figure 7 : Résultats du test de perception de Lindblom et Studdert-Kennedy (1967). Les pourcentages de réponses [ɪ] sont indiqués par les points et les pourcentages de réponses [ʊ], par les croix. Tirée de Lindblom et Studdert-Kennedy (1967).	26
Figure 8 : Diagrammes bi-formantiques F_2 / F_1 présentant les trajectoires moyennes de 8 voyelles de l'anglais américain en contextes variées (à gauche) et en contexte neutre /hVd/ (à droite) produites par des femmes. Le symbole phonétique correspond à la position finale de la voyelle. Adaptée de Hillenbrand et Nearey (1999) et Hillenbrand et coll. (2001).....	28
Figure 9 : Exemple d'une courbe ajustée à la trajectoire d'un formant au moyen d'une transformée en cosinus discrète. Tirée de Morrison (2013).	30
Figure 10 : Carte du Québec (tronquée) indiquant nos points d'enquête. Adaptée de la carte <i>Le Québec, carte muette</i> du Ministère des Ressources naturelles et de la Faune. © Gouvernement du Québec, 2006.	33
Figure 11 : Diagramme boîtes à moustaches présentant la durée des voyelles en fonction du contexte et de la classe phonémique. Les astérisques indiquent des données extrêmes.....	44
Figure 12 : Diagramme boîtes à moustaches présentant la durée des voyelles en fonction du contexte et de la ville d'origine des locuteurs. Les astérisques indiquent des données extrêmes.....	45
Figure 13 : Durée moyenne des voyelles en fonction de la ville d'origine des locuteurs et du contexte consonantique.....	47
Figure 14 : Diagramme boîtes à moustaches présentant la durée des voyelles fermées en (VK) suivies d'une consonne voisée produites par les locuteurs de Saguenay en fonction du sexe. Les astérisques indiquent des données extrêmes.	48

Figure 15 : Diagramme boîtes à moustaches présentant la durée des variantes en fonction du contexte consonantique. Les astérisques indiquent des données extrêmes.	49
Figure 16 : Diagramme boîtes à moustaches présentant la durée des voyelles allongées en fonction de la consonne suivante.....	51
Figure 17 : Diagrammes biformantiques F_1 / F_2 et F_2 / F_3 normalisés présentant les 1343 occurrences analysées à 25 % de leur durée regroupées en fonction des classes vocaliques. Les ellipses de dispersion englobent 86,5 % des occurrences d'une même classe vocalique.....	52
Figure 18 : Diagrammes biformantiques 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.	53
Figure 19 : Diagrammes biformantiques F_1 / F_2 et F_2 / F_3 normalisés présentant les 1343 occurrences analysées à 75 % de leur durée regroupées en fonction des classes vocaliques. Les ellipses de dispersion englobent 86,5 % des occurrences d'une même classe vocalique.	54
Figure 20 : Rapport entre les fréquences formantiques (F_1 , F_2 et F_3) masculines et féminines moyennes pairées par ville, par classe vocalique et par point de mesure.	55
Figure 21 : Diagramme biformantique F_1 / F_2 (en Hertz) présentant les valeurs féminines et masculines moyennes (calculées à 50 % de la durée des voyelles) des classes vocaliques (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK), ou (uR).	56
Figure 22 : Trajectoire temporelle moyenne du F_2 de la classe vocalique (y#) en fonction de la ville d'origine des locuteurs.....	57
Figure 23 : Trajectoire temporelle moyenne du F_1 de la classe vocalique (yK) en fonction de la ville d'origine des locuteurs.....	58
Figure 24 : Diagramme biformantique F_1 / F_2 normalisés présentant l'évolution moyenne des différentes classes vocaliques, tous locuteurs confondus.....	60
Figure 25 : Diagrammes de dispersion F_1 / F_2 et F_0 / F_2 (en Hertz) 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'une même classe phonémique, soit les occurrences situées à 2 écarts-types ou moins de la moyenne.	65

Conventions et abréviations

Dans le présent document, les transcriptions phonétiques sont placées entre crochets. Les phonèmes apparaissent entre barres obliques. Conformément à la tradition en sociolinguistique variationniste (Labov et coll., 1972 : 39), les classes vocaliques sont placées entre parenthèses. L'alphabet phonétique international est utilisé dans tous les cas (voir l'Annexe 1). Le dièse (#) indique une frontière de mot. La lettre *C* (en majuscule) correspond à une consonne. La lettre *R* (en majuscule) est employée de manière générique pour désigner les consonnes réputées allongeantes en langue française : /ʁ/, /r/, /z/, /ʒ/ et le groupe /ʁʁ/ (Malmberg, 1968 : 88). La lettre *K* (en majuscule) désigne quant à elle une consonne qui n'est pas réputée allongeante. Enfin, la lettre *V* (en majuscule) désigne génériquement une voyelle.

Par souci de cohérence, lorsque nous désignons des voyelles ou des consonnes, nous utilisons autant que possible des termes se rapportant à leurs caractéristiques articulatoires, et ce, même lorsque les auteurs cités emploient des termes se rapportant plutôt à leurs traits phonologiques ou à leurs propriétés perceptives (pourvu que nous nous référions bien aux mêmes sons de la langue). Par exemple, le terme *fermée* sera préféré au terme *haute*, le terme *voisée*, au terme *sonore*, etc.

Abréviations

F ₀	fréquence fondamentale
F ₁	fréquence centrale du premier formant
F ₂	fréquence centrale du deuxième formant
F ₃	fréquence centrale du troisième formant
FQ	français québécois
QC	la ville de Québec
RN	la ville de Rouyn-Noranda
SG	la ville de Saguenay
s.o.	sans objet, le test statistique n'a pu être appliqué (postulats non vérifiés) ou il n'y avait pas lieu de l'appliquer
VISC	<i>vowel-inherent spectral change</i> , dynamique spectrale intrinsèque des voyelles

Remerciements

Cher Monsieur Arnaud, si vous saviez comme j'ai cherché les mots pour vous dire merci... Mais je crois bien qu'il n'en existe pas d'assez puissants pour exprimer ma gratitude. Les années que j'ai passées au labo ont été parmi les plus stimulantes et heureuses de ma vie. Et c'est grâce à vous. Grâce à la confiance que vous m'avez accordée, à votre reconnaissance, votre intégrité, votre rigueur. Vous êtes des personnes qui auront le plus influencé le cours de mon existence et j'en suis privilégiée. Quand j'ai choisi de travailler avec vous, j'avais déjà la plus grande estime pour vous et pourtant, j'étais loin de m'imaginer à quel point cette collaboration allait être fructueuse. Merci pour votre encadrement, votre écoute, votre sens de l'humour et pour toutes ces discussions interminables accrochées à la poignée de porte. En toute amitié, salut!

Monsieur Dolbec, je vous serai à jamais reconnaissante de m'avoir initiée à la phonétique. Qui aurait cru, à l'époque, que ça me mènerait aussi loin? Je vous remercie de m'avoir permis de participer à PHONO, d'avoir été présent au labo, de m'avoir conseillée et d'avoir accepté d'évaluer mon travail. C'est un honneur d'avoir pu bénéficier de votre regard critique, de votre expérience et de votre culture scientifique magistrale.

Madame Roy, merci d'avoir évalué mon projet de maîtrise et mon mémoire. J'ai énormément de respect pour vos compétences, votre professionnalisme, votre authenticité, pour *vous*. Merci pour tous les fous rires à Hong Kong et merci de m'avoir ouvert (à moi, mes plantes, mon frigo, ma bibliothèque et mes lubies) la porte de votre labo pour le doctorat.

Monsieur Poiré, un grand merci à vous aussi d'avoir évalué mon mémoire. C'est sans doute un peu maladroit de ma part de dire ça, mais votre rapport m'a fait chaud au cœur, sachez-le.

Je tiens également à remercier Mme Hélène Crépeau, du Service de consultation statistique de l'Université Laval pour son expertise redoutable, mais aussi pour son enthousiasme et tout le savoir-faire qu'elle a accepté de me transmettre.

Je ne remercierai jamais assez les professeurs et chargées de cours en linguistique de l'Université du Québec à Chicoutimi grâce à qui j'ai acquis une solide formation et profité d'un environnement d'étude hors du commun. Un grand merci, de plus, à mes collègues (Josiane, Guylaine, Sophie, Guillaume, Félix, Marie-Hélène, Maude, Elena et j'en passe) avec qui j'ai tant partagé.

Merci à mon ancienne patronne Chantal de même qu'à toute l'équipe de Québec solidaire en Abitibi-Témiscamingue (Jaquy, Guy, France, Johanne et tous les autres) de m'avoir montré, avant même le début de mes études universitaires, à voir grand et à travailler avec diligence.

À mes proches, parents ou amis, je vous dis simplement merci. Sans vous, je ne serais jamais devenue qui je suis aujourd'hui. Votre présence et votre soutien ont contribué à ma réussite plus que vous ne le croyez.

Je mentirais si je disais que cette recherche a été rendue *possible* grâce au soutien financier du Conseil de recherche en sciences humaines du Canada et du Fonds de recherche québécois – Société et culture. Ma passion pour la phonétique est telle qu'elle aurait inévitablement trouvé son chemin... Mais comme on va loin quand la route est belle ! Jamais je n'aurais pu m'investir comme je l'ai fait dans ma recherche si ce n'avait été des bourses que ces deux organismes m'ont octroyées. Et je suis fière de vivre dans une société qui valorise encore l'excellence, le savoir et la recherche fondamentale.

Enfin, merci aux /y/ d'être éminemment plus coopératifs que les /i/ et les /u/. Merci en outre à Mme Germaine, et à feu M. Tsiboutsky pour leur présence indéfectible et leur résilience admirable...

Introduction

Les voyelles fermées /i/, /y/ et /u/ peuvent être réalisées de multiples façons en français québécois (désormais FQ). Parmi ces réalisations, trois se distribuent de manière complémentaire sous l'accent. Il s'agit des variantes brèves tendues, apparaissant en syllabe ouverte (V#) comme dans *pie* [pi], des variantes relâchées, présentes en syllabe fermée par une consonne non allongeante (VK) comme dans *pic* [pɪk], et des variantes allongées, attendues en syllabe fermée par une consonne allongeante (VR) comme dans *pire* [pi:ɾ] ou [pɪiɾ].

Les caractéristiques acoustiques distinguant ces trois types de variantes contextuelles restent en partie à découvrir. Comme nous le préciserons au Chapitre 1, les études acoustiques ayant porté sur les voyelles fermées du FQ se sont rarement intéressées aux variantes allongées; et lorsqu'elles s'y sont attachées, elles n'ont pas adopté un cadre expérimental permettant de comparer systématiquement ces variantes aux variantes tendues et relâchées. Deux interrogations plus précises subsistent par ailleurs. La première concerne le timbre des variantes allongées. Ces dernières sont-elles tendues, comme le laissent entendre Gendron (1966), Vinay (1973) et Paradis (1985)? Sont-elles tendues devant /v/, /z/, /ʒ/, mais relâchées lorsque suivies de /ɾ/ (Prairie, 1976)? Présentent-elles seulement une variation de timbre (Santerre et Millo, 1978) ou sont-elles (parfois ou toujours) diphtonguées (Dumas, 1974; Yaeger, 1979; Côté, 2010; Leblanc, 2012)? La seconde interrogation porte sur le rapport de durée entre les variantes tendues et relâchées. Gendron (1966), Martin (2002) et Poliquin (2006) rapportent en effet que les tendues sont plus longues que les relâchées, alors que nous sommes parvenus au résultat inverse dans une étude préliminaire (Arnaud et coll., 2011). On ignore également si les voyelles fermées tendues, relâchées et allongées sont marquées d'une variation diatopique en FQ (ce qui pourrait éventuellement expliquer certaines divergences présentes dans la littérature).

L'objectif de la présente étude est donc de caractériser acoustiquement les variantes tendues, relâchées et allongées des voyelles fermées du FQ sous l'accent, tout en explorant la possibilité d'une variation diatopique. Pour y parvenir, de nombreux indices acoustiques (fréquence centrale, amplitude et largeur de bande des formants, durée, intensité, fréquence fondamentale...) et plusieurs méthodes d'analyse (l'approche traditionnelle qui consiste à relever la fréquence centrale des deux premiers formants dans la portion centrale de la voyelle, la normalisation et la prise en compte de la dynamique spectrale), que nous présenterons au Chapitre 2, s'offraient à nous.

Le choix de notre cadre expérimental nous a amenée à adopter la méthodologie qui sera détaillée au Chapitre 3. Nous avons analysé la parole de 30 locuteurs, autant d'hommes que de femmes, étudiants ou

diplômés universitaires, âgés de 20 à 31 ans, provenant à parts égales des villes de Saguenay, de Québec et de Rouyn-Noranda et ayant le FQ comme langue maternelle. Au total, 1350 occurrences de /i/, de /y/ et de /u/, également réparties entre les contextes (V#), (VK) et (VR), ont été considérées. Ces occurrences sont extraites de mots monosyllabiques produits isolément lors d'une tâche de lecture oralisée. La durée de chaque occurrence a été relevée, puis la fréquence fondamentale (désormais F_0), et la fréquence centrale des trois premiers formants (désormais F_1 , F_2 et F_3) ont été estimées à 25 %, 50 % et 75 % de cette durée. Les valeurs fréquentielles brutes en Hertz et normalisées en scores-Z ont été utilisées. Des ANOVAs et des analyses discriminantes ont été menées afin de répondre aux objectifs de recherche présentés ci-haut.

La présentation et l'analyse des résultats seront divisées en trois sections principales. La première sera consacrée à la durée des variantes tendues, relâchées et allongées des voyelles fermées. La seconde traitera quant à elle de leurs caractéristiques spectrales. Finalement, la troisième section fournira une évaluation de l'efficacité de différentes combinaisons d'indices dans la différenciation acoustique de ces voyelles. Nous concluons notre mémoire par une discussion où seront exposées les limites de notre étude et les perspectives de recherches ultérieures que cette dernière appelle.

Chapitre 1 : Problématique et objectifs

1.1 Variantes phonétiques des voyelles fermées en français québécois

S'il est des variantes prototypiques des voyelles fermées /i/, /y/ et /u/, ce sont sans doute celles qu'on retrouve en syllabe ouverte : les variantes brèves tendues [i], [y] et [u]. Aussi est-ce généralement par opposition à ces dernières que sont décrites les autres réalisations possibles. Les recherches en phonétique et en phonologie ont en effet établi que différents phénomènes sont susceptibles d'affecter les voyelles fermées en FQ. Nous en proposons un bref survol dans les sections qui suivent.¹

1.1.1 Dévoisement

Une voyelle est dite dévoisée (ou désonorisée) lorsqu'elle est produite sans vibration des cordes vocales. En FQ, les voyelles fermées brèves, tendues et non accentuées sont particulièrement sujettes au dévoisement (Dumas, 1978 : 128), quoique des exemples de relâchées dévoisées sont également présentés dans la littérature (Gendron, 1959 : 104; Dumas, 1987 : 103). Le phénomène peut se produire lorsque la voyelle est en contact avec au moins une consonne non voisée (Brent, 1971 : 36; Dumas, 1978 : 128), notamment si cette dernière suit la voyelle (Cedergren et Simoneau, 1985 : 107). Le dévoisement est spécialement fréquent lorsque la voyelle se situe entre deux consonnes non voisées (Cedergren et Simoneau, 1985 : 104; Vinay, 1973 : 352). Par exemple :

(1) *député* → [depÿte]

1.1.2 Réduction et syncope

En FQ, les voyelles fermées peuvent également s'abrèger, au point parfois de s'amuir. Tout comme dans le cas du dévoisement, la réduction et la syncope peuvent survenir lorsque les voyelles sont brèves, tendues et non accentuées. Ces deux phénomènes semblent affecter plus fréquemment les /i/ que les /y/, et les /y/ que les /u/ (Dumas, 1978 : 130). L'environnement consonantique favorisant la réduction et la syncope est toutefois différent (au moins en partie) de celui favorisant le dévoisement. En effet, la réduction et la syncope se produisent notamment lorsque la voyelle est précédée de /ʁ/, de //, d'une consonne fricative ou nasale, ou encore lorsqu'elle est située entre une occlusive et une fricative (Walker, 1984 : 75). Par exemple :

(2) *déguisé* → [degʒe] voire [degze]

¹ Pour des informations et des références complémentaires, nous invitons le lecteur ou la lectrice à consulter le site internet PHONO (Dolbec et Paradis, 1998).

1.1.3 Relâchement

Le contraste entre les voyelles fermées tendues et relâchées se manifeste en FQ par une position du dos de la langue plus basse pour les relâchées (Dalton, 2011a, 2011b).² Lorsqu'une voyelle fermée perd sa tension,³ son timbre tend vers celui de la voyelle mi-fermée homorganique (Dumas, 1974 : 32) et se centralise légèrement (Brent, 1971 : 34; Vinay, 1973 : 351; Paradis, 1985 : 95). Le relâchement se produit en syllabe fermée par une consonne non allongéante (c'est-à-dire toute autre consonne que /ɛ/, /v/, /z/, /ʒ/ et le groupe /vɛ/).⁴⁻⁵ Il est systématique en syllabe finale de mot, et facultatif ailleurs (Dumas, 1974 : 32). Par exemple :

(3) *moustache* → [mustaʃ] ou [mʊstaʃ]

De plus, Dumas (1974 : 33) rapporte que le relâchement est facultatif en syllabe ouverte, dans les dérivés morphologiques issus de radicaux présentant le relâchement en syllabe fermée. Par exemple :

(4) *vite* → [vit]
☞ *vitesse* → [vites]

Dumas (1974 : 33, 38) mentionne également que le relâchement peut survenir en syllabe ouverte par harmonisation vocalique, lorsque, dans le même mot, la syllabe suivante contient une voyelle fermée relâchée ou diphtonguée.⁶ Par exemple :

(5) *pilule* → [pɪlʏl]
toujours → [tuʒɔ̃ɥɛʒ]

1.1.4 Allongement

Les voyelles fermées sont obligatoirement allongées en syllabe fermée par une consonne allongéante⁴ (/ɛ/, /v/, /z/, /ʒ/ et le groupe /vɛ/) en fin de mot (Dumas, 1974 : 18). Par exemple :

(6) *luge* → [lyːʒ]

² Dans d'autres langues, le relâchement peut aussi se traduire par un avancement de la racine de la langue (Ladefoged et Maddieson, 1996 : 302).

³ D'après Reighard (1986 : 290) et Côté (2010 : 27), les voyelles fermées relâchées constitueraient plutôt des phonèmes à part entière. Déchaine (1990) et Hannahs (1989) (tels que rapportés dans Dawn Armstrong, 1999 : 21) proposent quant à eux que les voyelles fermées du FQ sont relâchées par défaut et se tendent en syllabe ouverte.

⁴ Boulanger (1986) (entre autres) rapporte que dans certaines régions, notamment en Beauce, mais aussi au Saguenay–Lac-Saint-Jean, les voyelles fermées peuvent ne pas être allongées devant /v/, /z/ et /ʒ/, mais relâchées. Nous n'avons cependant décelé qu'une seule occurrence de ce type dans notre corpus (voir le point 3.1 pour plus de détails).

⁵ Marchal (1980 : 85) rapporte que devant les occlusives sonores /b/, /d/ et /g/, les voyelles fermées s'allongent légèrement, ce qui les soustrairait au relâchement (Marchal, 1980 : 163).

⁶ Nous reviendrons sur cette particularité au point 1.2.

Les voyelles fermées s'allongent parfois même en syllabe ouverte non finale, dans les dérivés morphologiques issus de radicaux présentant l'allongement en syllabe fermée (Dumas 1974 : 19). Par exemple :

- (7) *pire* → [pi:ʁ]
☞ *empirer* → [ɑ̃pi:ʁe]

1.1.5 Diphtongaison

D'après Dumas (1974 : 28) et Côté (2010 : 54), les voyelles fermées allongées sont obligatoirement diphtonguées sous l'accent. Leur noyau est relâché et leur appendice, tendu.⁷ Par exemple :

- (8) *pire* → [pi̯iʁ]

Mais le timbre des voyelles fermées allongées en FQ est sujet à débat. Ainsi, Gendron (1966) et Dionne (1971), dans leurs études pionnières en phonétique de corpus au Québec, ne relevaient aucune occurrence vocalique diphtonguée (qu'il s'agisse ou non d'une voyelle fermée). Soulignons néanmoins que les locuteurs de Gendron (1966) étaient des Canadiens qui vivaient à Paris au moment de l'enquête. Or, l'influence du français de France, réputé non diphtonguant (Delattre, 1953, 1963), sur les productions des locuteurs de Gendron (1966), pourrait expliquer ces résultats. Quant à Dionne (1971 : v), elle mentionne tout de même que si les voyelles québécoises ne sont, selon elle, pas diphtonguées, leur timbre lui semble plus instable qu'en français de France. À la lumière d'études plus récentes, « [n]ul ne songerait à contester l'importance de la diphtongaison dans les dialectes français du Québec et de l'Acadie » (Santerre et coll., 1985 : 33), et pourtant, le cas spécifique des voyelles fermées est toujours discuté. D'après Vinay (1973 : 351), les voyelles fermées allongées sont tendues; ce que corrobore Paradis (1985 : 95). Santerre et Millo (1978 : 174) abondent dans le même sens en apportant néanmoins la nuance suivante : « high vowels [...] do not present diphthongized variants but only a variation in quality, as in *rire* /riʁ/ → [ri̯iʁ], for example ». Prairie (1976 : 13-15) affirme pour sa part que les voyelles fermées sont tendues en syllabe fermée par /v/, /z/ et /ʒ/, mais qu'elles sont relâchées (et elles sont transcrites avec un appendice fermant sous l'accent) devant la consonne /ʁ/. Selon Walker (1984 : 68), le noyau des voyelles fermées allongées diphtonguées (quelle que soit la consonne allongeante impliquée) se relâche afin de maximiser le contraste avec l'appendice. Quant à Santerre (1971) (dont une partie du corpus a aussi été utilisée par Dionne, 1971), il relève de façon sporadique des occurrences de voyelles fermées diphtonguées. De même, Yaeger (1979) décèle quelques /u/ allongés diphtongués dans son corpus. Plus récemment, Leblanc (2012 : 48) constate de son côté que 14 % des /i/, 43 % des /y/ et 49 % des /u/ allongés de son corpus ont été perçus comme diphtongués par trois juges

⁷ Le noyau des diphtongues est parfois transcrit par le timbre tendu et l'appendice, par la semi-consonne homorganique.

expérimentés. Bref, les positions sur la question de la diphtongaison des voyelles fermées sont très variées, allant de la négation à l'obligation en passant par la nuance, la restriction et l'attestation.

Si la diversité des corpus et des méthodologies employés dans les études susmentionnées⁸ explique sans doute une part de cette discordance, une autre part trouve peut-être son origine dans le fait que la diphtongaison se définit comme une variation de timbre (et corollairement comme une variation spectrale) au sein d'une même voyelle, et qu'il n'existe pas de seuil précis au-delà duquel une telle variation est catégoriquement admise comme étant une diphtongaison. De plus, ce phénomène peut être envisagé sur deux plans : celui de la perception (une variation de timbre) et celui de la production (une variation spectrale). Or, Santerre et coll. (1985 : 38) suggèrent que ces deux perspectives ne s'accordent pas toujours, puisqu'ils affirment que « les diphtongues québécoises, si elles ne sont pas toujours perçues, sont en tout cas bien produites, articulatoirement et acoustiquement ».

1.2 Choix des variantes à l'étude

Comme le note McLaughlin (1986 : 21), en observant l'inventaire des réalisations des voyelles fermées possibles en FQ, on remarque qu'en syllabe non accentuée, /i/, /y/ et /u/ peuvent, selon le contexte, être affectés par différents phénomènes dont l'application est facultative. Les variantes présentes sous l'accent, au contraire, se distribuent catégoriquement dans trois contextes complémentaires : les variantes brèves tendues apparaissent en syllabe ouverte (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).⁹ C'est donc à ces trois derniers types de variantes (sous l'accent) que nous avons choisi de nous intéresser dans le cadre du présent travail. Les variantes canoniques apparaissent entre autres comme d'incontournables points de comparaison. Par ailleurs, Dumas (1974 : 38) propose une même motivation phonologique à la diphtongaison (consécutives à l'allongement) et au relâchement, affirmant que le trait de tension inhérent aux voyelles fermées doit être résolu en syllabe fermée. En syllabe fermée par une consonne non allongeante, le relâchement permettrait de résoudre la tension, alors qu'en syllabe fermée par une consonne allongeante, la tension serait résolue par la diphtongaison. Selon lui, c'est ce qui expliquerait qu'une voyelle puisse se relâcher par harmonisation avec une voyelle diphtonguée.¹⁰ Walker (1984 : 71) suggère lui aussi que les phénomènes de relâchement et de diphtongaison des voyelles fermées sont liés. Il considère que le relâchement est une lénition (qui met en relief la brièveté de la voyelle) et la diphtongaison, une fortition (qui augmente la perceptibilité de la longueur). Ces deux phénomènes accentueraient donc le contraste entre les temps forts et les temps faibles qui rythment l'énoncé.

⁸ Ces études seront exposées plus en détail au point 1.3.2.

⁹ Nous ne tiendrons pas compte ici des mots, somme toute marginaux, faisant exception à cette distribution, comme *cool* [kuw] ou *quiz* [kwiz] (Côté, 2010 : 53).

¹⁰ Voir point 1.1.3.

1.3 État de la question

Les voyelles fermées du FQ ont fait l'objet de nombreux travaux scientifiques. Dans cette partie, nous présenterons plus spécifiquement les études acoustiques ayant traité des variantes tendues, relâchées et allongées de /i/, /y/ et /u/ en FQ, bien que cela n'exclut pas quelques références à des études phonologiques et auditives. Après avoir fait le point sur l'état actuel des connaissances concernant la durée de ces variantes, nous consacrerons une seconde section aux caractéristiques spectrales de ces mêmes segments.

1.3.1 Durée

Par définition, la durée des variantes allongées des voyelles fermées du FQ est supérieure à celle des variantes canoniques, soit les variantes brèves tendues. La durée des variantes relâchées par rapport à celle des variantes allongées paraît également faire l'unanimité. Par comparaison aux variantes allongées, Walker (1984 : 54) décrit les voyelles fermées en syllabe fermée par une consonne non allongeante comme « short and lax ». De même, Dumas (1974 : 38) et Prairie (1976 : 13) stipulent que seules les voyelles fermées « brèves » en syllabe fermée se relâchent. Du reste, dans ces trois études phonologiques, le rapport de durée entre les variantes relâchées et tendues n'est pas explicité. Cela dit, aucun indice dans les transcriptions ne permet de supposer une différence de durée entre ces deux modalités de réalisation. Pourtant, Gendron (1966 : 41) conclut à la suite de ses analyses auditives et expérimentales qu'« en finale absolue, les *i*, *y* et *u*,¹¹ tout en étant perçus comme brefs, sont physiologiquement plus longs que dans la syllabe fermée par une consonne non-allongeante ». Martin (2002 : 77) et Poliquin (2006 : 28) constatent également que les voyelles fermées de leurs corpus respectifs sont plus longues en syllabe ouverte qu'en syllabe fermée par une consonne non allongeante. C'est donc dire que d'après ces deux auteurs, les voyelles fermées brèves tendues du FQ sont plus longues que les relâchées. Ce n'est toutefois pas le résultat auquel nous sommes parvenus dans le cadre d'une étude préliminaire où nous avons calculé que la durée des variantes tendues ne représente que 59 % à 63 % de la durée des relâchées (Arnaud et coll., 2011 : 247). Bref, il semble y avoir consensus quant au fait que les voyelles fermées allongées sont plus longues que les tendues et les relâchées, mais comment expliquer les divergences observées concernant le rapport de durée entre les relâchées et les tendues ?

Nous remarquons tout d'abord que les méthodes d'analyse employées par Dumas (1974), Prairie (1976) et Walker (1984), d'une part, et par Martin (2002), Poliquin (2006) et Arnaud et coll. (2011), d'autre part, diffèrent considérablement. Les premiers se sont fondés sur leurs impressions, alors que les seconds ont analysé

¹¹ L'alphabet phonétique utilisé par Gendron (1966) n'était pas reproductible ici.

acoustiquement des voyelles issues de corpus.¹² Ainsi, s'il existe une différence de durée systématique entre les voyelles fermées tendues et relâchées du FQ, peut-être est-elle trop faible pour être perçue, et sans doute n'est-elle pas pertinente d'un point de vue phonologique. Reste encore à expliquer l'écart entre les résultats de Martin (2002) et de Poliquin (2006), d'une part, et ceux de Arnaud et coll. (2011), d'autre part. Les principales différences méthodologiques entre ces études ont trait à l'environnement consonantique, à la structure syllabique et à la position syntaxique des occurrences, de même qu'à l'origine géographique des locuteurs. Dans Martin (2002), les mots contenant les voyelles fermées sont variés, comportent une, deux ou trois syllabes et ont été produits de manière isolée par des locuteurs provenant de six régions du Québec. Dans Poliquin (2006), des locuteurs montréalais ont produit des voyelles fermées au sein de mots mono- et bisyllabiques eux-mêmes insérés dans une phrase porteuse. En syllabe fermée, la consonne finale est toujours /t/. Dans Arnaud et coll. (2011), les voyelles, produites par des locuteurs de la ville de Saguenay, sont insérées dans des logatomes de forme /bV/ ou /bVb/ se trouvant dans une phrase porteuse. Or, il est envisageable que ces facteurs aient un impact sur les durées relevées. Poiré (1994 : 145) observe d'ailleurs que le rapport de durée entre les /i/ tendus et relâchés du locuteur qu'il a enregistré varie selon la consonne qui suit les /i/ relâchés (les /y/ et les /u/ n'étaient pas représentés dans son corpus).

1.3.2 Caractéristiques spectrales

Les premières descriptions acoustiques des voyelles fermées du FQ nous sont fournies par Gendron (1966), Santerre (1971) et Dionne (1971). Gendron (1966 : 31, 34, 38), dont les analyses acoustiques ont porté uniquement sur sa propre parole, révèle que les variantes relâchées des voyelles fermées ont un F_1 plus élevé (corrélât acoustique d'une plus grande aperture) que celui des variantes tendues, sans pour autant atteindre celui des voyelles mi-fermées. De plus, les /y/ et les /u/ relâchés apparaissent comme légèrement centralisés (F_2 plus bas pour /y/, plus élevé pour /u/) par rapport à leurs contreparties tendues. Santerre (1971) a pour sa part enregistré deux locuteurs montréalais, et l'un d'eux a constitué à lui seul la portion québécoise du corpus de Dionne (1971). Les analyses de ces deux derniers auteurs, semblables à celles de Gendron (1966), consistaient à relever les valeurs de F_1 et de F_2 de chaque occurrence en un seul point du continuum acoustique, là où les formants étaient visiblement le plus stables. Or, un tel cadre d'analyse, statique, peut difficilement rendre compte de la dynamique spectrale d'éventuelles voyelles diphtonguées. Il faut tout de même reconnaître que Santerre (1971) est allé au-delà du seul relevé des fréquences formantiques à l'état stable en procédant à un examen visuel de l'ensemble du spectre des occurrences analysées. Il est d'ailleurs le seul à avoir détecté des occurrences vocaliques diphtonguées, dont certaines de voyelles fermées allongées. Il a donc été en mesure de décrire qualitativement la structure spectrale de ces voyelles

¹² Globalement, la méthodologie de Gendron (1966) combine ces deux types d'analyses (en plus de certaines analyses articulatoires), mais comme la façon dont ont été analysées les durées vocaliques n'est pas explicitée, nous ne pouvons nous y attarder davantage.

diphthonguées, sans toutefois que cela ne permette une comparaison systématique entre les différentes variantes des voyelles fermées.

Les avancées technologiques en analyse acoustique du signal de parole ont permis aux chercheurs qui ont poursuivi le travail effectué par ces précurseurs de considérer plus d'un point de mesure par occurrence vocalique. Ainsi, Yaeger (1979), qui a analysé la parole de sept locuteurs montréalais, a estimé les valeurs de F_1 et de F_2 dans le noyau et dans l'appendice des voyelles qu'elle jugeait comme étant diphthonguées¹³ (mais pas des autres voyelles). Paradis (1983; 1985) a appliqué la même méthodologie à un corpus récolté auprès de 43 locuteurs de Chicoutimi et de Jonquière. Martin (2002) a quant à lui relevé les valeurs de F_1 et de F_2 à 25 % et à 75 %¹⁴ de la durée de toutes les voyelles perçues diphthonguées de son corpus; corpus constitué des enregistrements de 12 locuteurs natifs de diverses régions du Québec. Cependant, le corpus de Martin (2002) ne contenait aucune voyelle fermée en contexte allongeant ni aucune voyelle fermée brève perçue diphthonguée. Paradis (1983; 1985) n'a pas non plus identifié de voyelles fermées diphthonguées dans son corpus et seuls quelques /u/ allongés ont été considérés comme diphthongués par Yaeger (1979). Par conséquent, la dynamique spectrale des voyelles fermées allongées demeure largement ignorée.

MacKenzie et Sankoff (2009), quant à elles, ont relevé les valeurs de F_1 et de F_2 dans le noyau et l'appendice de toutes les voyelles en contexte allongeant du corpus montréalais qu'elles ont exploité.¹⁵ Elles ont ainsi pu démontrer que la différence de F_1 entre ces deux points était statistiquement significative dans le cas des voyelles fermées, quoique nettement inférieure à celle des autres voyelles. Cela dit, leur méthodologie s'appuie sur un paradigme expérimental semblable à ceux de Yaeger (1979), Paradis (1983; 1985) et Martin (2002) en ce que le relevé des mesures n'a pas été effectué aux mêmes endroits selon les contextes. MacKenzie et Sankoff (2009) utilisent deux points de mesure en (VR), mais un seul en (VK) et en (V#). Bien qu'elles comparent les valeurs du noyau des voyelles en (VR) à celles de l'unique point de mesure des voyelles dans les autres contextes, on ne peut savoir si ces données sont véritablement comparables, si elles correspondent à la même réalité acoustique dans les différents contextes.

Pour sa part, Poliquin (2006 : 33), qui a enregistré 5 locuteurs de Montréal, a relevé les valeurs de F_1 et de F_2 en trois points de chaque voyelle : « The first point was defined by the second peak in the acoustic waveform

¹³ Plus précisément, les deux points de mesure correspondaient aux points d'inflexion des deux premiers formants (Yaeger, 1979 : 14). Étaient jugées comme diphthonguées les voyelles fermées postérieures présentant une variation de 100 Hz sur F_1 et les autres voyelles présentant une variation de 200 Hz sur F_2 entre leur noyau et leur appendice, pourvu qu'elles soient effectivement perçues diphthonguées (Yaeger, 1979 : 95).

¹⁴ Martin (2002) n'assimile donc pas ses points de mesure au noyau et à l'appendice de ces voyelles. Il les définit simplement de façon à « retra[cer] [...] le déplacement en cours lors d'une diphthongaison » (Martin, 2002 : 77).

¹⁵ Les notions de *noyau* et d'*appendice* impliquent généralement qu'il soit question de diphthongues, alors que MacKenzie et Sankoff (2009) les appliquent à toute voyelle longue ou allongée. Elles définissent donc le noyau d'une voyelle comme « the initial steady state after any formant transitions out of the preceding consonant » et l'appendice comme « the final steady state before any transitions into the following consonant » (MacKenzie et Sankoff, 2009 : 92).

of the vowel, while the third and last point was located on the second to last peak of the waveform. The second point [...] corresponded to the mid point between the first and third points. ». Toutefois, les mesures obtenues ont par la suite été moyennées pour chaque occurrence, occultant ainsi toute dynamique spectrale. Par ailleurs, tout comme le corpus de Martin (2002), celui de Poliquin (2006) ne contenait pas de voyelles fermées allongées.

Les auteurs cités jusqu'à présent dans cette section n'ont estimé que la fréquence centrale des deux premiers formants des voyelles analysées, hormis Martin (2002) et Poliquin (2006), qui en ont également calculé la durée. Or, comme nous le verrons plus en détail au point 2.2, de nombreuses études indiquent que la fréquence fondamentale (F_0), la fréquence centrale des formants supérieurs (F_3 , voire F_4) et la durée sont des paramètres acoustiques qui participent aussi à la détermination du timbre des voyelles. Les plus récentes études acoustiques sur les voyelles fermées du FQ ont d'ailleurs pris en compte plusieurs de ces paramètres.¹⁶ Dans le cadre d'une étude préliminaire (Arnaud et coll., 2011), nous avons relevé la F_0 , F_1 , F_2 , F_3 , F_4 (en trois points de mesures) et la durée de voyelles fermées tendues et relâchées produites par 12 locuteurs de Saguenay. Nous avons ainsi été en mesure d'observer que « lax and tense variants exhibit opposite spectral trajectories. Between the first and the second quarter of their production, tense variants seem to increase their degree of peripherality while lax variants appear to show an increase in centralization » (Arnaud et coll., 2011 : 246). Cependant, les variantes allongées n'étaient pas non plus représentées dans ce corpus initial. Soulignons tout de même qu'il s'agit, à notre connaissance, de la seule étude s'étant intéressée à la dynamique spectrale de voyelles brèves en FQ; la dynamique spectrale n'étant traditionnellement considérée que dans l'étude de voyelles longues ou allongées potentiellement ou effectivement diphtonguées (tel qu'établi à l'audition). Enfin, Leblanc (2012), qui a analysé les productions de 10 locuteurs de Saguenay et de Québec, a relevé la F_0 , F_1 , F_2 (en quatre points de mesure) et la durée des voyelles de son corpus qui avaient préalablement été perçues comme diphtonguées (ce qui incluait des voyelles fermées allongées) par trois juges expérimentés. Il a constaté que la majorité des /i/ et des /y/ diphtongués suivaient une trajectoire linéaire (dans un diagramme F_1-F_0 / F_2-F_1 en barks),¹⁷ alors que les /u/ pouvaient suivre deux trajectoires linéaires distinctes ou même inverser leur direction en cours d'émission (Leblanc, 2012 : 63-64). Les voyelles n'ayant pas été reconnues comme diphtonguées n'ont pas été analysées acoustiquement.

¹⁶ Notons tout de même que certaines études récentes, comme celles de Séguin (2010) et de Dalton (2011a, 2011b), n'ont relevé que les valeurs de F_1 et de F_2 au centre des voyelles qu'ils ont étudiés, et que leurs corpus ne contenaient pas non plus de voyelles fermées allongées. Sur le plan acoustique, leurs résultats ne font d'ailleurs que confirmer une élévation du F_1 résultant du relâchement des voyelles fermées.

¹⁷ Nous reviendrons sur cette notion au point 2.3.1.

1.4 Objectifs

Pour résumer, les variantes tendues, relâchées et allongées des voyelles fermées du FQ, lorsqu'elles sont sous l'accent, se distribuent de manière complémentaire : les tendues apparaissent en syllabe ouverte, les relâchées, en syllabe fermée par une consonne non allongeante, et les allongées, en syllabe fermée par une consonne allongeante. Les allongées se distinguent des autres variantes par leur longue durée. Leur timbre n'est cependant pas clairement défini. Les études s'étant intéressées à la question n'ayant pas comparé systématiquement les caractéristiques spectrales de toutes les allongées (perçues ou non comme diphtonguées) à celles des autres variantes, la question demeure peu documentée sur le plan acoustique. Quant aux relâchées, elles se distinguent des tendues par une ouverture (F_1 plus élevé) et une centralisation (F_2 plus bas pour /i/ et /y/, plus élevé pour /u/). Par ailleurs, en cours d'émission, les tendues se déplaceraient vers la périphérie du trapèze vocalique alors que les relâchées se centraliseraient. Le rapport de durée entre ces deux types de variantes reste toutefois à déterminer, sachant que des résultats contradictoires sont rapportés dans la littérature et qu'il est difficile d'identifier la source de ces divergences vu les nombreuses différences méthodologiques (contexte consonantique, origine géographique des locuteurs, etc.).

Notre objectif est donc de caractériser acoustiquement les variantes tendues, relâchées et allongées des voyelles fermées du FQ sous l'accent. De cet objectif général se dégagent trois sous-questions auxquelles nous tenterons de répondre :

- Les voyelles fermées allongées se caractérisent-elles par une importante variation spectrale (pouvant éventuellement être interprétée comme une diphtongaison) ou présentent-elles plutôt des caractéristiques acoustiques se comparant à celles des variantes tendues et relâchées ?
- Quel rapport de durée entretiennent les variantes brèves tendues et relâchées des voyelles fermées ?
- Les voyelles fermées tendues, relâchées et allongées sont-elles marquées d'une variation diatopique en FQ ?

Pour être en mesure d'atteindre ces objectifs, nous devons adopter un cadre expérimental qui nous permette, en premier lieu, de comparer systématiquement ces trois types de variantes afin de mettre au jour ce qui les distingue, tant sur le plan de la durée que des caractéristiques spectrales. Notre cadre expérimental devait également prendre en compte l'évolution des indices acoustiques en cours d'émission propre à caractériser les trajectoires des voyelles. Enfin, notre échantillon devait être constitué de locuteurs provenant de différentes régions du Québec.

Chapitre 2 : Cadre expérimental

L'onde sonore de la parole est un objet complexe possédant de très nombreuses caractéristiques, dont plusieurs peuvent entrer dans la description des voyelles. Nous devons donc sélectionner, parmi les multiples indices acoustiques proposés dans la littérature, ceux que nous allons utiliser pour caractériser au mieux les voyelles fermées tendues, relâchées et allongées du FQ.

2.1 Descriptions pionnières du spectre des voyelles

À notre connaissance, la plus ancienne description acoustique des voyelles est celle donnée par Reyher (1679, rapporté dans Pfitzinger et Niebuhr, 2011 : 160-161). Ne disposant pas, à l'époque, d'une technologie lui permettant d'observer quelque représentation visuelle d'une onde sonore, Reyher était tout de même parvenu à estimer à l'audition (plus précisément à l'écoute de voyelles chuchotées), en utilisant une notation musicale, un ton caractéristique pour chacune des voyelles de l'allemand.¹⁸ À partir du XIX^e siècle, d'autres chercheurs¹⁹ ont également tenté de déterminer la fréquence de certaines composantes spectrales des voyelles, en l'occurrence des *peculiar notes* ou *pitches* (Willis, 1830 : 248), des *resonances* (Wheatstone, 1837 : 19) ou encore des *sons partiels* ou *partial tones* (Von Helmholtz, 1863, traduit et reproduit dans Von Helmholtz, 1990 : 136 ; Bell, 1879 : 164) ; toutes ces dénominations renvoyant plus ou moins explicitement à la notion d'harmoniques renforcés par les cavités de résonance du conduit vocal. Les techniques employées par ces chercheurs allaient des diapasons qu'on faisait vibrer devant la bouche lors de la phonation (Von Helmholtz, 1863, traduit dans Von Helmholtz, 1990) à la synthèse vocale à l'aide de tuyaux d'orgue (Willis, 1830), en passant par le phonautographe, permettant d'observer la forme de l'onde sonore (soit les variations de pression en fonction du temps), et le phonographe, permettant de faire jouer à vitesse variable ou en sens inverse un son préenregistré (Bell, 1879). Au tout début du XX^e siècle, Bevier (1900) est parvenu à décomposer mathématiquement l'onde sonore de la parole en la série d'harmoniques qui la composent par une transformée de Fourier. Une coupe spectrale pouvait alors être dérivée en projetant l'amplitude des harmoniques sur leur fréquence (voir Figure 1). Toutefois, comme le souligne Mattingly (1999 : 1), les techniques de l'époque (voir notamment celles rapportées par Miller, 1916), pour le moins laborieuses, n'étaient pas encore tout à fait au point.

¹⁸ Traunmüller et Lacerda (1987 : 144) font remarquer que les fréquences établies par Reyher se révèlent proches de celles du second formant effectif (F_2' ou F_2''), qu'on peut aujourd'hui calculer pour ces voyelles. Nous reviendrons brièvement sur la notion de F_2' aux points 2.2 et 2.3.2.

¹⁹ Voir Mattingly (1999) et les trois premiers chapitres de Russell (1928, reproduit dans Russell, 1970) pour de plus amples repères historiques.

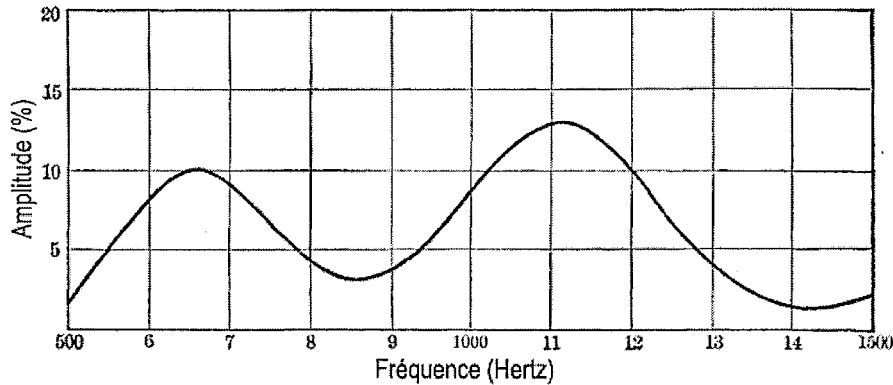


Figure 1 : Coupe spectrale d'un [ɑ]. Adaptée de Bevier (1900).

Enfin, au milieu du xx^e siècle, l'invention du spectrographe a permis aux phonéticiens d'obtenir une représentation visuelle de l'évolution temporelle du spectre des voyelles : le spectrogramme, qui est une projection de la fréquence des harmoniques sur l'axe du temps. L'amplitude des harmoniques y est généralement représentée par le degré de noirceur du trait. Comme l'expliquent Boite et coll. (2000 : 12-13), un spectrogramme à bande étroite (calculé à partir de fenêtres relativement longues du signal sonore - typiquement 30 ms) permet d'observer distinctement chacune des harmoniques du signal (qui apparaissent alors comme des lignes horizontales), tandis qu'un spectrogramme à bande large (calculé à partir de fenêtres plus courtes - par exemple 10 ms) met en évidence l'enveloppe spectrale (se présentant comme une version « lissée » d'un spectrogramme à bande étroite et révélant les zones de fréquences associées aux fréquences de résonance du conduit vocal).

Les premières observations spectrographiques de voyelles ont confirmé l'existence d'harmoniques – de groupes d'harmoniques, pour être plus exacte – plus intenses que les autres : les formants²⁰ (voir Figure 2).

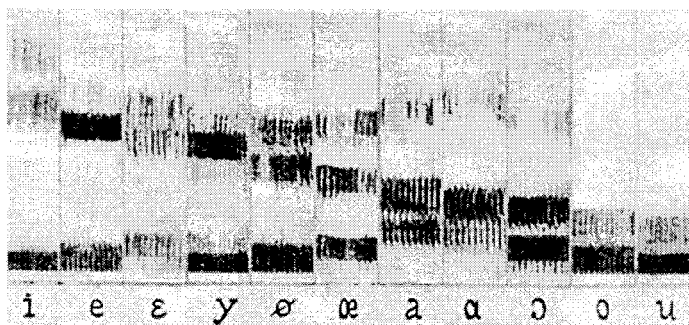


Figure 2 : Spectrogrammes à bande large de voyelles isolées. Tirée de Delattre (1948).

²⁰Delattre (1948) les appelait *formantes*. En anglais, le terme *bars* était parfois rencontré (par exemple dans Potter et Peterson, 1948). Potter et Steinberg (1950 : 807) précisent que « "formant" will mean a frequency region of energy concentration in the spectrum of a vowel sound. "Bars" are visible manifestations of such frequency regions as they appear in sound spectrograms ».

Ces observations ont de plus rapidement conduit les chercheurs à établir une corrélation²¹ entre la fréquence centrale du premier formant (celui dont la fréquence est la plus basse, F_1) d'une voyelle et son degré d'aperture, et entre la fréquence centrale du second formant (F_2) d'une voyelle et son degré d'antériorité (Joos, 1948 : 53 ; Delattre, 1948 : 482-483). Plus F_1 est élevé, plus la voyelle est ouverte ; plus F_2 est élevé, plus elle est postérieure, et vice-versa. L'aperture et l'antériorité étant depuis Hellwag (1781, reproduit dans Hellwag, 1886 : 41) regardées comme les traits articulatoires déterminants de l'identité des voyelles, F_1 et F_2 ont d'emblée été considérés comme les corrélats acoustiques essentiels (et même suffisants selon Delattre, 1948 : 478) du timbre des voyelles (une notion intuitive référant à la qualité spécifique d'un son indépendante de sa hauteur, de sa durée ou de son intensité, et qui peut être transcrite par un symbole de l'alphabet phonétique international). L'étude de Delattre et coll. (1952) a renforcé cette conception en démontrant qu'il était possible de synthétiser, à partir de seulement deux formants, des voyelles identifiables lors d'un test de perception.

C'est ainsi que dans les décennies qui ont suivi, et jusqu'aujourd'hui, une grande majorité des études phonétiques ont adopté un paradigme d'analyse fondé sur la seule prise en compte de F_1 et de F_2 mesurés dans l'état stable des voyelles. Labov (2006 : 500) explique d'ailleurs que ces deux paramètres se sont révélés suffisamment robustes pour permettre l'observation de changements linguistiques en cours et de variations phonétiques corrélées aux caractéristiques sociales des locuteurs.

Cela étant dit, si on s'intéresse non plus aux tendances observables à l'échelle d'une communauté linguistique, mais bien à la catégorisation de chaque occurrence vocalique, on constate rapidement que F_1 et F_2 mesurés au centre de l'état stable ne suffisent pas à circonscrire l'identité des voyelles. En effet, les fréquences formantiques d'occurrences d'une même catégorie phonémique varient grandement en fonction du sexe du locuteur, de son âge, du contexte consonantique, du débit de parole, etc. Si bien que dans un diagramme F_1 / F_2 , « [c]onsiderable overlapping of areas [representing different vowels] is indicated » (Peterson et Barney, 1952 : 182). La Figure 3 illustre de manière éloquente cet état de fait.

²¹Traunmüller et Lacerda (1987 : 144) affirment que Jung (1926), même sans spectrogramme, avait été le premier à relever cette corrélation.

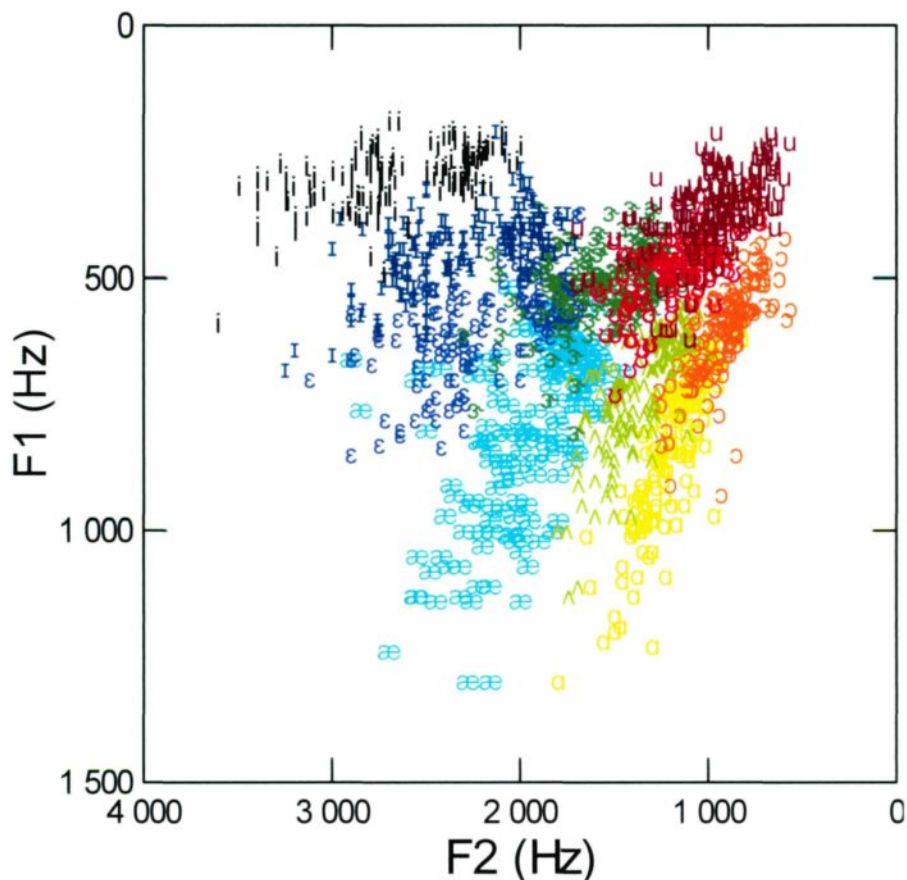


Figure 3 : Diagramme de dispersion F_1/F_2 de 1520 occurrences des dix monophthongues de l'anglais américain, produites en contexte [hVd] par 33 hommes, 28 femmes et 15 enfants. D'après les données de Peterson et Barney (1952).

Pourtant, les occurrences comprises dans ces zones de chevauchement demeurent identifiables (c'est-à-dire qu'elles sont reconnues, pour l'essentiel, comme appartenant à la catégorie phonémique attendue). Quels indices acoustiques supplémentaires, alors, favorisent l'identification de ces segments ?

2.2 Indices acoustiques supplémentaires

Comme F_1 et F_2 ne semblaient pas suffire à la définition de l'identité des voyelles, la contribution d'autres indices acoustiques a été envisagée. Parmi ceux-ci, mentionnons la fréquence centrale des formants supérieurs (F_3 , voire F_4). Déjà, Peterson et Barney (1952 : 182) notaient que la voyelle rothique [ə] se distinguait nettement des autres voyelles de l'anglais américain par son F_3 très bas. De nombreuses études ont par la suite confirmé que « [v]ariations in the frequency of F_3 indicate the degree of r-coloring: the lower the F_3 , the greater the degree of rhoticity » (Ladefoged, 2003 : 149). F_3 a de plus été associé à l'arrondissement des voyelles. Lindblom et Sundberg (1971), en utilisant la synthèse vocale articulatoire (qui consiste à dériver les caractéristiques spectrales d'un son correspondant à une configuration donnée du conduit vocal en calculant les fréquences de résonance de ce dernier), ont en effet remarqué que

l'arrondissement des lèvres abaissait la fréquence centrale de tous les formants, mais que ce phénomène affectait de manière particulièrement prononcée F_3 . Comme l'explique Vaissière (2007 : 65), « [a]s the effective length of the front cavity increases in the passage from French /i/ to /y/, F_3 , a front-cavity resonance, lowers and a change of cavity affiliation takes place ». Ainsi, alors que le F_3 et le F_4 de /i/ sont très près l'un de l'autre, c'est le F_2 et le F_3 qui sont rapprochés dans le cas des /y/. Pour cette raison, les voyelles /i/ et /y/ sont dites focales ; c'est-à-dire qu'il y a focalisation de l'énergie spectrale due à la proximité de deux formants (Schwartz et coll., 1993 : 415). La prise en considération de F_3 et de F_4 permet donc de rendre compte de la focalisation de ces voyelles (Gendrot et coll., 2008).

La largeur de bande des formants (l'intervalle de fréquences de forte amplitude que ceux-ci constituent) a également été avancée comme indice susceptible de participer à la discrimination des voyelles. House et Stevens (1956) ont entre autres constaté que l'ajout d'une résonance nasale augmentait la largeur du premier formant de voyelles anglaises synthétiques, et que les caractéristiques acoustiques de ces stimuli entraînaient bien une perception de nasalité chez des auditeurs. Pour ce qui est du timbre des voyelles orales, des tests de perception utilisant des stimuli synthétiques ont montré que la manipulation de la largeur de bande des formants ne modifiait pas, ou très peu (Klatt, 1982 ; Carlson et coll., 1979 ; Beddor et Hawkins, 1991), la perception que les auditeurs avaient de ces voyelles ; ce qui a mené Rosner et Pickering (1994 : 165) à conclure que « [v]owel identification [...] appears insensitive to changes in formant bandwidths as such. Altered categorization seems to occur only when bandwidth manipulations make formant peaks difficult to detect in the stimulus spectra ».

La contribution de l'amplitude des formants à la définition de l'identité des voyelles a également été examinée. Par exemple, Delattre (1954 : 106) remarque que lors du passage d'une voyelle orale française à son équivalent nasal, la zone fréquentielle correspondant au premier formant de la voyelle orale perd considérablement en intensité. L'importance de cette baisse d'intensité dans la perception de la nasalité a été confirmée par les tests de perception qu'il a menés à l'aide de voyelles synthétiques. Cela dit, si on exclut les voyelles nasales, l'importance de l'amplitude (ou de l'amplitude relative) des formants dans la perception des voyelles ne fait pas l'unanimité. Ainsi, les auditeurs de Miller (1953) perçoivent une différence de timbre lorsque l'amplitude des formants de voyelles synthétiques est modifiée, mais pas ceux de Klatt (1982). Bladon et Fant (1978), Chistovich et Lublinskaya (1979) de même que Mantakas et coll. (1986) soutiennent quant à eux que l'amplitude relative des formants joue un rôle non négligeable dans l'intégration spectrale. Le concept d'intégration spectrale a été formulé à la suite d'études (notamment celle de Delattre et coll., 1952) qui révélaient que le second formant de voyelles synthétiques à deux formants ne correspondait pas toujours au second formant des voyelles naturelles correspondantes (voir Figure 4).

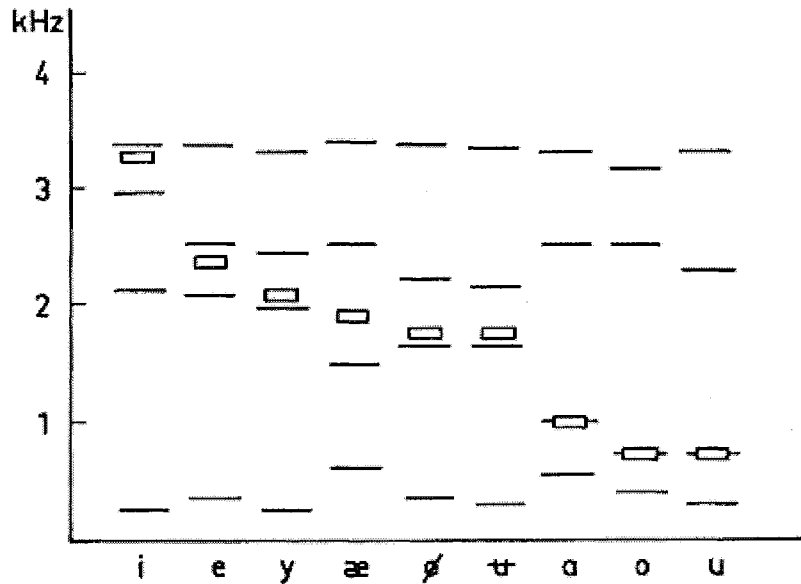


Figure 4 : Position des quatre premiers formants de voyelles naturelles (indiqués par les lignes) et du second formant des voyelles synthétiques correspondantes (représenté par un rectangle). Tirée de Carlson et coll. (1970).

Il a donc été postulé que lorsque deux pics spectraux (ou formants) sont suffisamment rapprochés, ils sont intégrés par le système auditif humain comme une seule et même prééminence spectrale (correspondant au F_2 de voyelles synthétiques à deux formants). L'amplitude relative des formants ainsi moyennés influencerait la position de cette prééminence spectrale perçue (nommée *centre de gravité* ou *second formant effectif*²² selon les théories). Ces conclusions sont cependant contestées (voir Rosner et Pickering, 1994 : 139-143). Citons par exemple Assmann (1991), qui ne parvient pas à reproduire les résultats de Chistovich et Lublinskaya (1979) lorsque des stimuli synthétiques à 6 formants (plus proches de voyelles naturelles) sont utilisés. Enfin, l'amplitude relative des formants forme, sur une coupe spectrale, un angle d'inclinaison rendant compte du rapport global entre l'amplitude des composantes de basse fréquence du spectre, et celles de plus haute fréquence (voir Figure 5). Cette inclinaison se nomme *penne spectrale* (ou *spectral tilt*).

²² Nous reviendrons brièvement sur la notion de second formant effectif, ou F_2 , au point 2.3.2.

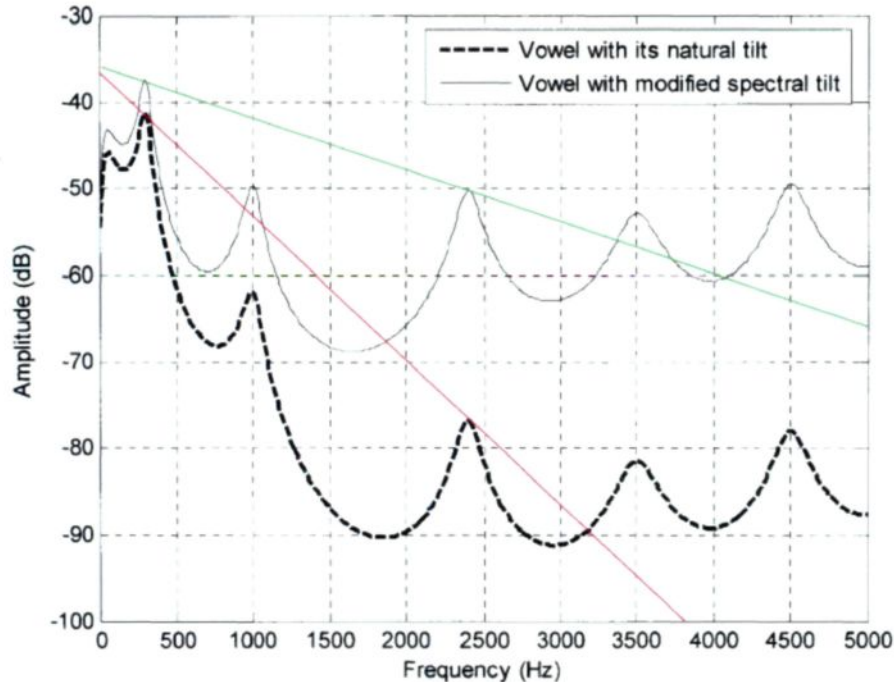


Figure 5 : Exemple de spectre d'un [i] synthétique présentant une forte pente spectrale (en rouge) et d'une version modifiée de ce spectre où la pente a été atténuée (en vert). Adaptée de Kale (2006).

Sur ce point, il n'y a pas non plus de consensus au sein de la communauté scientifique. Par exemple, les résultats du test de perception mené par Klatt (1982) indiquent que les auditeurs ne sont pas sensibles aux modifications de la pente spectrale de voyelles synthétiques. Pourtant, Ito et coll. (2001) rapportent qu'en l'absence du second formant, la pente spectrale de stimuli synthétiques permet aux auditeurs japonais d'identifier les voyelles malgré tout et que même en présence de F_2 , la perception des auditeurs est influencée par la pente spectrale. Ces résultats contradictoires s'expliquent possiblement, du moins en partie, par les caractéristiques des stimuli présentés aux auditeurs. En effet, l'influence de la pente spectrale paraît dépendre de nombreux facteurs. Kiefe et Kluender (2005) remarquent que la modification de la pente spectrale de voyelles anglaises synthétiques influe sur la perception de celles-ci, mais seulement s'il s'agit de monophongues. Kale (2006) observe pour sa part que la perception de ses sujets anglophones n'ayant pas de trouble de l'audition est influencée par la pente spectrale de stimuli synthétiques uniquement lorsque la fréquence centrale du deuxième formant est supérieure à 1200 Hz. Finalement, la modification de la pente spectrale de voyelles anglaises naturelles ne semble pas avoir d'impact sur la perception de celles-ci (Hillenbrand, 2004).

Un autre indice acoustique fréquemment employé dans la description des voyelles est la fréquence fondamentale, ou F_0 ; c'est-à-dire le nombre de cycles d'oscillation qui composent l'onde sonore de la parole par unité de temps (généralement la seconde). La F_0 est reconnue comme étant le corrélât acoustique de la

hauteur perçue de la voix, aussi appelée *tonie* (Martin, 1996 : 135). Cela dit, elle est également réputée jouer un certain rôle dans la perception du timbre des voyelles. De nombreuses études (Miller, 1953; Slawson, 1967; Fujisaki et Kawashima, 1968; Ainsworth, 1971; Ainsworth, 1975) ont en effet montré que la modification de la F_0 de voyelles synthétiques affecte de manière plus ou moins marquée la perception de leur timbre. Plus précisément, la distance, ou le rapport, entre la fréquence fondamentale et la fréquence centrale de certains formants semble être corrélée à la perception de traits articulatoires comme le degré d'aperture ou d'antériorité et l'arrondissement (Traunmüller, 1981; Syrdal et Gopal, 1986; Traunmüller et Lacerda, 1987; Hirahara et Kato, 1992; Ménard et coll., 2002). Nous y reviendrons au point 2.3.2. Quoi qu'il en soit, depuis Whalen et Levitt (1995), il est admis que la F_0 intrinsèque des voyelles fermées est universellement plus élevée que celle des voyelles ouvertes (Maddieson, 1999). Cette tendance est d'ailleurs attestée en français de France (Di Cristo, 1985 : 87) et en FQ (Lavoie, 1995 : 72). Cependant, la F_0 des voyelles tendues ne semble pas se différencier de celle des relâchées en allemand (Fischer-Jørgensen, 1990; Pape et Mooshammer, 2006; Schneeberg et Schlüßler, 2006), mais en FQ, Lavoie (1995 : 72) remarque que dans certaines circonstances, la F_0 des voyelles fermées tendues semble plus élevée que celle des relâchées.²³

La durée des voyelles est quant à elle conditionnée par l'environnement consonantique, la position dans l'énoncé (notamment par rapport à l'accent), la structure syllabique, le débit de parole, mais également par les propriétés intrinsèques des voyelles.²⁴ Ces propriétés peuvent être de nature articulatoire ou phonologique. Parmi les contraintes biomécaniques qui affectent la durée des voyelles françaises, on retrouve celles imputables à la nasalité, au degré d'aperture et éventuellement au degré d'antériorité. Ainsi, Ouellet (1992 : 16) explique que les voyelles nasales françaises sont plus longues que leurs pendants orales en raison du temps que prendrait le voile du palais à s'abaisser. En ce qui concerne le degré d'aperture, la tendance universelle - bien attestée en français (Di Cristo, 1985; Ouellet, 1992; Morasse, 1995) - est que plus une voyelle est ouverte, plus elle est longue, possiblement à cause du temps pris par la langue et la mâchoire pour s'abaisser (Maddieson, 1999). Enfin, Ouellet (1978 : 719) et Ouellet (1992 : 162) remarquent qu'en FQ, les voyelles postérieures semblent très légèrement plus longues que les antérieures. Sur le plan phonologique, les voyelles dites *longues par nature*, ou *historiquement longues* sont plus longues que les voyelles brèves. En français, cette longueur intrinsèque ne se manifeste qu'en syllabe fermée. Les voyelles françaises /ø o ɑ ē œ õ ā/ sont considérées comme longues par nature. En FQ, le /ɛ:/ présent dans des mots comme « fête » (par opposition au /ɛ/ bref de « faites ») intègre également cette catégorie (Côté, 2010).

²³ Notons que dans le travail de Lavoie (1995), les voyelles fermées dites tendues étaient toutes situées en contexte allongeant.

²⁴ Nous renvoyons le lecteur ou la lectrice à la synthèse de Klatt (1976), au chapitre IV de Di Cristo (1985), à la thèse de Ouellet (1992), ainsi qu'au mémoire de Morasse (1995) pour les explications concernant les facteurs extrinsèques de variation de la durée vocalique.

En dernier lieu, mentionnons que l'intensité des voyelles semble, toutes autres choses étant égales par ailleurs, différer d'une classe vocalique à l'autre. À notre connaissance cependant, cette variation demeure peu étudiée et semble encore mal comprise en français (Di Cristo, 1985; Ouellon et coll., 1993; Dolbec et Rogers, 1996).

La prise en compte supplémentaire de paramètres acoustiques autres que F_1 et F_2 permet donc de mieux différencier les classes vocaliques. Cela dit, la manière même d'envisager la prise en compte des indices acoustiques constitue en soi une façon d'aborder la question de la classification des voyelles. Deux grands paradigmes expérimentaux se sont notamment structurés en ce sens. Sans s'opposer et sans être mutuellement exclusifs, ces paradigmes reposent néanmoins sur des fondements bien distincts. Le premier consiste à opérer des modifications aux fréquences formantiques (et éventuellement à la F_0) dans le but de réduire la variation interindividuelle liée aux caractéristiques physiologiques des locuteurs. Il s'agit de la normalisation vocalique. Le second postule qu'une large part de l'information discriminante résiderait dans l'évolution des indices acoustiques en cours d'émission, c'est-à-dire dans la dynamique spectrale des voyelles.

2.3 Normalisation vocalique

Les procédures de normalisation peuvent être abordées de deux points de vue : celui de la perception²⁵ et celui de la production des voyelles. En ce qui concerne la perception, Johnson (1990 : 230) définit la normalisation comme « a hypothetical process in which interspeaker vowel variability is reduced in order that perceptual vowel identification may then be performed by reference to relative vowel quality rather than to the absolute values of the acoustic parameters of vowels ». Dans l'étude de la production des voyelles cependant, les procédures de normalisation sont envisagées comme des opérations ayant pour objectif de réduire spécifiquement la variabilité d'origine physiologique (liée à la forme et aux proportions du conduit vocal, et donc en grande partie au sexe et à l'âge du locuteur) tout en préservant l'information phonémique, dialectale et sociolinguistique contenue dans le signal sonore de la parole (Thomas et Kendall, 2007).

Clopper (2009 : 1431) souligne que « [v]owel normalization algorithms can be classified in several ways, based on the amount and type of information that is required and the type of metric that is used to calculate the normalized values from a data set ». L'une des distinctions les plus couramment opérées entre les procédures de normalisation est sans doute la distinction intrinsèque / extrinsèque. Les procédures qui puisent toute l'information nécessaire à la normalisation d'une valeur fréquentielle à l'intérieur de la seule occurrence d'où provient cette valeur sont qualifiées d'intrinsèques (*vowel-intrinsic*). Celles qui requièrent de l'information provenant d'autres occurrences sont dites extrinsèques (*vowel-extrinsic*). Dans le cadre de ce travail, nous

²⁵ Voir notamment Johnson (2005) à ce sujet.

jugions pertinent d'affiner cette distinction et de présenter les différentes procédures en les classant selon les trois catégories suivantes : les échelles psycho-acoustiques (intrinsèques), les objets et espaces psycho-acoustiques (intrinsèques), et les opérations extrinsèques. Mentionnons que les opérations extrinsèques ne sont généralement pas considérées comme représentatives des mécanismes psycho-acoustiques de la perception humaine de la parole (ce qui n'est d'ailleurs pas leur objectif) car un auditeur est capable d'identifier une voyelle isolée produite par un locuteur inconnu, donc sans avoir accès à d'autres informations que celles contenue dans l'occurrence en question. Cela dit, la connaissance préalable de la voix du locuteur facilite l'identification (Strange et coll., 1976; Assmann et coll., 1982; Mullenix et Pisoni, 1988). Bref, si les informations extrinsèques ne sont pas suffisantes pour expliquer la perception des voyelles, elles semblent néanmoins jouer un rôle dans ce processus lorsqu'elles sont disponibles dans la mémoire.

2.3.1 Échelles psycho-acoustiques

Comme le soulignent Thomas et Kendall (2007) ainsi que Fabricius et coll. (2009 : 417), certaines transformations, parfois envisagées comme des procédures de normalisation, sont d'abord des opérations de conversion : il s'agit de passer d'une échelle de mesure physique et linéaire de la fréquence (le nombre de cycles d'oscillation par seconde, ou *Hertz*) à une échelle psycho-acoustique censément représentative de l'audition humaine. Différentes échelles ont ainsi été proposées afin de rendre compte de divers aspects de la perception de la fréquence (consonance, hauteur, résolution de la fréquence, etc.). Ces opérations ne sont donc pas fondamentalement destinées à éliminer la variabilité d'origine physiologique, bien que leur potentiel normalisateur ait souvent été éprouvé (Syrdal et Gopal, 1986; Hillenbrand et Gayvert, 1993; Adank, 2003; Calamai, 2005; Volín et Studenovský, 2007; Clopper, 2009; Flynn et Foulkes, 2011) et que certaines d'entre elles soient intégrées à d'autres procédures de normalisation. Les principales opérations de conversion rencontrées en sciences phonétiques sont :

- le logarithme décimal ou népérien de la fréquence en Hertz²⁶;
- les échelles musicales : demi-ton, ton, octave, etc. ;
- l'échelle des mels, proposée par Stevens et Volkman (1940) ;
- l'échelle de Koenig (1949) ;
- l'échelle des barks, proposée par Zwicker (1961) ;
- l'échelle des *equivalent rectangular bandwidth-rate* (ou *ERB-rate*), proposée par Moore et Glasberg (1983).

²⁶ Il s'agit en fait de l'approximation la plus simple de la pseudo loi de Weber-Fechner, qui postule que la sensation varie comme le logarithme de l'excitation.

2.3.2 Objets et espaces psycho-acoustiques

D'autres procédures intrinsèques impliquent quant à elles la création d'un objet ou d'un espace psycho-acoustique. Elles ne consistent donc pas à faire correspondre à chaque mesure fréquentielle un équivalent normalisé, mais bien à construire de nouveaux indices psycho-acoustiques à partir d'au moins deux composantes spectrales. Ainsi, il a depuis longtemps été suggéré d'exprimer la fréquence de la F_0 et des formants par leur position relative dans le spectre plutôt que par leur position absolue. On peut entre autres calculer le ratio entre deux formants (F_3 / F_2 , F_3 / F_1 ou F_2 / F_1) ou entre F_1 et la F_0 (Lloyd, 1890; Potter et Steinberg, 1950; Peterson, 1951; Miller, 1989; Monahan et Idardi, 2010). Comme l'illustre la Figure 6, le rapport entre F_1 et F_2 semble spécifique à certains groupes de classes vocaliques.

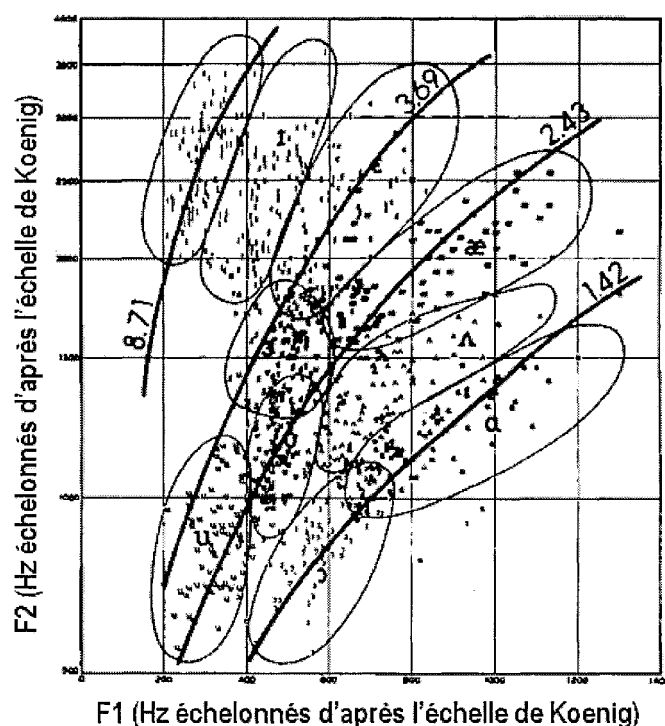


Figure 6 : Diagramme bi-formantique F_2 / F_1 présentant les voyelles de l'anglais américain d'après les données de Peterson et Barney (1952). Les lignes courbes indiquent les ratios entre F_2 et F_1 des différentes classes vocaliques. Adaptée de Miller (1989).

On peut également calculer la distance, sur une échelle tonotopique, telle que celle des barks, entre deux formants ou entre un formant et la F_0 .²⁷ Comme on l'a vu au point 2.2, la distance entre F_1 et la F_0 a été associée à la perception du degré d'aperture dans un dialecte bavarois (Traunmüller, 1981), en anglais

²⁷ Exprimer la position relative des composantes spectrales par leur ratio ou leur distance revient mathématiquement au même lorsqu'une échelle à base logarithmique (comme le sont toutes les échelles psycho-acoustiques présentées au point 2.3.1) est employée, puisque : $\log(x) - \log(y) = \log(x/y)$.

(Syrdal et Gopal, 1986), en suédois et en turc (Traunmüller et Lacerda, 1987), de même qu'en français de France (Ménard, 2002; Ménard et coll., 2002). Les distances entre F_2 et F_1 , entre F_3 et F_2 , ou encore entre F_2 et F_0 ont tour à tour été proposées comme corrélat acoustique du degré d'antériorité perçu en suédois (Fant, 1983), en anglais (Syrdal et Gopal, 1986), en japonais (Hirahara et Kato, 1992) et en français (Ménard, 2002; Ménard et coll., 2002). La distance entre F_2 et F_1 a de plus été associée à la perception de la rhoticité en anglais (Heselwood, 2009). Finalement, les distances entre F_2 et F_1 , de même qu'entre F_2 et F_0 concourraient possiblement à la perception de l'arrondissement en suédois et en turc (Traunmüller, 1981).

Le second formant effectif (F'_2), déjà évoqué au point 2.2, est un autre objet psycho-acoustique qui a été avancé pour rendre compte de la perception des voyelles. Cette prééminence spectrale qui réunirait en un seul percept F_2 , F_3 , voire F_4 aurait « a greater specificational power than F_2 alone and [...] appears to be the better parameter to be used together with F_1 in a two-dimensional representation » (Fant, 1978, réédité dans Fant, 2007 : 210). F'_2 permettrait également de mieux distinguer les voyelles arrondies des voyelles non arrondies, du moins en français (Mantakas et coll., 1986; Ménard, 2002; Ménard et coll., 2002).

D'autres propositions d'objets et d'espaces psycho-acoustiques ont par ailleurs été formulées; propositions qui, à notre connaissance, sont restées relativement confidentielles, comme le *auditory-oriented vowel diagram* de Fant (1969), la *dimensionless constant l* de Traunmüller et Lacerda (1987), de même que les cinq *functional dimensions* de Traunmüller (1987). L'espace bidimensionnel d'Iri (1959) a néanmoins connu un certain engouement plusieurs années après sa publication (Sussman, 1986; Thomas, 1997; Anderson, 1999; Wolfram et coll., 2011).

2.3.3 Opérations extrinsèques

Enfin, certaines opérations, extrinsèques, consistent à faire correspondre à chaque fréquence formantique un équivalent normalisé (par exemple Gerstman, 1968; Lobanov, 1971; Nordström et Lindblom, 1975; Fant, 1975; Nearey, 1977; Watt et Fabricius, 2002; Labov et coll., 2006). Au contraire des opérations de conversion cependant, ces procédures comportent au moins une autre variable que la valeur fréquentielle de départ. Elles impliquent en effet le calcul de données statistiques (maxima, minima, moyennes, écarts-types, etc.) à partir desquelles sont exprimées les valeurs fréquentielles à normaliser : égalisation des étendues, application de facteurs d'échelle, expression des valeurs par leur distance à la moyenne, etc. Par exemple, Gerstman (1968) suggère, pour chaque locuteur, de rééchelonner les valeurs de F_1 et de F_2 de façon à ce qu'elles soient comprises entre 0 et 999; c'est-à-dire que la valeur normalisée 0 soit attribuée à la plus basse fréquence de chaque formant et 999, à sa valeur la plus élevée; les autres valeurs étant ensuite linéairement réparties entre ces deux extrêmes. Lobanov (1971), de son côté, calcule les scores-Z des valeurs fréquentielles pour chaque formant d'un même locuteur.

2.4 Dynamique spectrale

Les procédures de normalisation que nous venons d'exposer brièvement²⁸ ont toutes été développées à partir de l'approche expérimentale traditionnelle, statique, qui se fonde sur le fait que l'information acoustique permettant de distinguer les voyelles les unes des autres est contenue en un point unique – un instant t – dans la partie centrale de l'état stable des occurrences. Mais comme le remarquent Andruski et Nearey (1992 : 390) : « in coarticulated vowels, formants often do not reach the steady-state values of their isolated counterparts, and frequently do not contain any region which can be described as "steady state" ». Qui plus est, de nombreuses études sur la perception des voyelles ont révélé que même en l'absence de cet état stable (dans des voyelles dont la portion centrale a été remplacé par du silence), les auditeurs parvenaient, dans la majorité des cas, à correctement identifier les voyelles produites (Strange et coll., 1976, 1983; Rakerd et Verbrugge, 1987; Murphy et coll., 1989; Andruski et Nearey, 1992; Jenkins et Strange, 1987, 1999; etc.). Or, si l'information contenue dans l'état stable n'est pas essentielle à l'identification des voyelles, où se situent les autres indices acoustiques impliqués dans cette identification ?

Au milieu du XX^e siècle, Joos (1948 : 102) offrait déjà une piste de réponse en remarquant qu'une certaine évolution des formants en cours de production semblait spécifique à une catégorie phonémique donnée (n'étant pourtant pas une diphtongue) et indépendante du contexte. Quatre ans plus tard, Peterson et Barney (1952 : 184) suggéraient à leur tour que « the complex acoustical patterns represented by the words are not adequately represented by a single section, but require a more complex portrayal. The initial and final influences often shown in the bar movements of the spectrogram are of importance here ». Tiffany (1953) observa par la suite que des auditeurs identifiaient généralement avec plus d'acuité des voyelles entières que des voyelles dont la seule portion centrale avait été conservée. C'est finalement tout un courant de pensée qui s'organisa autour de l'idée que la dynamique spectrale des voyelles joue un rôle clé dans la détermination de leur timbre.

2.4.1 Undershoot compensation theory

La *undershoot compensation theory*, issue des travaux de Lindblom et Studdert-Kennedy (1967), considère la dynamique spectrale impliquée dans la reconnaissance des voyelles comme dépendante de la coarticulation. Elle prédit que même si la cible (la configuration formantique canonique d'une classe vocalique) n'est pas atteinte lors de la production d'une voyelle en contexte, le système perceptif parviendra à compenser l'écart entre les configurations formantiques atteinte et visée. L'ampleur et la direction de cette compensation seraient déterminées par deux facteurs : la position relative du locus des consonnes adjacentes (c'est-à-dire le point d'où semblent provenir ou vers lequel semblent tendre les formants d'une voyelle) et la pente des

²⁸ Ces procédures sont présentée de manière plus détaillée dans notre étude préliminaire dédiée à la normalisation (Sigouin, à paraître, voir Annexe 2).

transitions C-V et V-C. Ainsi, si la fréquence du locus est inférieure à celles de la configuration formantique atteinte, le système perceptif présume que cette configuration est inférieure à celle qui était visée et surévalue alors les fréquences perçues. De même, si les transitions C-V et V-C sont rapides, les fréquences perçues sont surévaluées pour pallier l'écart entre configurations visée et atteinte qu'entraîne normalement un débit de parole rapide. Lindblom et Studdert-Kennedy (1967) sont parvenus à ces conclusions en synthétisant des voyelles sur le continuum /ɪ/-/ʊ/. Ils ont ensuite présenté ces voyelles dans trois contextes (/jVj/, /wVw/ et /#V#/). Les résultats du test de perception ont révélé que la position de la frontière entre /ɪ/ et /ʊ/ se déplaçait en fonction du contexte consonantique (voir Figure 7).

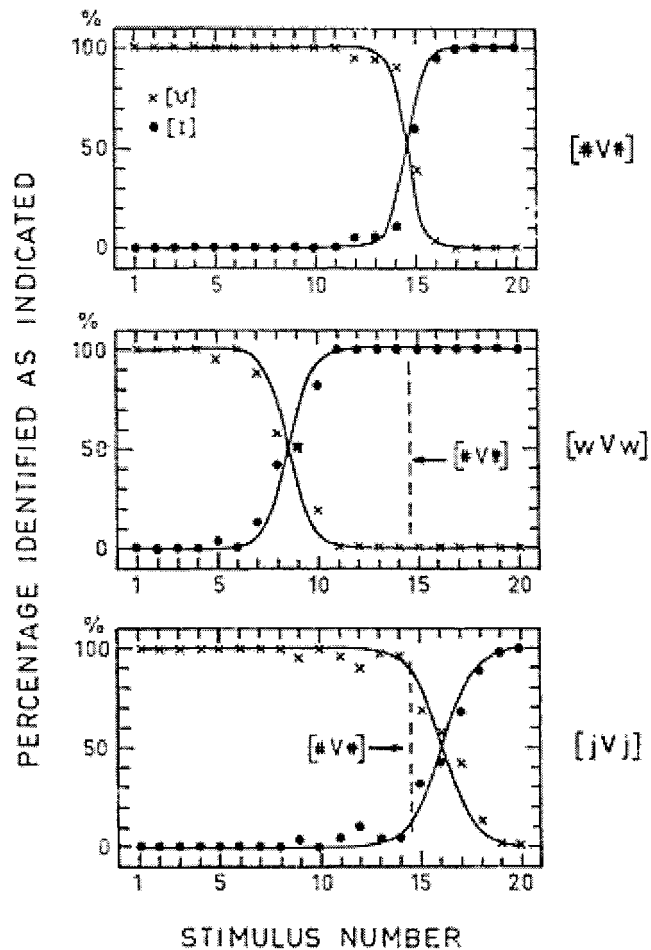


Figure 7 : Résultats du test de perception de Lindblom et Studdert-Kennedy (1967). Les pourcentages de réponses [ɪ] sont indiqués par les points et les pourcentages de réponses [ʊ], par les croix. Tirée de Lindblom et Studdert-Kennedy (1967).

En réduisant la durée des voyelles des stimuli de 200 ms à 100 ms (ce qui a pour effet d'augmenter la pente des transitions formantiques), le déplacement des frontières était encore plus marqué.

2.4.2 *Dynamic cospecification theory*

La *dynamic cospecification theory* postule quant à elle que les transitions C-V et V-C sont en elles-mêmes pertinentes pour l'identification des voyelles, sans leur attribuer la propriété de révéler quelque cible sous-jacente. Dit autrement, « the dynamic regions convey vowel information that is complementary to, and distinct from, target information » (Verbrugge et Rakerd, 1987 : 40). Selon Strange et coll. (1976), l'information impliquée dans la reconnaissance des voyelles contenue dans ces transitions serait intrinsèquement coarticulatoire puisque des voyelles en contexte sont mieux identifiées que des voyelles isolées. Verbrugge et Rakerd (1987) rejettent d'autre part l'idée d'une cible extrapolée ou restaurée à partir des transitions C-V et V-C, car même dans un contexte où une hypothétique cible ne pourrait être correctement déduite (dans des syllabes de type /CVC/ dont le centre a non seulement été remplacé par du silence, mais dont les portions initiale et finale ont de plus été produites par des locuteurs de sexes opposés aux configurations formantiques bien différentes), les auditeurs parviennent à identifier les voyelles aussi bien que si une cible avait pu être extrapolée.

2.4.3 *Vowel-inherent spectral change (VISC)*

Le terme *vowel-inherent spectral change* (VISC), proposé par Nearey et Assmann (1986 : 1297), « refers to the relatively slowly varying changes in formant frequencies associated with vowels themselves, even in the absence of consonantal context ». Sur le plan perceptif, de nombreux travaux mettent en évidence le rôle de la dynamique spectrale des voyelles et dénotent son caractère intrinsèque.²⁹ Par exemple, des voyelles isolées (donc exemptes de toute influence contextuelle) sont mieux identifiées lorsqu'elles sont présentées entières que lorsque leurs portions initiale et finale (renseignant sur la dynamique spectrale) ont été retranchées (Tiffany, 1953 ; Assmann et coll., 1982) ou que leur noyau et leur appendice ont été inversés (Nearey et Assmann, 1986). Andruski et Nearey (1992) ont, quant à eux, soumis à des auditeurs des voyelles isolées et des voyelles en contexte /hVd/ dont le centre avait été retiré et dont les portions initiale et finale n'avaient pas été produites par le même locuteur (*hybrid silent centre syllables*). Les erreurs d'identification se sont révélées similaires pour les voyelles isolées et les voyelles en contexte, ce qui suggère que les auditeurs « are using essentially the same type of information in essentially the same way to identify both kinds of stimuli » (Andruski et Nearey, 1992 : 390). Cette information serait donc propre à la voyelle (le seul élément commun aux deux types de stimuli) et non tributaire du contexte consonantique comme le proposaient les tenants de la *undershoot compensation theory* et de la *dynamic cospecification theory*.

Cette apparente résistance de la dynamique spectrale des voyelles aux effets de la coarticulation est d'ailleurs attestée sur le plan acoustique. Hillenbrand et coll. (2001) remarquent en effet que les trajectoires moyennes

²⁹ Voir Hillenbrand (2013) pour une revue plus complète.

de voyelles de l'anglais américain extraites de contextes consonantiques variés sont approximativement les mêmes qu'en contexte neutre /hVd/ dans 5 cas sur 8 (voir Figure 8).

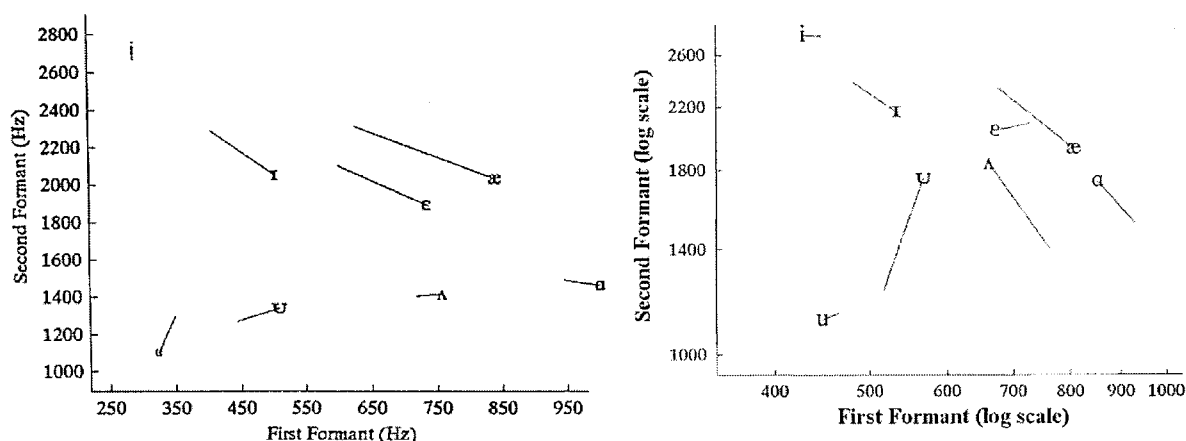


Figure 8 : Diagrammes bi-formantiques F_2 / F_1 présentant les trajectoires moyennes de 8 voyelles de l'anglais américain en contextes variés (à gauche) et en contexte neutre /hVd/ (à droite) produites par des femmes. Le symbole phonétique correspond à la position finale de la voyelle. Adaptée de Hillenbrand et Nearey (1999) et Hillenbrand et coll. (2001).

Les résultats de leurs analyses discriminantes montrent qu'en dépit des effets du contexte, la prise en compte de la dynamique spectrale améliore la classification des occurrences en fonction de leur classe phonémique (Hillenbrand et coll., 2001 : 758).

Sur la Figure 8, on peut également observer que les voyelles /ɪ/, /ɛ/ et /æ/, pourtant considérées comme des monophthongues, tant d'un point de vue phonétique que phonologique, présentent des trajectoires substantielles. Cela dit, la dynamique spectrale ne serait peut-être pas aussi essentielle à la séparabilité acoustique des monophthongues qu'à celle des diphtongues (Watson et Harrington, 1999 : 463). Cependant, elle pourrait accentuer la distinction entre les voyelles tendues et relâchées, qui montrent des trajectoires opposées en anglais (Jenkins et Strange, 1999 : 444), mais aussi en FQ (Arnaud et coll., 2011 : 246), et qui « differ in the time at which the target occurs relative to the vowel onset and offset » (Watson et Harrington, 1999 : 465). Enfin, certains changements linguistiques en cours (comme le passage du statut de diphtongue à celui de monophthongue) et des différences géolinguistiques pourraient être révélés par la dynamique spectrale mieux que par les caractéristiques spectrales statiques, à tout le moins en néerlandais (Adank, Van Hout et Smits 2004) et en anglais américain (Fox et Jacewicz, 2009; Jacewicz et coll., 2011; Jacewicz et Fox, 2012; Jacewicz et Fox, 2013).

Plus formellement, Nearey et Assmann (1986), s'inspirant de Pols (1977), ont proposé trois modèles pour rendre compte de la dynamique spectrale d'un point de vue acoustique :

- L'hypothèse de la double cible (*dual-target* ou *onset plus offset hypothesis*)³⁰;
- L'hypothèse de la cible et de la pente (*target plus slope* ou *onset plus slope hypothesis*);
- L'hypothèse de la cible et de l'orientation (*target plus direction* ou *onset plus direction hypothesis*).

Ces trois hypothèses « agree that the initial formant frequencies are perceptually relevant for vowel identification, but disagree on what additional cues are relevant » (Morrison et Nearey, 2007 : EL15). L'hypothèse de la double cible postule que les fréquences formantiques initiales et finales³¹ sont les éléments clefs de l'identification des voyelles. Nearey et Assmann (1986) l'ont formulée en se fondant sur Pols (1977 : 103), qui soulignait que Cohen (1971 : 279) était parvenu, en néerlandais, à synthétiser des « perfectly acceptable diphthongs by a succession of two steady-state vocoidlike sounds whose formant positions closely reflected those of the heard initial and final parts of the fragmented natural diphthongs ». L'hypothèse de la cible et de la pente met plutôt de l'avant les fréquences formantiques initiales et la vitesse d'évolution des formants (la pente, en particulier celle de F_2). Comme le remarquait également Pols (1977 : 103), ces deux paramètres seraient effectivement constants, peu importe le débit de parole, pour cinq diphtongues de l'anglais américain (Gay, 1968 : 1570). Quant à l'hypothèse de la cible et de l'orientation, elle propose que les fréquences formantiques initiales et l'orientation de l'évolution des formants (dans un espace multiformantique) sont déterminantes. Pols (1977 : 131) est parvenu à cette conclusion en soumettant des occurrences de diphtongues néerlandaises pour l'identification à des auditeurs puis en comparant les matrices de confusion obtenues avec les caractéristiques acoustiques des stimuli.

La dynamique spectrale des voyelles a également été estimée à l'aide de divers autres paramètres. Par exemple, Ferguson et Kewley-Port (2002) quantifient la dynamique spectrale par la distance euclidienne (la longueur du vecteur) entre le début et la fin d'une voyelle projetée dans un espace bi-formantique F_1 / F_2 . Ils observent ainsi qu'en anglais, la dynamique spectrale des voyelles est plus importante en parole claire (articulée distinctement) qu'en parole conversationnelle. Outre la longueur du vecteur, Fox et Jacewicz (2009) calculent également la longueur de la trajectoire des voyelles (telle que définie par six points de mesures plutôt que deux) dans un espace bi-formantique F_1 / F_2 , ainsi que la vitesse d'évolution de la trajectoire (ou

³⁰ Gottfried et coll. (1993 : 208) réservent l'appellation *dual-target hypothesis* au modèle impliquant que les deux cibles soient des états stables.

³¹ La position exacte des deux cibles varie d'une étude à l'autre. Par exemple, dans leur test de perception, Nearey et Assmann (1986 : 1298-1299) définissent le noyau et l'appendice comme étant les sections de 30 ms centrées respectivement à 24 % et à 64 % de la durée de la voyelle. Dans leur analyse acoustique cependant : « [i]nitial formant measures [...] were taken as early as possible in the vowel, provided the formant tracks were stable in the region and the overall amplitude of the section was within 15 dB of the peak amplitude section of the vowel. Final formant measures [...] were taken as late as possible, subject to the same criteria ».

spectral rate of change). Ils constatent que tous ces paramètres sont significativement influencés par le contexte consonantique, mais varient aussi en fonction de l'origine géographique des participants (des locuteurs de l'anglais américain). La longueur de la trajectoire permettrait cependant de déceler avec plus de précision les différences géolinguistiques que la longueur du vecteur. L'étude de Fox et Jacewicz (2009) révèle par ailleurs que la longueur de la trajectoire et la longueur du vecteur sont sensibles à la présence d'un accent d'insistance.

Enfin, d'autres chercheurs (Zahorian et Jagharghi, 1991, 1993; Watson et Harrington, 1999; Hillenbrand et coll., 2001) modélisent de façon plus complexe la trajectoire des formants, en calculant une courbe d'ajustement, par exemple au moyen d'une transformée en cosinus discrète (*Discrete Cosine Transform*, DCT) ou d'une équation polynomiale, comme l'illustre la Figure 9.

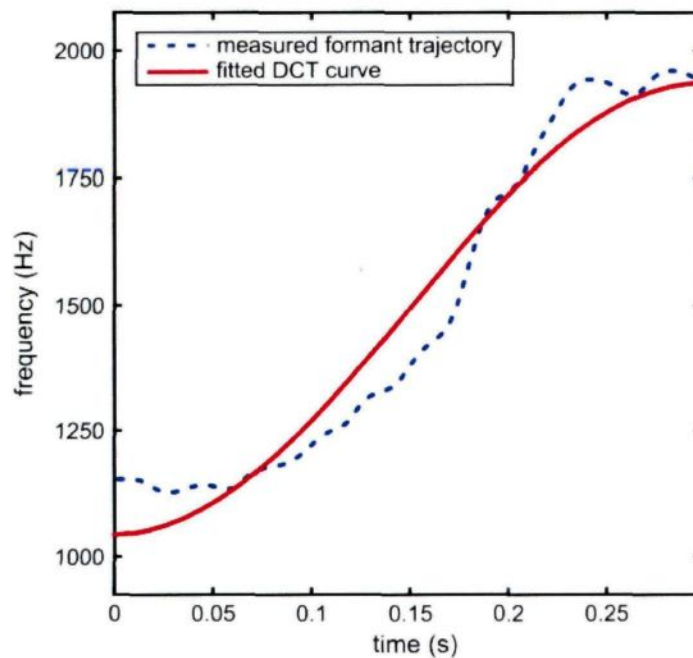


Figure 9 : Exemple d'une courbe ajustée à la trajectoire d'un formant au moyen d'une transformée en cosinus discrète. Tirée de Morrison (2013).

2.5 Cadre expérimental adopté

Comme nous l'avons mis en évidence au point 2.2, nous avons à notre disposition de nombreux indices acoustiques nous permettant de décrire les voyelles fermées du FQ. Pour des contraintes de temps évidentes, il nous était cependant nécessaire d'en faire une sélection. Nous avons choisi de nous concentrer sur des

indices acoustiques largement éprouvés, soit la fréquence fondamentale, la fréquence centrale des trois premiers formants³² et la durée.

Trois approches expérimentales, détaillées aux points 2.1, 2.3 et 2.4, s'offraient également à nous : 1) le paradigme traditionnel, qui consiste à relever les fréquences formantiques en Hertz en un point dans la portion centrale des voyelles ; 2) la normalisation vocalique ; et 3) la prise en compte de la dynamique spectrale. Ces trois approches comportent toutes des avantages. En effet, le paradigme traditionnel permet d'obtenir des données comparables à celles de nombreuses études antérieures. La normalisation vocalique, quant à elle, permet de traiter conjointement les productions vocaliques féminines et masculines. Enfin, la prise en compte de la dynamique spectrale paraît indispensable dans la description des voyelles diphtonguées et semble même utile dans la description des voyelles tendues et relâchées. Nous nous sommes proposé d'exploiter ces trois approches afin de caractériser acoustiquement les voyelles fermées tendues, relâchées et allongées du FQ ; et du même coup, de déterminer les paramètres acoustiques (procédant de l'une ou l'autre de ces approches) permettant de distinguer le plus efficacement ces variantes. Cela dit, ces approches se déclinent parfois sous plusieurs formes. Il nous a donc préalablement fallu opérer une sélection au sein des procédures de normalisation et des paramètres modélisant la dynamique spectrale.

En ce qui concerne la normalisation, de nombreux chercheurs (Hindle, 1978 ; Ferrari Disner, 1980 ; Miller et coll., 1980 ; Syrdal et Gopal, 1986 ; Miller, 1989 ; Hillenbrand et Gayvert, 1993 ; Ménard, 2002 et Ménard et coll., 2002 ; Adank, 2003 et Adank, Smits et Van Hout, 2004 ; Nguyen et Espesser, 2004 ; Calamai, 2005 ; Volín et Studenovský, 2007 ; Fabricius et coll., 2009 ; Clopper, 2009 ; Flynn et Foulkes, 2011) ont cherché à déterminer, dans plusieurs langues (anglais américain et britannique, allemand, norvégien, suédois, danois, français de France, néerlandais, italien, tchèque), les procédures les plus efficaces au regard de différents critères d'évaluation (réduction de la dispersion intraclasse, augmentation de la dispersion interclasses, réduction de l'écart entre les fréquences formantiques masculines et féminines, préservation de l'information phonémique, sociolinguistique ou dialectale, etc.). Cependant, aucune évaluation n'a, à notre connaissance, été réalisée en FQ. Or, l'étude de Ferrari Disner (1980) et celle de Nguyen et Espesser (2004) témoignent de l'efficacité variable des procédures de normalisation selon la langue, ou même la variété de langue, à laquelle elles sont appliquées. Nous avons donc procédé, dans le cadre d'une étude préliminaire (Sigouin, à paraître, voir Annexe 2), à une évaluation de différentes procédures de normalisation en FQ. Conformément au principal objectif de notre projet de mémoire, nos tests ont notamment porté sur la capacité des procédures de normalisation à améliorer la classification des occurrences au sein de leur classe vocalique (dans le cadre de

³² Au départ, nous souhaitions également intégrer la fréquence centrale du quatrième formant, mais il était souvent difficile de la détecter avec certitude. Aussi avons-nous décidé de ne pas l'inclure dans nos analyses finales. Les valeurs moyennes de F₄ sont toutefois données dans les tableaux de l'Annexe 7 pour information.

l'étude préliminaire, il s'agissait plus précisément de classes phonémiques, puisque les analyses ont été menées sur des voyelles isolées). Nos résultats indiquent que les données converties ou normalisées au moyen des opérations intrinsèques évaluées ne présentent virtuellement aucun avantage par rapport aux données brutes en Hertz au regard de ce critère d'évaluation (ce qui est la conclusion à laquelle sont également arrivés Hillenbrand et Gayvert, 1993 ainsi qu'Adank, 2003). Ces procédures n'auraient donc vraisemblablement rien apporté à la caractérisation acoustique des voyelles fermées du FQ. Au contraire, certaines opérations extrinsèques sont parvenues à améliorer la classification des occurrences au sein de leur classe vocalique. L'utilisation des scores-Z (tels que proposés par Lobanov, 1971) semble être la plus efficace à cet effet. En nous fondant sur ces résultats, et comme notre recherche ne se concentre pas sur les propriétés perceptives des voyelles, nous avons choisi d'employer la procédure extrinsèque des scores-Z pour normaliser les mesures spectrales traitées dans le cadre de ce mémoire.

En ce qui a trait à la dynamique spectrale, la *undershoot compensation theory* et la *dynamic cospecification theory* ont été élaborées sur la base d'études perceptives sans que des indices acoustiques précis n'y aient été associés. Nous avons donc privilégié la théorie du VISC. Parmi les modèles proposés, nous avons écarté l'ajustement de courbe puisque la supériorité de ce modèle complexe sur les trois premières hypothèses du VISC n'est pas établie (Hillenbrand et coll., 2001 : 759; Morrison, 2013 : 44). Quant aux paramètres préconisés par Fox et Jacewicz (2009), ils ont été utilisés pour quantifier la dynamique spectrale sans pour autant être proposés comme des modèles permettant de rendre compte globalement de l'identité des voyelles. C'est pourquoi nous nous sommes concentrée sur les trois premières hypothèses du VISC dans notre étude. Nous avons décidé de les tester toutes les trois puisque les résultats d'études comparatives ne favorisent pas incontestablement une hypothèse plus qu'une autre (Nearey et Assmann, 1986; Gottfried et coll., 1993; Morrison et Nearey, 2007; Arnaud et coll., 2011), quoique d'après Morrison (2013 : 45) l'hypothèse de la double cible serait, du moins en anglais, celle qui « provides a better account of the perceptually relevant aspects of VISC with respect to vowel identity ». Deux points de mesure, l'un situé dans la portion initiale de la voyelle, l'autre, dans sa portion finale, sont nécessaires pour formaliser ces trois hypothèses. Comme nous le mentionnions à la note de bas de page 31, la position de ces deux points varie d'une étude à l'autre. Pour notre part, nous avons choisi de les situer à 25 % et à 75 % de la durée des voyelles pour les mêmes raisons que Nearey et Assmann (1986 : 1298) :

The time points chosen were 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.

Chapitre 3 : Méthodologie

Afin de caractériser acoustiquement les voyelles fermées tendues, relâchées et allongées du FQ, nous avons analysé un corpus récolté auprès de locuteurs du FQ. Nous avons préalablement obtenu l'approbation du Comité d'éthique de la recherche de l'Université du Québec à Chicoutimi.

3.1 Échantillon et corpus

Nous avons analysé la parole de 30 locuteurs, 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, de Québec et de Rouyn-Noranda (voir Figure 10). Tous avaient le FQ comme langue maternelle.



Figure 10 : Carte du Québec (tronquée) indiquant nos points d'enquête. Adaptée de la carte *Le Québec, carte muette* du Ministère des Ressources naturelles et de la Faune. © Gouvernement du Québec, 2006.

Plus précisément, les participants devaient être nés dans la région où se trouve l'une de ces villes (le Saguenay–Lac-Saint-Jean, la région de la Capitale-Nationale ou l'Abitibi-Témiscamingue), avoir passé la majeure partie de leur vie dans la ville en question et y habiter depuis au moins deux ans au moment de l'enquête (laquelle s'est déroulée d'octobre 2010 à août 2012). La couverture géographique de notre étude aurait sans nul doute pu être différente, mais nous cherchions avant tout à nous assurer que nos résultats ne correspondent pas aux productions des locuteurs d'une seule région du Québec. De plus, la question de la variation géophonétique en FQ étant très peu documentée (Dolbec et Ouellon, 1999), il aurait été malaisé de justifier le choix des localités sur cette base. Nous avons plutôt privilégié les populations qui nous étaient les plus accessibles. Mentionnons que la région du Saguenay–Lac-Saint-Jean et celle de la Capitale-Nationale sont considérées comme faisant partie de l'aire des parlers de l'Est du Québec et que le Saguenay–Lac-Saint-Jean est généralement reconnu comme possédant un parler régional à l'intérieur de cette aire (Dolbec et Ouellon, 1999). L'Abitibi-Témiscamingue n'est pour sa part rattachée à aucune aire linguistique, étant une

Chapitre 3 : Méthodologie

Afin de caractériser acoustiquement les voyelles fermées tendues, relâchées et allongées du FQ, nous avons analysé un corpus récolté auprès de locuteurs du FQ. Nous avons préalablement obtenu l'approbation du Comité d'éthique de la recherche de l'Université du Québec à Chicoutimi.

3.1 Échantillon et corpus

Nous avons analysé la parole de 30 locuteurs, 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, de Québec et de Rouyn-Noranda (voir Figure 10). Tous avaient le FQ comme langue maternelle.



Figure 10 : Carte du Québec (tronquée) indiquant nos points d'enquête. Adaptée de la carte *Le Québec, carte muette* du Ministère des Ressources naturelles et de la Faune. © Gouvernement du Québec, 2006.

Plus précisément, les participants devaient être nés dans la région où se trouve l'une de ces villes (le Saguenay–Lac-Saint-Jean, la région de la Capitale-Nationale ou l'Abitibi-Témiscamingue), avoir passé la majeure partie de leur vie dans la ville en question et y habiter depuis au moins deux ans au moment de l'enquête (laquelle s'est déroulée d'octobre 2010 à août 2012). La couverture géographique de notre étude aurait sans nul doute pu être différente, mais nous cherchions avant tout à nous assurer que nos résultats ne correspondent pas aux productions des locuteurs d'une seule région du Québec. De plus, la question de la variation géophonétique en FQ étant très peu documentée (Dolbec et Ouellon, 1999), il aurait été malaisé de justifier le choix des localités sur cette base. Nous avons plutôt privilégié les populations qui nous étaient les plus accessibles. Mentionnons que la région du Saguenay–Lac-Saint-Jean et celle de la Capitale-Nationale sont considérées comme faisant partie de l'aire des parlers de l'Est du Québec et que le Saguenay–Lac-Saint-Jean est généralement reconnu comme possédant un parler régional à l'intérieur de cette aire (Dolbec et Ouellon, 1999). L'Abitibi-Témiscamingue n'est pour sa part rattachée à aucune aire linguistique, étant une

zone de peuplement mixte récent à laquelle très peu d'études se sont intéressées (quelques points d'enquête de l'Atlas linguistique de l'est du Canada de Dulong et Bergeron, 1980; Boissonneault, 1999; en partie Fast, 2002).

Les enregistrements de Saguenay et de Québec sont issus du corpus de Vincent Arnaud, professeur au Département des arts et lettres de l'Université du Québec à Chicoutimi.³³ Ils ont été réalisés en chambre anéchoïque, directement au format numérique (44 100 Hz, 16 bits), à l'aide d'un TASCAM HD-P2 et d'un microphone serre-tête SHURE SM10A amplifié au moyen d'une console de mixage BEHRINGER XENYX 802. Les locuteurs composant cet échantillon ont perçu une modique compensation financière pour leur participation, qui impliquait une séance unique d'enregistrement d'une durée approximative de 1 h 30. Une partie de la tâche demandée consistait à lire à haute voix, aussi naturellement que possible, une série de phrases porteuses pentasyllabiques et sémantiquement significatives contenant les mots cibles 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. L'échange ressemblait donc à ce qui suit :

Locuteur : *Les couturières cousent.*

Enquêtrice : *Elles ?*

Locuteur : *Cousent.*

Ce protocole nous permettait d'éviter nombre d'hésitations de la part des locuteurs quant au sens des mots cibles (grâce au contexte sémantique), en plus de soustraire ces mots à l'influence de la phrase porteuse (grâce à la répétition) et de prévenir un « effet de liste » (une prosodie répétitive associée à la lecture d'une liste de mots isolés).

À Rouyn-Noranda, où nous n'avions pas accès à une chambre anéchoïque, les enregistrements ont été effectués au domicile des locuteurs. Ces 10 témoins n'ont pas reçu de compensation financière pour leur participation. Du reste, la même méthodologie que dans le cadre de l'échantillon de Saguenay et de Québec a été appliquée.

Le choix de travailler sur de la parole dite *de laboratoire* (Xu, 2010) nous permettait de contrôler le nombre d'occurrences de chaque phonème et son contexte, et ainsi collecter des données hautement comparables d'un phonème à l'autre, d'un contexte à l'autre, d'un locuteur à l'autre. Et comme les locuteurs naïfs du FQ ne semblent pas porter de jugement négatif à l'endroit du relâchement (Dumas, 1987 : 95) — le remarquent-ils seulement? — et que la diphtongaison des voyelles fermées passe le plus souvent inaperçue (Santerre et

³³ Nous avons nous-même participé à l'élaboration de ce corpus à titre d'assistante de recherche. Le projet dont il découle, intitulé *Caractéristiques acoustiques des voyelles orales québécoises*, a été subventionné par le *Fonds de recherche québécois – Société et culture*.

coll., 1985 : 51), nous estimions qu'il était peu probable que le degré de formalité induit par ces conditions expérimentales conduise les locuteurs à adopter une prononciation inaccoutumée en ce qui concerne les phénomènes qui nous intéressent.

Des 523 mots cibles répétés isolément qui ont été produits une seule fois par chaque locuteur, nous en avons retenu 45 (voir Annexe 3). Les autres mots, qui contenaient l'une ou l'autre des voyelles orales du FQ et qui sont destinés à d'autres analyses, agissaient comme distracteurs. Un total de 1350 occurrences lexicales constitue donc notre corpus. Il s'agit, sauf exception, de noms et de verbes d'une syllabe dont le noyau correspond à la voyelle /i/, /y/ ou /u/. Ils se répartissent également entre les classes vocaliques (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK) et (uR),³⁴ telles que définies par le contexte.³⁵ Afin de faciliter la segmentation des voyelles (éviter l'affrication, la nasalisation, etc.), les mots sélectionnés commencent, dans la mesure du possible, par les consonnes /p/, /b/, /k/ ou /g/. Pour cette même raison, les voyelles en contexte (VK) sont, autant que le lexique français le permet, suivies de /p/, /b/, /t/, /d/, /k/ ou /g/. Les voyelles en contexte (VR) sont quant à elle suivies de /ʁ/, /v/, /z/ ou /ʒ/; chacune de ces consonnes apparaissant au moins une fois à la suite de chaque voyelle fermée /i/, /y/ et /u/.

Notons enfin que 7 occurrences ont été rejetées. Dans un cas, il s'agissait du mot « cousent », prononcé [kuz] par une locutrice de Saguenay. Comme nous nous intéressons à la variante allongée dans ce contexte et que la locutrice avait longuement et explicitement hésité sur la façon de prononcer ce mot, nous avons choisi de l'exclure. Dans les 6 autres cas, les voyelles se sont révélées acoustiquement inanalysables. Au final, c'est donc 1343 occurrences qui ont été analysées dans le cadre de ce mémoire.

3.2 Analyse acoustique

L'analyse acoustique a été réalisée à l'aide du logiciel PRAAT, version 5.3.x (Boersma et Weenink, 2013). La première étape de cette analyse consistait à segmenter les voyelles. De manière générale, les frontières ont été posées au passage par zéro³⁶ (là où la ligne de l'oscillogramme croise l'origine) le plus près du début et de la fin de la périodicité de l'onde sonore. Dans les cas où les extrémités de la voyelle n'étaient pas marquées par l'établissement ou la disparition de la périodicité, nous nous sommes fondée sur d'autres indices acoustiques, tels que l'établissement ou la disparition des formants, une barre d'explosion, une montée ou une chute rapide de l'intensité, la disparition ou l'apparition de bruits de friction, ou encore un changement notable dans la configuration formantique. Lorsque la transition entre la voyelle et la consonne était très graduelle

³⁴ Voir la section Conventions et abréviations.

³⁵ En nous fondant sur notre propre audition, nous nous sommes assurée que les variantes tendues, relâchées et allongées se retrouvaient effectivement en (V#), (VK) et (VR) respectivement.

³⁶ Nous avons vérifié que toutes les frontières étaient bien posées à un passage par zéro à l'aide d'un script PRAAT développé à cet effet (voir Annexe 4).

(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. La délimitation de chaque occurrence vocalique nous a ainsi permis de calculer la durée de chaque occurrence.

Par la suite, nous avons relevé la fréquence fondamentale et la fréquence centrale (en Hertz) des trois premiers formants de chaque occurrence à 25 %, 50 % et 75 % de leur durée. Les paramètres par défaut de PRAAT ont été utilisés pour la détection de la F_0 :

- Méthode d'analyse : autocorrélation
- Nombre maximal de candidats : 15
- Seuil du silence : 0,03
- Seuil du voisement : 0,45
- Coût de l'octave : 0,01
- Coût du saut d'octave : 0,35
- Coût voisé / non voisé : 0,14

Seules les valeurs du plancher et du plafond de détection de la F_0 ont parfois été ajustées. Notons que, lorsqu'elle était détectable, la F_0 de la voix craquée (ou *creaky voice*) n'a pas été exclue.

Les formants ont été détectés à l'aide de l'algorithme de codage prédictif linéaire (*linear predictive coding*, LPC) intégré à PRAAT. Les paramètres de détection suivants ont été maintenus constants :

- Méthode : Burg
- Pré-emphase à partir de : 50,0 Hz
- Largeur de la fenêtre : 0,025 s
- Intervalle dynamique : 30,0 dB

Le plafond de détection et le nombre de formants à détecter étaient ajustés pour chaque phonème, pour chaque locuteur, et modifiés au besoin si ces paramètres ne convenaient pas pour une occurrence donnée. Nous nous sommes donc assurée que l'algorithme de suivi des fréquences centrales formantiques détectait des pics spectraux qui correspondaient bien aux groupes d'harmoniques de forte amplitude visibles sur un spectrogramme à bandes larges. Dans le cas des occurrences de la voyelle /u/, il était généralement nécessaire d'adopter un réglage pour la détection de F_1 et de F_2 , et un second pour la détection de F_3 .

Une fois les paramètres de détection de la F_0 et des formants réglés, les mesures étaient relevées à chaque occurrence au moyen d'un script PRAAT développé par Vincent Arnaud (professeur à l'Université du Québec à Chicoutimi) que nous avons adapté (voir Annexe 5). En de très rares occasions, le relevé de certaines

mesures a été effectuée manuellement (c'est-à-dire en nous fondant sur une inspection visuelle du spectrogramme) lorsque la détection de la F_0 ou des formants était impossible au moyen des algorithmes (par exemple lorsque l'intensité d'un formant était très faible ou lorsque deux formants étaient très rapprochés). Évidemment, lorsque la détection était impossible, automatiquement et visuellement, la donnée était simplement rapportée manquante.

3.3 Normalisation

Les valeurs fréquentielles que nous avons relevées pour la F_0 , F_1 , F_2 et F_3 ont été normalisées en calculant leurs scores-Z, comme le propose Lobanov (1971). Pour ce faire, nous avons d'abord calculé, pour chaque individu, la F_0 , le F_1 , le F_2 et le F_3 moyens de chaque classe vocalique, tous points de mesure confondus. Appelons ces moyennes les *moyennes catégorielles individuelles*. À partir de ces moyennes catégorielles individuelles, nous avons ensuite calculé, pour chaque individu, la F_0 , le F_1 , le F_2 et le F_3 moyens (et les écarts-types associés), toutes classes vocaliques confondues. Appelons ces moyennes et ces écarts-types les *moyennes globales individuelles* et les *écarts-types globaux individuels*. Nous avons enfin calculé le score-Z de chaque valeur à l'aide de la formule suivante :

$$(9) \quad (N_x)_i = \frac{(F_x)_i - (\bar{F}_x)_i}{(s_x)_i}$$

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 Hertz du x^e formant (ou de la F_0) de cette même occurrence ;

$(\bar{F}_x)_i$ correspond à la moyenne globale individuelle du x^e formant (ou de la F_0) de ce même locuteur ;

$(s_x)_i$ correspond à l'écart-type global individuel du x^e formant (ou de la F_0) de ce même locuteur.

Tous ces calculs ont été effectués au moyen d'un script PERL développé à cet effet (voir Annexe 6).

3.4 Dynamique spectrale

En ce qui concerne le VISC, les mesures de la F_0 , de F_1 , F_2 et F_3 prises à 25 % et à 75 % de la durée des voyelles constituent en elles-mêmes une paramétrisation du modèle de la double cible. Pour formuler le

modèle de la cible et de la pente, en plus de considérer les mesures relevées à 25 % de la durée des voyelles, nous avons calculé, pour chaque occurrence, la pente de chacune des composantes spectrales comme suit :

$$(10) \quad p_x = \frac{\Delta F_x}{\Delta t}$$

où : p_x correspond à la pente du x^e formant (ou de la F_0) d'une occurrence donnée ;

ΔF_x correspond à la différence entre les valeurs du x^e formant (ou de la F_0) de cette occurrence à 75 % et 25 % ;

Δt correspond au temps écoulé entre 25 % et 75 % de la durée de cette occurrence (ce qui équivaut à 50 % de la durée totale de l'occurrence).

Enfin, pour rendre compte du modèle de la cible et de l'orientation, outre les mesures prises à 25 % de la durée des voyelles, nous avons calculé l'orientation des mesures spectrales de chaque occurrence dans :

- un espace bidimensionnel dont les dimensions sont F_1 et F_2 ;
- un espace tridimensionnel dont les dimensions sont la F_0 , F_1 et F_2 ;
- un espace tridimensionnel dont les dimensions sont F_1 , F_2 et F_3 ;
- un espace quadridimensionnel dont les dimensions sont la F_0 , F_1 , F_2 et F_3 .

Précisons que l'orientation d'une composante spectrale dans un espace donné ne peut être calculée que si celle-ci est représentée dans cet espace (par exemple, seules les orientations de F_1 et de F_2 peuvent être calculées dans l'espace bidimensionnel F_1 / F_2). Cette orientation se calcule comme suit :

$$(11) \quad C_x^\alpha = \frac{\Delta F_x}{VL} \text{ et } VL = \sqrt{\sum \Delta F^2}$$

où : C_x^α correspond à l'orientation du x^e formant (ou de la F_0) d'une occurrence donnée dans un espace α dont le x^e formant (ou la F_0) constitue l'une des dimensions ;

ΔF_x correspond à la différence entre les valeurs du x^e formant (ou de la F_0) de cette occurrence à 75 % et 25 % ;

VL correspond à la longueur du vecteur de l'occurrence dans l'espace α ;

ΔF correspond à la différence entre les valeurs de chaque formant (ou de la F_0) de cette occurrence à 75 % et 25 %, pourvu que celui-ci constitue l'une des dimensions de l'espace α .

3.5 Analyses statistiques

Les statistiques descriptives (moyennes et écarts-types) présentées au chapitre suivant ont été réalisées à l'aide du logiciel SYSTAT 13. Les graphiques ont quant à eux été effectués à l'aide des logiciels SYSTAT 13, PASW STATISTICS 18 et MICROSOFT OFFICE EXCEL 2003.

Afin de tester l'effet de différentes variables indépendantes (sexe, origine géographique, classe phonémique, etc.) sur les caractéristiques acoustiques des voyelles fermées composant notre corpus, nous avons réalisé des ANOVAs au moyen du logiciel SAS 9.1. Nous avons fait affaire avec le Service de consultation statistique de l'Université Laval pour la conception de ces ANOVAs. Les moyennes individuelles des indices acoustiques par catégorie (telle que définie par le recoupement des variables indépendantes subordonnées au locuteur qui étaient intégrées au modèle de chaque ANOVA) étaient préalablement calculées afin d'éviter toute pseudorépétition. Par exemple, afin de tester l'effet du sexe et de l'origine géographique des locuteurs ainsi que de la classe phonémique /i/, /y/ ou /u/ et du contexte (V#), (VK) ou (VR) sur la durée des voyelles, nous avons d'abord calculé la durée moyenne des 5 occurrences produites par un même locuteur dans chacune des classes vocaliques (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK) et (uR). Notons que dans tous les cas, la matrice de variance-covariance utilisée était celle qui minimisait le critère d'Akaike. La méthode de Kenward-Roger a été employée pour calculer les degrés de liberté. Les postulats de normalité et d'homogénéité des résidus ont été vérifiés. Les comparaisons multiples *post hoc* ont été réalisées à l'aide de la méthode *Protected Fisher's Least Significant Difference*. Nous avons opté pour un seuil de signification statistique de 1 % ($\alpha=0,01$) puisqu'avec un seuil de 5 % (plus courant en sciences humaines), nous obtenions parfois des résultats statistiquement significatifs sur des écarts infimes d'un point de vue phonétique et difficilement interprétables.

La première ANOVA effectuée visait à tester l'effet du sexe, de l'origine géographique des locuteurs, de la classe phonémique /i/, /y/ ou /u/ et du contexte (V#), (VK) ou (VR) sur la durée des voyelles. Il s'agissait d'un modèle mixte d'ANOVA à mesures répétées dont les facteurs répétés étaient la classe phonémique et le contexte.

Une seconde ANOVA nous a permis de tester l'effet du sexe, de l'origine géographique des locuteurs et du contexte consonantique (tel que défini au point 4.1) sur la durée des voyelles. Il s'agissait d'un modèle mixte d'ANOVA à mesures répétées dont le facteur répété était le contexte consonantique.

Une troisième ANOVA nous a permis de tester l'effet du sexe, de l'origine géographique des locuteurs et de la consonne finale sur la durée des voyelles allongées. Il s'agissait d'un modèle mixte d'ANOVA à mesures répétées dont le facteur répété était la consonne finale.

Nous avons ensuite testé l'effet du sexe, de l'origine géographique et du point de mesure (situé à 25 %, 50 % ou 75 % de la durée de la voyelle) sur la F_0 , F_1 , F_2 et F_3 (en Hertz) de chaque classe vocalique prise individuellement en procédant à une série d'ANOVAs (une pour chaque indice acoustique). Il s'agissait de modèles mixtes d'ANOVA à mesures répétées dont le facteur répété était le point de mesure. Une série d'ANOVAs identiques ont été menées pour les données normalisées à l'aide des scores-Z.

Enfin, dans le but d'identifier les combinaisons d'indices acoustiques les plus pertinentes dans la distinction des classes vocaliques, nous avons procédé à une série d'analyses discriminantes quadratiques en utilisant le logiciel SYSTAT 13. Le discriminant était calculé sur les données provenant de la moitié des locuteurs de notre corpus (sexe et origine géographique le plus équilibrés possible) puis appliqué sur les données de l'autre moitié. Les probabilités de classification n'étaient pas établies en fonction des effectifs des classes, qui, de toute façon, étaient à peu de choses près égaux. La variable dépendante était la classe vocalique. Dans un premier temps, les combinaisons suivantes de prédicteurs, tenant compte des différentes hypothèses du VISC (telles que nous les avons paramétrisées), ont été soumises à des analyses discriminantes, avec et sans l'inclusion de la durée des voyelles :

- F_1 et F_2 à 25 % de la durée des voyelles, les pentes de F_1 et F_2 entre 25 et 75 % ;
- F_0 , F_1 et F_2 à 25 % de la durée des voyelles, les pentes de F_0 , F_1 et F_2 entre 25 et 75 % ;
- F_1 , F_2 et F_3 à 25 % de la durée des voyelles, les pentes de F_1 , F_2 et F_3 entre 25 et 75 % ;
- F_0 , F_1 , F_2 , F_3 à 25 % de la durée des voyelles, les pentes de F_0 , F_1 , F_2 et F_3 entre 25 et 75 % ;
- F_1 et F_2 à 25 % de la durée des voyelles, l'orientation de F_1 et F_2 entre 25 et 75 % dans un espace F_1 / F_2 ;
- F_0 , F_1 et F_2 à 25 % de la durée des voyelles, l'orientation de F_0 , F_1 et F_2 entre 25 et 75 % dans un espace $F_0 / F_1 / F_2$;
- F_1 , F_2 et F_3 à 25 % de la durée des voyelles, l'orientation de F_1 , F_2 et F_3 entre 25 et 75 % dans un espace $F_1 / F_2 / F_3$;
- F_0 , F_1 , F_2 , F_3 à 25 % de la durée des voyelles, l'orientation de F_0 , F_1 , F_2 et F_3 entre 25 et 75 % dans un espace $F_0 / F_1 / F_2 / F_3$.

Dans un second temps, les prédicteurs étaient tour à tour :

- F_1 et F_2 ;
- F_0 , F_1 et F_2 ;
- F_1 , F_2 et F_3 ;
- F_0 , F_1 , F_2 , F_3 .

Ces combinaisons de prédicteurs, présentés en Hertz et en scores-Z, avec et sans la durée des voyelles, provenaient des données relevées à :

- 50 % de la durée des voyelles;
- 25 % et 75 % de la durée des voyelles;
- 25 %, 50 % et 75 % de la durée des voyelles.

Notons que les analyses discriminantes sont, en phonétique, employées à titre exploratoire; c'est-à-dire qu'on ne vise pas à établir la signification statistique de leurs résultats car elles ne prennent pas en compte l'effet aléatoire de l'individu (c'est-à-dire la corrélation entre les occurrences produites par un même individu). Les résultats des analyses discriminantes ne sont donc pas à interpréter de façon absolue. Par contre, en comparant les matrices de confusion et les taux de classification correcte d'analyses discriminantes n'ayant pas les mêmes prédicteurs, on est en mesure d'évaluer l'efficacité relative de ces prédicteurs et l'impact que ceux-ci ont sur la classification.

Chapitre 4 : Résultats et analyse

Les résultats de nos analyses acoustiques et statistiques sont rapportés et commentés dans les pages qui suivent. L'examen de la durée des voyelles fermées de notre corpus fera l'objet d'une première section. Nous nous intéresserons ensuite aux caractéristiques spectrales de ces mêmes segments. Nous nous pencherons enfin plus spécifiquement sur le caractère discriminant de ces divers indices acoustiques (temporels ou spectraux).

4.1 Durée

Les durées moyennes de chacune des classes vocaliques (i#, (iK), (iR), (y#), (yK), (yR), (u#), (uK) et (uR) ont d'abord été calculées pour chacun des 30 locuteurs composant notre échantillon. Le Tableau 1 présente les moyennes de ces moyennes individuelles pour chaque classe vocalique en fonction de la ville d'origine et du sexe des participants. L'écart-type est placé entre parenthèses et l'effectif est de 5 individus par cellule.

Tableau 1 : Durée moyenne et écart-type (en ms) des différentes classes vocaliques selon le sexe et la ville des locuteurs

	Hommes			Femmes		
(V#)	/i/	/y/	/u/	/i/	/y/	/u/
RN	114 (23)	111 (19)	110 (17)	101 (21)	101 (22)	103 (13)
SG	100 (13)	99 (10)	99 (16)	98 (17)	102 (25)	102 (28)
QC	98 (16)	102 (15)	107 (16)	101 (18)	98 (21)	100 (15)
(VK)	/i/	/y/	/u/	/i/	/y/	/u/
RN	88 (10)	111 (18)	105 (11)	92 (9)	115 (15)	102 (15)
SG	102 (16)	128 (19)	120 (14)	119 (12)	160 (18)	147 (11)
QC	93 (29)	114 (22)	105 (25)	91 (14)	118 (14)	105 (9)
(VR)	/i/	/y/	/u/	/i/	/y/	/u/
RN	239 (36)	244 (53)	259 (35)	264 (64)	258 (54)	254 (63)
SG	262 (36)	252 (49)	240 (35)	290 (42)	287 (34)	286 (14)
QC	248 (71)	254 (70)	249 (62)	247 (44)	264 (41)	258 (43)

Dans un premier temps, nous avons testé l'effet du sexe, de l'origine géographique des locuteurs (Rouyn-Noranda, Saguenay ou Québec), de la classe phonémique (/i/, /y/ ou /u/) et du contexte ((V#), (VK) ou (VR)) sur la durée des occurrences vocaliques de notre corpus. Les résultats de cette première ANOVA indiquent tout d'abord que le sexe des locuteurs n'influence pas significativement la durée de ces voyelles. Il appert ensuite qu'en syllabe ouverte (V#), la durée des occurrences de /i/, de /y/ et de /u/ ne diffère pas non plus significativement. Il en est de même en syllabe fermée par une consonne allongeante (VR). En revanche, en

syllabe fermée par une consonne non allongante (VK), les différences de durée entre les occurrences appartenant à ces trois classes phonémiques se révèlent statistiquement significatives ($p < 0,0001$). Dans ce contexte, les occurrences de /i/ sont les plus courtes. Elles durent en moyenne 0,098 s. Les occurrences de /u/ durent quant à elles 0,114 s en moyenne. Les occurrences de /y/ apparaissent pour leur part comme étant les plus longues, avec une durée moyenne de 0,124 s (voir Figure 11).

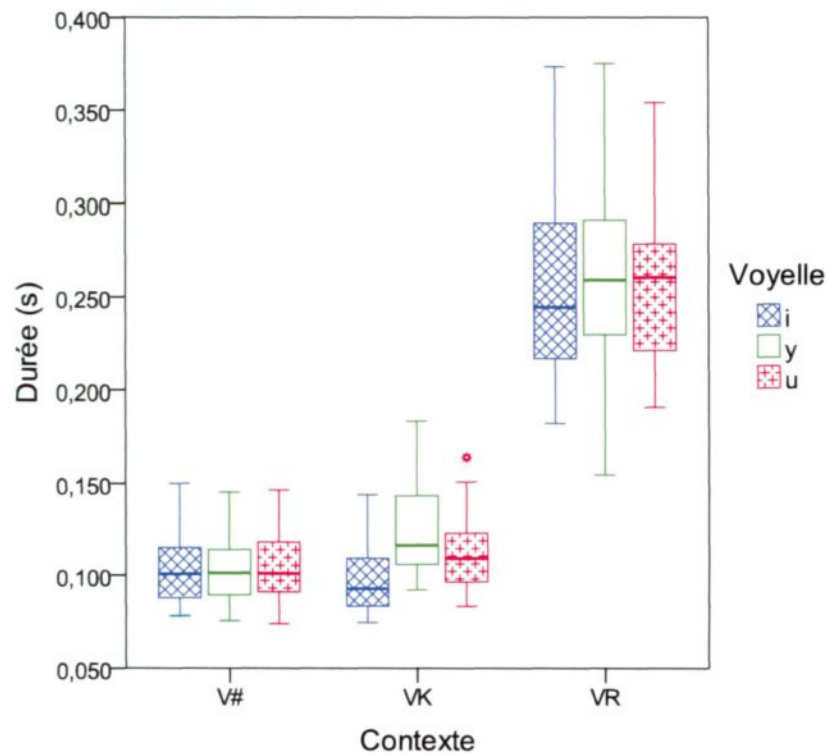


Figure 11 : Diagramme boîtes à moustaches présentant la durée des voyelles en fonction du contexte et de la classe phonémique. Les astérisques indiquent des données extrêmes.

Si ces différences sont potentiellement attribuables à la nature même des voyelles, il faut cependant garder à l'esprit que dans les contextes regroupés sous l'étiquette (VK), les consonnes finales sont variées et ne sont pas nécessairement les mêmes d'une classe phonémique à l'autre. Mentionnons d'autre part que nos résultats ne corroborent pas les conclusions de Ouellon (1978 : 719) et de Ouellet, (1992 : 162) selon lesquelles les voyelles fermées postérieures seraient très légèrement plus longues que les antérieures.

Cela étant dit, nous constatons sans surprise que, toutes classes phonémiques confondues, les segments en (VR) sont significativement ($p < 0,0001$) plus longs (0,259 s en moyenne) que ceux se trouvant en (V#) (0,103 s) et en (VK) (0,112 s). Dans ces deux derniers contextes, la durée des occurrences n'est généralement pas différente. Seules les occurrences de la voyelle /y/ en (V#) (durant 0,102 s en moyenne) s'avèrent significativement plus courtes ($p < 0,0001$) que les occurrences de cette même voyelle produites en

(VK) (0,124 s). Comme cette différence entre (V#) et (VK) ne se manifeste que dans le cas des /y/, il nous paraît hasardeux de l'attribuer aux timbres tendus et relâchés attendus dans ces contextes, ou encore à la structure syllabique. Nous soupçonnons plutôt que l'environnement consonantique (partiellement différent d'une classe vocalique à l'autre) puisse être en cause. Nous y reviendrons un peu plus loin.

Une différence régionale semble également émerger en (VK), comme l'illustre la Figure 12. Toutes classes phonémiques confondues, les occurrences (VK) produites par les locuteurs de Saguenay sont plus longues (durant 0,129 s en moyenne) que celles produites par les locuteurs de Québec (0,104 s) et de Rouyn-Noranda (0,102 s).

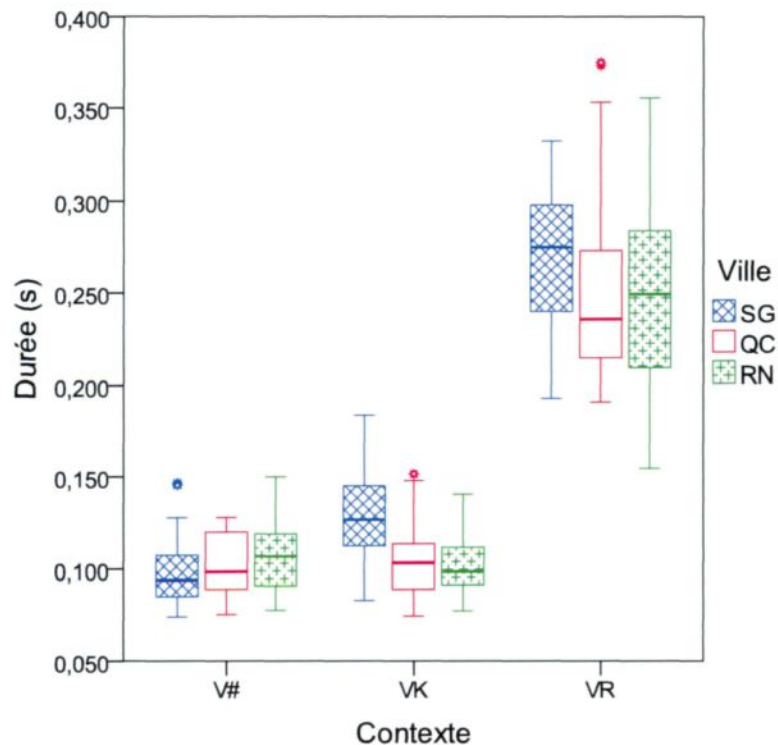


Figure 12: Diagramme boîtes à moustaches présentant la durée des voyelles en fonction du contexte et de la ville d'origine des locuteurs. Les astérisques indiquent des données extrêmes.

Cette différence n'est cependant pas statistiquement significative ($p=0,0186$) au seuil que nous nous sommes fixé ($\alpha=0,01$). Toutefois, elle est suffisante pour qu'à Saguenay, la différence de durée entre les occurrences en (V#) et celles en (VK), elle, soit significative ($p=0,0002$); ce qui n'est le cas ni à Québec, ni à Rouyn-Noranda.

Comme nous l'avons mentionné, l'interprétation de certains des résultats exposés plus haut demeure délicate en raison de l'éventuel impact de l'environnement consonantique (qui n'est pas parfaitement contrôlé dans

notre corpus) sur la durée des voyelles. Notre codage du contexte (reflétant la distribution des variantes tendues, relâchées et allongées des voyelles fermées du FQ rapportée dans la littérature) ne comporte en effet que deux grandes divisions : absence ou présence d'une consonne en coda et caractère allongeant ou non de cette consonne. Or, plusieurs autres caractéristiques des consonnes adjacentes aux voyelles sont susceptibles d'influencer leur durée. Le voisement des consonnes, notamment, aurait un certain effet allongeant sur les voyelles postposées et, à plus forte raison, sur les voyelles antéposées (voir notamment Ouellet, 1992 : 57, 65 pour un aperçu de la question). Nous nous sommes donc intéressée, dans un deuxième temps, à l'effet du voisement et du caractère allongeant des consonnes qui précèdent et qui suivent les voyelles sur leur durée. Pour ce faire, nous avons choisi de regrouper les occurrences de notre corpus dans 8 catégories pour lesquelles nous avons encore une fois calculé les durées moyennes individuelles avant de les soumettre à des tests statistiques. Ces catégories sont les suivantes :

- Les voyelles précédées d'une consonne non voisée, en syllabe ouverte (N.Voisée_#);
- Les voyelles précédées d'une consonne voisée, en syllabe ouverte (Voisée_#);
- Les voyelles précédées d'une consonne voisée et suivies d'une consonne voisée non allongeante (Voisée_Voisée);
- Les voyelles précédées d'une consonne voisée et suivies d'une consonne non voisée (Voisée_N.Voisée);
- Les voyelles précédées d'une consonne non voisée et suivies d'une consonne voisée non allongeante (N.Voisé_Voisée);
- Les voyelles précédées d'une consonne non voisée et suivies d'une consonne non voisée (N.Voisée_N.Voisée);
- Les voyelles précédées d'une consonne voisée et suivies d'une consonne allongeante (Voisée_Allongeante);
- Les voyelles précédées d'une consonne non voisée et suivies d'une consonne allongeante (N.Voisée_Allongeante).

Notons que ce codage reflète le voisement attendu des consonnes. Les résultats de cette seconde ANOVA nous permettent de préciser la différence régionale que nous observions dans le cas des voyelles produites en (VK). Comme on peut le constater sur la Figure 13, en (VK), seules les occurrences suivies d'une consonne voisée sont significativement plus longues chez les locuteurs de Saguenay que chez ceux de Québec et de Rouyn-Noranda ($p < 0,0001$), que celles-ci soient précédées d'une consonne voisée ou non voisée.

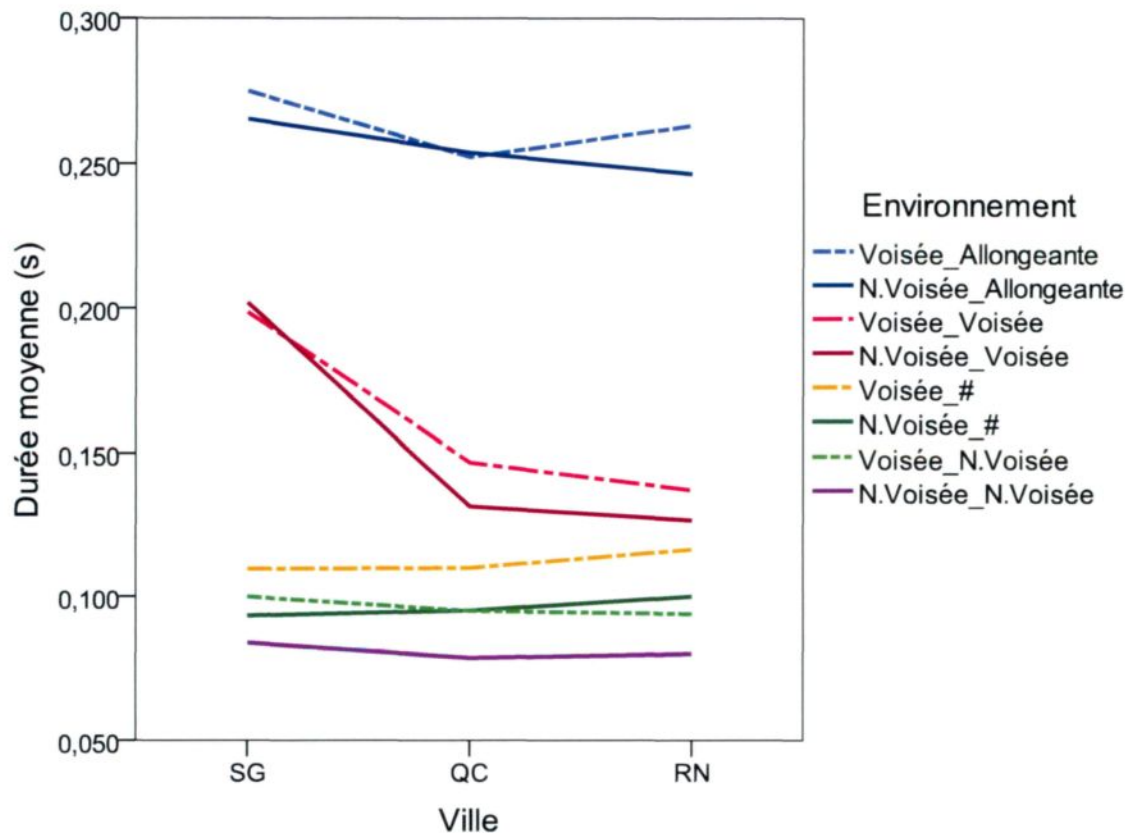


Figure 13 : Durée moyenne des voyelles en fonction de la ville d'origine des locuteurs et du contexte consonantique.

Les voyelles fermées en (VK) produites par les locuteurs saguenéens (à tout le moins ceux de notre corpus) seraient donc plus sensibles à l'effet allongeant du voisement de la consonne subséquente. Morasse (1995 : 21) mentionne d'ailleurs que : « [c]e phénomène varie toutefois en importance d'un idiome à l'autre ». Nos résultats suggèrent qu'une telle variation pourrait également exister au sein d'une même langue, selon la région d'origine des locuteurs. D'autre part, même si aucune différence significative liée au sexe des locuteurs n'a été détectée, nous sommes en droit de nous demander si, chez les locuteurs de Saguenay, cette sensibilité à l'effet allongeant du voisement de la consonne subséquente ne serait pas plus prononcée chez les femmes que chez les hommes. En effet, comme l'illustre la Figure 14, les voyelles fermées en (VK) suivies d'une consonne voisée produites par les femmes de Saguenay sont sensiblement plus longues que celles produites par les hommes de Saguenay. La faible taille des effectifs (5 hommes, 5 femmes) explique possiblement l'absence de signification statistique.

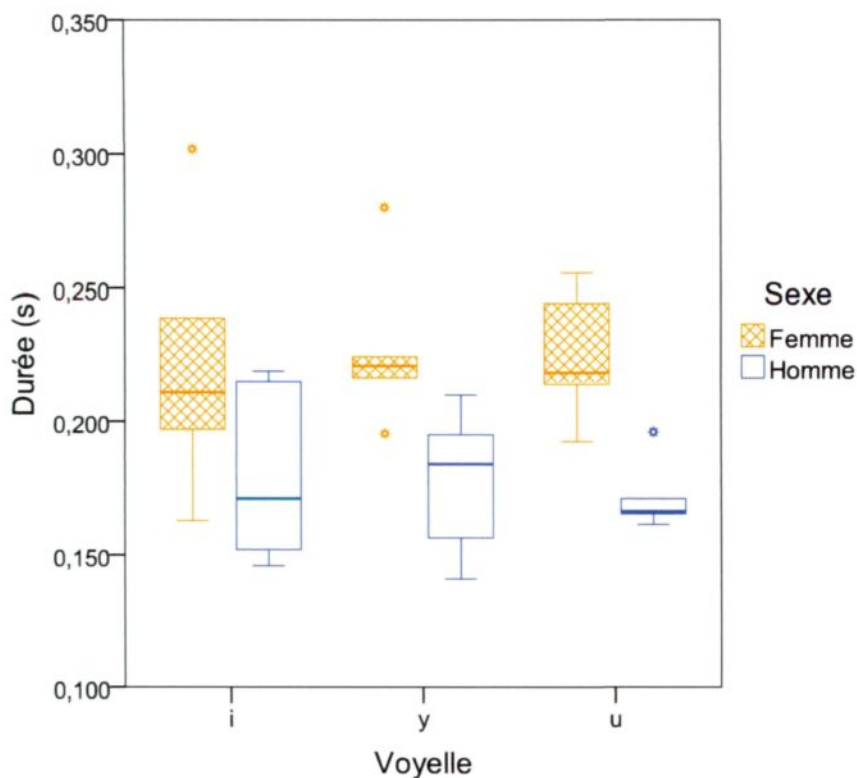


Figure 14 : Diagramme boîtes à moustaches présentant la durée des voyelles fermées en (VK) suivies d'une consonne voisée produites par les locuteurs de Saguenay en fonction du sexe. Les astérisques indiquent des données extrêmes.

En dépit de cette particularité saguenéenne, dans toutes les villes où nous avons enquêté, les rapports de durées entre les voyelles produites dans les 8 contextes consonantiques définis ci-haut restent les mêmes (voir Figure 15). En premier lieu, comme nous l'avons déjà mis en évidence (voir Figures 10 et 11), les voyelles fermées, toutes classes phonémiques confondues, sont significativement plus longues ($p < 0,0001$) en (VR) qu'en (V#) et en (VK). Cela dit, nous observons ici que parmi les occurrences produites en (VR), il n'y a pas de différence de durée significative en fonction du voisement de la consonne précédente. Il n'y a pas non plus de différence de durée significative entre les occurrences en syllabe fermée par une consonne non allongeante voisée, qu'elles soient précédées d'une consonne voisée ou d'une non-voisée. Par contre, les occurrences en syllabe fermée par une consonne non allongeante non voisée sont significativement plus longues lorsqu'elles sont précédées d'une consonne voisée que lorsqu'elles sont précédées d'une non-voisée ($p < 0,0001$). De même, les voyelles en (V#) sont significativement plus longues lorsqu'elles sont précédées d'une consonne voisée que lorsqu'elles sont précédées d'une non-voisée ($p < 0,0001$).

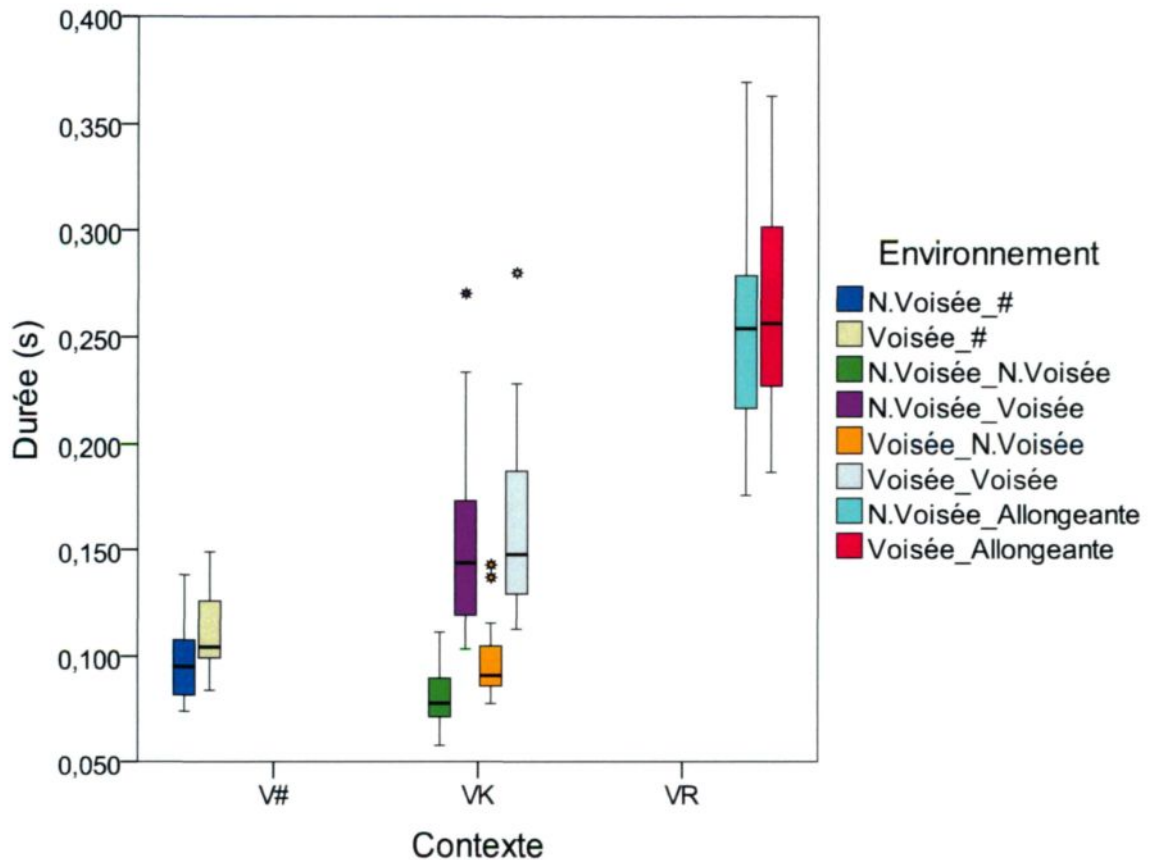


Figure 15 : Diagramme à boîtes à moustaches présentant la durée des variantes en fonction du contexte consonantique. Les astérisques indiquent des données extrêmes.

Bref, le voisement de la consonne antéposée ne semble avoir d'impact sur la durée des voyelles que lorsque la consonne postposée n'est pas elle-même voisée. Nous émettons l'hypothèse que ces tendances récurrentes puissent s'expliquer par ce que Di Cristo (1985 : 459) nomme un « critère d'inextensibilité ». C'est-à-dire que, dépendamment du contexte, les voyelles ne pourraient pas s'allonger au-delà d'un certain seuil. Dans le cas présent, ce seuil serait déjà atteint par l'effet allongeant de la consonne postposée.

En deuxième lieu, nous notons que les voyelles fermées en (V#) précédées d'une consonne voisée sont plus longues que les voyelles en (VK) précédées d'une consonne voisée et suivies d'une non-voisée ($p < 0,0001$). Similairement, les voyelles en (V#) précédées d'une consonne non-voisée sont plus longues que les voyelles en (VK) précédées et suivies d'une consonne non-voisée ($p < 0,0001$). Bref, à effet allongeant égal, les voyelles en (V#) sont plus longues que celles en (VK). Cependant, lorsque les voyelles en (VK) subissent un effet allongeant de la consonne postposée, leur durée dépasse celle des voyelles en (V#). Le rapport de durée entre les variantes tendues (attendues en (V#)) et les variantes relâchées (attendues en (VK)) semble donc largement tributaire du contexte consonantique.

D'après Straka (1959 : 279) et Gendron (1966 : 41), en raison de contraintes articulatoires, plus une voyelle fermée est longue, plus elle est fermée (ou tendue), et plus elle est brève, plus elle est ouverte (ou relâchée). Cette hypothèse peut expliquer les résultats que nous obtenons lorsque les voyelles relâchées sont suivies d'une consonne non-voisée (et sont alors plus courtes que leurs équivalents tendus). Néanmoins, nos résultats concernant les relâchées suivies d'une consonne voisée (plus longues que les tendues) la contredisent, ou, à tout le moins, impliquent qu'une autre source de variation intervienne. Marchal (1980 : 84) et Poiré (1994 : 145) suggèrent d'ailleurs à ce propos que les relâchées suivies d'une consonne non voisée sont soumises à un effet abrégant.

Ainsi, l'influence des consonnes adjacentes – plus précisément de leur voisement – que nous observons permet d'expliquer les résultats divergents rapportés par Martin (2002), Poliquin (2006) et Arnaud et coll. (2011) en ce qui concerne le rapport de durée entre les voyelles fermées tendues et relâchées. En effet, les relâchées du corpus de Martin (2002) se retrouvent principalement dans un environnement consonantique non voisé et celles du corpus de Poliquin (2006) sont systématiquement suivies de la consonne non voisée /t/. Au contraire, les relâchées analysées par Arnaud et coll. (2011) sont toutes précédées et suivies de la consonne voisée /b/. Les relâchées de Arnaud et coll. (2011) subissent donc un effet allongeant suffisant pour que leur durée dépasse celle des tendues (d'autant plus que les locuteurs sont tous originaires de Saguenay) ; ce qui n'est pas le cas des relâchées de Martin (2002) et de Poliquin (2006), qui ne subissent pas (ou peu) d'effet allongeant, voire qui sont soumises à l'effet abrégant de la consonne post-posée non voisée.

Dans un troisième et dernier temps, nous avons pris soin de tester l'effet de l'identité de la consonne finale sur la durée des voyelles en (VR), toutes classes phonémiques confondues (voir Figure 16). Les durées moyennes individuelles de ces voyelles devant /ʃ/, /v/, /z/ et /ʒ/ ont été calculées avant de procéder à une ANOVA. Les voyelles fermées produites en syllabe fermée par la consonne /v/ apparaissent significativement plus courtes ($p < 0,0001$) que les autres voyelles en (VR), quoique de peu. Encore une fois, l'éventuel effet de la consonne précédente (qui n'était pas parfaitement contrôlée) sur ce résultat reste à évaluer. On ne peut par ailleurs pas exclure que les transitions entre la voyelle et la consonne allongeante, variant sensiblement selon l'identité de ladite consonne, aient entraîné un biais dans notre technique de segmentation.

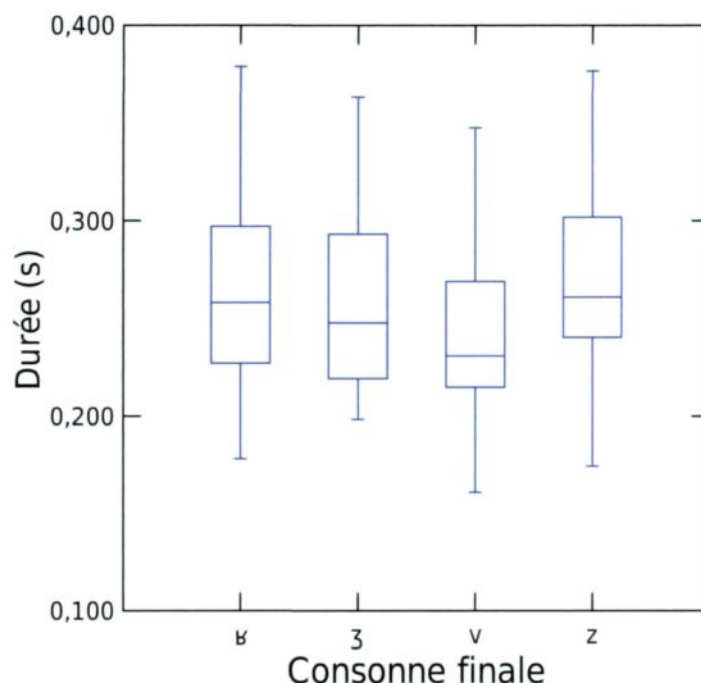


Figure 16 : Diagramme boîtes à moustaches présentant la durée des voyelles allongées en fonction de la consonne suivante.

4.2 Caractéristiques spectrales

Une description sommaire des variantes tendues, relâchées et allongées des voyelles fermées du FQ peut déjà être donnée à l'observation des Figures 17, 18 et 19, qui présentent les 1343 occurrences de notre corpus à 25 %, 50 % et 75 % de leur durée projetées dans des diagrammes biformantiques F_1 / F_2 et F_2 / F_3 . Les valeurs normalisées en scores-Z ont été utilisées afin de présenter conjointement les productions féminines et masculines. Nous nous sommes préalablement assurée que les mêmes tendances se retrouvaient en Hertz chez les hommes et les femmes.

En observant les diagrammes F_1 / F_2 , nous remarquons premièrement que les occurrences appartenant aux classes phonémiques /i/, /y/ et /u/ ne se distinguent à peu près pas en fonction de F_1 ; ce qui est loin d'être étonnant sachant qu'elles présentent le même degré d'aperture et que F_1 est un corrélat acoustique important de ce trait articulaire (Delattre, 1948 : 482). C'est plutôt sur F_2 , associé au degré d'antériorité (Delattre, 1948 : 483), et sur F_3 , associé à l'arrondissement (Lindblom et Sundberg, 1971 : 1176), que les trois voyelles fermées se distinguent. Comme on peut le voir sur les diagrammes F_2 / F_3 , /u/, qui est postérieur, possède un F_2 plus bas que les voyelles antérieures /i/ et /y/, tandis que /i/, non arrondi, possède un F_3 plus élevé que /y/ et /u/, qui eux sont arrondis. Cette distinction entre les classes phonémiques est particulièrement nette à 25 % de la durée des voyelles (Figure 17).

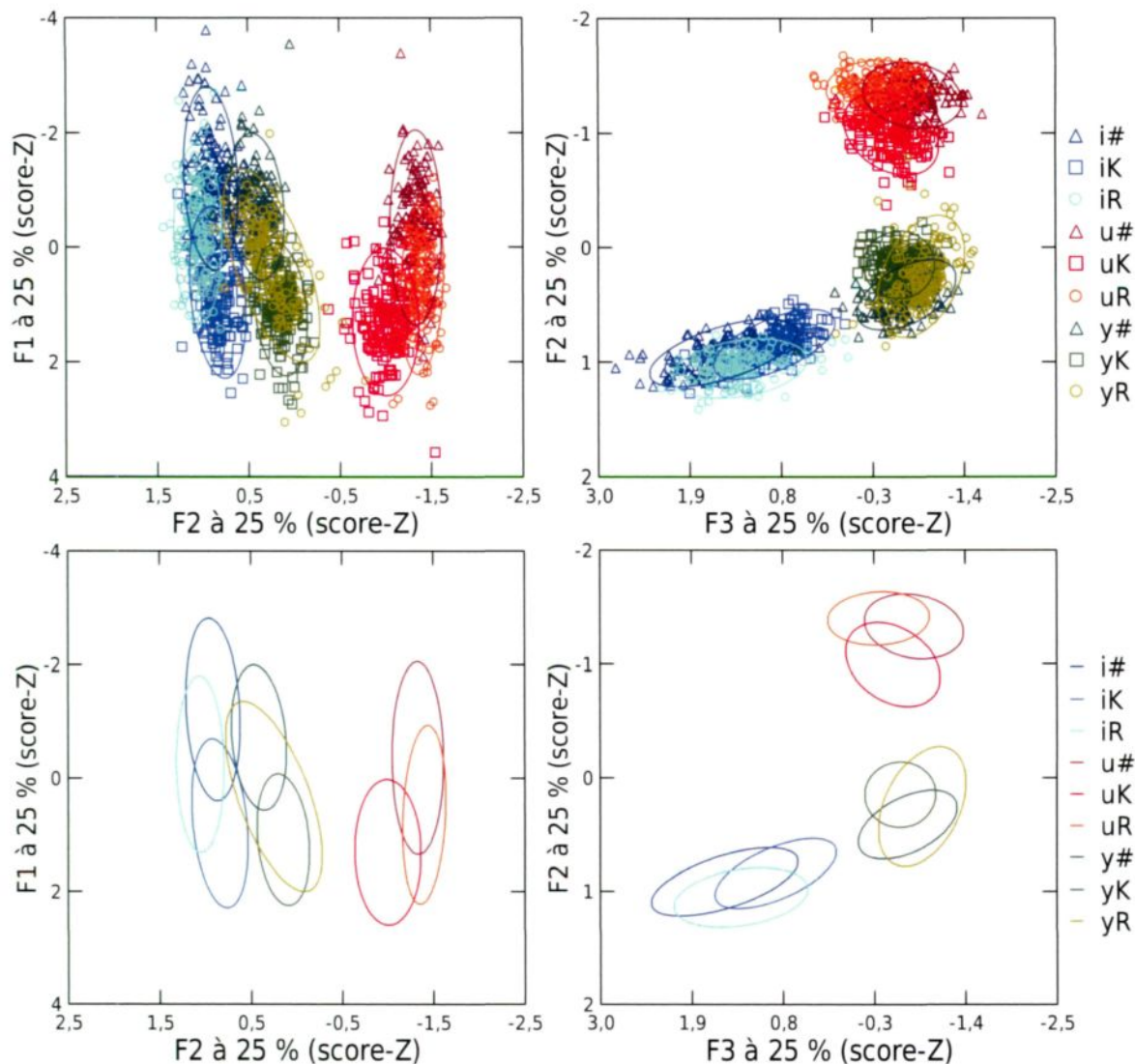


Figure 17 : Diagrammes bifurcés F_1 / F_2 et F_2 / F_3 normalisés présentant les 1343 occurrences analysées à 25 % de leur durée regroupées en fonction des classes vocaliques. Les ellipses de dispersion englobent 86,5 % des occurrences d'une même classe vocalique.

Au contraire, on observe sur les diagrammes F_1 / F_2 que les variantes tendues, relâchées et allongées d'une même classe phonémique se différencient maximale sur F_1 et présentent des trajectoires temporelles différentes. Le déplacement des ellipses entre les Figures 17, 18 et 19 met en lumière ces trajectoires (la Figure 24, à la p. 60, offre une synthèse des trajectoires moyennes). À 25 % de leur durée (Figure 17), les variantes tendues ont le F_1 le plus bas, les relâchées ont le F_1 le plus élevé et les allongées se situent entre les deux. En cours d'émission (donc en passant à la Figure 18 puis à la Figure 19), les allongées se tendent (ou se ferment) au point où leur F_1 rejoint partiellement celui des tendues. Les relâchées, quant à elles, se centralisent (F_2 plus élevé pour /u/, plus faible pour /i/ et /y/). Notons que les différentes variantes d'une même classe phonémique ne se différencient jamais aussi nettement les unes des autres que les différentes classes phonémiques entre elles.

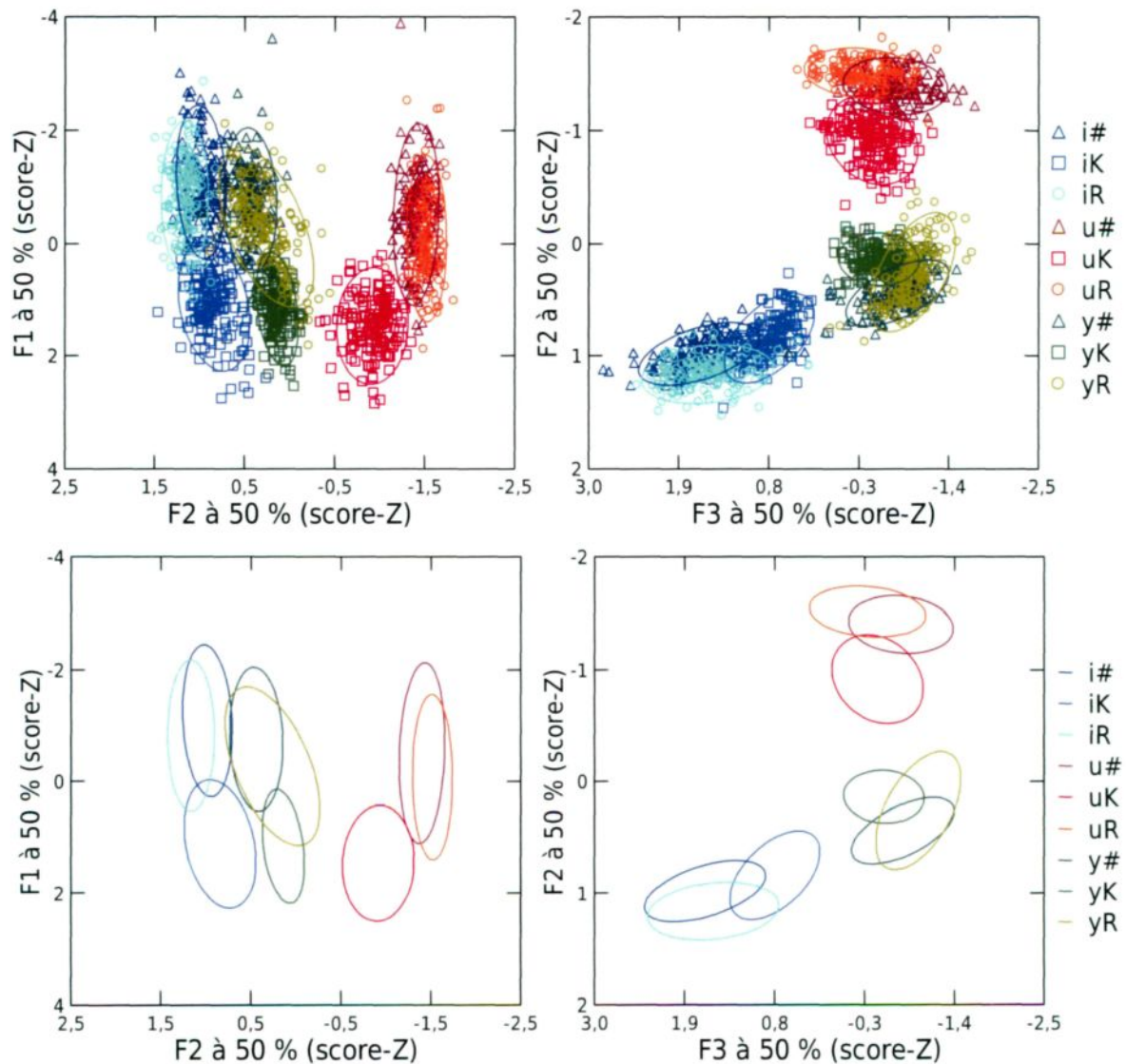


Figure 18 : Diagrammes bifurmantiques F_1 / F_2 et F_2 / F_3 normalisés présentant les 1343 occurrences analysées à 50 % de leur durée regroupées en fonction des classes vocaliques. Les ellipses de dispersion englobent 96,5 % des occurrences d'une même classe vocalique.

Les ellipses de dispersion plus importantes à 75 % de la durée des voyelles (Figure 19) témoignent possiblement de l'effet des consonnes finales, qui, rappelons-le, sont plus variées dans notre corpus que les consonnes initiales.

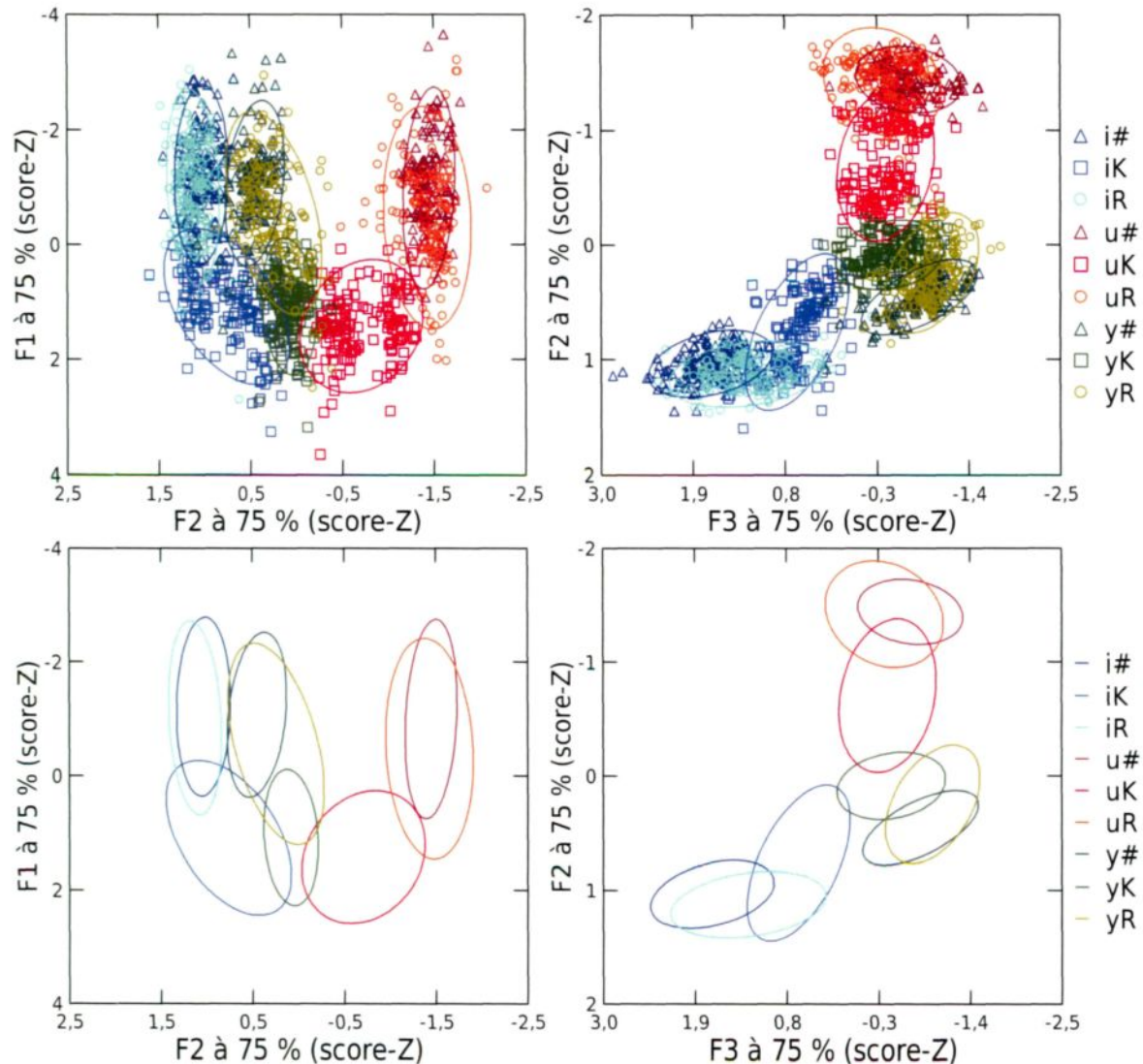


Figure 19 : Diagrammes bifurmantiques F_1 / F_2 et F_2 / F_3 normalisés présentant les 1343 occurrences analysées à 75 % de leur durée regroupées en fonction des classes vocaliques. Les ellipses de dispersion englobent 86,5 % des occurrences d'une même classe vocalique.

Nous avons ensuite procédé à un examen plus approfondi de nos données. Les tableaux en Annexe 7 présentent par région, par sexe, par classe vocalique et par point de mesure les valeurs moyennes (en Hertz) de la F_0 et de F_1 , F_2 et F_3 calculés à partir des moyennes individuelles. L'écart-type est indiqué entre parenthèses et l'effectif est de 5 individus par cellule. Nous avons d'abord cherché à savoir si les caractéristiques spectrales, telles que mesurées en Hertz,³⁷ de chacune des classes vocaliques ($i\#$), (iK), (iR), ($y\#$), (yK), (yR), ($u\#$), (uK) et (uR) différaient significativement selon le sexe et l'origine géographique des locuteurs, et si l'évolution de ces caractéristiques spectrales en cours d'émission était significative. Les résultats détaillés des ANOVAs menées à cet effet sont rapportés dans l'Annexe 8. D'emblée, on constate que

³⁷ Les données normalisées en scores-Z seront traitées un peu plus loin.

la fréquence fondamentale et les fréquences formantiques en Hertz sont, hormis de rares exceptions, significativement influencées par le sexe du locuteur. En moyenne, d'après nos données, la F_0 des hommes représente 61 % de celle des femmes. Si la principale cause de cet écart est sans doute d'ordre physiologique (Titze, 1989), il n'est pas exclu qu'une part de cet écart relève d'habitudes sociales spécifiques à la langue ou variété de langue (voir notamment Pépiot, 2013 : 27 pour un survol de la question). Chose certaine, les habitudes de nos locuteurs de Saguenay, de Québec et de Rouyn-Noranda ne diffèrent pas, puisque le rapport entre la F_0 des hommes et des femmes est de 61 % dans chacune de ces villes.

Les rapports entre les formants masculins et féminins, toutes classes vocaliques confondues, peuvent quant à eux être exprimés approximativement par les droites de régression illustrées sur la Figure 20. Les coefficients de détermination (R^2) mesurent le degré d'association entre deux variables. La valeur de ce coefficient est toujours comprise entre 0 et 1. Plus elle est élevée, plus l'association est forte (moins les points divergent de la droite de régression). Dans le cas présent, les coefficients de détermination indiquent que le rapport entre les fréquences formantiques féminines et masculines moyennes, tel qu'exprimé par les droites de régression, est très constant, particulièrement pour F_2 et F_3 .

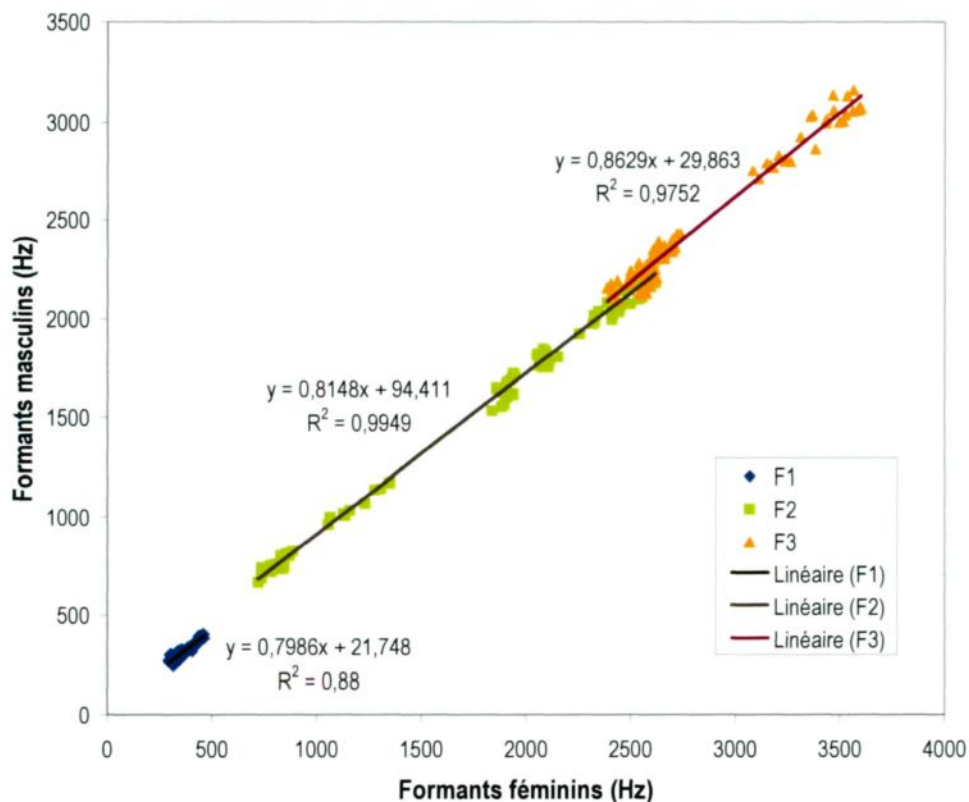


Figure 20 : Rapport entre les fréquences formantiques (F_1 , F_2 et F_3) masculines et féminines moyennes pairées par ville, par classe vocalique et par point de mesure.

Cela dit, comme l'illustre la Figure 21, et comme l'ont déjà rapporté certains chercheurs (par exemple Fant, 1966 : 22; Calliope, 1989 : 86; Simpson, 2009 : 626), la différence entre les productions masculines et féminines varie quelque peu selon la classe vocalique. D'après nos données, plus la valeur d'un formant est élevée, plus l'écart entre les fréquences masculines et féminines est important.³⁸ Autrement dit, plus une voyelle est ouverte (ou relâchée), plus l'écart entre le F_1 des hommes et des femmes est grand; et plus une voyelle est antérieure, plus l'écart entre le F_2 des hommes et des femmes est grand. Comme nous ne travaillons ni sur le même ensemble de classes vocaliques, ni sur la même langue ou variété de langue que Fant (1966), Calliope (1989) et Simpson (2009), une comparaison de nos résultats respectifs serait stérile. Johnson (2005 : 379) souligne en effet que l'écart entre les fréquences formantiques masculines et féminines varie considérablement d'une langue à l'autre, voire d'une variété de langue à l'autre.

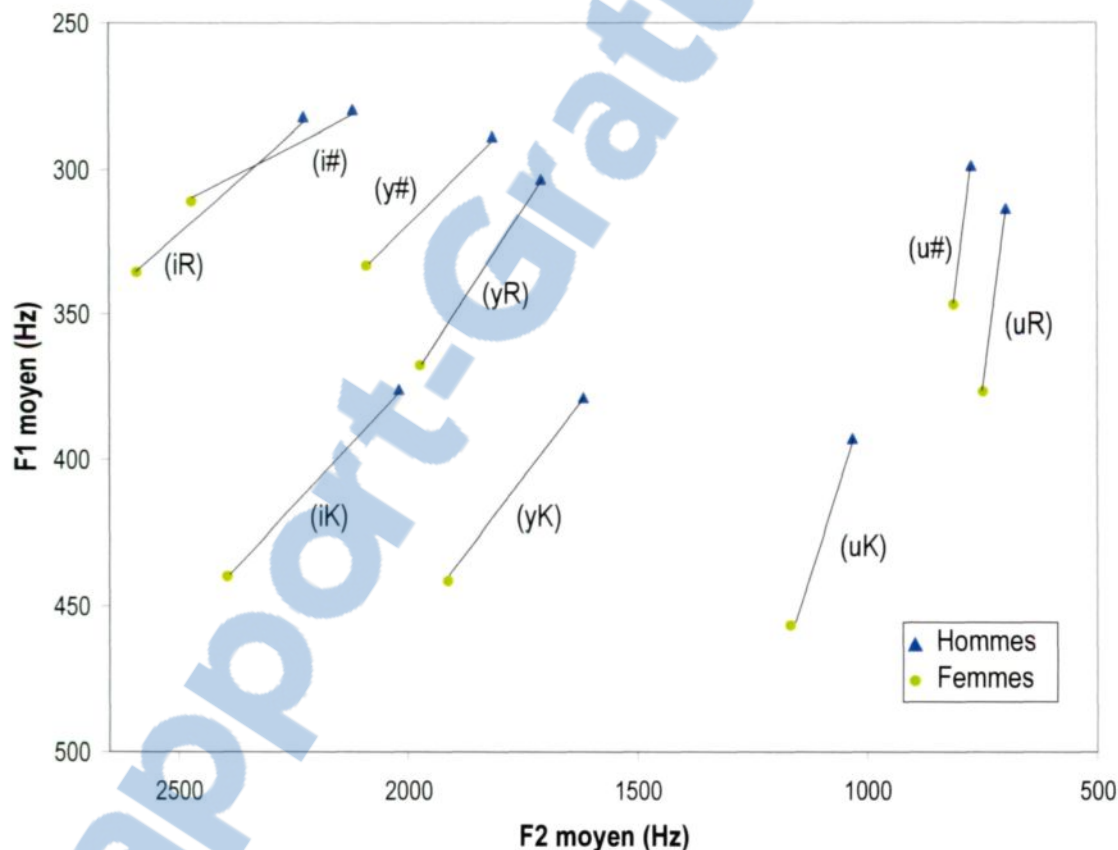


Figure 21 : Diagramme bifonnantique F_1 / F_2 (en Hertz) présentant les valeurs féminines et masculines moyennes (calculées à 50 % de la durée des voyelles) des classes vocaliques (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK), ou (uR).

³⁸ Ainsi donc, le rapport entre les fréquences formantiques masculines et féminines est exprimé de manière plus directe par une droite de régression que par un simple quotient (du moins en ce qui concerne nos données). En effet, alors que le quotient entre les fréquences masculines et féminines varie d'une classe vocalique à une autre, l'équation de la droite de régression demeure la même, grâce au poids de l'ordonnée à l'origine. C'est-à-dire que plus la fréquence est basse, plus l'ordonnée à l'origine réduit le rapport exprimé par l'équation.

En ce qui concerne la variation diatopique, les résultats des ANOVAs indiquent que les voyelles fermées allongées produites par les locuteurs de Saguenay présentent un F_1 globalement plus bas que celles produites par les locuteurs de Québec et de Rouyn-Noranda; cette différence n'étant toutefois pas significative dans le cas des /u/. Cela confirme notre impression auditive initiale selon laquelle les voyelles fermées allongées seraient plus tendues à Saguenay qu'ailleurs. Il semble que la dynamique temporelle du F_2 des /y/ en syllabe ouverte (y#) soit également différente à Saguenay. Comme l'illustre la Figure 22, en moyenne, le F_2 des (y#) produits par les locuteurs de Saguenay augmente significativement en cours d'émission alors que ce n'est le cas ni à Québec, ni à Rouyn-Noranda.

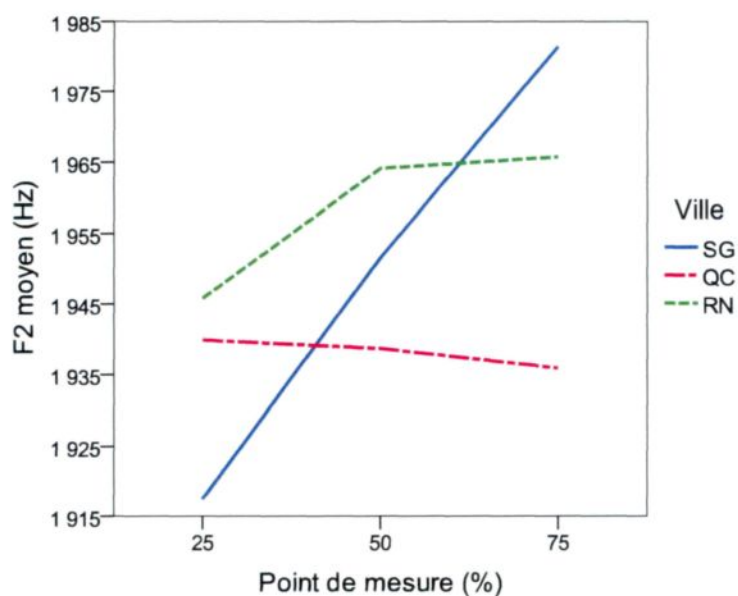


Figure 22 : Trajectoire temporelle moyenne du F_2 de la classe vocalique (y#) en fonction de la ville d'origine des locuteurs.

La Figure 23 présente quant à elle, pour chaque ville où nous avons enquêté, la dynamique temporelle moyenne du F_1 des /y/ produits en syllabe fermée par une consonne non allongeante (yK). À Saguenay et à Rouyn-Noranda, le F_1 moyen des (yK) ne varie pas significativement en cours d'émission alors qu'il chute de façon significative à Québec.

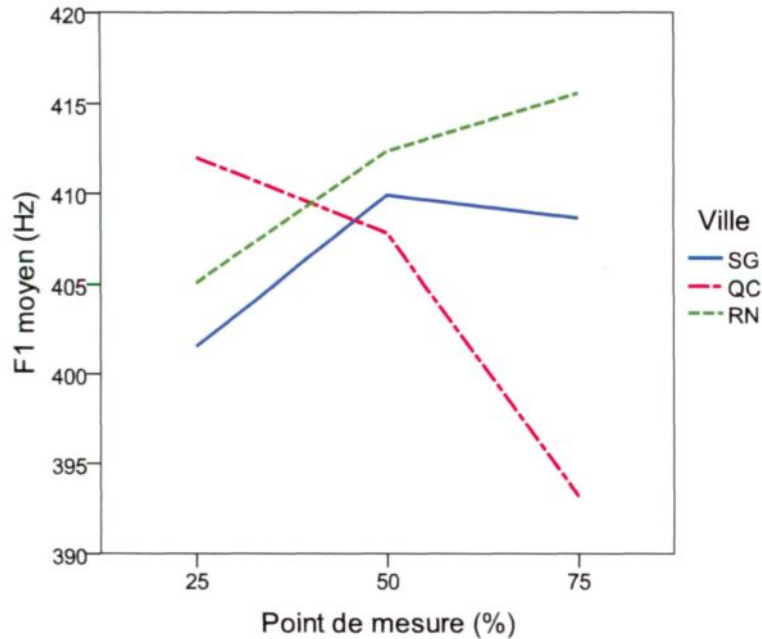


Figure 23 : Trajectoire temporelle moyenne du F₁ de la classe vocalique (yK) en fonction de la ville d'origine des locuteurs.

En ce qui a trait à la dynamique spectrale, les résultats des ANOVAs confirment les tendances que nous observons sur les Figures 17, 18 et 19, en plus d'en révéler de nouvelles, plus fines. Le Tableau 2 synthétise ces résultats. Une case grisée indique que, pour la classe vocalique concernée, aucune différence significative de la valeur de la composante spectrale en question n'est détectée entre l'un ou l'autre des trois points de mesure situés à 25 %, 50 % et 75 % de la durée des occurrences. Un astérisque indique que la dynamique temporelle de la composante n'est pas la même dans toutes les villes où nous avons enquêté (comme nous venons de le voir sur les Figures 22 et 23). Les signes « + » et « - » dénotent respectivement une augmentation ou une diminution significative de l'indice acoustique considéré entre 25 % et 75 % de la durée des occurrences (que cette évolution se manifeste entre 25 % et 50 %, entre 50 % et 75 %, ou les deux). Enfin, la série « +- » indique une augmentation de la valeur de la composante spectrale considérée entre 25 % et 50 % de la durée des occurrences suivie d'une diminution de cette valeur entre 50 % et 75 %.

Tableau 2 : Évolution des caractéristiques spectrales des différentes classes vocaliques

Classe vocalique	F ₀	F ₁	F ₂	F ₃
(i#)			+	+
(iK)	-	+	-	-
(iR)	-	-	+	+/-
(y#)		-	*	-
(yK)	-	*	-	+
(yR)	-	-		-
(u#)	-	-	-	+
(uK)		+	+	+
(uR)	-	-	-	

De prime abord, on constate que partout où la F₀ varie significativement entre 25 % et 75 % de la durée des voyelles, elle diminue ; ce qui n'est guère surprenant en français en fin d'énoncé déclaratif (Di Cristo, 1998 : 201). L'évolution temporelle de la F₀ n'apparaît donc pas comme une caractéristique distinctive des classes vocaliques étudiées. On remarque ensuite que les différences significatives concernant l'évolution temporelle de F₁ indiquent que les voyelles fermées tendues et allongées se tendent (ou se ferment) en cours d'émission, alors que les relâchées (à l'exception des (yK) dont nous avons discuté plus haut) ont tendance à se relâcher (ou à s'ouvrir) encore plus. En ce qui concerne la dynamique temporelle de F₂, les tendues et les allongées d'une même classe phonémique montrent également des comportements similaires, qui s'opposent à celui des relâchées : alors que les relâchées se centralisent, les tendues et les allongées se déplacent en périphérie du trapèze vocalique. Pour leur part, les variations temporelles de F₃ semblent moins systématiques. On note toutefois que le F₃ des relâchées arrondies augmente alors que celui des relâchées non arrondies diminue en cours d'émission.

Quoi qu'il en soit, les trajectoires opposées des voyelles fermées tendues et relâchées de notre corpus dans un espace F₁ / F₂ sont conformes à ce qui avait déjà été rapporté en FQ (Arnaud et coll., 2011 : 246) et en anglais américain (Nearey et Assmann, 1986 : 1300 ; Hillenbrand et coll., 2001 : 757). Soulignons également que bien que les voyelles fermées tendues et relâchées présentent une dynamique spectrale significative, les trajectoires des variantes allongées sont de loin les plus importantes en moyenne (voir Figure 24).

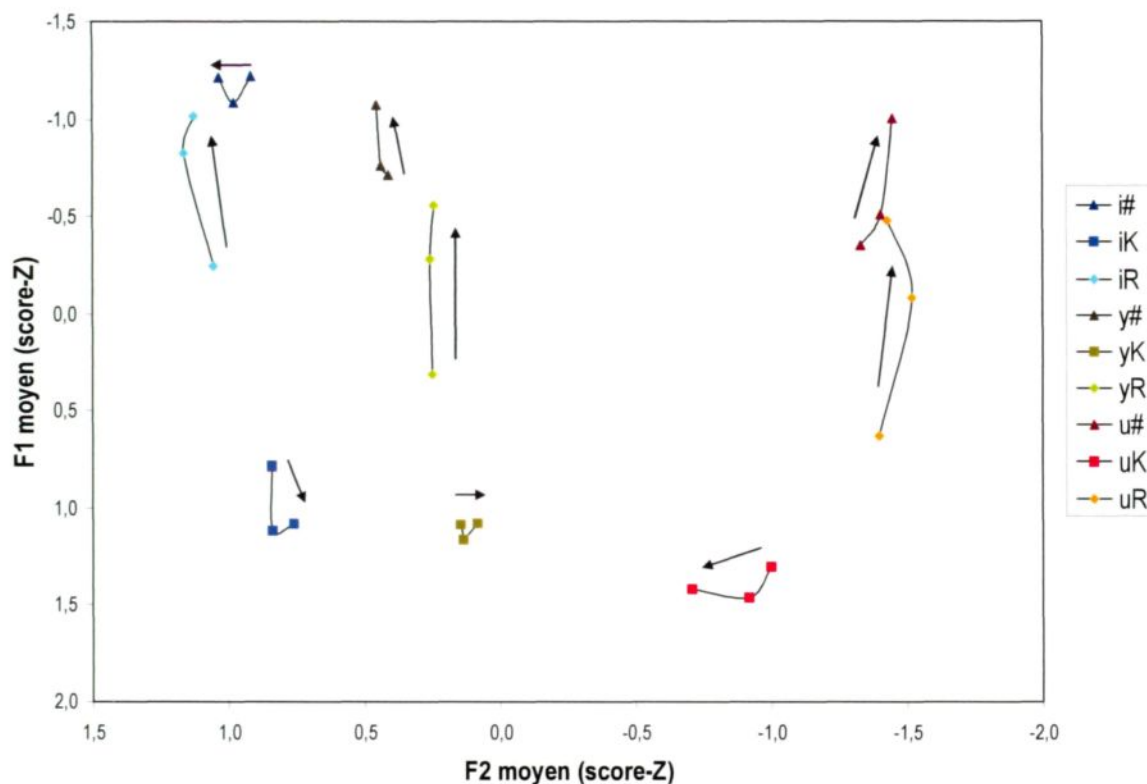


Figure 24 : Diagramme bifurmantique F_1 / F_2 normalisés présentant l'évolution moyenne des différentes classes vocales, tous locuteurs confondus.

Les ANOVAs dont nous avons rapporté les résultats jusqu'à présent portaient sur les valeurs en Hertz des composantes spectrales. Lorsque les mêmes tests sont réalisés sur les valeurs fréquentielles normalisées en scores-Z (voir Annexe 8), la fréquence fondamentale et les fréquences formantiques des hommes et des femmes ne se distinguent plus. Mentionnons également que pour certaines classes vocales (notamment (iR), (yK) et (uR)), il était impossible de tester l'effet des différentes variables indépendantes sur la F_0 au moyen d'une ANOVA car la normalité n'était pas rencontrée.³⁹ En effet, la F_0 varie peu d'une classe vocale à l'autre. En la normalisant, la variation interindividuelle se trouve fortement réduite. Dès lors, la distribution de la F_0 se concentre fortement autour de la valeur moyenne. Ainsi, le coefficient d'aplatissement (*kurtosis*) est trop élevé pour que la distribution puisse être considérée comme normale.

Hormis l'élimination des différences entre les productions féminines et masculines, les résultats des ANOVAs portant sur les données normalisées correspondent généralement à ceux des ANOVAs portant sur les données brutes en Hertz. En ce qui concerne l'évolution des caractéristiques spectrales en cours d'émission,

³⁹ Le recours à un test ne requérant pas la normalité n'était par ailleurs pas envisageable puisqu'à notre connaissance, il n'existe aucun test non paramétrique pouvant prendre en compte à la fois des effets fixes (comme celui du sexe des locuteurs) et des effets aléatoires (comme la corrélation qui existe entre les occurrences produites par un même individu).

seule la diminution de F_0 des (u#), qui était significative en Hertz, ne l'est plus en scores-Z. Autrement, bien que les résultats des comparaisons *post hoc* entre 25 % et 50 % ou entre 50 % et 75 % de la durée des occurrences diffèrent parfois, l'évolution des composantes spectrales entre 25 % et 75 % (rapportée dans le Tableau 2) demeure inchangée lorsque les données sont exprimées en scores-Z plutôt qu'en Hertz. Mentionnons enfin que l'augmentation en cours d'émission du F_1 des occurrences de (yK) produites par les locuteurs de Rouyn-Noranda se révèle statistiquement significative en scores-Z alors qu'elle ne l'était pas en Hertz (voir Figure 23 ci-haut).

C'est justement sur le plan de la variation régionale que surgissent le plus de disparités. Par exemple, le F_1 des /i/ et des /y/ allongés est significativement plus faible à Saguenay qu'à Québec et qu'à Rouyn-Noranda lorsque les données en Hertz sont considérées, alors qu'en utilisant les données normalisées en scores-Z, le F_1 de toutes les voyelles allongées (/u/ inclus) est plus faible à Saguenay qu'à Québec, mais pas qu'à Rouyn-Noranda. Toujours en scores-Z, plusieurs différences régionales qui n'étaient pas détectées en Hertz sont mises au jour :

- Le F_2 des (i#) est significativement plus faible à Saguenay qu'à Rouyn-Noranda ;
- Le F_1 des (iK) est significativement plus élevé à Saguenay qu'à Rouyn-Noranda ;
- Le F_2 des (iK) est significativement plus faible à Saguenay qu'à Québec ;
- Le F_1 des (y#) est significativement plus élevé à Saguenay qu'à Québec ;
- Le F_2 des (yR) est significativement plus élevé à Saguenay qu'à Québec et qu'à Rouyn-Noranda.

Ces différences sont toutefois ténues. Par exemple, le F_1 moyen des (y#) est de 305 Hz à Saguenay et de 298 Hz à Québec. La normalisation des données à l'aide des scores-Z peut-elle vraiment rendre un écart moyen de 7 Hz significatif (au-delà de l'acception statistique du terme)? En fait, en adoptant un seuil de signification de 0,1 % ($\alpha=0,001$) lors de l'utilisation des données normalisées en scores-Z, nous obtiendrions des résultats beaucoup plus similaires à ceux obtenus à partir des données en Hertz avec un seuil de signification de 1 %. Aucune des nouvelles différences régionales que nous venons de lister ne serait alors significative. Il en va de même de l'interaction entre le sexe des locuteurs et l'évolution temporelle du F_3 des (y#) qui apparaît lorsque les fréquences sont normalisées. Avec un seuil de signification fixé à 1 % et lorsque les données sont normalisées, aucune évolution temporelle significative du F_3 des occurrences de (y#) produites par les femmes n'est détectée, tandis que la valeur de cet indice diminue significativement pour les hommes. Si la valeur du F_3 des (y#) est maintenue en Hertz avec un seuil de signification de 1 %, ou bien si elle est normalisée en score-Z avec un seuil de signification de 0,1 %, on note alors une diminution de F_3 , peu importe le sexe des locuteurs.

4.3 Analyses discriminantes

Jusqu'à présent, dans ce chapitre, nous nous sommes intéressée séparément à la durée des voyelles et à leurs caractéristiques spectrales. Par ailleurs, les ANOVAs que nous avons menées afin de tester l'effet de diverses variables indépendantes sur les caractéristiques spectrales des voyelles portaient sur une seule composante spectrale et sur une seule classe vocalique à la fois. Dans la présente section, nous chercherons cette fois à évaluer la contribution de *tous* les indices acoustiques considérés à la *différentiation* des voyelles fermées tendues, relâchées et allongées du FQ. Ce faisant, nous tenterons de circonscrire les diverses *combinaisons* d'indices acoustiques permettant de distinguer le plus efficacement ces 9 classes vocaliques. Afin d'atteindre cet objectif, nous avons choisi d'employer à titre exploratoire des analyses discriminantes (voir le commentaire méthodologique à ce propos à la p. 41).

Une analyse discriminante est un test statistique qui établit l'appartenance des cas qui lui sont soumis (ici des occurrences vocaliques) à un groupe (dans ce cas-ci à l'une des classes vocaliques considérées) en fonction des informations (ou prédicteurs) dont elle dispose à propos de ces cas (ici les combinaisons d'indices acoustiques que nous avons définies au point 3.5). Pour ce faire, l'analyse compare la valeur des indices acoustiques des occurrences de l'échantillon testé (dont elle ignore la classe vocalique) à la valeur des indices acoustiques des occurrences de l'échantillon d'entraînement (dont elle connaît la classe vocalique). La classification ainsi obtenue est ensuite confrontée à celle établie par le chercheur de façon à calculer dans quelle proportion ces deux classifications correspondent. Plus le taux de classification correcte est élevé, plus les indices acoustiques pris en considération dans l'analyse discriminante permettent de différencier efficacement les occurrences issues de groupes différents.

Le Tableau 3 présente les taux de classification correcte obtenus lorsque différentes combinaisons de prédicteurs sont fournies. On remarque tout d'abord que les données normalisées en scores-Z différencient plus efficacement les classes vocaliques (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK) et (uR) que les données brutes en Hertz, et ce, quels que soient les indices acoustiques et le nombre de points de mesure considérés. La position relative des voyelles dans l'espace vocalique de chaque locuteur (ici exprimée en scores-Z) semble donc un meilleur indicateur de leur identité que leur position absolue dans un espace en Hertz.

En outre, puisque notre corpus contient à la fois des voyelles réputées brèves et allongées, il apparaît que la prise en compte de la durée améliore également de manière systématique la classification.

Tableau 3 : Taux de classification correcte globaux (en %), résultats des analyses discriminantes visant à classer 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

		50 %		25, 75 %		25, 50, 75 %	
		Durée	Durée	Durée	Durée	Durée	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

Concernant la prise en compte de la dynamique spectrale, on peut constater, à la lecture du Tableau 3, que le relevé des indices acoustiques en deux points de mesure plutôt qu'en un seul permet d'améliorer la classification dans la plupart des 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 F₀ (en Hertz comme en scores-Z) est considérée. Ces derniers résultats sont conformes à ce que nous avons observé dans le cadre d'une étude préliminaire (Arnaud et coll., 2011). Jacewicz et Fox (2012) étaient également parvenus à la conclusion que deux points de mesures étaient nécessaires et suffisants pour différencier adéquatement les voyelles de l'anglais américain et mettre au jour plusieurs différences régionales.

Au-delà de ces taux de classification globaux, il est possible de déterminer plus finement l'impact des différents indices acoustiques sur la classification des occurrences en examinant les matrices de confusion issues des analyses discriminantes. Le Tableau 4 présente la matrice de confusion obtenue à la suite de l'analyse discriminante qui incluait comme prédicteurs les valeurs de F₁ et de F₂ (en Hertz) relevés à 50 % de la durée de la voyelle; durée qui n'était cependant pas considérée. Les lignes du tableau correspondent aux classes vocaliques attendues (attribuées par le chercheur) et les colonnes, aux classes vocaliques attribuées par l'analyse discriminante. Il appert qu'en ne considérant que ces deux indices acoustiques, les /u/ sont parfaitement distincts des /i/ et des /y/; ce qui n'est pas étonnant, sachant que les /u/, postérieurs, ont un F₂ nettement inférieur à celui des /i/ et des /y/, qui sont antérieurs. Les confusions qui subsistent se produisent

donc entre les /i/ et les /y/, et entre les variantes d'une même voyelle, particulièrement entre les tendues et les allongées.

Tableau 4 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F₁ et F₂ (en Hertz) relevés à 50 % de la durée

	i#	iK	iR	u#	uK	uR	y#	yK	yR	% correct
i#	34	4	29	0	0	0	8	0	0	45
iK	0	47	3	0	0	0	12	8	5	63
iR	36	1	33	0	0	0	5	0	0	44
u#	0	0	0	46	7	21	0	0	0	62
uK	0	0	0	4	63	3	0	4	0	85
uR	0	0	0	28	1	45	0	0	0	61
y#	5	14	2	0	0	0	47	0	7	63
yK	0	3	0	0	0	0	1	60	11	80
yR	4	16	1	0	0	0	26	13	15	20
Total	79	85	68	78	71	69	99	85	38	58

Comme l'indique le Tableau 5, l'ajout aux prédicteurs de F₃ (en Hertz) relevé à 50 % de la durée de la voyelle, entraîne une réduction considérable de la confusion entre les /i/ et les /y/. Ce résultat concorde avec la séparation de ces deux classes phonémiques observée sur les diagrammes F₂ / F₃ des Figures 17, 18 et 19, et confirme l'importance de F₃ dans la distinction d'arrondissement (Lindblom et Sundberg, 1971), du moins pour les voyelles fermées antérieures du français (Schwartz et coll., 1993; Gendrot et coll., 2008).

Tableau 5 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F₁, F₂ et F₃ (en Hertz) à 50 % de la durée

	i#	iK	iR	u#	uK	uR	y#	yK	yR	% correct
i#	43	3	29	0	0	0	0	0	0	57
iK	4	63	2	0	0	0	3	3	0	84
iR	25	4	42	0	0	0	4	0	0	56
u#	0	0	0	45	8	20	0	0	0	62
uK	0	0	0	4	68	1	0	1	0	92
uR	0	0	0	26	1	47	0	0	0	64
y#	0	3	1	0	0	0	50	0	21	67
yK	0	7	1	0	1	0	0	59	7	79
yR	0	3	0	0	0	0	29	16	27	36
Total	72	83	75	75	78	68	86	79	55	66

Fait surprenant, en ajoutant non pas F₃, mais la F₀ aux prédicteurs de cette analyse discriminante, une réduction tout aussi importante de la confusion entre les /i/ et les /y/ est obtenue (voir Tableau 6). Pourtant, les résultats rapportés dans les tableaux de l'Annexe 7 indiquent que la F₀ des /i/ ne diffère pas sensiblement de celle des /y/.

Tableau 6 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F_0 , F_1 et F_2 (en Hertz) à 50 % de la durée

	i#	iK	iR	u#	uK	uR	y#	yK	yR	% correct
i#	28	4	36	0	0	0	5	0	1	38
iK	1	62	6	0	0	0	2	2	1	84
iR	9	3	62	0	0	0	1	0	0	83
u#	0	0	0	46	3	25	0	0	0	62
uK	0	0	0	1	69	0	0	3	1	93
uR	0	0	0	24	5	44	0	0	0	60
y#	2	2	2	0	0	0	41	0	27	55
yK	0	3	0	0	1	0	1	58	12	77
yR	4	1	2	0	0	0	16	9	42	57
Total	44	75	108	71	78	69	66	72	84	68

Cependant, comme l'illustre la Figure 25, la prise en compte de la F_0 permet de distinguer en très grande partie les /i/ masculins des /y/ féminins (tous contextes confondus), qui autrement se chevauchent largement dans un diagramme F_1 / F_2 .

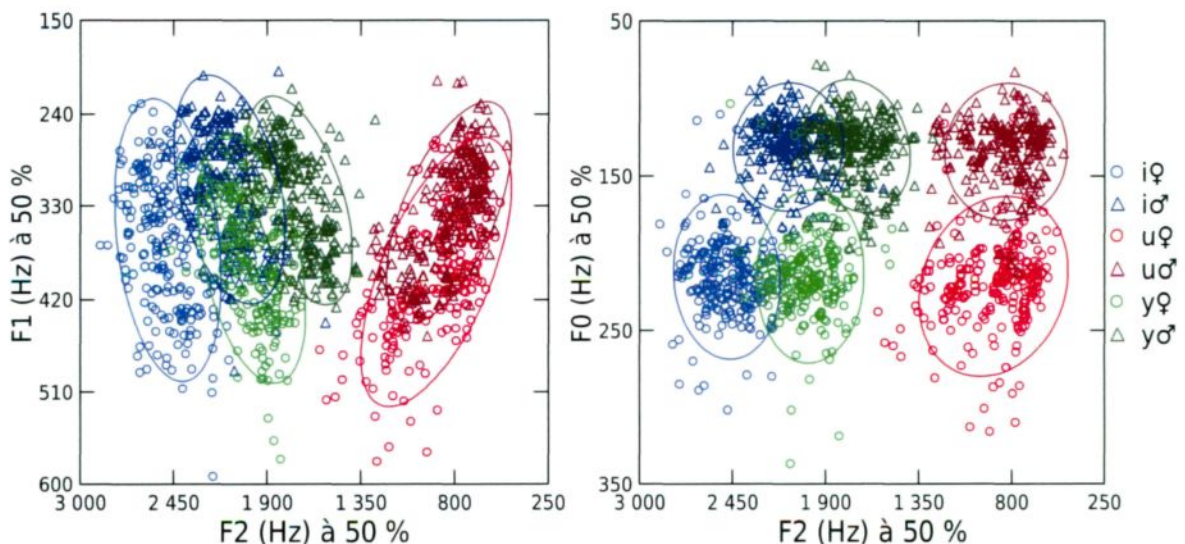


Figure 25 : Diagrammes de dispersion F_1 / F_2 et F_0 / F_2 (en Hertz) présentant les 1343 occurrences de notre corpus à 50 % de leur durée regroupées d'après leur classe phonémique et le sexe du locuteur. Les ellipses de dispersion englobent 86,5 % des occurrences d'une même classe phonémique, soit les occurrences situées à 2 écarts-types ou moins de la moyenne.

En comparant le Tableau 5 et le Tableau 7 qui suit, on constate que la prise en compte de la durée réduit principalement la confusion entre les variantes tendues et allongées d'un même phonème (les relâchées étant somme toute déjà bien distinguées quand la durée n'est pas considérée). Par exemple, 29 occurrences de (i#) sont identifiés comme étant des (iR) lorsque les seuls prédicteurs de l'analyse discriminante sont F_1 et F_2

F₂ d'origine sexuelle. La réduction de cette dispersion au moyen des scores-Z a donc un effet semblable à celui de l'ajout de la F₀ dans la distinction d'arrondissement.

Tableau 9 : Matrice de confusion de l'analyse discriminante ayant comme variable dépendante la classe vocalique et comme prédicteurs F₁ et F₂ (en scores-Z) relevés à 50 % de la durée

	i#	iK	iR	u#	uK	uR	y#	yK	yR	% correct
i#	45	2	23	0	0	0	5	0	0	60
iK	1	69	3	0	0	0	1	0	1	92
iR	17	2	56	0	0	0	0	0	0	75
u#	0	0	0	54	1	19	0	0	0	73
uK	0	0	0	0	74	0	0	0	0	100
uR	0	0	0	21	1	52	0	0	0	70
y#	0	0	0	0	0	0	57	3	15	76
yK	0	0	0	0	0	0	1	73	1	97
yR	1	1	0	0	0	0	37	5	31	41
Total	64	74	82	75	76	71	101	81	48	76

Jusqu'à présent la contribution de la dynamique spectrale à la discrimination acoustique des classes vocaliques a été évaluée en comparant les classifications obtenues lorsque les indices relevés en un seul point de mesure étaient inclus dans les prédicteurs à celles obtenues lorsque deux ou trois points de mesures étaient considérés. La prise en compte de deux points de mesures situés à 25 % et à 75 % de la durée des voyelles constitue une paramétrisation de l'hypothèse du VISC dite de la double cible. Cependant comme nous l'avons indiqué au Chapitre 2, nous avons également choisi de mettre à profit deux autres paramétrisations de la dynamique spectrale proposées dans la littérature : l'hypothèse de la cible et de la pente, et l'hypothèse de la cible et de l'orientation. Conformément à cet objectif, nous avons cherché à déterminer, au moyen d'analyses discriminantes, l'hypothèse du VISC qui permettait de rendre compte le plus efficacement de l'appartenance des occurrences de notre corpus à leur classe vocalique. Bref, nous avons testé l'efficacité des pentes formantiques et de l'orientation des formants, combinées aux fréquences formantiques initiales, comme prédicteurs de la classe vocalique. Les résultats rapportés dans le Tableau 10 indiquent que l'hypothèse de la cible et de l'orientation entraîne des taux de classification correcte systématiquement inférieurs à ceux obtenus avec les deux autres hypothèses. L'hypothèse de la cible et de l'orientation s'était d'ailleurs révélée légèrement moins efficace dans notre étude préliminaire (Arnaud et coll., 2011). L'hypothèse de la cible et de la pente semble pour sa part l'emporter sur l'hypothèse de la double cible lorsque la durée des voyelles n'est pas prise en compte, mais ces deux hypothèses parviennent à des taux de classification correcte sensiblement équivalents lorsque la durée est intégrée.

Tableau 10 : Taux de classification correcte globaux (en %), résultats des analyses discriminantes visant à classer les occurrences dans leur classe vocalique (i#), (iK), (iR), (y#), (yK), (yR), (u#), (uK), ou (uR) à partir des modèles du VISC

	25, 75 %		25 %, pente		25 %, orientation	
	Durée	Durée	Durée	Durée	Durée	Durée
F ₁ , F ₂ (Hz)	68	87	74	86	63	83
F ₀ , F ₁ , F ₂ (Hz)	76	87	78	88	73	86
F ₁ , F ₂ , F ₃ (Hz)	76	93	85	93	73	92
F ₀ , F ₁ , F ₂ , F ₃ (Hz)	80	91	84	92	77	91

En somme, les indices acoustiques que nous avons pris en considération (F₀, F₁, F₂, F₃ 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 une différenciation encore plus grande. Il n'est cependant pas nécessaire de combiner la totalité des ces indices et de ces approches pour parvenir, au moyen d'analyses discriminantes, à une classification optimale. Ainsi, les deux combinaisons de prédicteurs les plus efficaces à distinguer les classes vocaliques de notre corpus semblent être :

- F₁, F₂ et F₃ exprimés en scores-Z, relevés à 50 % de la durée des occurrences, et la durée vocalique ;
- F₁ et F₂ exprimés en scores-Z, relevés à 25 % et à 75 % de la durée des occurrences, et la durée vocalique.

Ces combinaisons de prédicteurs parviennent à un taux de classification correcte de 96 % en étant relativement économique. Il faut toutefois signaler que d'autres combinaisons, parfois plus économiques, parviennent à des taux du même ordre. Par exemple, on obtient un taux de classification de 94 % avec F₁ et F₂ exprimés en scores-Z, relevés à 50 % de la durée des occurrences, et la durée vocalique.

4.4 Synthèse récapitulative

Pour résumer, nos résultats suggèrent qu'en FQ, les voyelles fermées sont effectivement plus longues en (VR) qu'en (V#) et qu'en (VK). Le rapport de durée entre les variantes tendues et relâchées, attendues dans ces deux derniers contextes, semble quant à lui tributaire de l'environnement consonantique (plus précisément du voisement des consonnes adjacentes) plus que du timbre des voyelles. Ainsi, les relâchées sont plus longues que les tendues lorsqu'elles sont suivies d'une consonne voisée. Autrement, les tendues précédées d'une consonne voisée sont plus longues que les relâchées précédées d'une consonne voisée et suivies d'une non-voisée, et les tendues précédées d'une consonne non voisée sont plus longues que les relâchées précédées et suivies d'une consonne non voisée (les tendues précédées d'une consonne non voisée

présentant une durée équivalente à celle des relâchées précédées d'une consonne voisée et suivies d'une consonne non-voisée). Le sexe des locuteurs et la classe phonémique des occurrences n'ont sensiblement pas d'influence sur la durée des voyelles. Cette dernière varie cependant en fonction de l'origine géographique des locuteurs. Plus précisément, les voyelles fermées relâchées suivies d'une consonne voisée produites par les locuteurs de Saguenay sont significativement plus longues que celles produites par les locuteurs de Rouyn-Noranda et de Québec.

Nos résultats concernant les caractéristiques spectrales confirment que les /u/, postérieurs, ont un F_2 nettement plus bas que celui des voyelles antérieures /i/ et /y/. Le F_2 des /y/ est également plus bas que celui des /i/, mais tel qu'on pouvait s'y attendre, c'est principalement par leur F_3 plus élevé que les /i/, non arrondis, se différencient des voyelles arrondies /y/ et /u/. Au sein d'une même classe phonémique, les occurrences en (V#) paraissent effectivement être les plus tendues, avec un F_1 très bas. De même, les occurrences en (VK) semblent être les plus relâchées, avec un F_1 élevé. Le F_1 des occurrences en (VR) se situe quant à lui entre celui des tendues et des relâchées. En cours d'émission, les variantes tendues se déplacent légèrement vers la périphérie du trapèze vocalique, alors que les relâchées se centralisent. Les allongées présentent pour leur part les trajectoires les plus importantes, leur F_1 diminuant en cours d'émission, au point de se confondre en partie avec celui des tendues. À Saguenay, les voyelles fermées allongées (à l'exception peut-être des /u/) semblent plus tendues (F_1 plus bas) qu'à Rouyn-Noranda et à Québec.

Les résultats de nos analyses discriminantes indiquent que la seule prise en compte de F_1 et de F_2 relevés à 50 % de la durée des voyelles est suffisante pour distinguer les /u/ des /i/ et des /y/. La prise en compte supplémentaire de la F_0 ou de F_3 réduit pour sa part principalement la confusion entre les /i/ et les /y/. L'ajout de la durée des voyelles aux prédicteurs permet de réduire la confusion entre les variantes tendues et allongées (les relâchées étant déjà relativement bien distinguées). La normalisation des valeurs fréquentielles au moyen des scores-Z améliore systématiquement la classification des occurrences, quels que soient les indices acoustiques et le nombre de points de mesure considérés. Cela dit, plus les prédicteurs sont nombreux et moins cette amélioration est importante. La prise en compte de la dynamique spectrale améliore également la classification, mais cet avantage est moindre, voire disparaît, lorsque la durée des voyelles est également considérée. Enfin, parmi les trois hypothèses du VISC, celle de la cible et de la pente semble être la plus efficace à rendre compte de l'appartenance des occurrences de notre corpus à leur classe vocalique (notamment lorsque la durée des occurrences n'est pas prise en compte) et l'hypothèse de la cible et de l'orientation, la moins efficace.

Chapitre 5 : Discussion conclusive

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. Ainsi, 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. De même, une généralisation de nos résultats à l'ensemble des locuteurs de Rouyn-Noranda, de Saguenay et de Québec n'est pas immédiatement envisageable. Rappelons que notre échantillon n'est composé que de 30 locuteurs; ce qui, certes, constitue un effectif plus important que dans la plupart des études référencées au Chapitre 1, mais demeure modeste. Les productions de ces 30 locuteurs ne sont par ailleurs pas nécessairement représentatives de celles de l'ensemble des populations locales dont ils sont issus. Au mieux donnent-elles une indication sur quelques usages phonétiques de jeunes universitaires (qui ne correspondent d'ailleurs peut-être pas nécessairement au même groupe social d'une région à l'autre). Qui plus est, les caractéristiques de ces locuteurs auraient pu être contrôlées plus finement (origine géographique et occupation des parents, connaissance d'une langue seconde, orientation sexuelle, etc.). Il était cependant difficile d'être plus sélective en ce qui concerne le recrutement, d'une part parce que cette étude s'appuie aux deux tiers sur l'exploitation d'un corpus existant et d'autre part, en raison de certaines caractéristiques démographiques de l'Abitibi-Témiscamingue, notamment la taille relativement restreinte du bassin de population et la forte mobilité étudiante. Rappelons également que notre corpus a été récolté à la fois dans des conditions très formelles (situation expérimentale en chambre anéchoïque) 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.

Nos résultats appellent également un examen plus approfondi de l'effet de l'environnement consonantique sur les caractéristiques acoustiques des voyelles fermées du FQ. Nous avons effectivement effleuré la question en mettant au jour l'impact du voisement des consonnes adjacentes sur la durée des occurrences analysées. Cela dit, l'influence de nombreuses autres caractéristiques des consonnes adjacentes (lieu d'articulation, mode articulaire, labialisation, etc.) sur la durée des voyelles est attestée en FQ (voir notamment Ouellet, 1992 et Morasse, 1995). De plus, ces facteurs contextuels sont susceptibles d'avoir une incidence sur les fréquences formantiques des voyelles de diverses langues (Stevens et House, 1963; Stevens et coll., 1966; Pols, 1977; Van Bergem, 1994; Hillenbrand et coll., 2001; Strange et coll., 2007; Chládková et coll., 2011; Van Leussen et coll., 2011), dont le français (Chafcouloff et coll., 1989; Strange et coll., 2007). Le rôle de la position des voyelles dans le mot et dans l'énoncé (notamment par rapport à l'accent) sur nos résultats reste également à élucider, sachant que les voyelles de notre corpus sont toutes extraites de mots monosyllabiques répétés isolément.

Cela dit, en dépit de l'impact attendu des diverses consonnes adjacentes sur les caractéristiques spectrales des voyelles fermées de notre corpus, nous remarquons que les variantes tendues, relâchées et allongées se distinguent par leurs trajectoires à l'intérieur du trapèze vocalique. Il semble donc que la dynamique spectrale permettant de distinguer ces variantes leur soit propre et ne dépende pas strictement 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. À cet effet, l'étude des caractéristiques acoustiques des exceptions à la distribution des variantes tendues, relâchées et allongées des voyelles fermées du FQ (*quiz, beat, cool, etc.*) serait fort intéressante, puisqu'elle permettrait de départager en partie l'effet du contexte de celui du timbre. Notons par ailleurs que nos résultats confirment la pertinence de la prise en compte de la dynamique spectrale même dans la description de voyelles brèves n'étant jamais rapportées comme diphtonguées, en l'occurrence les voyelles fermées tendues et relâchées en FQ.

En ce qui concerne nos analyses acoustiques, elles reposent sur la prise en compte de la durée des voyelles ainsi que de la F_0 et de F_1 , F_2 et F_3 (en Hertz et normalisés en scores-Z) relevés à 25 %, 50 % et 75 % de cette durée. Or, comme on l'a vu au Chapitre 2, d'autres indices acoustiques auraient pu être exploités (F_4 , largeur de bande des formants, amplitude relative des formants, etc.) en vue d'obtenir une description plus riche des voyelles fermées en FQ. En particulier, la prise en compte de la fréquence centrale du quatrième formant, réputée participer à la distinction entre les /i/ et les /y/ (Schwartz et coll., 1993 ; Gendrot et coll., 2008), aurait sans doute été pertinente. Faute de temps, nous avons également choisi de ne pas mettre à profit de nombreuses procédures de normalisation (Sigouin, à paraître, voir Annexe 2) et paramétrisations de la dynamique spectrale. Mentionnons au passage qu'il existe d'autres représentations du signal sonore de la parole que le spectrogramme (et le codage linéaire prédictif pour la détection des formants). Par exemple, les coefficients cepstraux (*mel-frequency cepstral coefficients*, ou MFCC) sont couramment utilisés en reconnaissance de la parole depuis Davis et Mermelstein (1980). Le cochléagramme, une représentation basée sur la sensibilité tonotopique de la cochlée, a également été proposé par Lyon (1982). Il va de soi que la prise en considération de ces diverses techniques pourrait faire l'objet d'une étude à part entière.

Sur le plan statistique, l'établissement d'un seuil de signification comporte toujours une part d'arbitraire. Nous accueillons donc avec réserve certains résultats statistiquement significatifs, notamment ceux des ANOVAs portant sur les caractéristiques spectrales normalisées, qui nous semblent plus optimistes que de raison dans certains cas. D'autre part, vu la complexité de notre échantillon, nous n'avons pas vérifié, au moyen d'ANOVAs, si les composantes spectrales différaient significativement d'une classe vocalique à l'autre. Par

exemple, nous ne sommes pas en mesure de déterminer si les voyelles fermées tendues ont une F_0 plus élevée que les relâchées, comme le suggère Lavoie (1995 : 72). Afin de vérifier les caractéristiques les plus distinctives des classes vocaliques à l'étude, nous avons plutôt opté pour des analyses discriminantes. À ce propos, si notre échantillon s'y était prêté (si sa taille avait été suffisante), il aurait été intéressant de s'attarder aux différences régionales au moyen de ces tests. Il est en effet possible, en choisissant l'origine géographique du locuteur comme variable dépendante, de déterminer les indices acoustiques présentant le plus de variation diatopique, donc ceux qui s'avèrent les meilleurs prédicteurs de l'origine géographique (Adank, 2003; Adank, Smits et Van Hout, 2004). Nous aurions ainsi pu vérifier si la normalisation des données en scores-Z entraînait une perte d'information pertinente d'un point de vue sociolinguistique. Une alternative, exploitée entre autres par Jacewicz et Fox (2012), consiste à conserver la classe vocalique comme variable dépendante mais à calculer le discriminant sur une portion de l'échantillon d'une région en particulier, puis à l'appliquer à la portion restante de cet échantillon et aux échantillons d'autres régions. En comparant les matrices de confusion obtenues, on peut alors déterminer les classes vocaliques présentant les différences les plus marquées d'une région à l'autre, soient celles dont les occurrences sont le moins bien identifiées lorsque le discriminant est calculé sur l'échantillon provenant d'une tierce région.

Pour conclure, rappelons que l'objectif de la présente étude était de caractériser acoustiquement les variantes tendues, relâchées et allongées des voyelles fermées /i/, /y/ et /u/ en français québécois (FQ), qui, d'après la littérature, apparaissent respectivement en syllabe ouverte (V#), en syllabe fermée par une consonne non allongeante (VK) et en syllabe fermée par une consonne allongeante (VR) lorsqu'elles sont accentuées. Nos analyses ont ainsi fait état de différents indices, disponibles dans le signal sonore, permettant de différencier les classes vocaliques établies à partir de ces contextes et confirment la distribution complémentaire parfaite des variantes tendues, relâchées et allongées. Cependant, nous ne saurions trop insister sur le fait que ces indices ne sont pas nécessairement ceux sur lesquels repose effectivement la perception de ces voyelles. Par exemple, une procédure de normalisation extrinsèque comme les scores-Z n'explique certainement pas comment des auditeurs parviennent à correctement identifier des voyelles produites par des locuteurs qu'ils n'avaient jamais entendus auparavant. Cette mise en garde étant faite, sommes-nous parvenue à répondre à nos trois sous-questions de recherche ?

Les voyelles fermées allongées se caractérisent-elles par une importante variation spectrale (pouvant éventuellement être interprétée comme une diphtongaison) ou présentent-elles plutôt des caractéristiques acoustiques se comparant à celles des variantes tendues et relâchées ?

Nos résultats indiquent que les voyelles fermées allongées ont un F_1 plus élevé que celui des tendues, mais plus bas que celui des relâchées. En cours d'émission, le F_1 des allongées diminue, au point de rejoindre en

partie celui des tendues. Les allongées présentent d'ailleurs les trajectoires les plus importantes dans un diagramme F_1 / F_2 . Les relâchées, contrairement aux allongées et aux tendues, se centralisent en cours d'émission. Bref, les allongées s'apparentent plus aux tendues qu'aux relâchées, tout en se distinguant par un F_1 plus élevé et une trajectoire plus importante. Il n'est pas du ressort de la présente étude de déterminer si cette trajectoire peut être interprétée ou non comme une diphtongaison, mais elle ouvre la porte à une recherche ultérieure, dans la lignée de celle de Leblanc (2012), qui ferait le lien entre les caractéristiques acoustiques des voyelles fermées allongées et la perception de leur timbre.

Quel rapport de durée entretiennent les variantes brèves tendues et relâchées des voyelles fermées ?

Le voisement des consonnes adjacentes, en particulier de la consonne postposée, a vraisemblablement un effet allongeant sur la durée des voyelles fermées tendues et relâchées. À effet allongeant égal, les tendues seraient légèrement plus longues que les relâchées. Cependant, si on considère que les consonnes non voisées ont un effet abrégé sur la durée de ces voyelles, il devient impossible de déterminer avec certitude le rapport de durée entre les tendues, présentes en syllabe ouverte, et les relâchées, présentes en syllabe fermée, puisque ces variantes n'apparaissent jamais dans des contextes consonantique strictement équivalents quant au voisement.

Les voyelles fermées tendues, relâchées et allongées sont-elles marquées d'une variation diatopique en FQ ?

Deux principales différences régionales ont été mises au jour dans nos données. La première concerne la durée des voyelles fermées relâchées, qui, lorsqu'elles sont suivies d'une consonne voisée, sont plus longues à Saguenay qu'à Québec et qu'à Rouyn-Noranda. Les relâchées produites par les locuteurs saguenéens semblent donc plus sensibles à l'effet allongeant de la consonne voisée postposée. La seconde porte sur les caractéristiques spectrales des voyelles fermées allongées. Le F_1 de ces variantes est plus bas, en moyenne, à Saguenay qu'à Québec et qu'à Rouyn-Noranda. Cela dit, presque aucune différence significative n'émerge entre les voyelles produites par les locuteurs de Québec et celles produites par les locuteurs de Rouyn-Noranda. Cependant, il serait pour le moins hasardeux d'établir un rapprochement global entre les pratiques phonétiques en usage à Québec (considérées comme caractéristiques de celles d'un parler de l'Est du Québec) et celles en usage à Rouyn-Noranda (une zone de peuplement récent mixte) sur la seule base de cet aspect phonétique et sans autres points de comparaison au Québec. En terminant, soulignons que bien que le phonétisme du français québécois tende à s'uniformiser (Dolbec et Ouellon, 1999 : 26), nous sommes effectivement parvenue, en dépit du jeune âge de nos locuteurs et malgré les conditions très formelles dans lesquelles notre corpus a été récolté, à identifier quelques usages phonétiques ne se retrouvant pas dans toutes les régions où nous avons enquêté. Il n'est donc peut-être pas trop tard pour qu'advienne cette « grande enquête sociophonétique », fondée sur des « données primaires récentes et comparables » et visant

à « donner un portrait plus précis et plus fidèle de la réalité actuelle des prononciations sur l'ensemble du territoire québécois » qu'appelaient de leurs vœux Dolbec et Ouelton (1999 : 27).

Bibliographie

- Adank, Patti (2003), *Vowel Normalization: A Perceptual-Acoustic Study of Dutch Vowels*, Wageningen, Ponsen & Looijen, 227 p.
- Adank, Patti, Roeland van Hout et Roel Smits (2004), « An acoustic description of the vowels of Northern and Southern Standard Dutch », *Journal of the Acoustical Society of America*, vol. 116, n° 3, p. 1729-1738.
- Adank, Patti, Roel Smits et Roeland van Hout (2004), « A comparison of vowel normalization procedures for language variation research », *Journal of the Acoustical Society of America*, vol. 116, n° 5, p. 3099-3107.
- Ainsworth, William A. (1971), « Perception of synthesized isolated vowels and h_d words as a function of fundamental frequency », *Journal of the Acoustical Society of America*, vol. 49, n° 4, part. 2, p. 1323-1324.
- (1975), « Intrinsic and extrinsic factors in vowel judgements » dans Fant, Gunnar et Mark A. A. Tatham (dir.), *Auditory Analysis and Perception of Speech*, Londres, Academic Press, p. 103-113.
- Anderson, Bridget L. (1999), « Source-language transfer and vowel accommodation in the patterning of Cherokee English /ai/ and /oi/ », *American Speech*, vol. 74, n° 4, p. 339-368.
- Andruski, Jean E. et Terrance M. Nearey (1992), « On the sufficiency of compound target specification of isolated vowels and vowels in /bVb/ syllables », *Journal of the Acoustical Society of America*, vol. 91, n° 1, p. 390-410.
- Arnaud, Vincent, Caroline Sigouin et Johanna-Pascale Roy (2011), « Acoustic description of Quebec French high vowels: First results », *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, p. 244-247.
- Assmann, Peter F. (1991), « The perception of back vowels: Centre of gravity hypothesis », *The Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, vol. 43, n° 3, p. 423-448.
- Assmann, Peter F., Terrance M. Nearey et John T. Hogan (1982), « Vowel identification: Orthographic, perceptual, and acoustic aspects », *Journal of the Acoustical Society of America*, vol. 71, n° 4, p. 975-989.
- Beddor, Patrice Speeter et Sarah Hawkins (1991), « The influence of spectral prominence on perceived vowel quality », *Haskins Laboratories Status Report on Speech Research*, vol. 105/106, p. 187-214.
- Bell, Alexander Graham (1879), « Vowel theories », *The American Journal of Otology*, vol. 1, p. 163-180.
- Van Bergem, Dick R. (1994), « A model of coarticulatory effects on the schwa », *Speech Communication*, vol. 14, n° 2, p. 143-162.
- Bevier, Louis (1900), « The acoustic analysis of the vowels from the phonographic record », *Physical Review*, vol. 4, n° 10, p. 193-201.
- Bladon, R. Anthony W. et Gunnar Fant (1978), « A two-formant model and the cardinal vowels », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 19, n° 1, p. 1-8.
- Boersma, Paul et David Weenink (2013), *Praat: Doing Phonetics by Computer*, version 5.3.x. [Logiciel] Disponible au <http://www.praat.org>
- Boissonneault, Chantal (1999), *Le français de l'Abitibi : caractéristiques phonétiques et origine socio-géographique des locuteurs*, mémoire de maîtrise, Université Laval, 128 p.

- Boite, René, Hervé Bourlard, Thierry Dutoit, Joël Hancq et Henri Leich (2000), *Traitement de la parole*, collection Électricité, Lausanne, Presses polytechniques et universitaires romandes, 488 p.
- Boulanger, Aline (1986), « Les parlers en [IZ] », *Revue québécoise de linguistique théorique et appliquée*, vol. 5, n° 4, p. 129-142.
- Brent, Edmund (1971), *Canadian French: A Synthesis*, thèse de doctorat, Cornell University, 163 p.
- Calamai, Silvia (2005), « Intrinsic vowel normalization: Comparing different procedures (data from Tuscan Italian) », *Italian Journal of Linguistics / Rivista di Linguistica*, vol. 17, n° 2, p. 211-270.
- Calliope (1989), *La parole et son traitement automatique*, Paris, Masson, 718 p.
- Carlson, Rolf, Björn Granström et Gunnar Fant (1970), « Some studies concerning perception of isolated vowels », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 11, n° 2-3, p. 19-35.
- Carlson, Rolf, Björn Granström et Dennis H. Klatt (1979), « Vowel perception: The relative perceptual salience of selected acoustic manipulations », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 20, n° 3-4, p. 73-83.
- Cedergren, Henrietta J. et Louise Simoneau (1985), « La chute des voyelles hautes en français de Montréal : "As-tu entendu la belle syncope ?" » dans *Les tendances dynamiques du français parlé à Montréal*, tome 1, collection Langue et société, Québec, Gouvernement du Québec, Office de la langue française, p. 57-144.
- Chafcouloff, Michel, Alain Marchal et Thami Benkirane (1989), « Coarticulatory patterns in stop sequences », *Proceedings of the First European Conference on Speech Communication and Technology, Eurospeech 1989*, Paris, vol. 2, p. 304-307.
- Chistovich, Ludmilla A. et Valentina V. Lublinskaya (1979), « The "center of gravity" effect in vowel spectra and critical distance between the formants: Psychoacoustical study of the perception of vowel-like stimuli », *Hearing Research*, vol. 1, n° 3, p. 185-195.
- Chládková, Kateřina, Paola Escudero et Paul Boersma (2011), « Context-specific acoustic differences between Peruvian and Iberian Spanish vowels », *Journal of the Acoustical Society of America*, vol. 130, n° 1, p. 416-428.
- Clopper, Cynthia G. (2009), « Computational methods for normalizing acoustic vowel data for talker differences », *Language and Linguistics Compass*, vol. 3, n° 6, p. 1430-1442.
- Cohen, Antonie (1971), « Diphthongs, mainly Dutch » dans Hammerich, Louis Leonor, Roman Jakobson et Eberhard Zwirner (dir.), *Form and Substance: Phonetic and Linguistic Papers Presented to Eli Fischer-Jørgensen*, Odense, Akademisk Forlag, p. 277-289.
- Côté, Marie-Hélène (2010), « La longueur vocalique devant consonne allongeante en contexte final et dérivé en français laurentien » dans LeBlanc, Carmen, France Martineau et Yves Frenette (dir.), *Vues sur les français d'ici*, collection Les voies du français, Québec, Les Presses de l'Université Laval, p. 49-75.
- Di Cristo, Albert (1985), *De la microprosodie à l'intonosyntaxe*, Aix-en-Provence, Publications de l'Université d'Aix, 850 p.
- (1998), « Intonation in French » dans Hirst, Daniel et Albert Di Cristo (dir.), *Intonation Systems: A Survey of Twenty Languages*, Cambridge, Cambridge University Press, p. 195-218.
- Dalton, Will (2011a), « An ultrasound imaging study of the tense-lax distinction in Canadian French vowels », *Acoustique canadienne*, vol. 39, n° 3, p. 168-169.

- (2011b), « Articulation of tense and lax vowels in Canadian French: Findings from ultrasound imaging », communication orale présentée le 22 juin 2011 au colloque *Phonetics and Phonology in Iberia 2011*, Tarragona. Résumé consulté le 10 juin 2013 au http://wwwa.urv.cat/deaa/PaPI2011/scientific_program.html.
- Davis, Steven B. et Paul Mermelstein (1980), « Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences », *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, n° 4, p. 357-366.
- Dawn Armstrong, Susan (1999), *Stress and weight in Québec French*, mémoire de maîtrise, University of Calgary, 139 p.
- Déchaine, Rose-Marie (1990), *Stress and weight gain*, inédit, Amherst, University of Massachusetts.
- Delattre, Pierre (1948), « Un triangle acoustique des voyelles orales du français », *The French Review*, vol. 21, n° 6, p. 477-484.
- (1953), « Les modes phonétiques du français », *The French Review*, vol. 27, n° 1, p. 59-63.
- (1954), « Les attributs acoustiques de la nasalité vocalique et consonantique », *Studia Linguistica*, vol. 8, n° 1-2, p. 103-109.
- (1963), « Voyelles diphtonguées et voyelles pures », *The French Review*, vol. 37, n° 1, part. 1, p. 64-76.
- Delattre, Pierre, Alvin M. Liberman, Franklin S. Cooper et Louis J. Gerstman (1952), « An experimental study of the acoustic determinants of vowel color: Observations on one- and two-formant vowels synthesized from spectrographic patterns », *Word*, vol. 8, n° 3, p. 195-210.
- Dionne, Marie-Cécile (1971), *Analyse acoustique comparée des voyelles orales françaises et québécoises*, mémoire de maîtrise, Université de Montréal, 218 p.
- Dolbec, Jean et Conrad Ouellon (1999), « Peut-on distinguer des variétés phonétiques en français québécois ? », *Dialangue*, vol. 10, p. 17-28.
- Dolbec, Jean et Claude Paradis (1998), *PHONO, Principales caractéristiques phonétiques du français québécois*. [Site internet] Consulté le 26 mars 2013 sur <http://phono.uqac.ca>
- Dolbec, Jean et Susan Rogers (1996), « Caractéristiques microprosodiques de durée et d'intensité en lecture et en conversation semi-dirigée » dans Dolbec, Jean et Marise Ouellet (dir.), *Recherches en phonétique et en phonologie au Québec*, Québec, Centre international de recherche en aménagement linguistique, Université Laval, p. 18-35.
- Dulong, Gaston et Gaston Bergeron (1980), *Le parler populaire du Québec et de ses régions voisines. Atlas linguistique de l'est du Canada*, Québec, Office de la Langue Française, Ministère des Communications, documentation québécoise, 10 vol.
- Dumas, Denis (1974), « Durée vocalique et diphtongaison en français québécois », *Cahier de linguistique*, vol. 4, p. 13-55.
- (1978), *Phonologie des réductions vocaliques en français québécois*, thèse de doctorat, Université de Montréal, 183 p.
- (1987), *Nos façons de parler : les prononciations en français québécois*, Sillery, Presses de l'Université du Québec, 155 p.
- Fabricius, Anne H., Dominic Watt et Erza Johnson (2009), « A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics », *Language Variation and Change*, vol. 21, n° 3, p. 413-435.

- Fant, Gunnar (1966), « A note on vocal tract size factors and non-uniform F-pattern scalings », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 7, n° 4, p. 22-30.
- (1969), « Distinctive features and phonetic dimensions », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 10, n° 2-3, p. 1-18.
- (1975), « Non-uniform vowel normalization », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 16, n° 2-3, p. 1-19.
- (1978), « Vowel perception and specification », *Rivista italiana di acustica*, vol. 2, p. 69-87.
- (1983), « Feature analysis of Swedish vowels - a revisit », *Speech Transmission Laboratory Quarterly Progress and Status Report*, vol. 24, n° 2-3, p. 1-19.
- (2007), *Speech Acoustics and Phonetics: Selected Writings*, Dordrecht, Kluwer Academic Publishers, 320 p.
- Fast, Anicka (2002), *Evolving Linguistic Standards in Quebec: A Sociolinguistic Analysis of Production, Perception and Evaluation*, mémoire inédit de premier cycle, Université McGill, 104 p.
- Ferguson, Sarah Hargus et Diane Kewley-Port (2002), « Vowel intelligibility in clear and conversational speech for normal-hearing and hearing-impaired listeners », *Journal of the Acoustical Society of America*, vol. 112, n° 1, p. 259-271.
- Ferrari Disner, Sandra (1980), « Evaluation of vowel normalization procedures », *Journal of the Acoustical Society of America*, vol. 67, n° 1, p. 253-261.
- Fischer-Jørgensen, Eli (1990), « Intrinsic F₀ in tense and lax vowels with special reference to German », *Phonetica*, vol. 47, p. 99-140.
- Flynn, Nicholas et Paul Foulkes (2011), « Comparing vowel formant normalization methods », *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, p. 683-686.
- Fox, Robert Allen et Ewa Jacewicz (2009), « Cross-dialectal variation in formant dynamics of American English vowels », *Journal of the Acoustical Society of America*, vol. 126, n° 5, p. 2603-2618.
- Fujisaki, Hiroya et Takako Kawashima (1968), « The roles of pitch and higher formants in the perception of vowels », *IEEE Transactions on Audio and Electroacoustics*, vol. 16, n° 1, p. 73-77.
- Gay, Thomas (1968), « Effect of speaking rate on diphthong formant movements », *Journal of the Acoustical Society of America*, vol. 44, n° 6, p. 1570-1573.
- Gendron, Jean-Denis (1959), « Désonorisation des voyelles en franco-canadien », *Journal of the Canadian Linguistics Association*, vol. 5, n° 2, p. 99-108.
- (1966), *Tendances phonétiques du français parlé au Canada*, Québec, Paris, Les Presses de l'Université Laval, Klincksieck, 254 p.
- Gendrot, Cédric, Martine Adda-Decker et Jacqueline Vaissière (2008), « Les voyelles /i/ et /y/ du français : focalisation et variations formantiques », *Actes des XXVII^{es} Journées d'étude sur la parole*, Avignon, p. 205-208.
- Gerstman, Louis J. (1968), « Classification of self-normalized vowels », *IEEE Transactions on Audio and Electroacoustics*, vol. AU-16, n° 1, p. 78-80.
- Gottfried, Michael, James D. Miller et Donald J. Meyer (1993), « Three approaches to the classification of American English diphthongs », *Journal of Phonetics*, vol. 21, p. 205-229.
- Hannahs, Stephen J. (1989), *High vowel variation in Quebec French*, version révisée inédite d'une communication présentée à l'Annual Meeting of the Linguistic Society of America (Washington, 1989), Newark, University of Delaware.

- Hellweg, Christoph Friedrich (1781), *De Formatione Loquelae*, thèse de doctorat, Eberhard Karls Universität Tübingen, 37 p.
- Hellweg, Christoph Friedrich (1886), *De Formatione Loquelae*, Heilbronn, Verlag von Gebrüder Henninger, 60 p.
- Von Helmholtz, Hermann Ludwig Ferdinand (1863), *Die Lehre von den Tonempfindungen als physiologische Grundlage für die Theorie der Musik*, Brunswick, Friedrich Vieweg und Sohn, 600 p.
- (1990), *Théorie physiologique de la musique fondée sur l'étude des sensations auditives*, Paris, Gabay, 544 p.
- Heselwood, Barry (2009), « Rhoticity without F3: Lowpass filtering, F1-F2 relations and the perception of rhoticity in NORTH-FORCE, START and NURSE words », *Leeds Working Papers in Linguistics and Phonetics*, vol. 14, p. 49-64.
- Hillenbrand, James M. (2004), « Effects of spectral tilt on the perception of naturally spoken vowels », *Journal of the Acoustical Society of America*, vol. 116, n° 4, p. 2546-2547.
- (2013), « Static and dynamic approaches to vowel perception » dans Morrison, Geoffrey Stewart et Peter F. Assmann (dir.), *Vowel Inherent Spectral Change*, Berlin, Heidelberg, Springer-Verlag, p. 9-30.
- Hillenbrand, James M., Michael J. Clark et Terrance M. Nearey (2001), « Effects of consonant environment on vowel formant patterns », *Journal of the Acoustical Society of America*, vol. 109, n° 2, p. 748-763.
- Hillenbrand, James M. et Robert T. Gayvert (1993), « Vowel classification based on fundamental frequency and formant frequencies », *Journal of Speech and Hearing Research*, vol. 36, n° 4, p. 694-700.
- Hillenbrand, James M. et Terrance M. Nearey (1999), « Identification of resynthesized /hVd/ utterances: Effects of formant contour », *Journal of the Acoustical Society of America*, vol. 105, n° 6, p. 3509-3523.
- Hindle, Donald (1978), « Approaches to vowel normalization in the study of natural speech » dans Sankoff, David (dir.), *Linguistic Variation: Models and Methods*, New York, San Francisco et Londres, Academic Press, p. 161-171.
- Hirahara, Tatsuya et Hiroaki Kato (1992), « The effect of F0 on vowel identification » dans Tokhura, Yoh'ichi, E. Vatikiotis-Bateson et Y. Sagisaka (dir.), *Speech Perception, Production and Linguistic Structure*, Tokyo, Ohmsha Ltd., p. 89-112.
- House, Arthur S. et Kenneth N. Stevens (1956), « Analog studies of the nasalization of vowels », *Journal of Speech and Hearing Disorders*, vol. 21, n° 2, p. 218-232.
- Iri, Masao (1959), « A mathematical method in phonetics with a special reference to the acoustical structure of Japanese vowels », *Gengo Kenkyu*, n° 35, p. 23-30.
- Ito, Masashi, Jun Tsuchida et Masafumi Yano (2001), « On the effectiveness of whole spectral shape for vowel perception », *Journal of the Acoustical Society of America*, vol. 110, n° 2, p. 1141-1149.
- Jacewicz, Ewa, Robert A. Fox et Joseph Salmons (2011), « Vowel change across three age groups of speakers in three regional varieties of American English », *Journal of Phonetics*, vol. 39, n° 4, p. 683-693.
- Jacewicz, Ewa et Robert Allen Fox (2012), « The effects of cross-generational and cross-dialectal variation on vowel identification and classification », *Journal of the Acoustical Society of America*, vol. 131, n° 2, p. 1413-1433.
- (2013), « Cross-dialectal differences in dynamic formant patterns in American English vowels » dans Morrison, Geoffrey Stewart et Peter F. Assmann (dir.), *Vowel Inherent Spectral Change*, Berlin, Heidelberg, Springer-Verlag, p. 177-198.

- Jenkins, James J. et Winifred Strange (1987), « Identification of "hybrid" vowels in sentence context », *Journal of the Acoustical Society of America*, vol. 82, n° suppl. 1, p. S82.
- (1999), « Perception of dynamic information for vowels in syllable onsets and offsets », *Perception & Psychophysics*, vol. 61, n° 6, p. 1200-1210.
- Johnson, Keith (1990), « Contrast and normalization in vowel perception », *Journal of Phonetics*, vol. 18, p. 229-254.
- (2005), « Speaker normalization in speech perception » dans Pisoni, David B. et Robert E. Remez (dir.), *The Handbook of Speech Perception*, Oxford, Blackwell Publishing, p. 363-389.
- Joos, Martin (1948), « Acoustic phonetics », *Language*, vol. 24, n° 2, Language monograph No. 23, p. 5-131+133-136.
- Jung, Heinrich (1926), « Die neueren Vokaltheorien », *Physikalische Zeitschrift*, vol. 27, p. 716-723.
- Kale, Sushrut Suresh (2006), *Effect of Variations in Global and Local Spectral Properties of a Steady State Vowel Spectrum on Speech Perception by Cochlear Implant Users*, mémoire de maîtrise, University of Southern California, 37 p.
- Kieffe, Michael et Keith R. Kluender (2005), « The relative importance of spectral tilt in monophthongs and diphthongs », *Journal of the Acoustical Society of America*, vol. 117, n° 3, part. 1, p. 1395-1404.
- Klatt, Dennis H. (1976), « Linguistic uses of segmental duration in English: Acoustic and perceptual evidence », *Journal of the Acoustical Society of America*, vol. 59, n° 5, p. 1208-1221.
- (1982), « Prediction of perceived phonetic distance from critical-band spectra: A first step », *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, p. 1278-1281.
- Koenig, Winfried (1949), « A new frequency scale for acoustic measurements », *Bell Labs Record*, n° 27, p. 299-301.
- Labov, William (2006), « A sociolinguistic perspective on sociophonetic research », *Journal of Phonetics*, vol. 34, n° 4, p. 500-515.
- Labov, William, Sharon Ash et Charles Boberg (2006), *The Atlas of North American English: Phonetics, Phonology and Sound Change*, Berlin, Mouton de Gruyter, 318 p.
- Labov, William, Malcah Yaeger et Richard Steiner (1972), *A Quantitative Study of Sound Change in Progress*, Report on National Science Foundation, Contract NSF-GS-3287, University of Pennsylvania, 337 p.
- Ladefoged, Peter (2003), *Phonetic Data Analysis: An Introduction to Fieldwork and Instrumental Techniques*, Malden, Blackwell Publishing, 196 p.
- Ladefoged, Peter et Ian Maddieson (1996), *The Sounds of the World's Languages*, Oxford, Blackwell Publishers Ltd, 425 p.
- Lavoie, Julie (1995), *La fréquence intrinsèque des voyelles en français québécois*, mémoire de maîtrise, Université Laval, 128 p.
- Leblanc, Guillaume (2012), *Étude des voyelles orales susceptibles d'être diphtonguées en français québécois*, mémoire de maîtrise, Université Laval, 134 p.
- Van Leussen, Jan-Willem, Daniel Williams et Paola Escudero (2011), « Acoustic properties of dutch steady-state vowels: Contextual effects and a comparison with previous studies », *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, p. 1194-1197.
- Lindblom, Björn E. F. et Michael Studdert-Kennedy (1967), « On the role of formant transitions in vowel recognition », *Journal of the Acoustical Society of America*, vol. 42, n° 4, p. 830-843.

- Lindblom, Björn E. F. et Johan E. F. Sundberg (1971), « Acoustical consequences of lip, tongue, jaw, and larynx movement », *Journal of the Acoustical Society of America*, vol. 50, n° 4, part. 2, p. 1166-1179.
- Lloyd, Richard J. (1890), *Some Researches into the Nature of Vowel-Sound*, Liverpool, Turner and Dunnett, 45 p.
- Lobanov, Boris M. (1971), « Classification of Russian vowels spoken by different speakers », *Journal of the Acoustical Society of America*, vol. 49, n° 2B, p. 606-608.
- Lyon, Richard F. (1982), « A computational model of filtering, detection, and compression in the cochlea », *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, vol. 7, p. 1282-1285.
- MacKenzie, Laurel et Gillian Sankoff (2010), « A quantitative analysis of diphthongization in Montreal French », *University of Pennsylvania Working Papers in Linguistics: Selected Papers from NAWAV 37*, vol. 15, n° 2, p. 91-100.
- Maddieson, Ian (1999), « Phonetic universals » dans Hardcastle, William J. et John Laver (dir.), *The Handbook of Phonetic Sciences*, Blackwell Publishing, Blackwell Reference Online.
- Malmberg, Bertil (1968), *La phonétique*, 7^e édition, collection Que sais-je ?, Paris, Presses universitaires de France, 128 p.
- Mantakas, Marios, Jean-Luc Schwartz et Pierre Escudier (1986), « Modèle de prédiction du "deuxième formant effectif" F'2 et application à l'étude de la labialité des voyelles du français », *Actes des 15^{es} Journées d'études sur la parole*, Aix-en-Provence, p. 157-161.
- Marchal, Alain (1980), *Les sons et la parole*, Montréal, Guérin, 182 p.
- Martin, Pierre (1996), *Éléments de phonétique avec application au français*, Québec, Les Presses de l'Université Laval, 253 p.
- (2002), « Le système vocalique du français du Québec. De l'acoustique à la phonologie », *La linguistique*, vol. 38, n° 2, p. 71-88.
- Mattingly, Ignatius G. (1999), « A short history of acoustic phonetics in the U.S. », *Proceedings of the 14th International Congress of Phonetic Sciences*, San Francisco, p. 1-6.
- McLaughlin, Anne (1986), « Une (autre) analyse de la distribution des variantes des voyelles hautes en français montréalais », *Revue québécoise de linguistique théorique et appliquée*, vol. 5, n° 4, p. 21-60.
- Ménard, Lucie (2002), *Production et perception des voyelles au cours de la croissance du conduit vocal : variabilité, invariance et normalisation*, thèse de doctorat, Institut de la communication parlée, Grenoble, 483 p.
- Ménard, Lucie, Jean-Luc Schwartz, Louis-Jean Boë, Sonia Kandel et Nathalie Vallée (2002), « Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood », *Journal of the Acoustical Society of America*, vol. 111, n° 4, p. 1892-1905.
- Miller, Dayton Clarence (1916), *The Science of Musical Sounds*, New York, The MacMillan Company, 286 p.
- Miller, James D. (1989), « Auditory-perceptual interpretation of the vowel », *Journal of the Acoustical Society of America*, vol. 85, n° 5, p. 2114-2134.
- Miller, James D., A. Maynard Engebretson et N. Rao Vemula (1980), « Vowel normalization: Differences between vowels spoken by children, women, and men », *Journal of the Acoustical Society of America*, vol. 68, n° suppl. 1, p. S33.
- Miller, Roger L. (1953), « Auditory tests with synthetic vowels », *Journal of the Acoustical Society of America*, vol. 25, n° 1, p. 114-121.

- Monahan, Philip J. et William J. Idsardi (2010), « Auditory sensitivity to formant ratios: Toward an account of vowel normalization », *Language and Cognitive Processes*, vol. 25, n° 6, p. 808-839.
- Moore, Brian C. J. et Brian R. Glasberg (1983), « Suggested formulae for calculating auditory-filter bandwidths and excitation patterns », *Journal of the Acoustical Society of America*, vol. 74, n° 3, p. 750-753.
- Morasse, Hélène (1995), *Variations intrinsèques et co-intrinsèques de durée vocalique en français québécois*, mémoire de maîtrise, Université du Québec à Chicoutimi, 141 p.
- Morrison, Geoffrey Stewart (2013), « Theories of vowel inherent spectral change » dans Morrison, Geoffrey Stewart et Peter F. Assmann (dir.), *Vowel Inherent Spectral Change*, série Modern Acoustics and Signal Processing, Berlin, Heidelberg, Springer-Verlag, p. 31-47.
- Morrison, Geoffrey Stewart et Terrance M. Nearey (2007), « Testing theories of vowel inherent spectral change », *Journal of the Acoustical Society of America*, vol. 122, n° 1, p. EL15-EL22.
- Mullennix, John W. et David B. Pisoni (1988), « Detailing the nature of talker normalization in speech perception », *Research on Speech Perception*, Progress Report n° 14, Bloomington, Speech Research Laboratory, Department of Psychology, Indiana University, p. 289-305.
- Murphy, Walter D., Sandra L. Shea et Richard N. Aslin (1989), « Identification of vowels in "vowelless" syllables by 3-year-olds », *Perception & Psychophysics*, vol. 46, n° 4, p. 375-383.
- Nearey, Terrance M. (1977), *Phonetic Feature Systems for Vowels*, thèse de doctorat, University of Connecticut, 200 p.
- Nearey, Terrance M. et Peter F. Assmann (1986), « Modeling the role of inherent spectral change in vowel identification », *Journal of the Acoustical Society of America*, vol. 80, n° 5, p. 1297-1308.
- Nguyen, Noël et Robert Espesser (2004), « Méthodes et outils pour l'analyse acoustique des systèmes vocaliques », *Bulletin PFC (Phonologie du français contemporain)*, vol. 3, p. 77-85.
- Nordström, Per-Erik et Björn E. F. Lindblom (1975), « A normalization procedure for vowel formant data », *Proceedings of the 8th International Congress of Phonetic Sciences (unpublished)*, Leeds, article 212.
- Ouellet, Marise (1992), *Systématique des durées segmentales dans les syllabes en français du Québec et de France*, thèse de doctorat, Université de Montréal, 411 p.
- Ouillon, Conrad (1978), *L'enchaînement des voyelles suivies de consonnes occlusives en français*, thèse de doctorat, Université Laval, 739 p.
- Ouillon, Conrad, Jean Dolbec et Marise Ouellet (1993), « La question de l'intensité en français québécois », *Actes du XV Congrès international des linguistes*, Québec, vol. 2, p. 75-78.
- Pape, Daniel et Christine Mooshammer (2006), « Intrinsic F0 differences for German tense and lax vowels », *Proceedings of the 7th International Seminar on Speech Production*, Ubatuba, p. 271-278.
- Paradis, Claude (1983), « La diphtongaison : stabilité et changement dans le système vocalique du français de Chicoutimi-Jonquièr », *Protée*, vol. 11, n° 2, p. 43-53.
- (1985), *An acoustic study of variation and change in the vowel system of Chicoutimi and Jonquièr (Québec)*, thèse de doctorat, University of Pennsylvania, 326 p.
- Pépiot, Erwan (2013), *Voix de femmes, voix d'hommes : différences acoustiques, identification du genre par la voix et implications psycholinguistiques chez les locuteurs anglophones et francophones*, thèse de doctorat, Université Paris 8, 316 p.
- Peterson, Gordon E. (1951), « The phonetic value of vowels », *Language*, vol. 27, n° 4, p. 541-553.
- Peterson, Gordon E. et Harold L. Barney (1952), « Control methods used in a study of the vowels », *Journal of the Acoustical Society of America*, vol. 24, n° 2, p. 175-184.

- Pfützing, Hartmut R. et Oliver Niebuhr (2011), « Historical development of phonetic vowel systems: The last 400 years », *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, p. 160-163.
- Poiré, François (1994), « Relâchement et abrègement des voyelles : phonétique ou phonologie? », *Actes du symposium technique de la Semaine canadienne d'acoustique 1994*, Ottawa, p. 145-146.
- Poliquin, Gabriel Christophe (2006), *Canadian French Vowel Harmony*, thèse de doctorat, Harvard University, 328 p.
- Pols, Louis C. W. (1977), *Spectral Analysis and Identification of Dutch Vowels in Monosyllabic Words*, thèse de doctorat, Free University of Amsterdam, 152 p.
- Potter, Ralph K. et Gordon E. Peterson (1948), « The representation of vowels and their movements », *Journal of the Acoustical Society of America*, vol. 20, n° 4, p. 528-535.
- Potter, Ralph K. et John C. Steinberg (1950), « Toward the specification of speech », *Journal of the Acoustical Society of America*, vol. 22, n° 6, p. 807-820.
- Prairie, Michel (1976), *Sur la structure du mot phonologique en français de Montréal*, mémoire de maîtrise, Université du Québec à Montréal, 128 p.
- Rakerd, Brad et Robert R. Verbrugge (1987), « Evidence that the dynamic information for vowels is talker independent in form », *Journal of Memory and Language*, vol. 26, p. 558-563.
- Reighard, John (1986), « Une analyse concrète du système vocalique du français montréalais », *Revue québécoise de linguistique théorique et appliquée*, vol. 5, n° 4, p. 281-308.
- Reyher, Samuel (1679), *Mathesis mosaica*, Kiel, Joachim Reumann, 31 p.
- Rosner, Burton S. et John B. Pickering (1994), *Vowel Perception and Production*, Oxford, Oxford University Press, 431 p.
- Russell, George Oscar (1928), *The vowel. Its Physiological Mechanism as Shown by X-Ray*, Columbus, The Ohio State University Press, 353 p.
- (1970), *The vowel. Its Physiological Mechanism as Shown by X-Ray*, College Park, McGrath Publishing Co., 353 p.
- Santerre, Laurent (1971), *Les voyelles orales dans le français parlé à Montréal*, thèse de doctorat, Université de Strasbourg, 642 p.
- Santerre, Laurent, Simon-Pierre Dufour et Stéphane McDuff (1985), « La perception de la diphtongaison : son importance dans les grands corpus », *Revue québécoise de linguistique théorique et appliquée*, vol. 4, n° 4, p. 33-53.
- Santerre, Laurent et Jean Millo (1978), « Diphtongization in Montreal French » dans Sankoff, David (dir.), *Linguistic Variation: Models and Methods*, New York, Academic Press, p. 173-184.
- Schneeberg, Jennifer et Bente Schläppler (2006), « Relations between intrinsic f₀, voice quality and the tenseness contrast for German vowels », *Arbeitsberichte des Instituts für Phonetik und digitale Sprachverarbeitung der Universität Kiel*, vol. 37, p. 20-38.
- Schwartz, Jean-Luc, Denis Beautemps, Christian Abry et Pierre Escudier (1993), « Inter-individual and cross-linguistic strategies for the production of the [i] vs. [y] contrast », *Journal of Phonetics*, vol. 21, n° 4, p. 411-425.
- Séguin, Marie-Claude (2010), *Categorisation acoustique des voyelles moyennes antérieures arrondies en français laurentien*, thèse de doctorat, Université d'Ottawa, 180 p.

- Sigouin, Caroline (à paraître), « Une évaluation de différentes procédures de normalisation en français québécois », *Actes du XXVI^e colloque des Journées de linguistique*, Université Laval, Québec.
- Simpson, Adrian P. (2009), « Phonetic differences between male and female speech », *Language and Linguistics Compass*, vol. 3, n° 2, p. 621-640.
- Slawson, A. Wayne (1967), « Vowel quality and musical timbre as functions of spectrum envelope and fundamental frequency », *Journal of the Acoustical Society of America*, vol. 43, n° 1, p. 87-101.
- Stevens, Kenneth N. et Arthur S. House (1963), « Perturbation of vowel articulations by consonantal context: An acoustical study », *Journal of Speech and Hearing Research*, vol. 6, n° 2, p. 111-128.
- Stevens, Kenneth N., Arthur S. House et Allan P. Paul (1966), « Acoustical description of syllabic nuclei: An interpretation in terms of a dynamic model of articulation », *Journal of the Acoustical Society of America*, vol. 40, n° 1, p. 123-132.
- Stevens, Stanley Smith et John E. Volkmann (1940), « The relation of pitch to frequency: A revised scale », *The American Journal of Psychology*, vol. 53, n° 3, p. 329-353.
- Straka, Georges (1959), « Durée et timbre vocalique », *Zeitschrift für Phonetik und allgemeine Sprachwissenschaft*, vol. 12, n° 1/4, p. 276-300.
- Strange, Winifred, James J. Jenkins et Thomas L. Johnson (1983), « Dynamic specification of coarticulated vowels », *Journal of the Acoustical Society of America*, vol. 74, n° 3, p. 695-705.
- Strange, Winifred, Robert R. Verbrugge, Donald P. Shankweiler et Thomas R. Edman (1976), « Consonant environment specifies vowel identity », *Journal of the Acoustical Society of America*, vol. 60, n° 1, p. 213-224.
- Strange, Winifred, Andrea Weber, Erika S. Levy, Valeriy Shafiro, Miwako Hisagi et Kanae Nishi (2007), « Acoustic variability within and across German, French, and American English vowels: phonetic context effects », *Journal of the Acoustical Society of America*, vol. 122, n° 2, p. 1111-1129.
- Sussman, Harvey M. (1986), « A neuronal model of vowel normalization and representation », *Brain and Language*, vol. 28, n° 1, p. 12-23.
- Syrdal, Ann K. et Hundrai S. Gopal (1986), « A perceptual model of vowel recognition based on the auditory representation of American English vowels », *Journal of the Acoustical Society of America*, vol. 79, n° 4, p. 1086-1100.
- Thomas, Erik R. (1997), « A rural/metropolitan split in the speech of Texas Anglos », *Language Variation and Change*, vol. 9, n° 3, p. 309-332.
- Thomas, Erik R. et Tyler Kendall (2007), *NORM: The Vowel Normalization and Plotting Suite*. Consulté le 7 avril 2013 sur <http://ncslaap.lib.ncsu.edu/tools/norm/index.php>
- Tiffany, William R (1953), « Vowel recognition as a function of duration, frequency modulation and phonetic context », *Journal of Speech and Hearing Disorders*, vol. 18, n° 3, p. 289-301.
- Titze, Ingo R. (1989), « Physiologic and acoustic differences between male and female voices », *Journal of the Acoustical Society of America*, vol. 85, n° 4, p. 1699-1707.
- Trautmüller, Hartmut (1981), « Perceptual dimension of openness in vowels », *Journal of the Acoustical Society of America*, vol. 69, n° 5, p. 1465-1475.
- (1987), « Some aspects of the sounds of speech sounds » dans Schouten, Marten Egbertus Hendrik (dir.), *The psychophysics of speech perception*, Dordrecht, Martinus Nijhoff, p. 293-305.
- Trautmüller, Hartmut et Francisco Lacerda (1987), « Perceptual relativity in identification of two-formant vowels », *Speech Communication*, vol. 6, p. 143-157.

- Vaissière, Jacqueline (2007), « Area functions and articulatory modeling as a tool for investigating the articulatory, acoustic, and perceptual properties of sounds across languages » dans Solé, Maria-Josep, Patrice Speeter Beddor et Manjari Ohala (dir.), *Experimental Approaches to Phonology*, Oxford, Oxford University Press, p. 54-71.
- Verbrugge, Robert R. et Brad Rakerd (1987), « Evidence of talker-independent information for vowels », *Haskins Laboratories Status Report on Speech Research*, vol. SR-89/90, p. 39-55.
- Vinay, Jean-Paul (1973), « Le français en Amérique du Nord : problèmes et réalisations », *Current Trends in Linguistics*, vol. 10, p. 323-406.
- Volín, Jan et David Studenovský (2007), « Normalization of czech vowels from continuous read texts », *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, p. 185-190.
- Walker, Douglas C. (1984), *The Pronunciation of Canadian French*, Ottawa, University of Ottawa Press, 148 p.
- Watson, Catherine I. et Jonathan Harrington (1999), « Acoustic evidence for dynamic formant trajectories in Australian English vowels », *Journal of the Acoustical Society of America*, vol. 106, n° 1, p. 458-468.
- Watt, Dominic et Anne H. Fabricius (2002), « Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1~F2 plane », *Leeds Working Papers in Linguistics and Phonetics*, vol. 9, p. 159-173.
- Whalen, Douglas H. et Andrea G. Levitt (1995), « The universality of intrinsic F₀ for vowels », *Journal of Phonetics*, vol. 23, p. 349-366.
- Wheatstone, Charles (1837), « Willis on reed organ sounds », *The London and Westminster Review*, n° 11 et 54, p. 14-22.
- Willis, Robert (1830), « On the vowel sounds, and on reed organ-pipes », *Transactions of the Cambridge Philosophical Society*, vol. 3, p. 231-268.
- Wolfram, Walt, Erik R. Thomas et Elaine W. Green (2011), « The regional context of earlier African American speech: Evidence for reconstructing the development of AAVE », *Language*, vol. 29, n° 3, p. 315-355.
- Xu, Yi (2010), « In defense of lab speech », *Journal of Phonetics*, vol. 38, n° 3, p. 329-336.
- Yaeger, Malcah (1979), *Context-determined variation in Montreal French vowels*, thèse de doctorat, University of Pennsylvania, 353 p.
- Zahorian, Stephen A. et Amir Jalali Jagharghi (1991), « Speaker normalization of static and dynamic vowel spectral features », *Journal of the Acoustical Society of America*, vol. 90, n° 1, p. 67-75.
- (1993), « Spectral-shape features versus formants as acoustic correlates for vowels », *Journal of the Acoustical Society of America*, vol. 94, n° 4, p. 1966-1982.
- Zwicker, Eberhard (1961), « Subdivision of the audible frequency range into critical bands (Frequenzgruppen) », *Journal of the Acoustical Society of America*, vol. 33, n° 2, p. 248.

Annexe 2 : Sigouin (à paraître)

UNE ÉVALUATION DE DIFFÉRENTES PROCÉDURES DE NORMALISATION EN FRANÇAIS QUÉBÉCOIS

Caroline Sigouin
Université du Québec à Chicoutimi

Résumé

L'objectif de la présente étude est de déterminer, parmi une sélection de procédures de normalisation, lesquelles parviennent efficacement, en français québécois, à 1) rendre compte de l'identité phonémique des voyelles et à 2) réduire la variation acoustique liée au sexe du locuteur. Pour ce faire, la fréquence centrale des quatre premiers formants et la fréquence fondamentale ont été relevées à 50 % de la durée des voyelles /i y u e ø o ε œ α/ produites isolément (deux fois) par 20 locuteurs, 10 hommes et 10 femmes, originaires de la ville de Saguenay. Les valeurs fréquentielles ont été normalisées au moyen de dix procédures. L'efficacité de ces dernières a été évaluée à l'aide d'analyses discriminantes. La procédure d'Iri (1959) s'avère moins efficace que les données brutes en Hertz au regard du premier des deux critères d'évaluation susmentionnés. La conversion en barks (Traunmüller, 1990) se révèle quant à elle aussi efficace que les données brutes au regard de ces deux critères. Enfin, les procédures préconisées par Bladon et coll. (1984), Gerstman (1968), Lobanov (1971), Nearey (1977), Ménard (2002), Labov et coll. (2006) et Watt et Fabricius (2002) se montrent au moins aussi efficaces que les données brutes; les résultats variant selon le critère d'évaluation, les indices acoustiques considérés et la procédure employée. La procédure de Gerstman (1968), celle de Lobanov (1971), et la *Constant Ratio Hypothesis 2* de Nearey (1977) apparaissent comme les plus efficaces.

Abstract

The goal of the present study is to evaluate the efficiency of a selection of speaker normalization procedures 1) to maintain phonemic distinctions between Quebec French vowels and 2) to reduce gender-induced formants variability. Recordings were made for 20 speakers (10 men, 10 women) from Saguenay town. Isolated vowels /i y u e ø o ε œ α/ were produced twice by each speaker. Fundamental frequency and the first four formants central frequencies were extracted at 50% of the vowel duration. Raw Hertz values were normalized by using ten different procedures that have been evaluated using discriminant analyses. Iri (1959)'s procedure appears to be less effective than raw Hertz values for the first evaluation criterion indicated above while Bark-converted (Traunmüller, 1990) and raw Hertz values are equally effective for both criteria. Procedures preconized by Bladon et al. (1984), Gerstman (1968), Lobanov (1971), Nearey (1977), Ménard (2002), Labov et al. (2006) and Watt & Fabricius (2002) show to be equally or more effective than raw Hertz values, depending on evaluation criterion, acoustic features involved and normalization procedure applied. By and large, Gerstman (1968)'s procedure, Lobanov (1971)'s procedure and Nearey (1977)'s *Constant Ratio Hypothesis 2* reveal to be the most efficient.

Introduction

Les fréquences centrales des deux premiers formants (désormais F_1 et F_2) d'une voyelle estimées au centre de son état stable sont depuis longtemps considérées comme les corrélats acoustiques essentiels de son identité (Delattre, 1948); F_1 apparaissant comme un indice robuste du degré d'aperture, et F_2 , du degré d'antériorité (Joos, 1948 : 53; Delattre, 1948 : 482-483). Cependant, les fréquences formantiques varient considérablement en fonction de multiples facteurs tant linguistiques qu'extralinguistiques (contexte consonantique, débit de parole, sexe, âge, origine géographique du locuteur, etc.). Adank (2003 : 5) souligne ainsi que dans un diagramme F_1 / F_2 , « it can be observed that areas representing different vowels show considerable overlap ». Comment expliquer, alors, que les occurrences comprises dans ces zones de chevauchement soient reconnues, pour la plupart, comme appartenant à la classe phonémique attendue? Certes, la prise en compte de paramètres supplémentaires comme la durée, la fréquence fondamentale (F_0), les troisième et quatrième formants (F_3 et F_4) ou encore l'évolution des formants au cours du temps peut résoudre en partie ce problème (Hillenbrand et coll., 1995). Malgré tout, une variabilité formantique non négligeable, induite notamment par les proportions dissemblables des conduits vocaux féminins, masculins et enfantins, subsiste entre les voyelles produites par des locuteurs d'âges et de sexes différents. Conséquemment, diverses procédures de normalisation ont été proposées afin de réduire la variation inhérente aux différences physiologiques tout en préservant l'information phonémique, dialectale et sociolinguistique, et éventuellement en modélisant le processus de perception des voyelles⁴⁰ (Thomas et Kendall, 2007).

Les procédures de normalisation

Les procédures de normalisation diffèrent les unes des autres par « the amount and type of information that is required and the type of metric that is used to calculate the normalized values from a data set » (Clopper, 2009 : 1431). Dans les lignes qui suivent, nous proposons un aperçu de différents types de procédures de normalisation. Les procédures évaluées spécifiquement dans le cadre de la présente étude seront détaillées dans la section consacrée à la méthodologie.

⁴⁰ Voir notamment Johnson [2005] à ce sujet.

Comme le soulignent Thomas et Kendall (2007) ainsi que Fabricius et coll. (2009 : 417), certaines transformations, parfois envisagées comme des procédures de normalisation, sont avant tout des opérations de conversion : il s'agit d'exprimer la fréquence non plus à l'aide d'une échelle de mesure physique et linéaire (le nombre de cycles d'oscillation par seconde, ou *Hertz*), mais par une échelle psycho-acoustique censément représentative de la perception auditive humaine. Différentes échelles ont ainsi été proposées afin de rendre compte de divers aspects de la perception de la fréquence (consonance, hauteur, résolution de la fréquence, etc.). Ces opérations ne sont donc pas fondamentalement destinées à éliminer la variabilité d'origine physiologique, bien que leur potentiel normalisateur ait souvent été éprouvé (Syrdal et Gopal, 1986; Hillenbrand et Gayvert, 1993; Adank, 2003; Calamai, 2005; Volín et Studenovský, 2007; Clopper, 2009; Flynn et Foulkes, 2011) et que certaines d'entre elles soient intégrées à d'autres procédures de normalisation (comme les ratios formantiques, les distances tonotopiques et le second formant effectif que nous présenterons plus loin). Les principales opérations de conversion rencontrées en sciences phonétiques sont :

- le logarithme décimal ou népérien de la fréquence en Hertz⁴¹;
- les échelles musicales (octave, ton, demi-ton...);
- l'échelle des mels (Stevens et Volkman, 1940);
- l'échelle de Koenig (Koenig, 1949);
- l'échelle des barks (Zwicker, 1961);
- l'échelle du *equivalent rectangular bandwidth-rate* (ERB-rate, Moore et Glasberg, 1983).

Toutes ces formules de conversion sont appliquées uniformément aux valeurs fréquentielles de départ sans qu'aucune autre variable n'entre en jeu. Notons toutefois que Bladon et coll. (1984) suggèrent de normaliser l'échelle des barks en ajoutant 1 bark aux fréquences formantiques masculines ou en soustrayant 1 bark aux fréquences formantiques féminines.

D'autres opérations impliquent quant à elles la création d'un objet ou d'un espace psycho-acoustique. Elles ne consistent donc pas à faire correspondre à chaque fréquence formantique un équivalent normalisé, mais bien à construire de nouveaux indices psycho-

⁴¹ Il s'agit en fait de l'approximation la plus simple de la pseudo-loi de Weber-Fechner, qui postule que la sensation varie comme le logarithme de l'excitation.

acoustiques à partir des composantes spectrales issues d'une même occurrence. Ces objets sont le plus souvent mis en relation avec la perception d'un trait articulatoire des voyelles (aperture, antériorité, arrondissement, etc.). On peut ainsi calculer le ratio (Lloyd, 1890; Potter et Steinberg, 1950; Miller, 1989) ou la distance, sur une échelle tonotopique comme les barks (Traunmüller, 1981; Syrdal et Gopal, 1986), entre deux formants ou entre un formant et la F_0 ; ou encore le second formant effectif (F'_2). F'_2 est une proéminence spectrale perçue qui réunirait F_2 , F_3 , voire F_4 (Delattre et coll., 1952; Bladon et Fant, 1978; Mantakas et coll., 1986; Ménard, 2002). Iri (1959) a également élaboré un espace bidimensionnel qui a récemment connu un regain d'intérêt (Sussman, 1986; Thomas, 1997; Anderson, 1999; Wolfram et coll., 2011).

Enfin, d'autres procédures de normalisation consistent pour leur part à faire correspondre à chaque fréquence formantique un équivalent normalisé (Gerstman, 1968; Lobanov, 1971; Nordström et Lindblom, 1975; Fant, 1975; Nearey, 1977; Watt et Fabricius, 2002; Labov et coll., 2006). Au contraire des opérations de conversion, cependant, ces procédures comportent au moins une autre variable que la valeur fréquentielle de départ. Elles impliquent en effet le calcul de données statistiques (maxima, minima, moyennes, écarts-types, etc.) par rapport auxquelles sont exprimées les valeurs fréquentielles à normaliser : rééchelonnement des valeurs entre les extrêmes, calcul d'un facteur d'échelle, estimation de la taille du conduit vocal, distance entre la valeur à normaliser et une moyenne, etc. Ces données statistiques peuvent être calculées à partir des valeurs fréquentielles brutes provenant ou non d'une même occurrence, d'un même locuteur ou d'un même indice acoustique.

L'évaluation des procédures de normalisation

S'interrogeant devant autant de procédures différentes, de nombreux chercheurs (Hindle, 1978; Ferrari Disner, 1980; Miller et coll., 1980; Syrdal et Gopal, 1986; Miller, 1989; Hillenbrand et Gayvert, 1993; Ménard, 2002; Adank, 2003; Nguyen et Espesser, 2004; Calamai, 2005; Volín et Studenovský, 2007; Fabricius et coll., 2009; Clopper, 2009; Flynn et Foulkes, 2011) ont cherché à déterminer, dans plusieurs langues et variétés de langue (anglais américain et britannique, allemand, norvégien, suédois, danois, français de France,

néerlandais, italien, tchèque), les procédures les plus efficaces au regard de différents critères d'évaluation (réduction de la dispersion intraclasse, augmentation de la dispersion interclasses, réduction de l'écart entre les fréquences formantiques masculines et féminines, préservation de l'information phonémique, sociolinguistique ou dialectale, etc.). Bien qu'aucune de ces études n'ait porté sur la totalité des procédures de normalisation disponible dans la littérature, les résultats proposés indiquent que les opérations de conversion s'avèrent moins efficaces que les objets et espaces psycho-acoustiques, qui s'avèrent eux-mêmes moins efficaces que les procédures impliquant le calcul de données statistiques. Johnson (2005 : 370) remarque d'ailleurs que « most of the practically useful vowel normalization algorithms require that summary statistics be derived over a full set of vowels for each talker ». Parmi les procédures de ce dernier groupe, celle de Nordström et Lindblom (1975) ne s'est pas montrée particulièrement efficace (Adank, 2003; Clopper, 2009; Flynn et Foulkes, 2011). Celle de Fant (1975) n'a pas été testée à notre connaissance, mais elle s'apparente fortement à celle de Nordström et Lindblom (1975). Les autres procédures de ce groupe ont toutes démontré une grande efficacité.

D'autre part, mentionnons qu'une combinaison de différents objets psycho-acoustiques est parfois préconisée. Ainsi, Ménard (2002) et Ménard et coll. (2002) proposent d'utiliser F_0 - F_1 , F_2 - F_1 et F'_2 (en barks) pour rendre respectivement compte du degré d'aperture, du lieu d'articulation et de l'arrondissement. Par ailleurs, l'échelle normalisée des barks proposée par Bladon et coll. (1984) s'est également révélée intéressante (Clopper, 2009).

Problématique et objectif

Les procédures de normalisation vocalique sont nombreuses et variées. Si leur efficacité a souvent été évaluée, elle l'a rarement été en français (Ménard, 2002; Nguyen et Espesser, 2004), et, à notre connaissance, elle ne l'a jamais été en français québécois. Or, l'étude de Ferrari Disner (1980) témoigne de l'efficacité variable des procédures de normalisation selon la langue à laquelle elles sont appliquées. L'objectif de la présente étude est donc de déterminer, parmi une sélection de procédures de normalisation, lesquelles parviennent efficacement, en français québécois, à rendre compte de l'identité phonémique des voyelles, de même qu'à réduire la variation acoustique liée au sexe du locuteur.

Méthodologie

À la lumière des résultats des études précédentes, nous avons tout d'abord opéré une sélection parmi les multiples procédures proposées dans la littérature. Nos analyses ont ainsi porté sur dix procédures courantes ayant déjà fait leurs preuves ou n'ayant pas encore fait l'objet d'une évaluation.

L'échelle des barks

L'échelle des barks rend compte de la relation entre la fréquence (en Hertz) et les bandes critiques (des étendues de fréquences en deçà desquelles l'oreille humaine ne semble pas percevoir de différence de hauteur). De nombreuses équations ont été proposées afin d'effectuer la conversion des Hertz en barks. Nous retiendrons pour notre part celle de Trautmüller (1990) (voir la formule 1 de l'annexe 1). Bien que l'échelle des barks, comme procédure de normalisation, se soit rarement révélée efficace (Syrdal et Gopal, 1986), nous avons choisi de l'intégrer à nos analyses afin de tester au moins une échelle psycho-acoustique, et parce qu'en outre, cette échelle est intégrée à d'autres procédures de normalisation.

L'échelle des barks normalisée (Bladon et coll., 1984)

Bladon et coll. (1984) suggèrent de normaliser l'échelle des barks en ajoutant 1 bark aux fréquences formantiques masculines ou en soustrayant 1 bark aux fréquences formantiques féminines. En ce qui nous concerne, nous avons choisi d'ajouter 1 bark aux valeurs masculines, plutôt que de soustraire 1 bark aux valeurs féminines, afin d'éviter d'obtenir des valeurs négatives (voir la formule 2 de l'annexe 1).

Un ensemble d'objets psycho-acoustiques (Ménard, 2002)

Ménard (2002) propose pour sa part d'utiliser la distance tonotopique (en barks) entre F_0 et F_1 , et entre F_2 et F_1 , de même que le second formant effectif (F'_2) pour rendre respectivement compte du degré d'aperture, du lieu d'articulation et de l'arrondissement des voyelles (voir les formules 3, 4 et 5 de l'annexe 1).

La procédure d'Iri (1959)

De son côté, Iri (1959) calcule la distance logarithmique entre la fréquence formantique à normaliser et la moyenne géométrique des trois premiers formants de l'occurrence dont elle est tirée. Iri (1959) multiplie cette distance par dix, mais les auteurs qui, à sa suite, ont employé cette procédure (Sussman, 1986; Thomas, 1997; Anderson, 1999; Wolfram et coll., 2011), n'ont pas conservé cette constante multiplicative. À des fins de comparaison, la formule employée dans la présente contribution ne comporte pas non plus cette constante (voir la formule 6 de l'annexe 1).

La procédure de Gerstman (1968)

Gerstman (1968) propose quant à lui de rééchelonner les valeurs de F_1 et de F_2 de chaque locuteur de façon à ce qu'elles soient comprises entre 0 et 999. La valeur normalisée 0 est alors attribuée à la plus basse fréquence de chaque formant, et 999, à sa valeur la plus élevée. Les autres valeurs sont linéairement réparties entre ces deux valeurs extrêmes (voir la formule 7 de l'annexe 1).

Les scores-Z (Lobanov, 1971)

La procédure de Lobanov (1971) consiste à calculer le nombre d'écarts-types qui séparent chaque fréquence formantique de la moyenne des formants de même ordre d'un locuteur donné. Cela revient à exprimer les fréquences par leur score-Z (voir la formule 8 de l'annexe 1).

La Constant Ratio Hypothesis 1 (Nearey, 1977)

La *Constant Ratio Hypothesis 1* publiée par Nearey (1977) traduit la position des fréquences formantiques par leur distance à la moyenne des deux premiers formants de l'ensemble des voyelles d'un locuteur donné dans un espace logarithmique (voir la formule 9 de l'annexe 1). Notons que Nearey (1977) a développé cette procédure en ne considérant que les deux premiers formants des voyelles. Lorsque d'autres indices acoustiques (comme la F_0 et F_3) sont considérés, certains chercheurs ont décidé d'inclure leurs valeurs dans le calcul de la moyenne (Adank et coll., 2004). Quant à nous, à la suite de quelques examens préliminaires suggérant que la version originale de cette procédure

était plus efficace, nous avons choisi de calculer la moyenne sur F_1 et F_2 seulement, et ce, même lorsque la F_0 ou F_3 étaient normalisés. Nous avons ainsi exprimé la fréquence des F_0 et des F_3 par rapport à la moyenne de F_1 et de F_2 , et non par rapport à la moyenne de F_0 , F_1 , F_2 et F_3 .

La Constant Ratio Hypothesis 2 (Nearey, 1977)

La *Constant Ratio Hypothesis 2*, également proposée par Nearey (1977), est une procédure identique à la *Constant Ratio Hypothesis 1*, si ce n'est que la moyenne individuelle par rapport à laquelle est exprimée la fréquence formantique à normaliser est calculée sur les formants de même ordre seulement (voir la formule 10 de l'annexe 1). Par exemple, la moyenne des F_1 d'un locuteur est calculée afin de normaliser les fréquences des F_1 , mais la moyenne des F_2 est également calculée afin de normaliser les fréquences des F_2 . Il en va de même pour la F_0 et F_3 .

La procédure de l'Atlas of North American English (Labov et coll., 2006)

La procédure de normalisation employée par Labov et coll. (2006) dans *The Atlas of North American English* consiste à calculer le rapport entre la moyenne du logarithme des fréquences formantiques de toutes les voyelles d'un locuteur, et celle des fréquences formantiques de toutes les voyelles de l'ensemble des locuteurs. Ce rapport, auquel on applique une fonction exponentielle afin de repasser en unités linéaires, est ensuite employé comme facteur k de normalisation à appliquer aux formants de toutes les occurrences d'un locuteur i (voir la formule 11 de l'annexe 1). Tout comme la *Constant Ratio Hypothesis 1* de Nearey (1977), cette procédure a été développée en ne prenant en compte que les deux premiers formants des voyelles. Nous avons également choisi de calculer la moyenne sur F_1 et F_2 seulement, et ce, même lorsque la F_0 ou F_3 étaient normalisés.

La procédure de Watt et Fabricius (2002)

Cette dernière procédure consiste à calculer le rapport entre les fréquences formantiques à normaliser et le centroïde, que les auteurs abrègent S , du système vocalique de chaque locuteur. Le centroïde constitue une moyenne des coordonnées des trois extrémités du triangle vocalique (normalement /i/, /u/ et /a/) dans un diagramme F_1 / F_2 . Si le F_1 et le F_2

moyens de /i/ sont intégrés directement dans le calcul du centroïde, de même que le F_1 de /a/, il n'en va pas de même des autres fréquences formantiques. En effet, puisque la réalisation de /u/ est, selon les auteurs, trop instable, le F_1 et le F_2 moyens de ce dernier sont extrapolés suivant le raisonnement suivant : le F_1 de /u/ doit être semblable à celui du /i/, et son F_2 est très près de son F_1 , au point de se confondre avec celui-ci. Ainsi, $F_2^{/u/} = F_1^{/u/} = F_1^{/i/}$. Quant au F_2 de /a/, il n'est souvent pas assez central, ce qui cause une déformation du bas du triangle normalisé (Thomas et Kendall, 2007). $F_2^{/a/}$ peut donc être intégré directement au calcul, mais préférablement, comme le suggèrent Fabricius et coll. (2009), être établi comme étant la moyenne des F_2 de /i/ et /u/. Nous avons intégré la recommandation de Fabricius et coll. (2009) dans nos calculs (voir la formule 12 de l'annexe 1). Comme cette procédure repose fondamentalement sur la représentation bidimensionnelle F_1 / F_2 des voyelles, nous ne l'avons appliquée qu'aux deux premiers formants des voyelles du corpus exploité.

Le corpus exploité lors de cette expérimentation est issu d'un projet de recherche de Vincent Arnaud (Université du Québec à Chicoutimi) subventionné par le Fonds québécois de recherche sur la société et la culture, et qui est intitulé « Caractéristiques acoustiques des voyelles orales québécoises »⁴². Ce corpus est constitué des voyelles /i y u e ø o ε œ α/ produites isolément (au moins deux fois chacune) par 20 locuteurs, autant d'hommes que de femmes provenant de la ville de Saguenay. Ces locuteurs, âgés de 20 à 29 ans, avaient le français québécois comme langue maternelle et poursuivaient des études universitaires au moment de l'enquête, menée entre 2010 et 2012. Les enregistrements ont eu lieu en chambre anéchoïque et ont été réalisés directement au format numérique (44 100 Hz, 16 bits) à l'aide d'un TASCAM HD-P2 et d'un microphone serre-tête SHURE SM10A amplifié au moyen d'une console de mixage BEHRINGER XENYX 802. La tâche demandée aux participants était de lire à haute voix des phrases du type « Dans [tel mot], tu dis [telle voyelle] ». Les différentes enquêtrices ayant pris part aux enregistrements simulaient alors le fait qu'elles n'avaient pas compris la voyelle finale, et demandaient au locuteur de la répéter de façon isolée. Les échanges se présentaient donc comme suit :

⁴² Nous remercions chaleureusement M. Arnaud pour ses innombrables relectures et pour nous avoir donné accès à son corpus, de même que Sophie Diotte, Vicky Tremblay et Francesca Bouchard, pour leur participation à la constitution de ce dernier.

Locuteur : Dans *pie*, tu dis *i*.

Enquêtrice : Tu dis?

Locuteur : *i*.

Ce sont ces répétitions isolées, totalisant 405 occurrences, qui ont été analysées. Nous avons établi à l'audition la classe phonémique des voyelles produites.

Les analyses acoustiques ont été réalisées à l'aide du logiciel PRAAT (version 5.3.x). La fréquence fondamentale et la fréquence centrale des quatre premiers formants ont été relevées (méthode LPC) à 50 % de la durée des voyelles, qui avaient préalablement été segmentées manuellement. Ces données ont par la suite été soumises à un script PERL que nous avons développé⁴³ et qui permet de normaliser automatiquement les valeurs fréquentielles contenues dans une base de données tabulée en fonction des procédures sélectionnées. Nous disposons donc, pour chaque occurrence, des valeurs brutes (en Hertz) de ses indices acoustiques et des valeurs normalisées selon les dix procédures décrites ci-dessus. Ces ensembles de valeurs ont finalement été soumis à des analyses discriminantes quadratiques avec rééchantillonnage *jackknife* afin de déterminer ceux qui s'avèrent les meilleurs prédicteurs de l'identité phonémique des voyelles (puisque l'on vise à conserver l'information phonémique), et les moins bons prédicteurs du sexe du locuteur les ayant produites (puisque l'on cherche à réduire l'écart entre les productions féminines et masculines). Une analyse discriminante est un test statistique qui établit l'appartenance des cas qui lui sont soumis (ici des occurrences vocaliques) à un groupe (à une classe phonémique ou au sexe du locuteur, par exemple) en fonction des informations (prédicteurs) dont elle dispose à propos de ces cas (ici des indices acoustiques). Pour ce faire, l'analyse compare les valeurs des indices acoustiques de chaque occurrence (dont elle ignore momentanément la classe phonémique ou le sexe du locuteur) à celles des autres occurrences (dont elle connaît la classe phonémique ou le sexe du locuteur). La classification « machine » ainsi obtenue est ensuite confrontée à celle établie par la chercheuse de façon à calculer dans quelle proportion (en pourcentage) ces deux classifications correspondent. Plus le taux de classification correcte est élevé, plus les indices acoustiques pris en considération dans l'analyse discriminante permettent de

⁴³ Ce script, réalisé sous la supervision de Jacques Ladouceur [Université Laval], est disponible gratuitement pour téléchargement à l'adresse suivante : <http://linglang.uqac.ca/recherche/laboratoire-phonetique>

différencier efficacement les occurrences issues de groupes différents. Sauf exception, chaque procédure a ainsi été évaluée lorsque différentes combinaisons d'indices acoustiques étaient considérées : F_1 et F_2 ; F_0 , F_1 et F_2 ; F_0 , F_1 , F_2 et F_3 . L'ensemble de procédures proposé par Ménard (2002) ne fournissant toutefois pas d'équivalents normalisés aux indices acoustiques, mais bien de nouveaux indices (intégrant F_0 , F_1 , F_2 , F_3 et F_4 dans leur calcul), nous comparerons les résultats obtenus par cette méthode aux résultats des autres procédures lorsque F_0 , F_1 , F_2 et F_3 sont pris en compte. Quant à la procédure de Watt et Fabricius (2002), elle ne permet de normaliser que F_1 et F_2 . Enfin, le F_1 , le F_2 et le F_3 normalisés à l'aide de la procédure d'Iri (1959) ne sont pas mathématiquement indépendants. Ils ne peuvent donc être soumis conjointement à une analyse discriminante. La procédure d'Iri (1959) a donc été évaluée seulement lorsque F_1 et F_2 et lorsque F_0 , F_1 et F_2 sont pris en compte.

Résultats et analyse

En premier lieu, nous avons évalué l'efficacité des différentes procédures sélectionnées à rendre compte de la classe phonémique des occurrences, donc à permettre la classification des occurrences en fonction de leur classe phonémique à partir des indices acoustiques normalisés. Ainsi, plus le taux de classification correct retourné par l'analyse discriminante est élevé et plus la procédure de normalisation est efficace à cet égard.

La Figure 1 présente les résultats obtenus lorsque F_1 et F_2 seulement sont pris en compte. D'emblée, soulignons que 90 % des occurrences sont correctement classées à partir des fréquences formantiques non normalisées. Conformément à ce que Hillenbrand et Gayvert (1993) avaient observé, la conversion en barks ne permet pas d'améliorer le taux de classification correcte des occurrences. Le recours aux barks normalisés (Bladon et coll., 1984) permet néanmoins une augmentation de 2 % de ce taux. Les fréquences normalisées par la procédure d'Iri (1959) obtiennent quant à elles un taux de classification correcte inférieur à celui des données brutes, soit 86 %. C'est donc dire qu'une partie de l'information phonémique a été perdue au cours du processus de normalisation. Les autres procédures ont entraîné une augmentation plus substantielle du taux de classification, qui est passé à 98 %, 99 % et même à 100 % avec les scores-Z de Lobanov (1971).

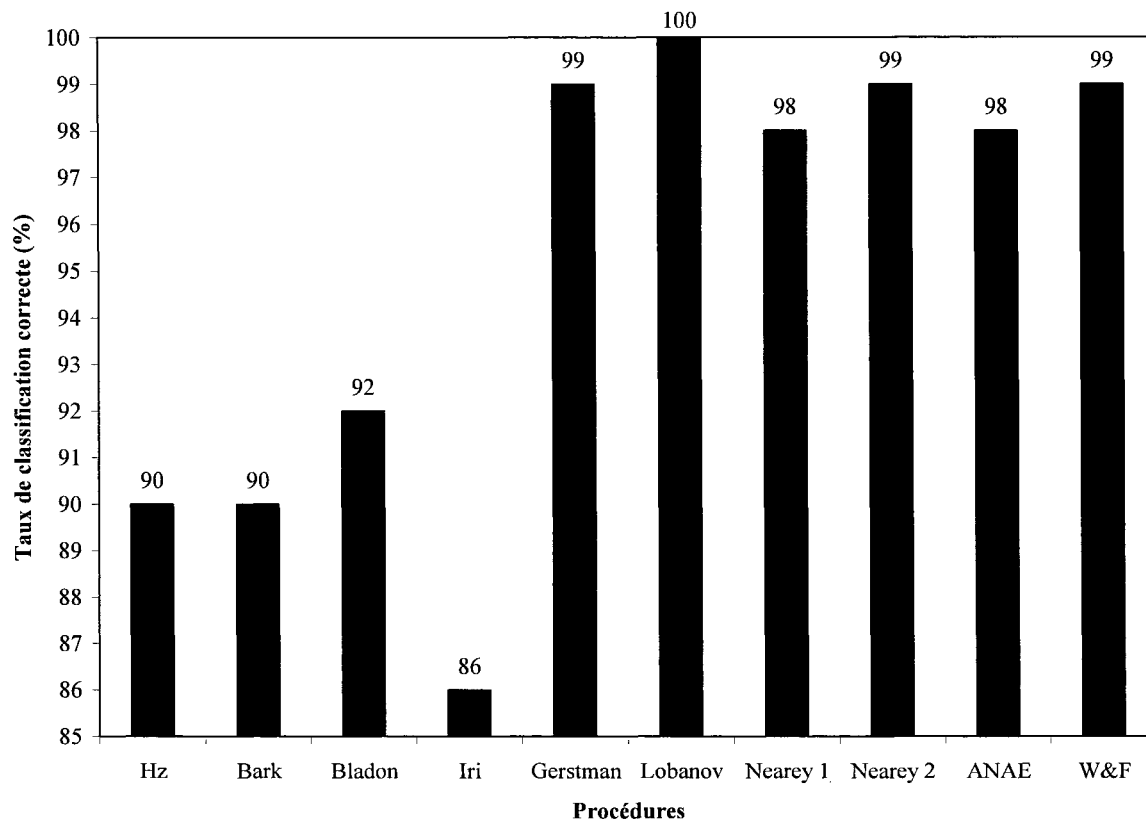


Figure 1 : Taux de classification correcte obtenus par les différentes procédures lorsque F_1 et F_2 sont pris en compte et que les occurrences sont classées d'après leur classe phonémique.

Comme l'illustre la Figure 2, lorsque la F_0 est également prise en compte, le taux de classification correcte s'élève à 97 % pour les données en Hertz, en barks et en barks normalisés (Bladon et coll., 1984). L'emploi des barks normalisés ne présente alors plus d'avantage. Quant à la procédure d'Iri (1959), elle obtient un taux de classification correcte plus élevé avec la prise en compte de la F_0 (90 %), mais ce taux est toujours inférieur à celui des données non normalisées. Les autres procédures plafonnent à 98 %, 99 % et 100 %.

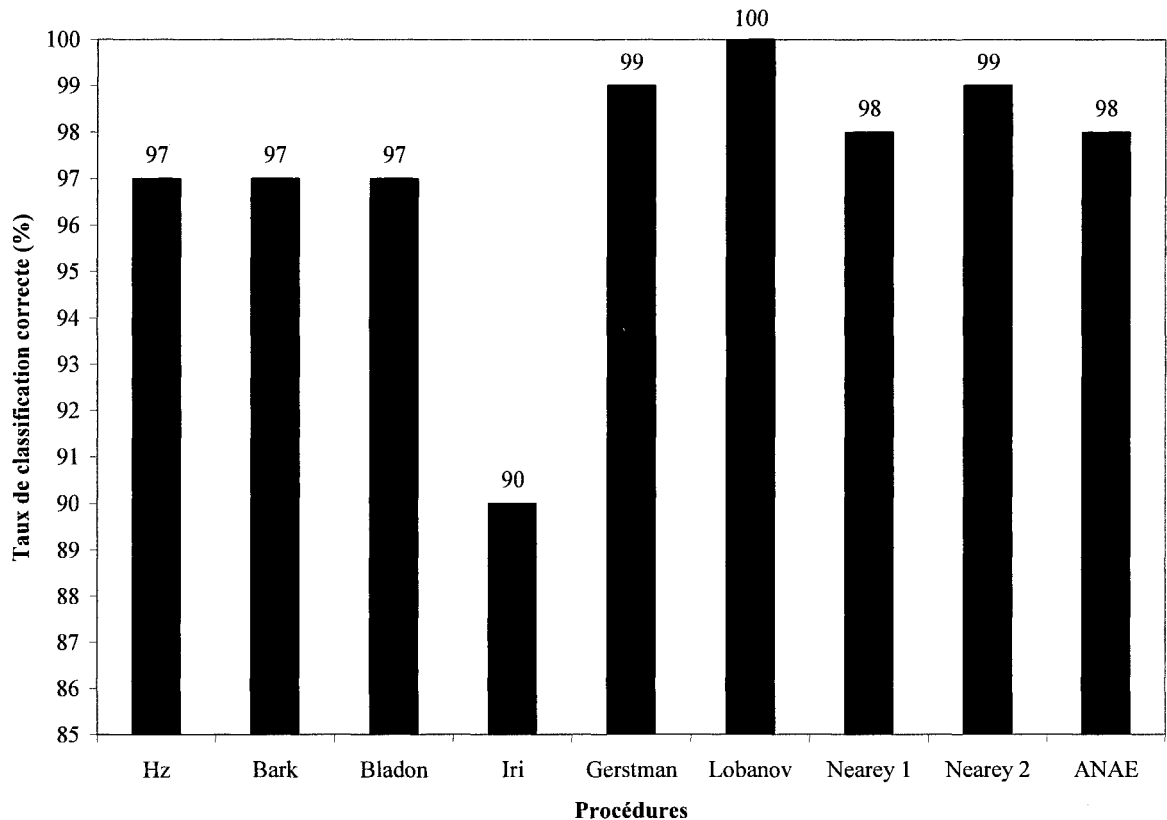


Figure 2 : Taux de classification correcte obtenus par les différentes procédures lorsque la F_0 , F_1 et F_2 sont pris en compte et que les occurrences sont classées d'après leur classe phonémique.

Des tendances similaires sont observées lorsque F_1 , F_2 et F_3 sont pris en considération (voir Figure 3). Le taux de classification correcte obtenu par les données en Hertz, en barks et en barks normalisés (Bladon et coll., 1984) est de 95 %, alors que celui des autres procédures⁴⁴ est de 99 % ou 100 %. La prise en compte de la F_0 semble donc au moins aussi pertinente que l'intégration de F_3 dans la discrimination des classes phonémiques soumises à l'analyse, comme en témoigne le taux de classification légèrement plus élevé obtenu pour les données brutes lorsque la F_0 , F_1 et F_2 sont considérés (97 %) plutôt que F_1 , F_2 et F_3 (95 %). Cependant, la majorité des procédures normalisent un peu plus efficacement F_3 (quatre procédures obtiennent un taux de classification correcte de 100 %) que F_0 (seule la procédure de Lobanov [1971] obtient 100 %).

⁴⁴ Rappelons que les résultats obtenus par la procédure d'Iri [1959] ne sont pas représentés ici puisque le F_1 , le F_2 et le F_3 normalisés à l'aide de cette procédure ne peuvent être soumis conjointement à une analyse discriminante.

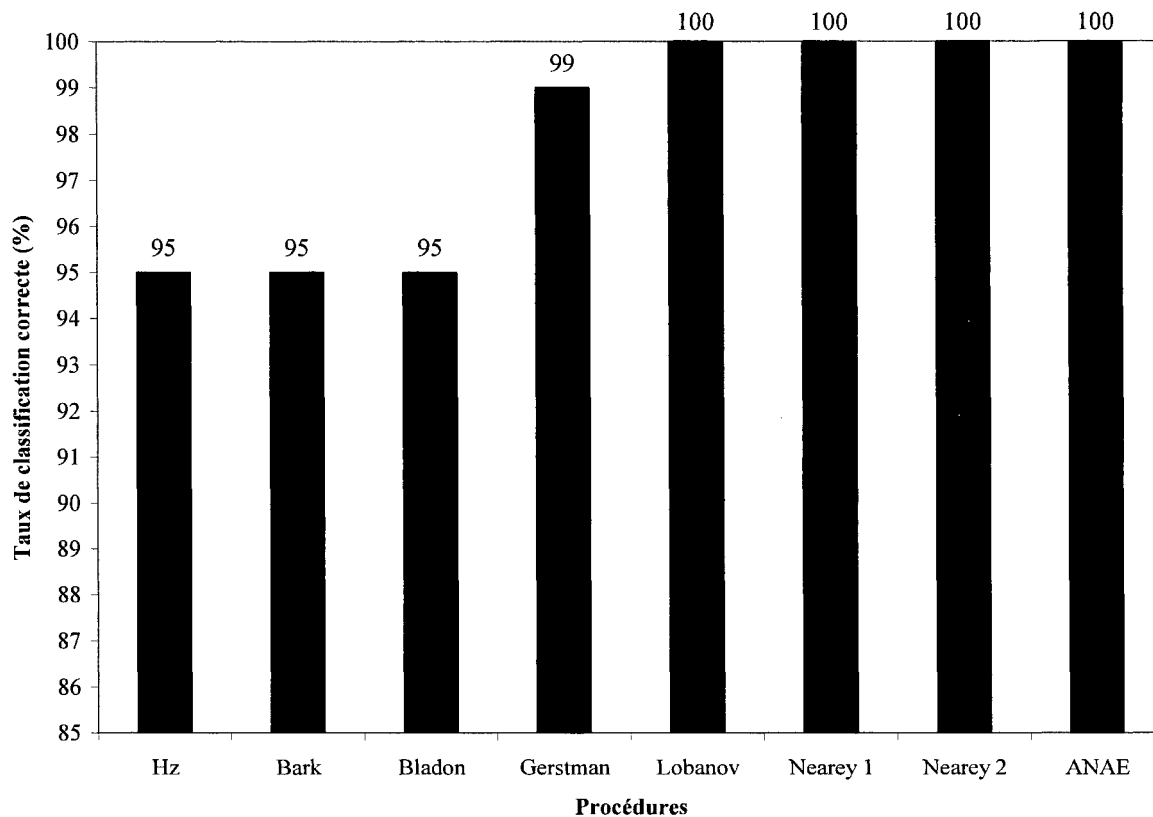


Figure 3 : Taux de classification correcte obtenus par les différentes procédures lorsque F_1 , F_2 et F_3 sont pris en compte et que les occurrences sont classées d'après leur classe phonémique.

La Figure 4, quant à elle, présente les résultats obtenus lorsque la F_0 , F_1 , F_2 et F_3 sont intégrés au modèle de classification. Les procédures considérées obtiennent alors un taux de classification correcte oscillant entre 98 % et 100 %. Même les données non normalisées obtiennent 99 % de classification correcte. On peut donc se demander si, dans le cas de voyelles isolées et lorsque ces quatre indices acoustiques sont pris en considération, il est nécessaire de recourir à une quelconque procédure de normalisation afin de rendre compte de la classe phonémique des voyelles.

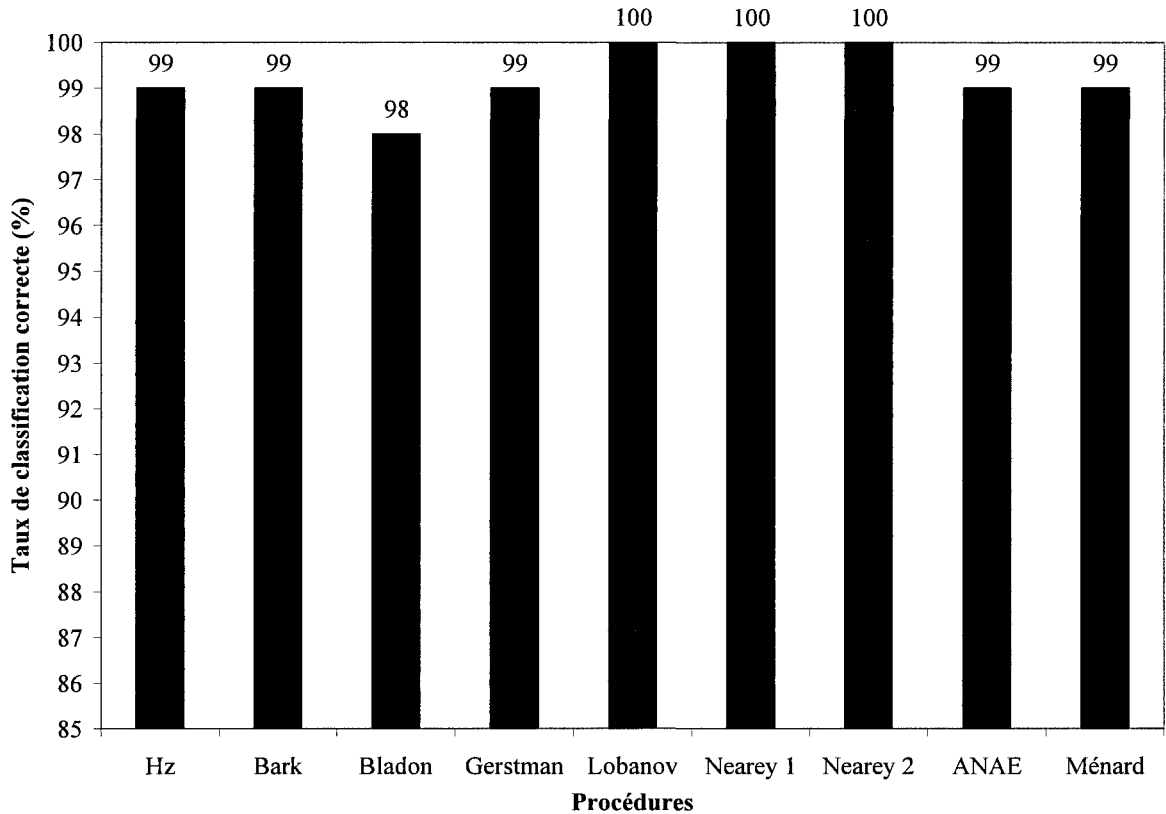


Figure 4 : Taux de classification correcte obtenus par les différentes procédures lorsque la F_0 , F_1 , F_2 et F_3 sont pris en compte et que les occurrences sont classées d'après leur classe phonémique. Dans le cas de la procédure de Ménard [2002], ce sont F_2 - F_1 , F_1 - F_0 et F'_2 [en barks] qui sont pris en compte.

En second lieu, nous avons évalué l'efficacité des différentes procédures sélectionnées à réduire l'écart entre les productions féminines et masculines, donc à rendre difficile une classification des occurrences vocaliques en fonction du sexe du locuteur ayant produit ces dernières. Plus le taux de classification correcte obtenu lors de l'analyse discriminante s'approche du seuil du hasard (50 %, puisqu'il y a une chance sur deux que le locuteur soit un homme et une chance sur deux qu'il soit une femme), plus la procédure est réputée efficace.

La Figure 5 présente les résultats obtenus lorsque seuls F_1 et F_2 sont pris en compte. Les données non normalisées, converties en barks ou en barks normalisés et celles obtenues par la procédure d'Iri (1959) obtiennent les taux de classification correcte les plus élevés. Dans 60 % à 61 % des cas, le sexe du locuteur est correctement déduit. Les autres procédures

parviennent à éliminer en tout ou en partie les différences entre les productions féminines et masculines.

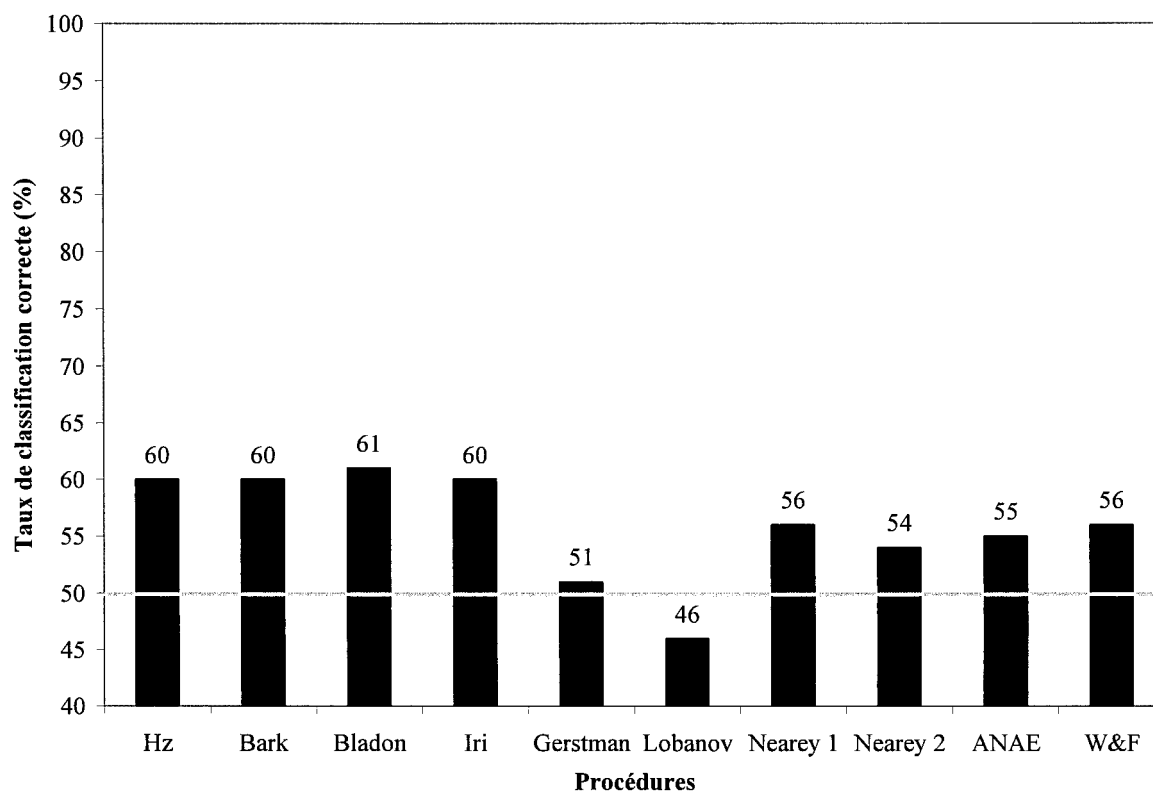


Figure 5 : Taux de classification correcte obtenus par les différentes procédures lorsque F_1 et F_2 sont pris en compte et que les occurrences sont classées d'après le sexe du locuteur qui les a produites. La ligne verte indique le seuil du hasard [une chance sur deux qu'il s'agisse d'un homme].

L'ajout de F_3 aux prédicteurs entraîne une légère augmentation du taux de classification correcte de toutes les procédures considérées (voir Figure 6). Le taux le plus bas se situe ainsi à 52 % pour la procédure de Gerstman (1968) et le taux le plus élevé, à 68 % pour les données brutes en Hertz.

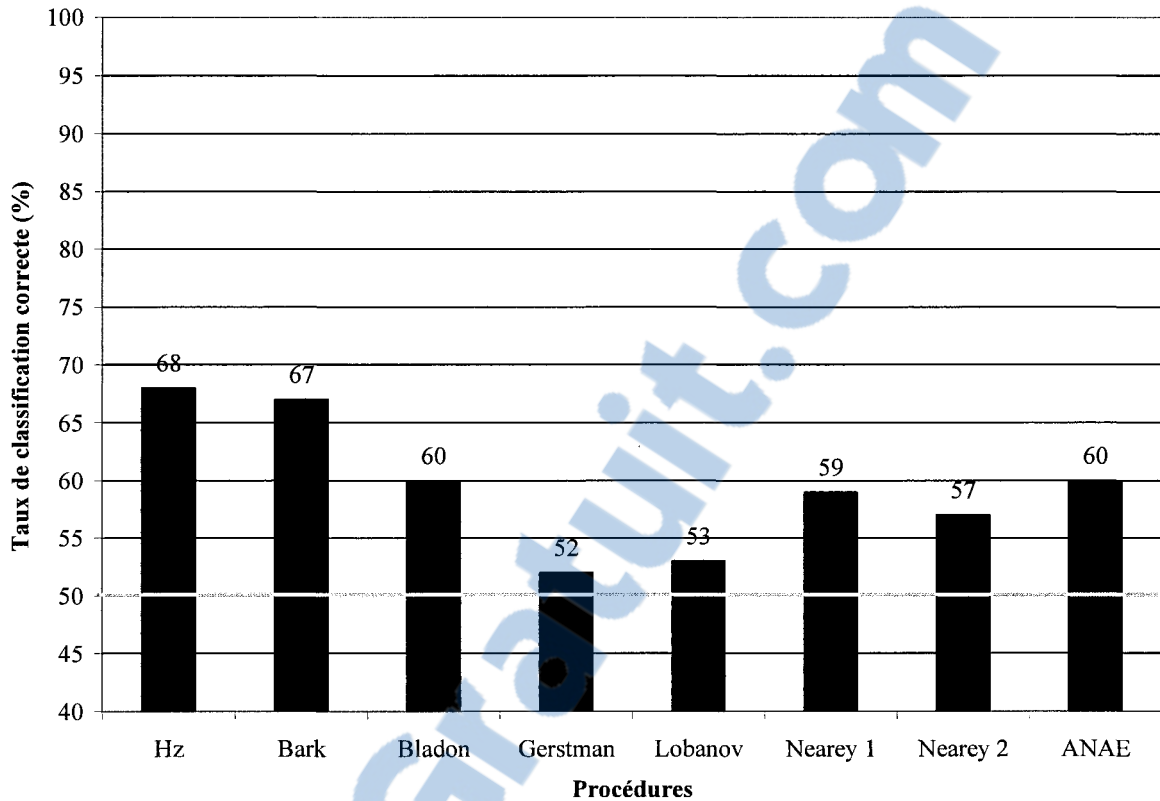


Figure 6 : Taux de classification correcte obtenus par les différentes procédures lorsque F_1 , F_2 et F_3 sont pris en compte et que les occurrences sont classées d'après le sexe du locuteur qui les a produites. La ligne verte indique le seuil du hasard [une chance sur deux qu'il s'agisse d'un homme].

Comme le mentionnent Hillenbrand et Clark (2009 : 1150), « the male–female difference in f_0 is proportionally much larger than the typical differences in formant frequencies ». Il n'est donc pas surprenant qu'avec la prise en compte de la F_0 , la quasi-totalité des occurrences ait été correctement classée en fonction du sexe du locuteur lorsque les données n'étaient pas normalisées ou étaient simplement converties en barks (voir Figure 7). Par ailleurs, la procédure de Labov et coll. (2006) et la *Constant Ratio Hypothesis 1* de Nearey (1977) échouent vraisemblablement à éliminer les différences de F_0 entre les productions féminines et masculines. Les procédures de Bladon et coll. (1984) et d'Iri (1959) parviennent quant à elle à éliminer une partie de ces différences, alors que la procédure de Gerstman (1968), les scores-Z de Lobanov (1971) et la *Constant Ratio Hypothesis 2* de Nearey (1977) les éliminent tout à fait.

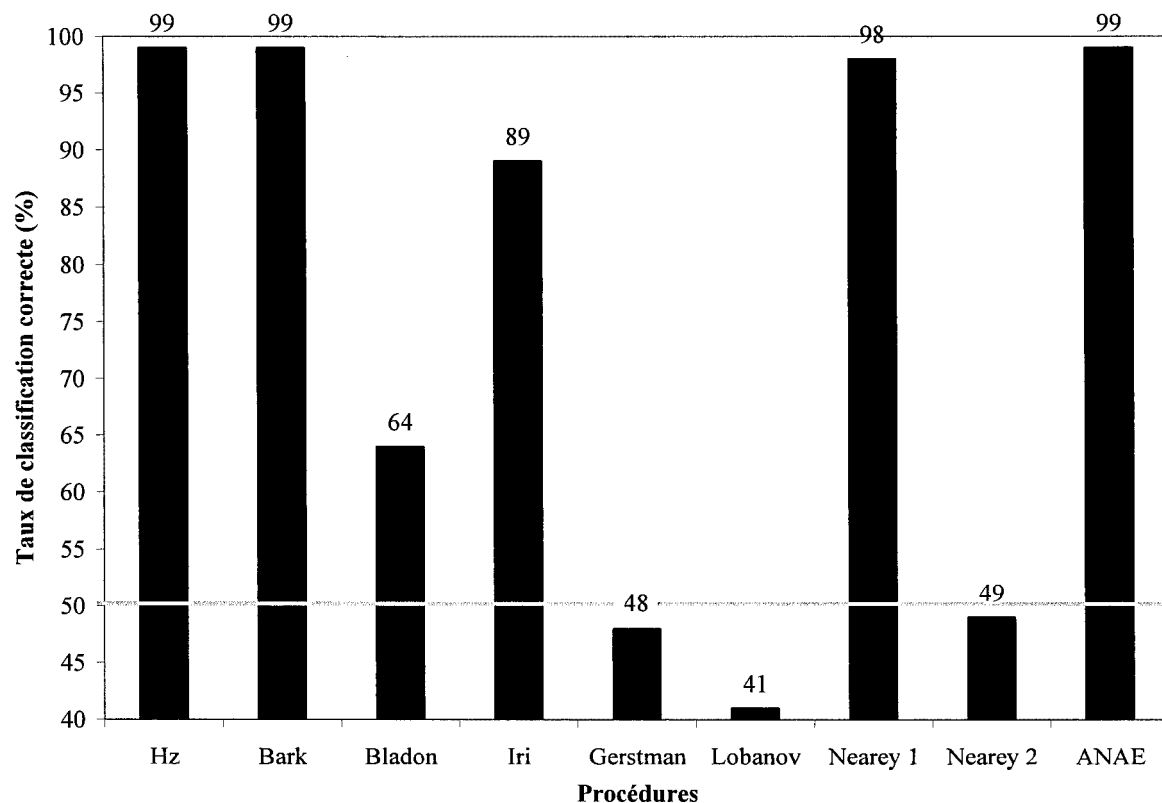


Figure 7 : Taux de classification correcte obtenus par les différentes procédures lorsque la F_0 , F_1 et F_2 sont pris en compte et que les occurrences sont classées d'après le sexe du locuteur qui les a produites. La ligne verte indique le seuil du hasard [une chance sur deux qu'il s'agisse d'un homme].

Ces résultats se maintiennent lorsque F_3 est également pris en compte (voir Figure 8). Quant à l'ensemble de procédures préconisé par Ménard (2002), qui n'était pas présent sur la Figure 7, il ne parvient qu'à éliminer une partie des différences entre les productions féminines et masculines.

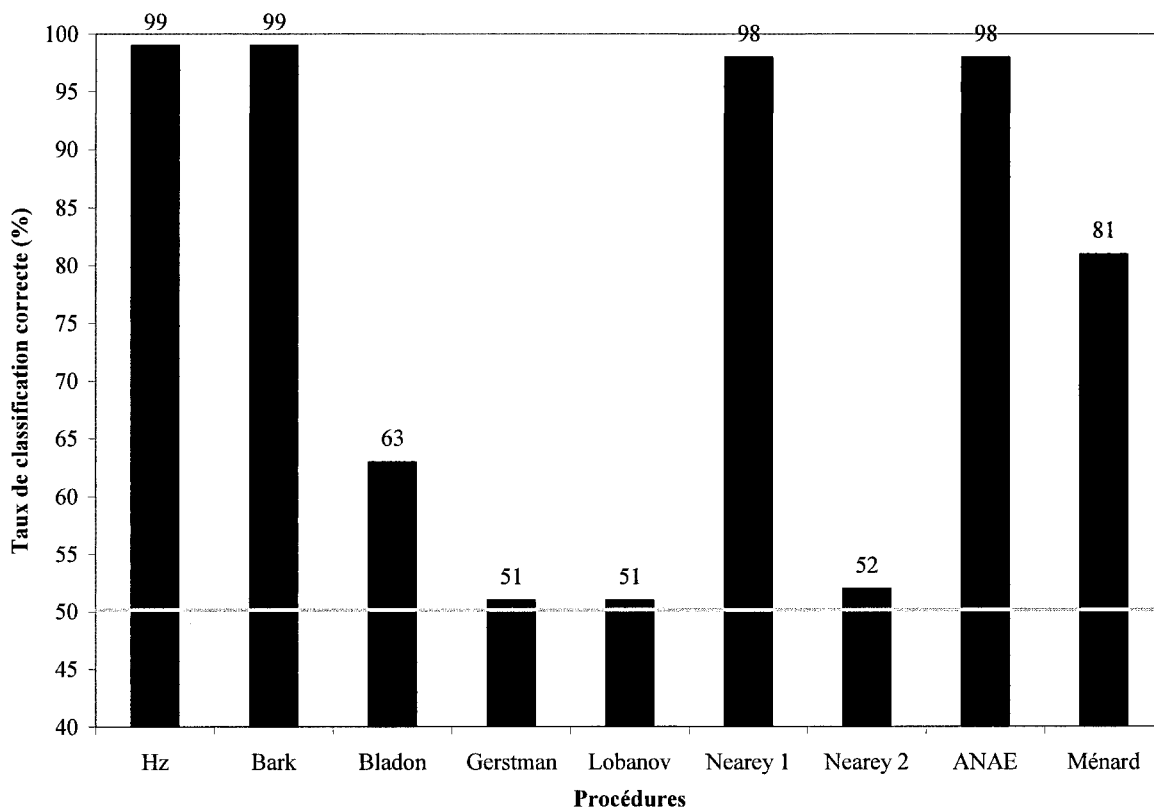


Figure 8 : Taux de classification correcte obtenus par les différentes procédures lorsque la F_0 , F_1 , F_2 et F_3 sont pris en compte et que les occurrences sont classées d'après le sexe du locuteur qui les a produites. La ligne verte indique le seuil du hasard [une chance sur deux qu'il s'agisse d'un homme].

Discussion conclusive

Les résultats obtenus dans la présente étude concordent globalement avec ceux exposés dans la littérature. Tout comme Adank et coll. (2004) et Flynn et Foulkes (2011), nous observons que les procédures les plus efficaces sont *vowel-extrinsic* (dont le calcul intègre des valeurs issues d'autres occurrences que celle à être normalisée), mais *formant-intrinsic* (dont le calcul n'intègre que des valeurs issues d'une même composante spectrale : formant ou fréquence fondamentale). À ce sujet, le cas des deux procédures de Nearey (1977), lorsque la F_0 est prise en compte, est éloquent. La *Constant Ratio Hypothesis 1*, qui est *formant-extrinsic*, ne parvient pas à éliminer les différences entre les productions féminines et masculines, alors que la *Constant Ratio Hypothesis 2*, qui elle, est *formant-intrinsic*, y parvient.

Au-delà de ces différences inhérentes au procédé mathématique utilisé, nous remarquons que l'efficacité des procédures varie selon la nature des indices acoustiques normalisés et leur nombre. Par exemple, la procédure employée par Labov et coll. (2006) dans l'*Atlas of North American English* parvient à réduire l'écart entre les productions féminines et masculines seulement si la F_0 n'est pas prise en compte. Certaines procédures développées avec l'objectif de ne normaliser que F_1 et F_2 ne semblent donc pas adaptées à la normalisation d'autres composantes spectrales, de plus en plus prises en considération dans l'étude acoustique des voyelles.

Nos résultats témoignent également de l'efficacité variable des procédures en fonction du critère d'évaluation envisagé. Ainsi, une procédure comme celle de Bladon et coll. (1984) n'entraîne pas d'amélioration substantielle de la classification en fonction de la classe phonémique, mais elle réduit sensiblement l'écart entre les productions féminines et masculines.

Notre étude laisse cependant un certain nombre de questions en suspens. Tout d'abord, la portée de nos conclusions est limitée, puisque notre échantillon, au demeurant restreint, ne provenait que d'une seule région et d'un seul groupe d'âge. La classe sociale n'était quant à elle contrôlée que par le niveau de scolarité des participants. Qui plus est, notre corpus ne contenait qu'une partie des voyelles du français québécois, produites isolément. Enfin, nous n'avons évalué ces quelques procédures de normalisation qu'à l'égard de deux critères. Or, dans nombre d'études, la plausibilité perceptuelle et la préservation des différences sociolinguistiques⁴⁵ sont aussi des conditions que doivent remplir les procédures de normalisation. Cela dit, même en ce qui concerne la réduction de l'écart entre les productions féminines et masculines, notre étude comporte certaines lacunes. En effet, il n'est pas exclu que la réduction mesurée soit en partie attribuable à d'autres facteurs que les proportions dissemblables des conduits vocaux féminins et masculins. Une part de ces différences peut être d'origine sociale.

⁴⁵ Adank [2003] a notamment évalué sur ce critère les procédures de normalisation qu'elle avait sélectionnées.

En dépit de ces limites, nos résultats indiquent en somme qu'il est désavantageux d'utiliser la procédure d'Iri (1959), que la conversion en barks est sans effet et que les autres procédures sont intéressantes dans une mesure variable. Les procédures de Gerstman (1968) et de Lobanov (1971), de même que la *Constant Ratio Hypothesis 2* de Nearey (1977), s'avèrent les plus efficaces et les plus robustes.

Bibliographie

Adank, Patti (2003), *Vowel Normalization: A Perceptual-Acoustic Study of Dutch Vowels*, Wageningen, Ponsen & Looijen.

Adank, Patti, Roel Smits, et Roeland van Hout (2004), « A comparison of vowel normalization procedures for language variation research », *Journal of the Acoustical Society of America*, 116 (5), p. 3099-3107.

Anderson, Bridget L. (1999), « Source-language transfer and vowel accommodation in the patterning of Cherokee English /ai/ and /oi/ », *American Speech*, 74 (4), p. 339-368.

Bladon, R. Anthony W., et Gunnar Fant (1978), « A two-formant model and the cardinal vowels », *Speech Transmission Laboratory Quarterly Progress and Status Report*, 19 (1), p. 1-8.

Bladon, R. Anthony W., Caroline G. Henton, et J. Brian Pickering (1984), « Towards an auditory theory of speaker normalization », *Language & Communication*, 4 (1), p. 59-69.

Calamai, Silvia (2005), « Intrinsic vowel normalization: Comparing different procedures (data from Tuscan Italian) », *Italian Journal of Linguistics / Rivista di Linguistica*, 17 (2), p. 211-270.

Clopper, Cynthia G. (2009), « Computational methods for normalizing acoustic vowel data for talker differences », *Language and Linguistics Compass*, 3 (6), p. 1430-1442.

Delattre, Pierre (1948), « Un triangle acoustique des voyelles orales du français », *The French Review*, 21 (6), p. 477-484.

Delattre, Pierre, Alvin M. Liberman, Franklin S. Cooper, et Louis J. Gerstman (1952), « An experimental study of the acoustic determinants of vowel color: Observations on one- and two-formant vowels synthesized from spectrographic patterns », *Word*, 8 (3), p. 195-210.

Fabricius, Anne H., Dominic Watt, et Erza Johnson (2009), « A comparison of three speaker-intrinsic vowel formant frequency normalization algorithms for sociophonetics », *Language Variation and Change*, 21 (3), p. 413-435.

- Fant, Gunnar (1975), « Non-uniform vowel normalization », *Speech Transmission Laboratory Quarterly Progress and Status Report*, 16 (2-3), p. 1-19.
- Ferrari Disner, Sandra (1980), « Evaluation of vowel normalization procedures », *Journal of the Acoustical Society of America*, 67 (1), p. 253-261.
- Flynn, Nicholas, et Paul Foulkes (2011), « Comparing vowel formant normalization methods », *Proceedings of the 17th International Congress of Phonetic Sciences*, Hong Kong, p. 683-686.
- Gerstman, Louis J. (1968), « Classification of self-normalized vowels », *IEEE Transactions on Audio and Electroacoustics*, AU-16 (1), p. 78-80.
- Hillenbrand, James M., et Michael J. Clark (2009), « The role of f0 and formant frequencies in distinguishing the voices of men and women », *Attention, Perception, & Psychophysics*, 71 (5), p. 1150-1166.
- Hillenbrand, James M., et Robert T. Gayvert (1993), « Identification of steady-state vowels synthesized from the Peterson and Barney measurements », *Journal of the Acoustical Society of America*, 94 (2), p. 668-674.
- Hillenbrand, James M., Laura A. Getty, Michael J. Clark, et Kimberlee Wheeler (1995), « Acoustic characteristics of American English vowels », *Journal of the Acoustical Society of America*, 97 (5), part. 1, p. 3099-3111.
- Hindle, Donald (1978), « Approaches to vowel normalization in the study of natural speech » dans David Sankoff (dir.), *Linguistic Variation: Models and Methods*, New York, San Francisco et Londres, Academic Press, p. 161-171.
- Iri, Masao (1959), « A mathematical method in phonetics with a special reference to the acoustical structure of Japanese vowels », *Gengo Kenkyu*, (35), p. 23-30.
- Johnson, Keith (2005), « Speaker normalization in speech perception » dans David B. Pisoni et Robert E. Remez (dir.), *The Handbook of Speech Perception*, Oxford, Blackwell Publishing, p. 363-389.
- Joos, Martin (1948), « Acoustic phonetics », *Language*, 24 (2), Language monograph No. 23, p. 1-136.
- Koenig, Winfried (1949), « A new frequency scale for acoustic measurements », *Bell Labs Record*, (27), p. 299-301.
- Labov, William, Sharon Ash, et Charles Boberg (2006), *The Atlas of North American English: Phonetics, Phonology and Sound Change*, Berlin, Mouton de Gruyter.

Lloyd, Richard J. (1890), *Some Researches into the Nature of Vowel-Sound*, Liverpool, Turner and Dunnett.

Lobanov, Boris M. (1971), « Classification of Russian vowels spoken by different speakers », *Journal of the Acoustical Society of America*, 49 (2B), p. 606-608.

Mantakas, Marios, Jean-Luc Schwartz, et Pierre Escudier (1986), « Modèle de prédiction du “deuxième formant effectif” F’2 et application à l’étude de la labialité des voyelles du français », *Actes des 15^{es} Journées d’étude sur la parole*, Aix-en-Provence, p. 157-161.

Miller, James D. (1989), « Auditory-perceptual interpretation of the vowel », *Journal of the Acoustical Society of America*, 85 (5), p. 2114-2134.

Miller, James D., A. Maynard Engbretson, et N. Rao Vemula (1980), « Vowel normalization: Differences between vowels spoken by children, women, and men », *Journal of the Acoustical Society of America*, 68 (suppl. 1), p. S33.

Moore, Brian C. J., et Brian R. Glasberg (1983), « Suggested formulae for calculating auditory-filter bandwidths and excitation patterns », *Journal of the Acoustical Society of America*, 74 (3), p. 750-753.

Ménard, Lucie (2002), *Production et perception des voyelles au cours de la croissance du conduit vocal : variabilité, invariance et normalisation*, thèse de doctorat (linguistique), Institut de la communication parlée, Grenoble.

Ménard, Lucie, Jean-Luc Schwartz, Louis-Jean Boë, Sonia Kandel et Nathalie Vallée (2002), « Auditory normalization of French vowels synthesized by an articulatory model simulating growth from birth to adulthood », *Journal of the Acoustical Society of America*, 111 (4), p. 1892-1905.

Nearey, Terrance M. (1977), *Phonetic Feature Systems for Vowels*, thèse de doctorat (linguistique), University of Connecticut.

Nguyen, Noël, et Robert Espesser (2004), « Méthodes et outils pour l’analyse acoustique des systèmes vocaliques », *Bulletin PFC (Phonologie du français contemporain)*, 3, p. 77-85.

Nordström, Per-Erik, et Björn E. F. Lindblom (1975), « A normalization procedure for vowel formant data », *Proceedings of the 8th International Congress of Phonetic Sciences (unpublished)*, Leeds, article 212.

Potter, Ralph K., et John C. Steinberg (1950), « Toward the specification of speech », *Journal of the Acoustical Society of America*, 22 (6), p. 807-820.

Stevens, Stanley Smith, et John E. Volkman (1940), « The relation of pitch to frequency: A revised scale », *The American Journal of Psychology*, 53 (3), p. 329-353.

- Sussman, Harvey M. (1986), « A neuronal model of vowel normalization and representation », *Brain and language*, 28 (1), p. 12-23.
- Syrdal, Ann K., et Hundrai S. Gopal (1986), « A perceptual model of vowel recognition based on the auditory representation of American English vowels », *Journal of the Acoustical Society of America*, 79 (4), p. 1086-1100.
- Thomas, Erik R. (1997), « A rural/metropolitan split in the speech of Texas Anglos », *Language Variation and Change*, 9 (3), p. 309-332.
- Thomas, Erik R., et Tyler Kendall (2007), *NORM: The Vowel Normalization and Plotting Suite*. Consulté le 16 mars 2011 sur <http://ncslaap.lib.ncsu.edu/tools/norm/index.php>
- Trautmüller, Hartmut (1981), « Perceptual dimension of openness in vowels », *Journal of the Acoustical Society of America*, 69 (5), p. 1465-1475.
- Volín, Jan, et David Studenovský (2007), « Normalization of czech vowels from continuous read texts », *Proceedings of the 16th International Congress of Phonetic Sciences*, Saarbrücken, p. 185-190.
- Watt, Dominic, et Anne H. Fabricius (2002), « Evaluation of a technique for improving the mapping of multiple speakers' vowel spaces in the F1~F2 plane », *Leeds Working Papers in Linguistics and Phonetics*, 9, p. 159-173.
- Wolfram, Walt, Erik R. Thomas, et Elaine W. Green (2011), « The regional context of earlier African American speech: Evidence for reconstructing the development of AAVE », *Language*, 29 (3), p. 315-355.
- Zwicker, Eberhard (1961), « Subdivision of the audible frequency range into critical bands (Frequenzgruppen) », *Journal of the Acoustical Society of America*, 33 (2), p. 248.

Annexe 1 – Formules et algorithmes

L'échelle des barks

$$1) \quad f_{bark} = \frac{26,81 \cdot f_{Hz}}{1960 \cdot f_{Hz}} - 0,53$$

$$f'_{bark} = f_{bark} + 0,15 \cdot (2 - f_{bark}) \quad \text{si} \quad f_{bark} < 2$$

$$f'_{bark} = f_{bark} + 0,22 \cdot (f_{bark} - 20,1) \quad \text{si} \quad f_{bark} > 20,1$$

où : f_{Hz} correspond à une fréquence donnée en Hertz;

f_{bark} correspond à cette fréquence en barks (préliminaire);

f'_{bark} correspond à cette fréquence en barks (finale).

L'échelle des barks normalisée (Bladon et coll., 1984)

$$2) \quad f''_{bark} = f'_{bark} + 1 \quad \text{si} \quad \text{locuteur masculin}$$

$$f''_{bark} = f'_{bark} \quad \text{si} \quad \text{locuteur féminin}$$

où : f'_{bark} correspond à une fréquence en barks d'après la formule 1;

f''_{bark} correspond à cette fréquence en barks normalisés.

Un ensemble d'objets psycho-acoustiques (Ménard, 2002)

Distance tonotopique entre le premier formant et la fréquence fondamentale :

$$3) \quad F_1 - F_0$$

où les fréquences proviennent de la même occurrence et sont exprimées en barks d'après la formule 1.

Distance tonotopique entre le premier et le second formant :

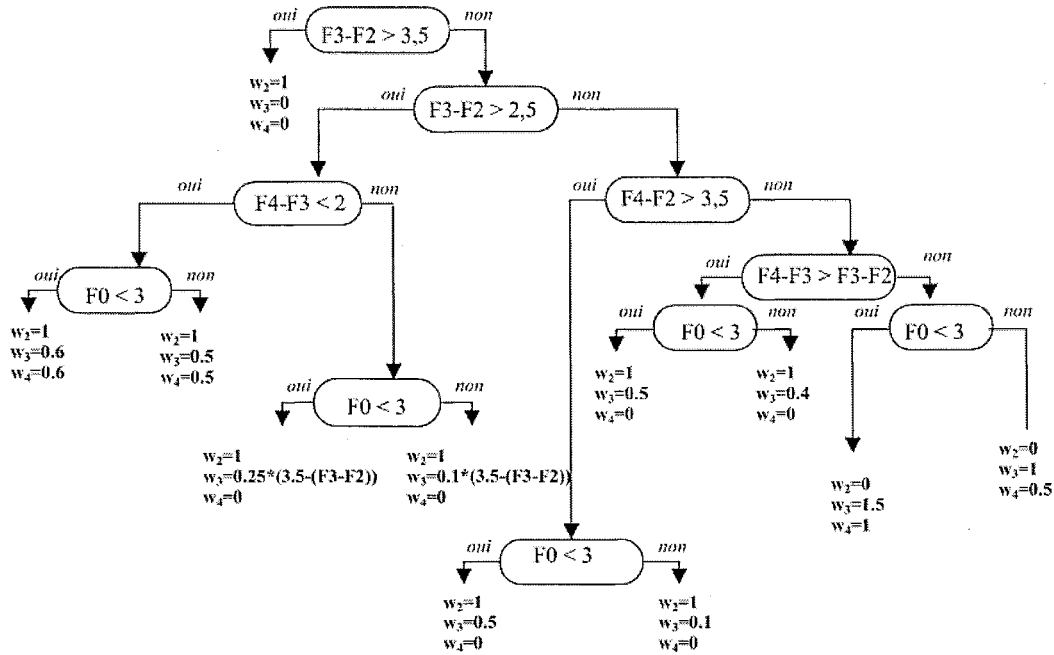
$$4) \quad F_2 - F_1$$

où les fréquences proviennent de la même occurrence et sont exprimées en barks d'après la formule 1.

Second formant effectif :

$$5) \quad F'_2 = \frac{F_2 \cdot w_2 + F_3 \cdot w_3 + F_4 \cdot w_4}{w_2 + w_3 + w_4}$$

où toutes les variables proviennent de la même occurrence, sont exprimées en barks d'après la formule 1 et sont calculées à l'aide de l'algorithme suivant, tiré de Ménard (2002) :



La procédure d'Iri (1959)

6)
$$N_x = \log \left(\frac{F_x}{\sqrt[3]{F_1 \cdot F_2 \cdot F_3}} \right)$$

où : N_x correspond à la fréquence normalisée du x^e formant;
 F_x correspond à la fréquence en Hertz du x^e formant;
 toutes les variables proviennent de la même occurrence.

La procédure de Gerstman (1968)

7)
$$(N_x)_i = 999 \cdot \frac{(F_x)_i - \min(F_x)_i}{\max(F_x)_i - \min(F_x)_i}$$

où : $(N_x)_i$ correspond à la fréquence normalisée du x^e formant d'une occurrence donnée du locuteur i ;
 $(F_x)_i$ correspond à la fréquence en Hertz du x^e formant de cette occurrence;
 $\min(F_x)_i$ correspond à la plus petite fréquence du x^e formant relevée pour le locuteur i , toutes occurrences confondues;

$\max(F_x)_i$ correspond à la fréquence la plus élevée du x^e formant relevée pour le locuteur i , toutes occurrences confondues.

Les scores-Z (Lobanov, 1971)

$$8) \quad (N_x)_i = \frac{(F_x)_i - \overline{(F_x)}_i}{(s_x)_i}$$

où : $(N_x)_i$ correspond à la fréquence normalisée du x^e formant d'une occurrence donnée du locuteur i ;
 $(F_x)_i$ correspond à la fréquence en Hertz du x^e formant de cette occurrence;
 $\overline{(F_x)}_i$ correspond à la moyenne du x^e formant du locuteur i , toutes occurrences confondues;
 $(s_x)_i$ correspond à l'écart-type du x^e formant du locuteur i , toutes occurrences confondues.

La Constant Ratio Hypothesis 1 (Nearey, 1977)

$$9) \quad (N_x)_i^z = \ln(F_x)_i^z - \frac{\sum \ln(F_{1-2})_i}{2 \cdot n_i}$$

où : $(N_x)_i^z$ correspond à la fréquence normalisée du x^e formant de l'occurrence z du locuteur i ;
 $(F_x)_i^z$ correspond à la fréquence en Hertz du x^e formant de l'occurrence z du locuteur i ;
 $(F_{1-2})_i$ correspond à la fréquence en Hertz du premier ou du second formant de l'une ou l'autre des occurrences du locuteur i ;
 n_i correspond au nombre d'occurrences du locuteur i .

La Constant Ratio Hypothesis 2 (Nearey, 1977)

$$10) \quad (N_x)_i^z = \ln(F_x)_i^z - \frac{\sum \ln(F_x)_i}{n_i}$$

où : $(N_x)_i^z$ correspond à la fréquence normalisée du x^e formant de l'occurrence z du locuteur i ;
 $(F_x)_i^z$ correspond à la fréquence en Hertz du x^e formant de l'occurrence z du locuteur i ;
 $(F_x)_i$ correspond à la fréquence en Hertz du x^e formant de l'une ou l'autre des occurrences du locuteur i ;
 n_i correspond au nombre d'occurrences du locuteur i .

La procédure de l'Atlas of North American English (Labov et coll., 2006)

$$11) \quad (N_x)_i = k \cdot (F_x)_i$$

$$k = \exp(G - S)$$

$$G = \frac{\sum \ln(F)}{m \cdot n}$$

$$S = \frac{\sum \ln(F)_i}{m \cdot n_i}$$

où : $(N_x)_i$ correspond à la fréquence normalisée du x^e formant d'une occurrence donnée du locuteur i ;
 $(F_x)_i$ correspond à la fréquence en Hertz du x^e formant de cette occurrence;
 (F) correspond à la fréquence en Hertz de l'un ou l'autre des formants de l'une ou l'autre des occurrences de l'un ou l'autre des locuteurs;
 $(F)_i$ correspond à la fréquence en Hertz de l'un ou l'autre des formants de l'une ou l'autre des occurrences du locuteur i ;
 n correspond au nombre total d'occurrences, tous locuteurs confondus;
 n_i correspond au nombre d'occurrences du locuteur i ;
 m correspond au nombre de formants considérés.

La procédure de Watt et Fabricius (2002)

$$12) \quad (N_x)_i = \frac{(F_x)_i}{(S_x)_i}$$

$$(S_x)_i = \frac{(\overline{F_x})_i^{/i/} + (\overline{F_x})_i^{/u/} + (\overline{F_x})_i^{/a/}}{3}$$

- où :
- $(N_x)_i$ correspond à la fréquence normalisée du x^e formant d'une occurrence donnée du locuteur i ;
 - $(F_x)_i$ correspond à la fréquence en Hertz du x^e formant de cette occurrence;
 - $(\overline{F_x})_i^{/i/}$ correspond à la moyenne des x^e formants de la voyelle /i/ du locuteur i ;
 - si $x = 1$, $(\overline{F_x})_i^{/u/}$ correspond au F_1 moyen de la voyelle /i/ du locuteur i ;
 - si $x = 2$, $(\overline{F_x})_i^{/u/}$ correspond au F_2 moyen de la voyelle /u/ du locuteur i ;
 - si $x = 1$, $(\overline{F_x})_i^{/a/}$ correspond au F_1 moyen de la voyelle /a/ du locuteur i ;
 - si $x = 2$, $(\overline{F_x})_i^{/a/}$ correspond au F_2 moyen des voyelles /i/ et /u/ du locuteur i ;
 - $x \in \{1,2\}$.

Annexe 3 : Liste des phrases et des mots cibles du corpus

(i#)	(iK)	(iR)
Ça pue quand il <u>chie</u> .	Il allait au <u>Bic</u> .	Tu lui fais la <u>bise</u> .
Il cueillait du <u>gui</u> .	On lui voit la <u>bite</u> .	Il boit du bon <u>kir</u> .
Il s'agit d'une <u>pie</u> .	J'ai vendu un <u>guide</u> .	On va faire une <u>pige</u> .
Tu dis le mot « <u>qui</u> ».	Ces choses viennent en <u>kit</u> .	Georgette est la <u>pire</u> .
J'ai gagné des <u>vies</u> .	Il grimpe ce haut <u>pic</u> .	Cette couleur est <u>vive</u> .

(y#)	(yK)	(yR)
J'ai compté un <u>but</u> .	Yvan fend des <u>bûches</u> .	Les rats craignent les <u>buses</u> .
Les jeunes parlent de <u>cul</u> .	J'ai monté une <u>butte</u> .	Hugo suit une <u>cure</u> .
J'aime la bière en <u>fût</u> .	Ils empilent des <u>cubes</u> .	On nettoie la <u>cuve</u> .
Ta plaie fait du <u>pus</u> .	J'ai créé cette <u>pub</u> .	Elle descend en <u>luge</u> .
J'ai perdu la <u>vue</u> .	Mon chat a des <u>puces</u> .	Luc est bien trop <u>pur</u> .

(u#)	(uK)	(uR)
Elles jouent dans la <u>boue</u> .	Ma chèvre aime ton <u>bouc</u> .	Il faudrait qu'on <u>bouge</u> .
Elle lui masse le <u>cou</u> .	Mon fréro me <u>boude</u> .	L'ourson perd sa <u>bourre</u> .
Ça l'a rendu <u>fou</u> .	J'ai très mal au <u>coude</u> .	J'ai fini mes <u>cours</u> .
Ce mets a bon <u>goût</u> .	J'en prendrais une <u>coupe</u> .	Les couturières <u>cousent</u> .
Vicky a des <u>poux</u> .	Il tombait des <u>gouttes</u> .	Ma grand-mère me <u>couve</u> .

Annexe 4 : Script PRAAT vérifiant que les frontières sont bien posées à un passage par zéro

Script intégré au menu contextuel du PRAAT lorsqu'un fichier sonore et une grille de transcription sont sélectionnés :

```
# Formulaire de sélection du palier de transcription à traiter

    beginPause ("Options")
    integer ("Numero du palier", 0)
    comment ("* Inscrivez 0 comme numéro de palier pour traiter la
TextGrid en entier.")
    comment ("Souhaitez-vous que le script déplace les
points/frontières ne passant pas par zéro au")
    comment ("passage par zéro le plus proche?")
    boolean ("Oui", 0)
    endPause ("Poursuivre", 1)

nom_du_fichier_sonore$ = selected$ ("Sound")
nom_de_la_textgrid$ = selected$ ("TextGrid")

    if nom_du_fichier_sonore$ <> nom_de_la_textgrid$
        beginPause ("ATTENTION!")
        comment ("La TextGrid et le fichier sonore sélectionnés ne
portent pas le même nom!")
        comment ("Si ces deux fichiers sont bien associés, cliquez
sur Poursuivre.")
        endPause ("Poursuivre", 1)
    elsif nom_du_fichier_sonore$ <> nom_de_la_textgrid$
    endif

Edit
    editor TextGrid 'nom_de_la_textgrid$'
    Show analyses... no no no no no 40
    endeditor

# Pour effectuer les passages par zéro

if oui = 1

# Écriture de l'entête du rapport dans la fenêtre Info

clearinfo
print Rapport du script'newline$'Voici les points/frontières qui ont été
déplacées (si vide, aucune frontière déplacée) :'newline$'newline$'
print Palier'tab$'Temps initial'tab$'Temps post-script'newline$'

# Sélection de la TextGrid

select TextGrid 'nom_de_la_textgrid$'

# Procédure si un seul palier à traiter
```



```

    if numero_du_palier <> 0
        intervalles_ou_points = Is interval tier... numero_du_palier
# Procédure s'il s'agit d'un palier d'intervalles

        if intervalles_ou_points = 1
# Placement du curseur sur le palier voulu

        nom_palier$ = Get tier name... numero_du_palier
        nom_fichier_palier$ = "TextGrid " + nom_palier$
        nb_intervalles = Get number of intervals... numero_du_palier
        editor TextGrid 'nom_de_la_textgrid$'
        Extract entire selected tier
        endeditor
        palier_actu$ = selected$ ()
        Remove
        while palier_actu$ <> nom_fichier_palier$
            editor TextGrid 'nom_de_la_textgrid$'
            Select next tier
            Extract entire selected tier
            endeditor
        palier_actu$ = selected$ ()
        Remove
        endwhile

# Déplacement de toutes les frontières au passage par zéro le plus proche
/ Écriture des déplacements dans le rapport

        select TextGrid 'nom_de_la_textgrid$'
        for num_intervalle from 2 to nb_intervalles
            debut_intervalle = Get start point...
numero_du_palier num_intervalle
            editor TextGrid 'nom_de_la_textgrid$'
            Move cursor to... 'debut_intervalle'
            temps_ini = Get cursor
            Move to nearest zero crossing
            temps_post = Get cursor
            if temps_ini <> temps_post
                print
'nom_palier$' 'tab$' 'temps_ini' 'tab$' 'temps_post' 'newline$'
            endif
        endeditor
        endfor
        editor TextGrid 'nom_de_la_textgrid$'
        Close
        endeditor

# Procédure s'il s'agit d'un palier de points

        elsif intervalles_ou_points = 0
# Placement du curseur sur le palier voulu

        nom_palier$ = Get tier name... numero_du_palier
        nom_fichier_palier$ = "TextGrid " + nom_palier$
        nb_points = Get number of points... numero_du_palier

```

```

        editor TextGrid 'nom_de_la_textgrid$'
        Extract entire selected tier
        endeditor
    palier_actu$ = selected$ ()
    Remove
        while palier_actu$ <> nom_fichier_palier$
            editor TextGrid 'nom_de_la_textgrid$'
            Select next tier
            Extract entire selected tier
            endeditor
            palier_actu$ = selected$ ()
            Remove
        endwhile

# Déplacement de tous les points au passage par zéro le plus proche /
# Écriture des déplacements dans le rapport

        select TextGrid 'nom_de_la_textgrid$'
        for num_point from 1 to nb_points
            temps = Get time of point... numero_du_palier num_point
            editor TextGrid 'nom_de_la_textgrid$'
            Move cursor to... temps
            temps_ini = Get cursor
            Move to nearest zero crossing
            temps_post = Get cursor
            if temps_ini <> temps_post
                print
'nom_palier$' 'tab$' 'temps_ini' 'tab$' 'temps_post' 'newline$'
            endif
            endeditor
        endfor
        editor TextGrid 'nom_de_la_textgrid$'
        Close
        endeditor
    endif

# Procédure si tous les paliers à traiter

    elsif numero_du_palier = 0
        nb_paliers = Get number of tiers
        select TextGrid 'nom_de_la_textgrid$'
        for num_palier from 1 to nb_paliers
            nom_palier$ = Get tier name... num_palier
            intervalles_ou_points = Is interval tier... num_palier

# Procédure s'il s'agit d'un palier d'intervalles
# Déplacement de toutes les frontières au passage par zéro le plus proche
# Écriture des déplacements dans le rapport

            if intervalles_ou_points = 1
                nb_intervalles = Get number of intervals... num_palier
                for num_intervalle from 2 to nb_intervalles
                    debut_intervalle = Get start point... num_palier
num_intervalle
                    editor TextGrid 'nom_de_la_textgrid$'
                    Move cursor to... 'debut_intervalle'

```

```

        temps_ini = Get cursor
        Move to nearest zero crossing
        temps_post = Get cursor
        if temps_ini <> temps_post
            print
'nom_palier$''tab$''temps_ini''tab$''temps_post''newline$'
            endif
        endeditor
        select TextGrid 'nom_de_la_textgrid$'
    endfor
    editor TextGrid 'nom_de_la_textgrid$'
    Select next tier
    endeditor

# Procédure s'il s'agit d'un palier de points
# Déplacement de tous les points au passage par zéro le plus proche /
Écriture des déplacements dans le rapport

        elsif intervalles_ou_points = 0
        nb_points = Get number of points... num_palier
        for num_point from 1 to nb_points
            temps = Get time of point... num_palier num_point
            editor TextGrid 'nom_de_la_textgrid$'
            Move cursor to... temps
            temps_ini = Get cursor
            Move to nearest zero crossing
            temps_post = Get cursor
            if temps_ini <> temps_post
                print
'nom_palier$''tab$''temps_ini''tab$''temps_post''newline$'
                endif
            endeditor
            select TextGrid 'nom_de_la_textgrid$'
        endfor
        editor TextGrid 'nom_de_la_textgrid$'
        Select next tier
        endeditor
    endif
endfor
editor TextGrid 'nom_de_la_textgrid$'
Close
endeditor
endif

# Pour signaler les points/frontières ne passant pas par zéro

elsif oui = 0

# Écriture de l'entête du rapport dans la fenêtre Info

clearinfo
print Rapport du script'newline$'Voici les points/frontières qui ne
passent pas par zéro (si vide, passent déjà tous par zéro)
:'newline$''newline$'
print Palier'tab$'Temps'tab$'Passage par zéro le plus proche'newline$'

# Sélection de la TextGrid

```

```

select TextGrid 'nom_de_la_textgrid$'

# Procédure si un seul palier à traiter

    if numero_du_palier <> 0
        intervalles_ou_points = Is interval tier... numero_du_palier

# Procédure s'il s'agit d'un palier d'intervalles

    if intervalles_ou_points = 1

# Placement du curseur sur le palier voulu

        nom_palier$ = Get tier name... numero_du_palier
        nom_fichier_palier$ = "TextGrid " + nom_palier$
        nb_intervalles = Get number of intervals... numero_du_palier
        editor TextGrid 'nom_de_la_textgrid$'
        Extract entire selected tier
        endeditor
        palier_actu$ = selected$ ()
        Remove
        while palier_actu$ <> nom_fichier_palier$
            editor TextGrid 'nom_de_la_textgrid$'
            Select next tier
            Extract entire selected tier
            endeditor
            palier_actu$ = selected$ ()
            Remove
        endwhile

# Signalement des frontières ne passant pas par zéro / Écriture des
déplacements dans le rapport

        select TextGrid 'nom_de_la_textgrid$'
            for num_intervalle from 2 to nb_intervalles
                debut_intervalle = Get start point...
numero_du_palier num_intervalle
                editor TextGrid 'nom_de_la_textgrid$'
                Move cursor to... 'debut_intervalle'
                temps_ini = Get cursor
                Move cursor to nearest zero crossing
                temps_post = Get cursor
                if temps_ini <> temps_post
                    print
'nom_palier$' 'tab$' 'temps_ini' 'tab$' 'temps_post' 'newline$'
                    endif
                endeditor
            endfor
        editor TextGrid 'nom_de_la_textgrid$'
        Close
        endeditor

# Procédure s'il s'agit d'un palier de points

        elsif intervalles_ou_points = 0

```

```

# Placement du curseur sur le palier voulu

nom_palier$ = Get tier name... numero_du_palier
nom_fichier_palier$ = "TextGrid " + nom_palier$
nb_points = Get number of points... numero_du_palier
    editor TextGrid 'nom_de_la_textgrid$'
    Extract entire selected tier
    endeditor
palier_actu$ = selected$ ()
Remove
    while palier_actu$ <> nom_fichier_palier$
        editor TextGrid 'nom_de_la_textgrid$'
        Select next tier
        Extract entire selected tier
        endeditor
        palier_actu$ = selected$ ()
        Remove
    endwhile

# Signalement des points ne passant pas par zéro / Écriture des
déplacements dans le rapport

select TextGrid 'nom_de_la_textgrid$'
for num_point from 1 to nb_points
    temps = Get time of point... numero_du_palier num_point
    editor TextGrid 'nom_de_la_textgrid$'
    Move cursor to... temps
    temps_ini = Get cursor
    Move cursor to nearest zero crossing
    temps_post = Get cursor
    if temps_ini <> temps_post
        print
'nom_palier$''tab$''temps_ini''tab$''temps_post''newline$'
    endif
    endeditor
endfor
editor TextGrid 'nom_de_la_textgrid$'
Close
endeditor
endif

# Procédure si tous les paliers à traiter

elsif numero_du_palier = 0
nb_paliers = Get number of tiers
select TextGrid 'nom_de_la_textgrid$'
for num_palier from 1 to nb_paliers
    nom_palier$ = Get tier name... num_palier
    intervalles_ou_points = Is interval tier... num_palier

# Procédure s'il s'agit d'un palier d'intervalles
# Signalement de toutes les frontières ne passant pas par zéro / Écriture
des déplacements dans le rapport

    if intervalles_ou_points = 1
        nb_intervalles = Get number of intervals... num_palier
        for num_intervalle from 2 to nb_intervalles

```

```

                                debut_intervalle = Get start point... num_palier
num_intervalle
                                editor TextGrid 'nom_de_la_textgrid$'
                                Move cursor to... 'debut_intervalle'
                                temps_ini = Get cursor
                                Move cursor to nearest zero crossing
                                temps_post = Get cursor
                                    if temps_ini <> temps_post
                                        print
'nom_palier$'tab$'temps_ini'tab$'temps_post'newline$'
                                        endif
                                endeditor
                                select TextGrid 'nom_de_la_textgrid$'
                                endfor
                                editor TextGrid 'nom_de_la_textgrid$'
                                Select next tier
                                endeditor

# Procédure s'il s'agit d'un palier de points
# Signalement de tous les points ne passant pas par zéro / Écriture des
déplacements dans le rapport

                                elsif intervalles_ou_points = 0
                                nb_points = Get number of points... num_palier
                                for num_point from 1 to nb_points
                                    temps = Get time of point... num_palier num_point
                                    editor TextGrid 'nom_de_la_textgrid$'
                                    Move cursor to... temps
                                    temps_ini = Get cursor
                                    Move cursor to nearest zero crossing
                                    temps_post = Get cursor
                                        if temps_ini <> temps_post
                                            print
'nom_palier$'tab$'temps_ini'tab$'temps_post'newline$'
                                        endif
                                    endeditor
                                    select TextGrid 'nom_de_la_textgrid$'
                                    endfor
                                    editor TextGrid 'nom_de_la_textgrid$'
                                    Select next tier
                                    endeditor
                                endif
                                endfor
                                editor TextGrid 'nom_de_la_textgrid$'
                                Close
                                endeditor
                            endif
                        endif

# Écriture de la clôture du rapport dans la fenêtre Info

print 'newline$'Fin du rapport

if oui = 1
    beginPause ("Prenez note que...")
    comment ("Les frontières ont été déplacées. N'oubliez pas de
sauvegarder votre TextGrid!")

```

```
        endPause ("Terminer", 1)
    elsif oui = 0
        beginPause ("Prenez note que...")
        comment ("Le rapport vous signale les frontières ne passant pas par
zéro.")
        comment ("Celles-ci n'ont pas été déplacées.")
        endPause ("Terminer", 1)
    endif
```

Annexe 5 : Script PRAAT relevant les mesures à chaque occurrence

Script à intégrer en touche de raccourcis dans la fenêtre d'édition de grille de transcription de PRAAT⁴⁶ :

```
nom_son$ = selected$ ("Sound")
nom_tg$ = selected$ ("TextGrid")
beg = Get starting point of interval
end = Get end point of interval
Select... beg end

duree = Get selection length

phone$ = Get label of interval

Select previous tier
mot$ = Get label of interval
mot$ = replace_regex$ (mot$, "_.*", "", 1)
mot$ = replace_regex$ (mot$, "(.)", "\\L\\1", 1)

if mid$ (nom_son$, 2, 1) = "_"
    nom_son$ = right$ (nom_son$, 5)
endif

region$ = left$ (nom_son$, 1)
sexe$ = mid$ (nom_son$, 2, 1)

contexte$ = right$ (phone$, 1)
voyelle$ = left$ (phone$, 1)

if contexte$ = "#"
    variante$ = "#"
elseif contexte$ = "R" or contexte$ = "v" or contexte$ = "z" or contexte$
= "Z"
    variante$ = "R"
else
    variante$ = "K"
endif

if left$ (mot$, 1) = "c" or left$ (mot$, 1) = "q"
    if left$ (mot$, 2) = "ch"
        gauche$ = "S"
    else
        gauche$ = "k"
    endif
else
    gauche$ = left$ (mot$, 1)
```

⁴⁶ Ce script a été conçu spécifiquement en fonction des conventions de notation adoptées dans le cadre de notre projet et des mots contenus dans notre corpus. Ainsi, le nom de fichier de la grille de transcription doit commencer par le pseudonyme du locuteur (par exemple : *CFCBB*), commençant lui-même par une lettre indiquant la région d'origine du locuteur, suivie d'une lettre indiquant son sexe. La grille de transcription doit comporter deux paliers. Sur le premier, le mot cible est segmenté et est étiqueté orthographiquement. Sur le second, la voyelle cible est segmentée et étiquetée d'après sa classe vocalique.


```

endif

#CHANGER LE RÉPERTOIRE EN FONTION DU LOCUTEUR

fichier_out$ = "C:\Users\carolinesigouin\Documents\CFCBB.txt"

titre$ = newline$
titre$ >> 'fichier_out$'

entete0$ = nom_son$ + tab$ + region$ + tab$ + sexe$ + tab$ + mot$ + tab$
+ gauche$ + tab$ + voyelle$ + tab$ + contexte$ + tab$ + variante$ +
newline$
entete0$ >> 'fichier_out$'

entete1$ = "début (s)" + tab$ + "fin (s)" + tab$ + "durée (s)" + newline$
entete1$ >> 'fichier_out$'

tempo1$ = "'beg:5'" + tab$ + "'end:5'" + tab$ + "'duree:5'" + newline$
tempo1$ = replace$(tempo1$, ".", ",", 0)
tempo1$ >> 'fichier_out$'

point_analyse1 = ((duree*25)/100)+beg
point_analyse2 = ((duree*50)/100)+beg
point_analyse3 = ((duree*75)/100)+beg

Move cursor to... point_analyse1
fzero = Get pitch
fun = Get first formant
fdeux = Get second formant
ftrois = Get third formant
fquatre = Get fourth formant
fcinq = Get formant... 5
fsix = Get formant... 6

tempo2$ = "25" + tab$ + "'fzero:0'" + tab$ + "'fun:0'" + tab$ +
"'fdeux:0'" + tab$ + "'ftrois:0'" + tab$ + "'fquatre:0'" + tab$ +
"'fcinq:0'" + tab$ + "'fsix:0'" + newline$
tempo2$ = replace$(tempo2$, "--undefined--", "", 0)
tempo2$ = replace_regex$(tempo2$, "\t+$", "", 0)
tempo2$ >> 'fichier_out$'

Move cursor to... point_analyse2
fzero = Get pitch
fun = Get first formant
fdeux = Get second formant
ftrois = Get third formant
fquatre = Get fourth formant
fcinq = Get formant... 5
fsix = Get formant... 6

tempo2$ = "50" + tab$ + "'fzero:0'" + tab$ + "'fun:0'" + tab$ +
"'fdeux:0'" + tab$ + "'ftrois:0'" + tab$ + "'fquatre:0'" + tab$ +
"'fcinq:0'" + tab$ + "'fsix:0'" + newline$
tempo2$ = replace$(tempo2$, "--undefined--", "", 0)
tempo2$ = replace_regex$(tempo2$, "\t+$", "", 0)
tempo2$ >> 'fichier_out$'

```

```
Move cursor to... point_analyse3
```

```
fzero = Get pitch
```

```
fun = Get first formant
```

```
fdeux = Get second formant
```

```
ftrois = Get third formant
```

```
fquatre = Get fourth formant
```

```
fcing = Get formant... 5
```

```
fsix = Get formant... 6
```

```
tempo2$ = "75" + tab$ + "'fzero:0'" + tab$ + "'fun:0'" + tab$ +  
"'fdeux:0'" + tab$ + "'ftrois:0'" + tab$ + "'fquatre:0'" + tab$ +  
"'fcing:0'" + tab$ + "'fsix:0'" + newline$
```

```
tempo2$ = replace$(tempo2$, "--undefined--", "", 0)
```

```
tempo2$ = replace_regex$(tempo2$, "\t+$", "", 0)
```

```
tempo2$ >> 'fichier_out$'
```


Annexe 6 : Script Perl calculant les scores-Z

Format de la base de données tabulée :

Locuteur	Sexe	Variante	Clé de tri	F0	F1	F2	F3	F4
CFCBB	F	i#	1-25	224	401	2216	3368	3895
CFCBB	F	i#	1-50	219	270	2616	3741	4331
CFCBB	F	i#	1-75	243	341	2292	3571	4240
CFCBB	F	i#	2-25	243	372	2397	3651	4205

...

Script :

```
#!/usr/bin/perl

# Modules utilisés
use strict;
use Math::Trig;
use Statistics::Descriptive;
use utf8;

# Déclaration des variables
my $bd_initiale;
my @bd1;
my @bd2;
my @bd3;
my @bd4;
my @bd5;
my @bd6;
my $bd4;
my $bd6;
my $titres;
my $ligne1;
my $ligne2;
my $ligne3;
my $ligne4;
my $ligne5;
my $ligne6;
my @ligne1;
my @ligne2;
my @ligne3;
my @ligne4;
my @ligne5;
my @ligne6;
my @premier1;
my @premiere2;
my $locuteur_ref;
my $voyelle_ref;
my $stat;
my $intermed1;
my $intermed2;
my @f0;
my @f1;
my @f2;
```

```

my @f3;
my @f4;
my $moyf0;
my $moyf1;
my $moyf2;
my $moyf3;
my $moyf4;
my $etf0;
my $etf1;
my $etf2;
my $etf3;
my $etf4;

# Demande la base de données à traiter
print "Inscrivez le chemin d'accès et le nom de fichier (avec l'extension
.txt) de la base de données à normaliser puis appuyez sur
ENTER\.\n(Exemple :
C:\\Users\\votrenom\\Documents\\mesures_voyelles\\.txt)\n";
$bd_initiale = <STDIN>;
$bd_initiale =~ s/\\\\/\\/g;

# Ouverture de la base de données
open (BD, "< $bd_initiale") or die "Impossible d'ouvrir la base de
données.\n";
@bd1 = <BD>;
close BD;

# Traitement de la ligne de titre
$titres = shift(@bd1);
chomp ($titres);
print $titres, "\n";

# Tri de la base de données
foreach $ligne1 (@bd1){ # Interspersion des colonnes 3 et 1
    chomp ($ligne1);
    $ligne1 =~ s/,/\\/g; # Conversion des virgules en points
    @ligne1 = split("\t", $ligne1, -1);
    $intermed1 = $ligne1[2];
    $ligne1[2] = $ligne1[0];
    $ligne1[0] = $intermed1;
    $ligne1 = join("\t", @ligne1);
    push (@bd2, $ligne1);
}
@bd2 = sort(@bd2); # Tri en fonction de la voyelle
foreach $ligne2 (@bd2){ # Interspersion des colonnes 3 et 1
    chomp ($ligne2);
    $ligne2 =~ s/,/\\/g; # Conversion des virgules en points
    @ligne2 = split("\t", $ligne2, -1);
    $intermed2 = $ligne2[2];
    $ligne2[2] = $ligne2[0];
    $ligne2[0] = $intermed2;
    $ligne2 = join("\t", @ligne2);
    push (@bd3, $ligne2);
}
@bd3 = sort(@bd3); # Tri en fonction du locuteur

```

```

# Détermination de la voyelle et du locuteur initiaux
@premierel = split("\t", $bd3[0], -1);
$voyelle_ref = $premierel[2];
$locuteur_ref = $premierel[0];

# Stocke les valeurs des composantes pour chaque voyelle de chaque
locuteur
foreach $ligne3 (@bd3){
  chomp ($ligne3);
  $ligne3 =~ s/,/\./g; # Conversion des vigules en points
  @ligne3 = split("\t", $ligne3, -1);
  if ($ligne3[2] eq $voyelle_ref) { # Pour la voyelle en cours...
    push(@f0, $ligne3[4]) if ($ligne3[4]);
    push(@f1, $ligne3[5]) if ($ligne3[5]);
    push(@f2, $ligne3[6]) if ($ligne3[6]);
    push(@f3, $ligne3[7]) if ($ligne3[7]);
    push(@f4, $ligne3[8]) if ($ligne3[8]);
  }
  else{ # Lorsque la voyelle change... Calcul de la moyenne de chaque
composante pour la voyelle précédente
    $stat = Statistics::Descriptive::Full->new();
    $stat->add_data(@f0);
    $moyf0 = $stat->mean();
    $stat = Statistics::Descriptive::Full->new();
    $stat->add_data(@f1);
    $moyf1 = $stat->mean();
    $stat = Statistics::Descriptive::Full->new();
    $stat->add_data(@f2);
    $moyf2 = $stat->mean();
    $stat = Statistics::Descriptive::Full->new();
    $stat->add_data(@f3);
    $moyf3 = $stat->mean();
    $stat = Statistics::Descriptive::Full->new();
    $stat->add_data(@f4);
    $moyf4 = $stat->mean();
    $bd4 =
"$locuteur_ref\t$voyelle_ref\t$moyf0\t$moyf1\t$moyf2\t$moyf3\t$moyf4"; #
Stocke les résultats dans une base de données
    push(@bd4, $bd4);
    @f0 = (); # Remise à zéro des variables et réinitialisation
pour la voyelle en cours
    @f1 = ();
    @f2 = ();
    @f3 = ();
    @f4 = ();
    push(@f0, $ligne3[4]) if ($ligne3[4]);
    push(@f1, $ligne3[5]) if ($ligne3[5]);
    push(@f2, $ligne3[6]) if ($ligne3[6]);
    push(@f3, $ligne3[7]) if ($ligne3[7]);
    push(@f4, $ligne3[8]) if ($ligne3[8]);
    $voyelle_ref = $ligne3[2];
    $locuteur_ref = $ligne3[0];
  }
} # Calcul de la moyenne de chaque composante pour la dernière
voyelle
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f0);

```



```

$moyf0 = $stat->mean();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f1);
$moyf1 = $stat->mean();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f2);
$moyf2 = $stat->mean();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f3);
$moyf3 = $stat->mean();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f4);
$moyf4 = $stat->mean();
$bd4 =
"$locuteur_ref\t$voyelle_ref\t$moyf0\t$moyf1\t$moyf2\t$moyf3\t$moyf4"; #
Stoque les résultats dans une base de données
push(@bd4, $bd4);
@f0 = (); # Remise à zéro des variables
@f1 = ();
@f2 = ();
@f3 = ();
@f4 = ();

# Calcul de la moyenne et de l'écart-type de chaque composante pour
chaque locuteur, toutes voyelles confondues
@premiere2 = split("\t", $bd4[0], -1);
$locuteur_ref = $premiere2[0];
foreach $ligne4 (@bd4){
    chomp ($ligne4);
    $ligne4 =~ s/,/\./g; # Conversion des virgules en points
    @ligne4 = split("\t", $ligne4, -1);
    if ($ligne4[0] eq $locuteur_ref) { # Stockage de chaque
composante pour le locuteur en cours
        push(@f0, $ligne4[2]) if ($ligne4[2]);
        push(@f1, $ligne4[3]) if ($ligne4[3]);
        push(@f2, $ligne4[4]) if ($ligne4[4]);
        push(@f3, $ligne4[5]) if ($ligne4[5]);
        push(@f4, $ligne4[6]) if ($ligne4[6]);
    }
    else{ # Quand le locuteur change, calcul de la moyenne et de
l'écart-type de chaque composante du locuteur précédent
        $stat = Statistics::Descriptive::Full->new();
        $stat->add_data(@f0);
        $moyf0 = $stat->mean();
        $setf0 = $stat->standard_deviation();
        $stat = Statistics::Descriptive::Full->new();
        $stat->add_data(@f1);
        $moyf1 = $stat->mean();
        $setf1 = $stat->standard_deviation();
        $stat = Statistics::Descriptive::Full->new();
        $stat->add_data(@f2);
        $moyf2 = $stat->mean();
        $setf2 = $stat->standard_deviation();
        $stat = Statistics::Descriptive::Full->new();
        $stat->add_data(@f3);
        $moyf3 = $stat->mean();
        $setf3 = $stat->standard_deviation();
    }
}

```

```

        $stat = Statistics::Descriptive::Full->new();
        $stat->add_data(@f4);
        $moyf4 = $stat->mean();
        $setf4 = $stat->standard_deviation();
        $bd6 =
"$locuteur_ref\t$moyf0\t$setf0\t$moyf1\t$setf1\t$moyf2\t$setf2\t$moyf3\t$setf
3\t$moyf4\t$setf4";
        push (@bd6, $bd6);      # Stockage dans une base de données
        @f0 = ();      # Remise à zéro des variables et réinitialisation
pour le locuteur en cours
        @f1 = ();
        @f2 = ();
        @f3 = ();
        @f4 = ();
        push(@f0, $ligne4[2]) if ($ligne4[2]);
        push(@f1, $ligne4[3]) if ($ligne4[3]);
        push(@f2, $ligne4[4]) if ($ligne4[4]);
        push(@f3, $ligne4[5]) if ($ligne4[5]);
        push(@f4, $ligne4[6]) if ($ligne4[6]);
        $locuteur_ref = $ligne4[0];
    }
}      # Calcul de la moyenne et de l'écart-type de chaque composante du
dernier locuteur
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f0);
$moyf0 = $stat->mean();
$setf0 = $stat->standard_deviation();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f1);
$moyf1 = $stat->mean();
$setf1 = $stat->standard_deviation();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f2);
$moyf2 = $stat->mean();
$setf2 = $stat->standard_deviation();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f3);
$moyf3 = $stat->mean();
$setf3 = $stat->standard_deviation();
$stat = Statistics::Descriptive::Full->new();
$stat->add_data(@f4);
$moyf4 = $stat->mean();
$setf4 = $stat->standard_deviation();
$bd6 =
"$locuteur_ref\t$moyf0\t$setf0\t$moyf1\t$setf1\t$moyf2\t$setf2\t$moyf3\t$setf
3\t$moyf4\t$setf4";
push (@bd6, $bd6);      # Stockage dans une base de données
@f0 = ();      # Remise à zéro des variables
@f1 = ();
@f2 = ();
@f3 = ();
@f4 = ();

# Réouverture de la base de données initiale
open (BD, "< $bd_initiale") or die "Impossible d'ouvrir la base de
données.\n";
@bd5 = <BD>;

```



```

close BD;
shift (@bd5);

# Calcul des scores-Z et écriture des résultats
foreach $ligne5 (@bd5){
  chomp ($ligne5);
  $ligne5 =~ s/,/\./g; # Conversion des virgules en points
  @ligne5 = split("\t", $ligne5, -1);
  $locuteur_ref = $ligne5[0];

  foreach $ligne6 (@bd6){
    chomp ($ligne6);
    @ligne6 = split("\t", $ligne6, 1);
    if ($ligne6[0] eq $locuteur_ref){
      $moyf0 = $ligne6[1];
      $setf0 = $ligne6[2];
      $moyf1 = $ligne6[3];
      $setf1 = $ligne6[4];
      $moyf2 = $ligne6[5];
      $setf2 = $ligne6[6];
      $moyf3 = $ligne6[7];
      $setf3 = $ligne6[8];
      $moyf4 = $ligne6[9];
      $setf4 = $ligne6[10];
    }
  }

  if ($ligne5[4]){
    $ligne5[4] = sprintf("%.3f", ($ligne5[4]-$moyf0)/$setf0);
  }
  else{
    $ligne5[4] = "FAUX";
  }
  if ($ligne5[5]){
    $ligne5[5] = sprintf("%.3f", ($ligne5[5]-$moyf1)/$setf1);
  }
  else{
    $ligne5[5] = "FAUX";
  }
  if ($ligne5[6]){
    $ligne5[6] = sprintf("%.3f", ($ligne5[6]-$moyf2)/$setf2);
  }
  else{
    $ligne5[6] = "FAUX";
  }
  if ($ligne5[7]){
    $ligne5[7] = sprintf("%.3f", ($ligne5[7]-$moyf3)/$setf3);
  }
  else{
    $ligne5[7] = "FAUX";
  }
  if ($ligne5[8]){
    $ligne5[8] = sprintf("%.3f", ($ligne5[8]-$moyf4)/$setf4);
  }
  else{
    $ligne5[8] = "FAUX";
  }
}

```



```
        print join("\t", @ligne5), "\n";
    }

# Fin
print "Vous pouvez copier la base de données normalisées générée et
appuyer sur ENTER pour fermer le programme.\n(Sous Windows, un clic droit
sur la barre de titre de la fenêtre vous permet d'accéder au menu
\"modifier\" dans lequel se trouvent les commandes de sélection et de
copie.)\n";
<STDIN>;
exit;
```


Annexe 7 : Valeurs moyennes et écarts-types de la F_0 , de F_1 , F_2 et F_3 (en Hertz) calculées à partir des moyennes individuelles

Locutrices de Saguenay

Classe vocalique	Point de mesure	F_0 (Hz)	F_1 (Hz)	F_2 (Hz)	F_3 (Hz)	F_4 (Hz)
i#	25 %	225 (6)	315 (36)	2349 (70)	3497 (107)	4179 (58)
	50 %	224 (5)	320 (39)	2389 (78)	3560 (126)	4229 (67)
	75 %	223 (7)	308 (34)	2412 (92)	3600 (139)	4260 (76)
iK	25 %	226 (8)	428 (18)	2331 (54)	3235 (65)	4203 (60)
	50 %	224 (9)	443 (11)	2322 (35)	3151 (55)	4210 (65)
	75 %	218 (12)	447 (15)	2257 (44)	3084 (77)	4221 (58)
iR	25 %	221 (11)	346 (34)	2464 (89)	3515 (143)	4280 (160)
	50 %	210 (10)	316 (18)	2539 (134)	3595 (160)	4260 (176)
	75 %	197 (20)	318 (12)	2521 (123)	3434 (184)	4171 (155)
y#	25 %	218 (16)	335 (35)	2081 (80)	2567 (86)	3712 (81)
	50 %	218 (15)	339 (42)	2115 (62)	2563 (87)	3697 (94)
	75 %	217 (6)	323 (37)	2155 (61)	2597 (76)	3712 (88)
yK	25 %	219 (12)	431 (14)	1899 (49)	2605 (68)	3777 (57)
	50 %	216 (13)	439 (10)	1887 (40)	2648 (102)	3808 (67)
	75 %	211 (12)	438 (11)	1840 (50)	2699 (105)	3841 (54)
yR	25 %	222 (11)	372 (39)	2089 (71)	2565 (59)	3753 (86)
	50 %	211 (10)	347 (25)	2108 (72)	2578 (80)	3746 (82)
	75 %	205 (17)	335 (18)	2066 (90)	2543 (114)	3724 (69)
u#	25 %	228 (10)	359 (51)	832 (49)	2549 (120)	3914 (130)
	50 %	225 (8)	355 (53)	781 (65)	2559 (122)	3944 (163)
	75 %	220 (10)	333 (43)	786 (64)	2617 (98)	3975 (154)
uK	25 %	224 (9)	438 (19)	1066 (96)	2609 (111)	3794 (140)
	50 %	221 (6)	443 (9)	1130 (57)	2636 (120)	3784 (114)
	75 %	222 (18)	444 (5)	1306 (47)	2635 (136)	3818 (86)
uR	25 %	220 (13)	409 (23)	799 (47)	2691 (179)	3870 (154)
	50 %	211 (14)	364 (14)	722 (91)	2700 (185)	3922 (140)
	75 %	191 (19)	350 (19)	739 (137)	2663 (170)	3870 (122)

Locuteurs de Saguenay

Classe vocalique	Point de mesure	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₄ (Hz)
i#	25 %	140 (15)	267 (24)	2040 (115)	3000 (137)	3616 (141)
	50 %	139 (16)	267 (22)	2085 (111)	3057 (147)	3668 (137)
	75 %	136 (18)	260 (28)	2119 (102)	3073 (155)	3770 (152)
iK	25 %	135 (13)	347 (28)	1988 (147)	2811 (184)	3694 (194)
	50 %	133 (13)	387 (17)	1975 (129)	2790 (184)	3739 (190)
	75 %	126 (14)	387 (26)	1923 (124)	2752 (199)	3814 (242)
iR	25 %	140 (15)	296 (18)	2148 (107)	3006 (123)	3712 (139)
	50 %	122 (13)	260 (13)	2206 (91)	3081 (90)	3747 (137)
	75 %	116 (25)	242 (18)	2182 (101)	2996 (129)	3794 (209)
y#	25 %	137 (18)	289 (23)	1754 (128)	2213 (102)	3106 (172)
	50 %	138 (21)	281 (26)	1788 (130)	2191 (121)	3100 (171)
	75 %	134 (23)	260 (22)	1808 (134)	2167 (150)	3103 (168)
yK	25 %	131 (13)	372 (16)	1567 (115)	2313 (147)	3232 (156)
	50 %	127 (13)	381 (20)	1555 (93)	2347 (153)	3263 (170)
	75 %	119 (13)	380 (20)	1533 (93)	2379 (160)	3289 (179)
yR	25 %	139 (21)	304 (23)	1756 (155)	2197 (95)	3198 (156)
	50 %	123 (22)	277 (23)	1755 (125)	2133 (96)	3171 (151)
	75 %	103 (16)	265 (29)	1762 (113)	2132 (124)	3143 (144)
u#	25 %	141 (17)	295 (22)	806 (59)	2124 (102)	3279 (300)
	50 %	141 (17)	286 (17)	755 (48)	2150 (127)	3315 (275)
	75 %	135 (15)	257 (20)	721 (53)	2194 (103)	3339 (278)
uK	25 %	134 (10)	383 (7)	999 (103)	2350 (123)	3257 (198)
	50 %	133 (12)	395 (14)	1018 (84)	2373 (112)	3277 (174)
	75 %	124 (17)	399 (21)	1143 (67)	2382 (149)	3324 (165)
uR	25 %	144 (17)	314 (30)	738 (69)	2351 (108)	3281 (274)
	50 %	127 (16)	292 (26)	666 (51)	2342 (138)	3258 (257)
	75 %	108 (17)	274 (27)	695 (20)	2327 (162)	3206 (240)

Locutrices de Québec

Classe vocalique	Point de mesure	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₄ (Hz)
i#	25 %	228 (19)	297 (21)	2452 (87)	3526 (83)	4258 (195)
	50 %	221 (19)	304 (27)	2501 (76)	3590 (108)	4250 (195)
	75 %	217 (14)	294 (35)	2538 (94)	3586 (98)	4303 (170)
iK	25 %	214 (28)	427 (26)	2418 (86)	3264 (82)	4194 (154)
	50 %	210 (28)	438 (25)	2446 (76)	3242 (69)	4227 (143)
	75 %	198 (42)	425 (25)	2412 (71)	3182 (80)	4252 (148)
iR	25 %	225 (16)	377 (24)	2555 (55)	3383 (186)	4236 (169)
	50 %	199 (17)	352 (16)	2619 (68)	3435 (175)	4253 (168)
	75 %	180 (24)	343 (20)	2575 (81)	3312 (139)	4203 (111)
y#	25 %	232 (19)	324 (25)	2063 (74)	2541 (133)	3617 (71)
	50 %	222 (17)	324 (30)	2056 (94)	2502 (153)	3602 (59)
	75 %	216 (14)	304 (33)	2055 (90)	2502 (152)	3615 (50)
yK	25 %	229 (13)	451 (39)	1940 (65)	2576 (75)	3771 (137)
	50 %	222 (14)	443 (41)	1940 (67)	2599 (81)	3798 (142)
	75 %	211 (19)	428 (41)	1910 (73)	2637 (80)	3798 (148)
yR	25 %	225 (15)	425 (31)	1881 (100)	2438 (105)	3714 (74)
	50 %	204 (22)	388 (20)	1861 (105)	2388 (79)	3669 (86)
	75 %	177 (44)	366 (20)	1892 (82)	2407 (69)	3633 (124)
u#	25 %	238 (20)	346 (29)	887 (88)	2567 (54)	4116 (124)
	50 %	231 (21)	345 (29)	843 (97)	2590 (58)	4176 (108)
	75 %	218 (19)	316 (41)	845 (117)	2625 (113)	4088 (165)
uK	25 %	229 (20)	453 (28)	1159 (77)	2589 (67)	4046 (93)
	50 %	226 (19)	465 (40)	1232 (76)	2615 (75)	4043 (102)
	75 %	224 (18)	451 (41)	1350 (60)	2653 (80)	4039 (118)
uR	25 %	222 (7)	434 (23)	872 (74)	2723 (22)	4073 (102)
	50 %	202 (15)	395 (26)	783 (38)	2739 (63)	4079 (116)
	75 %	198 (29)	363 (31)	859 (51)	2703 (111)	4006 (160)

Locuteurs de Québec

Classe vocalique	Point de mesure	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₄ (Hz)
i#	25 %	139 (32)	262 (20)	2062 (137)	3038 (168)	3683 (210)
	50 %	137 (31)	275 (20)	2077 (147)	3060 (194)	3723 (215)
	75 %	131 (20)	272 (35)	2101 (128)	3070 (199)	3666 (269)
iK	25 %	130 (29)	351 (21)	2023 (117)	2800 (161)	3572 (162)
	50 %	128 (24)	358 (17)	2036 (111)	2799 (171)	3636 (178)
	75 %	125 (21)	353 (20)	1997 (118)	2767 (166)	3674 (177)
iR	25 %	132 (26)	323 (34)	2110 (122)	2864 (225)	3604 (135)
	50 %	125 (22)	294 (30)	2184 (127)	3016 (194)	3661 (133)
	75 %	110 (18)	287 (25)	2167 (155)	2927 (201)	3608 (204)
y#	25 %	133 (36)	283 (14)	1817 (41)	2281 (103)	3169 (141)
	50 %	130 (31)	282 (20)	1821 (41)	2241 (81)	3141 (162)
	75 %	125 (24)	269 (38)	1817 (46)	2181 (70)	3158 (138)
yK	25 %	136 (32)	373 (10)	1614 (54)	2273 (94)	3204 (187)
	50 %	134 (28)	373 (8)	1620 (54)	2303 (118)	3227 (194)
	75 %	130 (22)	359 (11)	1601 (57)	2327 (102)	3259 (211)
yR	25 %	134 (31)	352 (35)	1631 (166)	2194 (131)	3156 (225)
	50 %	127 (26)	317 (29)	1652 (111)	2159 (90)	3191 (220)
	75 %	112 (19)	304 (19)	1628 (105)	2179 (62)	3155 (188)
u#	25 %	143 (42)	295 (22)	826 (106)	2187 (84)	3258 (136)
	50 %	139 (39)	289 (25)	802 (115)	2195 (98)	3275 (140)
	75 %	133 (31)	269 (53)	750 (138)	2207 (115)	3405 (212)
uK	25 %	134 (32)	374 (19)	1033 (98)	2283 (131)	3241 (194)
	50 %	134 (28)	381 (18)	1069 (87)	2309 (137)	3240 (171)
	75 %	130 (23)	376 (21)	1172 (86)	2339 (116)	3281 (204)
uR	25 %	135 (32)	356 (25)	804 (88)	2428 (131)	3270 (178)
	50 %	128 (21)	322 (20)	739 (88)	2429 (137)	3265 (162)
	75 %	115 (15)	312 (15)	814 (91)	2405 (119)	3298 (157)

Locutrices de Rouyn-Noranda

Classe vocalique	Point de mesure	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₄ (Hz)
i#	25 %	231 (21)	306 (17)	2474 (121)	3469 (129)	4163 (166)
	50 %	233 (19)	309 (19)	2519 (126)	3534 (147)	4205 (150)
	75 %	225 (21)	300 (23)	2562 (140)	3564 (125)	4224 (129)
iK	25 %	217 (17)	422 (31)	2431 (106)	3207 (114)	4229 (131)
	50 %	216 (19)	437 (34)	2411 (109)	3168 (121)	4245 (149)
	75 %	207 (23)	437 (37)	2328 (135)	3111 (126)	4265 (191)
iR	25 %	217 (18)	367 (21)	2553 (125)	3369 (153)	4206 (69)
	50 %	199 (11)	339 (19)	2612 (113)	3466 (90)	4184 (38)
	75 %	177 (16)	333 (7)	2595 (105)	3360 (117)	4124 (161)
y#	25 %	220 (34)	339 (15)	2071 (131)	2530 (199)	3679 (183)
	50 %	219 (34)	337 (12)	2095 (141)	2502 (197)	3675 (177)
	75 %	213 (33)	318 (15)	2085 (137)	2493 (201)	3693 (191)
yK	25 %	222 (28)	435 (49)	1923 (91)	2565 (170)	3742 (198)
	50 %	220 (26)	443 (46)	1910 (98)	2605 (165)	3758 (205)
	75 %	215 (22)	438 (40)	1862 (107)	2631 (173)	3759 (212)
yR	25 %	216 (17)	388 (23)	1941 (151)	2418 (206)	3731 (203)
	50 %	203 (15)	368 (22)	1948 (119)	2432 (202)	3724 (178)
	75 %	181 (16)	357 (14)	1936 (120)	2440 (222)	3681 (159)
u#	25 %	240 (30)	358 (25)	866 (96)	2594 (158)	4026 (228)
	50 %	239 (31)	340 (24)	812 (98)	2576 (167)	4035 (182)
	75 %	234 (29)	306 (12)	741 (151)	2618 (123)	4095 (197)
uK	25 %	226 (29)	454 (47)	1056 (74)	2667 (187)	3943 (259)
	50 %	224 (27)	462 (50)	1137 (96)	2696 (183)	3921 (236)
	75 %	223 (30)	459 (38)	1278 (125)	2715 (198)	3902 (206)
uR	25 %	227 (19)	420 (27)	842 (57)	2659 (169)	3922 (219)
	50 %	213 (16)	370 (26)	738 (90)	2635 (154)	3904 (280)
	75 %	188 (13)	352 (17)	801 (146)	2662 (172)	3896 (296)

Locuteurs de Rouyn-Noranda

Classe vocalique	Point de mesure	F ₀ (Hz)	F ₁ (Hz)	F ₂ (Hz)	F ₃ (Hz)	F ₄ (Hz)
i#	25 %	140 (21)	295 (12)	2144 (89)	3062 (44)	3653 (143)
	50 %	140 (17)	298 (15)	2191 (107)	3136 (43)	3709 (147)
	75 %	135 (10)	298 (31)	2229 (95)	3164 (68)	3734 (110)
iK	25 %	135 (13)	365 (10)	2055 (82)	2830 (62)	3623 (39)
	50 %	133 (13)	382 (17)	2045 (87)	2781 (55)	3679 (59)
	75 %	129 (10)	388 (19)	2020 (105)	2709 (66)	3740 (110)
iR	25 %	133 (13)	315 (19)	2216 (138)	3038 (43)	3663 (80)
	50 %	121 (8)	293 (25)	2286 (126)	3140 (63)	3731 (77)
	75 %	110 (15)	284 (26)	2269 (130)	3030 (113)	3758 (120)
y#	25 %	135 (15)	313 (8)	1821 (147)	2230 (167)	3146 (129)
	50 %	135 (14)	303 (15)	1833 (162)	2215 (181)	3158 (136)
	75 %	130 (12)	296 (40)	1847 (188)	2204 (169)	3158 (146)
yK	25 %	139 (20)	376 (15)	1686 (93)	2251 (112)	3232 (132)
	50 %	136 (17)	382 (11)	1676 (95)	2275 (114)	3242 (153)
	75 %	132 (17)	393 (11)	1642 (75)	2303 (87)	3246 (153)
yR	25 %	136 (14)	339 (24)	1725 (118)	2140 (121)	3179 (148)
	50 %	123 (9)	316 (27)	1718 (131)	2095 (123)	3165 (162)
	75 %	109 (9)	311 (30)	1702 (130)	2094 (122)	3138 (175)
u#	25 %	142 (20)	330 (24)	818 (47)	2204 (151)	3345 (204)
	50 %	142 (18)	321 (24)	760 (80)	2246 (140)	3335 (219)
	75 %	137 (13)	306 (37)	744 (98)	2251 (123)	3342 (223)
uK	25 %	139 (16)	396 (11)	961 (136)	2305 (82)	3361 (178)
	50 %	139 (16)	402 (13)	1007 (124)	2337 (85)	3344 (185)
	75 %	134 (14)	405 (18)	1138 (93)	2358 (59)	3361 (190)
uR	25 %	138 (15)	347 (21)	734 (76)	2342 (122)	3398 (207)
	50 %	123 (10)	326 (25)	688 (85)	2392 (143)	3420 (255)
	75 %	111 (12)	311 (26)	758 (72)	2370 (115)	3446 (239)

Annexe 8 : Résultats des ANOVAs portant sur les caractéristiques spectrales

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (i#)

Variable indépendante	Variable dépendante	F	p	
Sexe (2 modalités)	F ₀	Hertz	181,49	<0,0001*
		Score-Z	0,47	0,5010
	F ₁	Hertz	10,18	0,0039*
		Score-Z	7,15	0,0133
	F ₂	Hertz	79,63	<0,0001*
		Score-Z	0,97	0,3357
	F ₃	Hertz	102,52	<0,0001*
		Score-Z	1,32	0,2626
Ville (3 modalités)	F ₀	Hertz	0,20	0,8165
		Score-Z	1,39	0,2677
	F ₁	Hertz	1,18	0,3252
		Score-Z	4,66	0,0196
	F ₂	Hertz	3,17	0,0600
		Score-Z	6,79	0,0046*
	F ₃	Hertz	0,09	0,9183
		Score-Z	1,37	0,2743
Point de mesure (3 modalités)	F ₀	Hertz	3,32	0,0543
		Score-Z	0,45	0,6432
	F ₁	Hertz	3,44	0,0405
		Score-Z	4,32	0,0256
	F ₂	Hertz	90,40	<0,0001*
		Score-Z	82,43	<0,0001*
	F ₃	Hertz	53,93	<0,0001*
		Score-Z	52,82	<0,0001*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (iK)

Variable indépendante	Variable dépendante	F	p	
Sexe (2 modalités)	F ₀	Hertz	133,25	<0,0001*
		Score-Z	1,33	0,2608
	F ₁	Hertz	77,58	<0,0001*
		Score-Z	0,61	0,4426
	F ₂	Hertz	97,34	<0,0001*
		Score-Z	2,64	0,1175
F ₃	Hertz	79,79	<0,0001*	
	Score-Z	0,02	0,8957	
Ville (3 modalités)	F ₀	Hertz	0,58	0,5660
		Score-Z	2,48	0,1046
	F ₁	Hertz	1,57	0,2296
		Score-Z	8,96	0,0012*
	F ₂	Hertz	2,39	0,1127
		Score-Z	8,34	0,0018*
F ₃	Hertz	0,35	0,7073	
	Score-Z	4,48	0,0223	
Point de mesure (3 modalités)	F ₀	Hertz	11,29	0,0004*
		Score-Z	9,57	0,0009*
	F ₁	Hertz	14,03	<0,0001*
		Score-Z	15,62	<0,0001*
	F ₂	Hertz	43,41	<0,0001*
		Score-Z	44,31	<0,0001*
F ₃	Hertz	59,74	<0,0001*	
	Score-Z	80,30	<0,0001*	

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (iR)

Variable indépendante	Variable dépendante	F	p	
Sexe (2 modalités)	F ₀	Hertz	217,29	<0,0001*
		Score-Z	s.o.	s.o.
	F ₁	Hertz	56,58	<0,0001*
		Score-Z	0,72	0,4046
	F ₂	Hertz	81,61	<0,0001*
		Score-Z	1,01	0,3259
	F ₃	Hertz	64,94	<0,0001*
		Score-Z	0,24	0,6263
Ville (3 modalités)	F ₀	Hertz	0,83	0,4501
		Score-Z	s.o.	s.o.
	F ₁	Hertz	7,51	0,0029*
		Score-Z	5,95	0,0079*
	F ₂	Hertz	1,33	0,2825
		Score-Z	0,23	0,8001
	F ₃	Hertz	1,69	0,2049
		Score-Z	2,97	0,0705
Point de mesure (3 modalités)	F ₀	Hertz	53,33	<0,0001*
		Score-Z	s.o.	s.o.
	F ₁	Hertz	76,71	<0,0001*
		Score-Z	70,33	<0,0001*
	F ₂	Hertz	28,95	<0,0001*
		Score-Z	31,55	<0,0001*
	F ₃	Hertz	41,42	<0,0001*
		Score-Z	71,10	<0,0001*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (y#)

Variable indépendante	Variable dépendante	F	p	
Sexe (2 modalités)	F ₀	Hertz	118,11	<0,0001*
		Score-Z	0,35	0,7083
	F ₁	Hertz	19,47	0,0002*
		Score-Z	0,16	0,6966
	F ₂	Hertz	44,60	<0,0001*
		Score-Z	0,83	0,3713
F ₃	Hertz	39,16	<0,0001*	
	Score-Z	cf. interactions	cf. interactions	
Ville (3 modalités)	F ₀	Hertz	0,02	0,9847
		Score-Z	0,35	0,7083
	F ₁	Hertz	1,61	0,2215
		Score-Z	6,07	0,0073*
	F ₂	Hertz	cf. interactions	cf. interactions
		Score-Z	cf. interactions	cf. interactions
	F ₃	Hertz	0,05	0,9473
		Score-Z	0,03	0,9668
Point de mesure (3 modalités)	F ₀	Hertz	5,08	0,0106
		Score-Z	3,03	0,0575
	F ₁	Hertz	13,54	<0,0001*
		Score-Z	13,45	<0,0001*
	F ₂	Hertz	cf. interactions	cf. interactions
		Score-Z	cf. interactions	cf. interactions
	F ₃	Hertz	11,81	<0,0001*
		Score-Z	cf. interactions	cf. interactions
Ville / Point de mesure	F ₂	Hertz	10,54	<0,0001*
		Score-Z	7,26	0,0001
Sexe / Point de mesure	F ₃	Hertz	s.o.	s.o.
		Score-Z	5,41	0,0076*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (yK)

Variable indépendante	Variable dépendante	F	p	
Sexe (2 modalités)	F ₀	Hertz	154,25	<0,0001*
		Score-Z	s.o.	s.o.
	F ₁	Hertz	41,92	<0,0001*
		Score-Z	2,06	0,1637
	F ₂	Hertz	107,73	<0,0001*
		Score-Z	6,09	0,0211
	F ₃	Hertz	49,63	<0,0001*
		Score-Z	1,36	0,2553
Ville (3 modalités)	F ₀	Hertz	0,43	0,6543
		Score-Z	s.o.	s.o.
	F ₁	Hertz	<i>cf. interactions</i>	<i>cf. interactions</i>
		Score-Z	<i>cf. interactions</i>	<i>cf. interactions</i>
	F ₂	Hertz	2,37	0,1150
		Score-Z	2,54	0,0995
	F ₃	Hertz	0,68	0,5157
		Score-Z	1,95	0,1646
Point de mesure (3 modalités)	F ₀	Hertz	14,61	<0,0001*
		Score-Z	s.o.	s.o.
	F ₁	Hertz	<i>cf. interactions</i>	<i>cf. interactions</i>
		Score-Z	<i>cf. interactions</i>	<i>cf. interactions</i>
	F ₂	Hertz	25,13	<0,0001*
		Score-Z	24,60	<0,0001*
	F ₃	Hertz	69,45	<0,0001*
		Score-Z	76,87	<0,0001*
Ville / Point de mesure	F ₁	Hertz	6,19	0,0004*
		Score-Z	7,68	<0,0001*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (yR)

Variable indépendante	Variable dépendante	F	p	
Sexe (2 modalités)	F ₀	Hertz	152,45	<0,0001*
		Score-Z	0,62	0,4385
	F ₁	Hertz	51,76	<0,0001*
		Score-Z	0,91	0,3486
	F ₂	Hertz	41,30	<0,0001*
		Score-Z	0,00	0,9738
	F ₃	Hertz	51,80	<0,0001*
		Score-Z	0,16	0,6914
Ville (3 modalités)	F ₀	Hertz	0,27	0,7693
		Score-Z	0,04	0,9584
	F ₁	Hertz	8,39	0,0017*
		Score-Z	9,77	0,0008*
	F ₂	Hertz	5,34	0,0120
		Score-Z	15,39	<0,0001*
	F ₃	Hertz	1,39	0,2680
		Score-Z	3,43	0,0491
Point de mesure (3 modalités)	F ₀	Hertz	32,93	<0,0001*
		Score-Z	50,53	<0,0001*
	F ₁	Hertz	80,89	<0,0001*
		Score-Z	75,57	<0,0001*
	F ₂	Hertz	0,48	0,6213
		Score-Z	0,55	0,5788
	F ₃	Hertz	5,22	0,0089*
		Score-Z	5,67	0,0062*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (u#)

Variable indépendante	Variable dépendante		F	p
Sexe (2 modalités)	F ₀	Hertz	117,42	<0,0001*
		Score-Z	0,48	0,4961
	F ₁	Hertz	16,70	0,0004*
		Score-Z	0,02	0,8882
	F ₂	Hertz	2,00	0,1700
		Score-Z	1,53	0,2282
	F ₃	Hertz	87,78	<0,0001*
		Score-Z	6,38	0,0186
Ville (3 modalités)	F ₀	Hertz	0,28	0,7562
		Score-Z	0,58	0,5688
	F ₁	Hertz	0,81	0,4558
		Score-Z	1,95	0,1646
	F ₂	Hertz	0,74	0,4890
		Score-Z	0,32	0,7303
	F ₃	Hertz	0,47	0,6324
		Score-Z	2,44	0,1086
Point de mesure (3 modalités)	F ₀	Hertz	7,81	0,0026*
		Score-Z	2,70	0,0882
	F ₁	Hertz	22,89	<0,0001*
		Score-Z	17,44	<0,0001*
	F ₂	Hertz	33,42	<0,0001*
		Score-Z	37,82	<0,0001*
	F ₃	Hertz	16,89	<0,0001*
		Score-Z	14,78	<0,0001*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (uK)

Variable indépendante	Variable dépendante		F	p
Sexe (2 modalités)	F ₀	Hertz	161,31	<0,0001*
		Score-Z	0,67	0,4209
	F ₁	Hertz	44,00	<0,0001*
		Score-Z	3,64	0,0684
	F ₂	Hertz	17,18	0,0004*
		Score-Z	0,01	0,9303
	F ₃	Hertz	46,22	<0,0001*
		Score-Z	1,34	0,2588
Ville (3 modalités)	F ₀	Hertz	0,14	0,8717
		Score-Z	1,59	0,2249
	F ₁	Hertz	0,85	0,4388
		Score-Z	2,85	0,0778
	F ₂	Hertz	2,00	0,1567
		Score-Z	3,24	0,0566
	F ₃	Hertz	0,39	0,6785
		Score-Z	2,20	0,1332
Point de mesure (3 modalités)	F ₀	Hertz	3,11	0,0638
		Score-Z	0,90	0,4198
	F ₁	Hertz	5,22	0,0089*
		Score-Z	9,37	0,0011*
	F ₂	Hertz	210,74	<0,0001*
		Score-Z	218,22	<0,0001*
	F ₃	Hertz	29,56	<0,0001*
		Score-Z	27,08	<0,0001*

Effet du sexe, de la ville d'origine et du point de mesure sur les caractéristiques spectrales de la variante (uR)

Variable indépendante	Variable dépendante		F	p
Sexe (2 modalités)	F ₀	Hertz	246,87	<0,0001*
		Score-Z	s.o.	s.o.
	F ₁	Hertz	75,25	<0,0001*
		Score-Z	3,91	0,0596
	F ₂	Hertz	4,90	0,0367
		Score-Z	0,02	0,8956
	F ₃	Hertz	39,64	<0,0001*
		Score-Z	1,56	0,2236
Ville (3 modalités)	F ₀	Hertz	0,00	0,9981
		Score-Z	s.o.	s.o.
	F ₁	Hertz	5,24	0,0129
		Score-Z	5,83	0,0086*
	F ₂	Hertz	3,61	0,0426
		Score-Z	2,41	0,1120
	F ₃	Hertz	0,67	0,5204
		Score-Z	2,47	0,1062
Point de mesure (3 modalités)	F ₀	Hertz	33,94	<0,0001*
		Score-Z	s.o.	s.o.
	F ₁	Hertz	80,80	<0,0001*
		Score-Z	89,41	<0,0001*
	F ₂	Hertz	18,69	<0,0001*
		Score-Z	29,39	<0,0001*
	F ₃	Hertz	1,30	0,2823
		Score-Z	1,24	0,2995

Rapport-Gratuit.com



X0404272 9