

Normalisation des contours intonatifs et étude de la variation régionale en français

Alice Bardiaux*, Piet Mertens**

* FNRS – Université catholique de Louvain,

** Katholieke Universiteit Leuven

<alice.bardiaux@uclouvain.be, piet.mertens@arts.kuleuven.be>

Abstract

This paper describes a novel approach to pitch contour normalization, relative to the observed pitch range of the speaker, and applies it to the study of regional variation of intonation contours in French. In a corpus of 12 French speakers from Belgium (Brussels and Liège), normalized pitch values and pitch intervals are obtained for the three last syllables (“clausules”) of intonation groups. For the identification of regional variants, a perceptual approach is used, in which the criterion for classifying speech productions is not the geographic origin of the speaker, but rather the judgment by experts about the markedness of particular intonation groups or syllables.

Mots clés : *prosodie, normalisation de f₀, mesures de f₀, contour intonatif, variation régionale, perception*

1. Introduction

L'étude du registre tonal et des variations mélodiques a retenu l'attention de nombreux chercheurs qui ont étudié les différences de fréquence fondamentale (f₀) tantôt globales (notamment Bissonnette 2003, Ménard 2003, Demers 2003), tantôt locales (par exemple Coquillon 2005, Miller 2007, Kaminskaïa 2009, Bardiaux 2014) entre des locuteurs originaires de différentes régions de la francophonie. L'intérêt des chercheurs pour la dimension prosodique de la f₀ vient probablement des nombreux stéréotypes « populaires » qui tournent autour des variétés régionales du type « l'accent du midi est chantant », « le français québécois est monotone », etc.

L'étude de la f₀ pose toutefois des difficultés méthodologiques de taille, trop peu souvent soulevées dans les études empiriques de la variation prosodique. Nous aborderons d'abord la question de l'échelle de hauteur choisie et de la normalisation des données en fonction de la tessiture du locuteur (cf. 2). Ensuite, nous définirons les mesures de f₀ à prendre pour caractériser le plus précisément possible le contour intonatif porté par les dernières syllabes du groupe intonatif, soit la configuration des cibles alignées (cf. 3). Nous

terminerons par la présentation de quelques résultats obtenus sur un corpus de locuteurs belges francophones de Bruxelles et Liège (cf. 4). Cette dernière partie sera l'occasion de proposer une nouvelle approche des variétés basée sur la catégorisation perceptive et formelle des données et d'identifier quelques corrélats acoustiques des traits perçus comme régionalement marqués par l'accent belge.

2. Quelle échelle et quelle normalisation pour l'étude de la f0 ?

L'analyse de la f0 pose deux questions : d'une part celle du choix de l'échelle appropriée pour la mesure des variations mélodiques dans la parole, d'autre part celle de leur normalisation en fonction des locuteurs.

Très souvent, les études portant sur la variation prosodique régionale utilisent une conversion des mesures de f0 en Hertz vers une autre échelle, éventuellement normalisée. Toutefois, bien que plusieurs échelles de hauteur normalisées aient été proposées (demi-tons, OME, ERB, etc.), aucune ne fait consensus dans la communauté des chercheurs (Hermes & Van Gestel 1991, Nolan 2003).

Dans la présente étude, l'échelle musicale en demi-tons a été retenue. La conversion de l'échelle linéaire en Hertz vers l'échelle logarithmique en demi-tons peut être considérée comme une normalisation des intervalles mélodiques, qui associe une même valeur à des intervalles mélodiques comparables sur le plan perceptif, quelle que soit la fréquence de départ de ces variations. Ceci permet une meilleure comparabilité des mesures, notamment entre les voix de locuteurs avec une tessiture différente (des femmes et des hommes, par exemple) ou au sein de la voix d'un seul locuteur mais dans différentes parties (inférieure, supérieure) de sa tessiture.

La formule suivante convertit l'écart entre deux fréquences f_1 et f_2 (en Hertz) en l'intervalle musical correspondant i , en demi-tons (semitones, ST) :

$$(1) \quad i = 12 * \log_2(f_2/f_1)$$

Pour utiliser les intervalles en demi-tons comme échelle absolue, on mesure chaque fréquence observée f comme l'intervalle qui sépare f d'une fréquence de référence constante f_{0_ref} (en Hertz). La hauteur correspondante sera indiquée p (pour *pitch*).

$$(2) \quad p = 12 * \log_2(f/f_{0_ref}) ; f_{0_ref} = 1\text{Hz}$$

La référence à 1Hz a été choisie afin d'obtenir uniquement des valeurs en demi-tons positives, ce qui facilitera l'interprétation des données.

L'échelle logarithmique en demi-tons rend mieux compte de la perception humaine de la parole, le rapport entre cette dernière et la

dimension acoustique de la f_0 étant non-linéaire¹. Elle a donc souvent été utilisée dans les études sur la perception de la parole et « peut se révéler intéressante pour comparer les fréquences utilisées par des locuteurs différents ou pour évaluer l'écart entre deux fréquences particulières » (Coquillon 2005, 108). Or, c'est précisément ce type d'observations qui nous intéressent dans le cadre de l'étude de la variation prosodique.

La deuxième question concernant l'analyse de la hauteur concerne la normalisation des données en fonction de la tessiture des locuteurs qui les ont produites. En effet, les locuteurs de notre corpus sont très différents, à plusieurs points de vue (sexe, âge, origine, milieu social, etc.), et ont été enregistrés dans deux situations de parole différentes. À ces différences sociodémographiques et contextuelles s'ajoutent d'autres différences personnelles (physiologiques) difficilement caractérisables et qui influencent directement et de manière trop importante la f_0 des locuteurs, en particulier leur tessiture et l'amplitude de leur registre, pour considérer que des mesures brutes (quelle que soit l'échelle) de f_0 soient directement comparables. Dès lors, après la conversion en demi-tons, l'étape suivante du traitement des données consiste en une normalisation des intervalles en fonction de la tessiture des locuteurs, suivant la formule suivante :

$$(3) \quad p_x \text{ normalisé} = 100 * ((p_x - p_{\text{bottom}}) / (p_{\text{top}} - p_{\text{bottom}}))$$

Où p_x correspond à la hauteur (ST rel 1Hz) en un point donné, p_{bottom} au plancher de la tessiture du locuteur et p_{top} au plafond de cette tessiture. Cette formule rejoint la méthode de normalisation de la f_0 proposée par Earle (1975) et rapportée par Ladd (2008). Notons que cette normalisation en fonction de la tessiture peut en principe s'appliquer à des valeurs de f_0 exprimées sur n'importe quelle échelle (Hertz, demi-tons, ERB, OME, etc.).

Le plancher et le plafond de la tessiture du locuteur ont été déterminés par les percentiles 2 et 98 des cibles tonales détectées par Prosogramme (Mertens 2003, 2004). Les valeurs de hauteur obtenues à la suite de cette transformation correspondent donc à une proportion et sont comprises, moyennant l'ajout du coefficient 100, entre 0 (plancher) et 100 (plafond). Dès lors, les valeurs de f_0 obtenues sont directement comparables entre les locuteurs étudiés, indépendamment de leurs caractéristiques morphologiques, sociales ou des paramètres situationnels de production (cf. Figure 1). De plus,

¹ La nature exacte du rapport entre la dimension acoustique de la f_0 (en Hz) et sa dimension perceptive (hauteur perçue de la parole) n'a pas été établie et est toujours débattue (cf. notamment les différentes échelles existantes pour rendre compte de cette dimension prosodique).

puisque les résultats obtenus seront exprimés en proportions (hauteur relative), il sera plus facile de les exploiter en vue d'une modélisation des variations de f_0 observées dans les données.

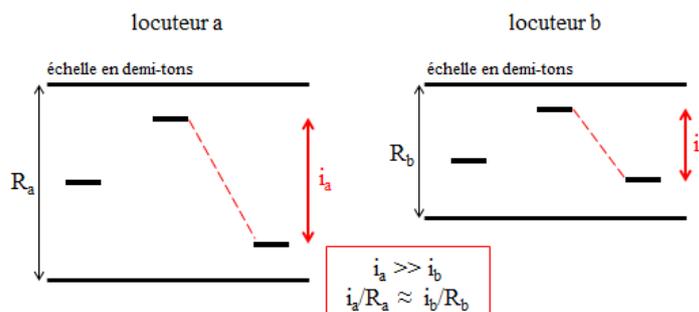


Figure 1 : Comparaison des écarts mélodiques au sein de registres (R_a et R_b) d'étendue différente en valeurs absolues (i_a et i_b) et en valeurs normalisées par rapport à l'étendue du registre (i_a/R_a et i_b/R_b). Pour les deux locuteurs, schématisation de la hauteur de trois syllabes finales d'un groupe intonatif

Sur la figure ci-dessus, l'écart mélodique entre les syllabes pénultième et finale du locuteur a n'est absolument pas comparable à celui produit par le locuteur b : il est plus ample et termine sur une cible plus basse. Par contre, si les deux écarts mélodiques sont rapportés à l'étendue du registre (R) du locuteur, ces deux écarts mélodiques deviennent tout à fait comparables : ils partent tous deux du haut du registre du locuteur (2^{e} quart supérieur) pour couvrir les $2/3$ du registre et terminer dans le dernier quart inférieur du registre. La même logique s'applique pour la hauteur des syllabes : la hauteur absolue de la syllabe finale du locuteur a est nettement inférieure à la hauteur absolue de la syllabe finale du locuteur b. Or, si on normalise cette hauteur par rapport au registre des locuteurs, les deux hauteurs deviennent similaires : elles se situent toutes deux à même distance du plancher du registre. Ces deux clausules, a priori trop différentes pour être comparées, sont désormais comparables.

3. Quelle(s) mesure(s) de f_0 pour les syllabes de la clausule ?

Pour caractériser la configuration mélodique d'un contour, la plupart des études en variation prosodique régionale réduisent l'information mélodique d'une syllabe à sa f_0 moyenne. Cette méthode présente un inconvénient majeur. Calculer la f_0 moyenne d'une syllabe ne permet pas de rendre compte d'un éventuel mouvement mélodique au sein de cette syllabe. Or, c'est justement cette information qui est l'objet de notre intérêt. En effet, nous partons du postulat selon lequel le contour intonatif est essentiellement localisé sur la dernière syllabe

des groupes prosodiques. Dans un tel cas de figure, prendre comme mesure la f_0 moyenne de la syllabe reviendrait à décrire de la même manière un contour montant, un contour descendant ou un contour statique (cf. Figure 2). On perd alors une information essentielle et déterminante puisqu'on ne rend pas compte de ce qu'il se passe à l'intérieur des syllabes.

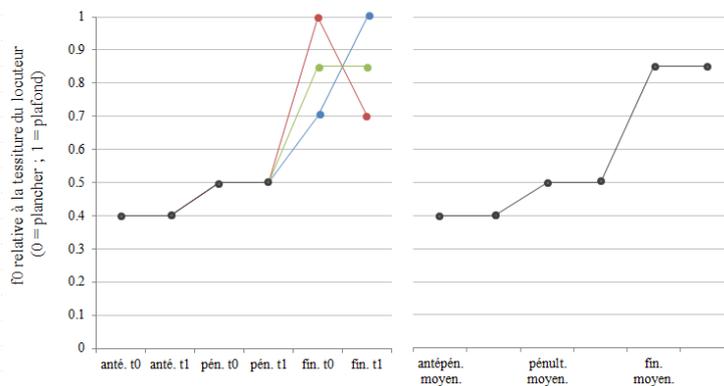


Figure 2 : Contours intonatifs montant, descendant, et statique : cibles de f_0 au début (t0) et à la fin (t1) de chaque syllabe de la clause (à gauche) et moyenne de f_0 pour chaque syllabe de la clause (à droite)

Pour chaque syllabe dans le corpus, l'outil Prosogramme (Mertens 2003, 2004) fournit la hauteur (ST rel 1 Hz) ainsi que la valeur normalisée en fonction de la tessiture du locuteur, et ceci pour le point initial et le point final du noyau syllabique. Ces données sont utilisées par un deuxième script « clause.praat » qui, pour chaque groupe intonatif (GI) dans le corpus, établit la configuration mélodique normalisée du contour à partir des valeurs des trois dernières syllabes dans le groupe. Les valeurs de hauteur initiale et finale des syllabes de la clause ont été choisies comme points de référence car « deux moments suffisent, d'abord parce qu'il n'y a pas de mouvements complexes (montant-descendant, par exemple), ensuite parce que la forme de la courbe entre les points de départ et d'arrivée (concave, convexe, linéaire) n'est pas distinctive. » (Mertens 1987, 84). Cette méthodologie nous permettra de décrire plus finement les phénomènes prosodiques perçus comme régionalement marqués et d'améliorer la comparabilité des clauses grâce à une meilleure catégorisation des contours intonatifs (cf. 4).

À partir de ces valeurs, le script « clause.praat » calcule également automatiquement la variation intrasyllabique normalisée des syllabes de la clause (Intra) et les écarts intersyllabiques entre le

point final d'une syllabe et le point initial de la syllabe suivante (Mov), moyennant les formules suivantes :

$$(4) \text{ Intra}_i = (p_{i \max} - p_{i \min}) / (p_{\text{top}} - p_{\text{bottom}}) ; \text{ pour une variation montante}$$

$$(5) \text{ Intra}_i = (p_{i \min} - p_{i \max}) / (f0_{\max} - f0_{\min}) ; \text{ pour une variation descendante}$$

Où Intra_i correspond à la variation intrasyllabique normalisée pour la syllabe i , $p_{i \max}$ au maximum de hauteur (ST rel 1Hz) du noyau de la syllabe i et $p_{i \min}$ au minimum de hauteur de ce même noyau. La variation intrasyllabique est donc calculée à partir du maximum et du minimum de hauteur dans le noyau syllabique, du caractère dynamique ou non de la variation perçue (seuil de glissando) et de sa pente. Les variations inférieures au seuil de glissando obtiennent une valeur nulle (0) et correspondent aux syllabes (perçues comme) statiques.

$$(6) \text{ Mov}_i = p_{i \text{ ini}} - p_{j \text{ fin}}$$

Où Mov_i correspond à l'écart mélodique entre la fin de la syllabe j précédente ($p_{j \text{ fin}}$) et le début de la syllabe i ($p_{i \text{ ini}}$), en demi-tons. Si la différence est négative, la cible initiale de la syllabe i est plus basse que la cible finale de la syllabe j : l'écart entre les deux syllabes est négatif, un mouvement descendant les sépare. À l'inverse, si la différence est positive, la cible initiale de la syllabe i est plus haute, un mouvement montant sépare les deux syllabes.

4. Quels corrélats acoustiques de f0 pour les traits régionaux ?

Les propositions méthodologiques concernant l'échelle de hauteur, la normalisation de la hauteur et les localisations utilisées dans la caractérisation du contour ont été testées sur un corpus de locuteurs belges francophones dans le cadre d'une étude de la variation prosodique régionale. Les résultats obtenus suivant ces propositions sont exposés sous ce point.

4.1. Corpus : protocole et locuteurs

Le corpus d'étude se fonde sur le protocole d'enquête du projet PFC² (Durand et al. 2009). Nous avons repris deux points d'enquête du corpus PFC : Bruxelles et Liège. Pour chaque ville, nous avons interrogé 3 hommes et 3 femmes, répartis entre deux tranches d'âge (entre 30 et 37 ans et entre 48 et 65 ans), sans tenir compte de leur origine sociale. Les locuteurs de notre corpus ont réalisé deux tâches : la lecture du texte « Le premier ministre ira-t-il à Beaulieu » et la lecture d'un dialogue (dorénavant « pseudo-dialogue »).

² <http://www.projet-pfc.net/>

4.2. *Transcription et annotation des domaines prosodiques*

Le corpus comporte les annotations alignées suivantes (sous Praat, Boersma & Weenink 2013) : transcription orthographique, alignement phonétique et syllabique (Goldman 2011), proéminence syllabique (Simon et al. 2008). Les hésitations et les interruptions ont été exclues des analyses.

Chaque syllabe a été annotée semi-automatiquement selon sa position et selon la nature du mot : syllabe initiale de mot plein (i), syllabe médiane de mot plein (m), syllabe finale de mot plein (f), syllabe de mot plein monosyllabique (1), syllabe de clitique (c), schwa post-tonique (@). Une tire reprenant les groupes accentuels (GA), constitués d'un mot plein et des clitiques associés (Mertens 2009, 26), a été créée automatiquement à partir de l'annotation de la structure accentuelle des syllabes.

À partir de cette tire GA, une tire générant les groupes intonatifs (GI, Mertens 2009) a été créée. Une frontière de GI a été créée automatiquement lorsqu'une proéminence correspondait à la syllabe finale d'un GA, ou lorsqu'une proéminence était détectée à la fois sur la syllabe pénultième du GA considéré et sur la syllabe initiale du GA suivant. Ce dernier cas de figure permet de tenir compte de l'impact rétroactif d'une proéminence initiale sur la perception d'une frontière de groupe prosodique (Astésano et al. 2012).

4.3. *Approche contrastive : définition des variétés à comparer*

Étudier la variation nécessite une approche contrastive où un ensemble de données est comparé à un autre ensemble afin d'identifier et de décrire ce qui les distingue. Ces ensembles sont généralement appelés « variétés ». La plupart des études portant sur la variation (quel que soit le niveau linguistique concerné) mobilise des critères extralinguistiques comme l'origine géographique, la classe sociale, l'âge, la profession, etc. (ou une combinaison plusieurs de ces critères) pour définir les variétés. L'inconvénient de cette méthode est qu'on présuppose une série de pratiques communes à un groupe de locuteurs réunis sur des critères extérieurs à leurs pratiques linguistiques.

L'approche qui a été choisie dans le cadre de cette étude opte pour une autre méthode de catégorisation puisque les productions sont regroupées sur la base de leur perception par des auditeurs. Le biais méthodologique d'une approche extralinguistique est ainsi évité puisque l'approche perceptive des variétés présuppose une série de pratiques communes dans un groupe de productions (vs. de locuteurs) qui ont été perçues et catégorisées de la même manière. Méthodologiquement, l'intérêt de la seconde approche apparaît

comme évident puisqu'elle permet non seulement de comparer des groupes qui sont pertinents perceptivement pour les locuteurs et reflètent leurs catégorisations et leurs représentations linguistiques mais également de ne pas postuler une homogénéité des façons de parler, que ce soit au niveau d'un pays, d'une région, d'une ville et même d'un individu.

Trois experts ont donc réalisé une annotation perceptive du marquage régional des trois dernières syllabes de chaque GI, correspondant à la clause de Carton (1983, 45). L'empan d'écoute était fixé entre 3 et 6 secondes ou 1 à 3 GI consécutifs, avec un maximum de 3 écoutes par segment afin d'aboutir à une annotation proche de la perception dans la parole continue. Les clauses ont été annotées M lorsqu'elles étaient perçues comme marquées régionalement et N lorsqu'elles étaient perçues comme non marquées régionalement. Les consignes aux experts définissaient ce qui était entendu par « régionalement marqué » comme « ce qui est différent de sa propre manière de parler et/ou dont on peut identifier l'origine géographique ; ce qui est perçu comme typique d'une manière de parler en Belgique ». Après une comparaison de l'annotation perceptive des trois experts³, seules les clauses perçues comme régionalement marquées par au moins deux experts sur les trois ont été annotées M (marquées) dans une tire de référence et considérées comme régionalement marquées dans les analyses.

Les clauses perçues comme régionalement marquées sont rares puisqu'elles ne représentent que 8,6% de l'ensemble des clauses. Cette catégorisation perceptive des clauses permet d'éviter que les mesures saillantes des clauses marquées ne s'estompent dans la masse des mesures des clauses non marquées et de mettre au jour des traits prosodiques distinguant les clauses perçues comme régionalement marquées de celles qui ne le sont pas. Ces traits régionaux, s'ils interviennent de manière sporadique dans la parole des locuteurs, peuvent néanmoins suffire à les stigmatiser (Léon 1993, Hambye & Francard 2008, 43, Francard 2010, 120).

4.4. *Catégorisation formelle*

Au-delà de la catégorisation des variétés, l'étude de la variation prosodique nécessite également une catégorisation linguistique fine afin de garantir la comparabilité des données : cette contrainte est

³ De manière générale, l'accord entre les 3 experts pour l'annotation du marquage régional perçu est très bon. Le degré d'accord général obtient un coefficient kappa de 0,84. Le degré d'accord est légèrement meilleur entre les juges AB et ACS (coefficient kappa = 0,96) qu'entre les juges AB et MCM d'une part (coefficient kappa = 0,78) et ACS et MCM d'autre part (coefficient kappa = 0,79).

particulièrement prégnante lorsqu'il s'agit du paramètre de la f_0 . En effet, nous partons de l'hypothèse selon laquelle le marquage régional n'est sans doute pas réalisé de la même manière dans les clauses selon qu'elles portent un contour montant, descendant, plat, situé dans le haut ou le bas du registre du locuteur. De plus, analyser conjointement les clauses portant des contours intonatifs très variés nous expose au risque de voir les différentes mesures de f_0 (hauteur et mouvement inter et intrasyllabique) se compenser les unes les autres entre les contours bas et hauts, montants et descendants, etc. Dès lors, il nous semble indispensable de procéder à une caractérisation des contours intonatifs portés par les clauses de notre corpus afin d'analyser séparément les clauses portant des contours intonatifs différents et d'assurer ce que nous appellerons désormais une « comparabilité formelle ».

Les clauses de notre corpus ont donc été catégorisées en fonction de la forme (montante, descendante, statique) et de la localisation au sein du registre (plancher, bas, milieu, haut, plafond) du contour intonatif porté par la syllabe finale du GI. La forme du contour est définie par la valeur obtenue pour la mesure Intra (cf. 3) et la localisation est déterminée par la hauteur de la cible finale de la dernière syllabe du GI. Nous proposons ici l'analyse d'une catégorie de clauses : les clauses portant un contour descendant situé au plancher du registre.

4.5. *Contribution du trait de hauteur au marquage régional*

Les clauses portant un contour descendant situé au plancher du registre et qui ont été perçues comme régionalement marquées (M) ont été comparées aux clauses présentant le même type de contour mais qui n'ont pas été perçues comme régionalement marquées (N). Les contrastes les plus importants qui peuvent être observés se situent au niveau de la syllabe pénultième. En effet, les cibles initiale et finale de la syllabe pénultième sont plus hautes dans les clauses M que dans les clauses N (44,33 contre 32,28, écart moyen de 12)⁴. Puisque les cibles initiales de la syllabe finale sont similaires dans les clauses M et N (respectivement 28,7 et 24,6), cette différence de hauteur de la syllabe pénultième entre clauses M et N induit un écart mélodique

⁴ Les données ont été analysées au moyen d'un modèle linéaire généralisé (à mesures répétées) avec les différentes mesures de f_0 (en valeurs normalisées) comme variables dépendantes (dans des modèles indépendants) et avec le marquage régional de la clause et le marquage régional de la syllabe pénultième comme prédicteurs. Seule la différence de hauteur entre la syllabe pénultième des clauses marquées et la syllabe pénultième des clauses non marquées est significative : χ^2 Wald (1, n = 38) = 5,121, p < 0,05.

entre syllabe pénultième et syllabe finale deux fois plus important dans les clausules M (-15,67 contre -7,69).

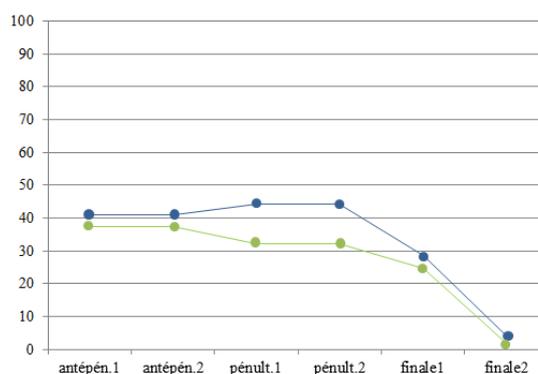


Figure 3 : Hauteur normalisée des cibles initiale (1) et finale (2) pour les trois syllabes des clausules M (ligne supérieure) et N (ligne inférieure)

De plus, l'écart entre la cible finale de la syllabe antépénultième et la cible initiale de la syllabe pénultième est positif (+3,3) dans les clausules M et négatif (-5,06) dans les clausules N. Dans les clausules M, si l'écart est moins important, il ne va pas dans le sens du contour final (descente) et contribue donc à rendre saillante la syllabe pénultième. À l'inverse, dans les clausules N, la syllabe pénultième est plus basse que la syllabe antépénultième et amorce donc, en quelque sorte, le contour final descendant.

Cette saillance mélodique de la syllabe pénultième par rapport à son contexte a été mentionnée notamment par Simon (2004) comme typique du français parlé en Belgique. Carton (1986, 253) avait également avancé l'importance des ruptures de f_0 à gauche et à droite de la syllabe pénultième, entraînant une discontinuité mélodique, dans la caractérisation des variations prosodiques régionales et sociales. C'est exactement cette configuration mélodique « en rupture » que nous retrouvons ici dans les clausules M.

5. Conclusion

Les propositions méthodologiques formulées dans le présent article ont permis de mettre en évidence les difficultés liées à l'étude de la fréquence fondamentale dans le cadre de la variation régionale. Le choix d'une échelle de hauteur, d'une méthode de normalisation de la hauteur et des localisations utilisées dans la caractérisation des contours intonatifs se révèlent déterminants pour la qualité, la pertinence et la portée des résultats. Notre réflexion gagnerait à être enrichie d'une comparaison avec d'autres méthodes de normalisation

de la hauteur (e. g. normalisation par z-score), testées avec différentes échelles de hauteur et sur des données plus variées.

Bibliographie

- Astésano, C., Bertrand R., Espesser, R. & Nguyen N. (2012). Perception des frontières et des proéminences en français. *Actes de la conférence conjointe JEP-TALN-RECITAL 1* (pp. 353-360). Grenoble.
- Bardiaux, A. (2014). *La prosodie de quelques variétés de français en Belgique. Analyse perceptive et acoustique*. Thèse de doctorat, Université catholique de Louvain, <http://dial.academielouvain.be/handle/boreal:143568>.
- Bissonnette, S. (2003). Le registre du lecteur de bulletins de nouvelles québécois et français : un reflet de l'idéal vocal des communautés linguistiques ? In M. Demers (éd.), *Registre et voix sociale* (pp. 17-34). Québec : Nota Bene.
- Boersma, P. & Weenink D. (2013). *Praat : doing phonetics by computer* (version 5.3.42) [computer program], téléchargé le 4 mars 2013 depuis <http://www.praat.org>
- Carton, F. (1983). Clausules rythmiques et mélodies des parlers normands. In *Dialectologie et littérature du domaine d'oïl occidental. Cahiers des Annales de Normandie* 15, 45-51.
- Carton, F. (1986). À la recherche d'intonations régionales. *Actes du Congrès International de Linguistique Romane* 6 (pp. 249-257). Aix-en-Provence : Presses de l'Université.
- Coquillon, A. (2005). *Caractérisation prosodique du parler de la région de Marseille*. Thèse de doctorat, Université d'Aix-Marseille.
- Demers, M. (2003). La voix du plus fort. Étude acoustique sur le registre vocal en tant qu'indicateur sociolectal et dialectal en français spontané. In M. Demers (éd.), *Registre et voix sociale* (pp. 79-124). Québec : Nota Bene.
- Durand, J., Laks B. & Lyche C. (2009). Le projet PFC : une source de données primaires structurées. In J. Durand, B. Laks & C. Lyche (éds), *Phonologie, variation et accents du français*. Paris : Hermès.
- Earle, M. A. (1975). An acoustic phonetic study of Northern Vietnamese Tones. *Speech Communications Research Laboratory*. Santa Barbara.
- Francard, M. (2010). Variation diatopique et norme endogène. Français et langues régionales en Belgique francophone. *Langue Française* 167/3, 113-126.
- Goldman, J.-P. (2011). EasyAlign : an automatic phonetic alignment tool under Praat. *Actes du colloque international InterSpeech*, septembre 2011, Florence, Italie.
- Hambye, P. & Francard M. (2008). Normes endogènes et processus identitaires. Le cas de la Wallonie romane. In C. Bavoux, L. F. Prudent & S. Wharton (éds), *Normes endogènes et plurilinguisme. Aires francophones, aires créoles* (pp. 45-60). Lyon : ENS éditions.
- Hermes, D. J. & van Gestel, J. C. (1991). The frequency scale of speech intonation. In *JASA* 90/1, 97-102.

- Kaminskaïa, S. (2009). *La variation intonative dialectale en français. Une approche phonologique*. München : LINCOM studies in French Linguistics.
- Ladd, R. D. (2008). *Intonational Phonology*. 2nd edition, Cambridge : Cambridge University Press.
- Léon, P. (1993). *Précis de phonostylistique. Parole et expressivité*. Paris : Nathan Université.
- Ménard, L. (2003). Le registre en tant que marqueur prosodique de la dialectalité : le cas du français du Québec et de la France. In M. Demers (éd.), *Registre et voix sociale* (pp. 35-58). Québec : Nota Bene.
- Mertens, P. (1987). *L'intonation du français. De la description à la reconnaissance automatique*. Thèse de doctorat, Katholieke Universiteit Leuven.
- Mertens, P. (2003). Le prosogramme : une transcription semi-automatique de la prosodie. In *Cahiers de l'Institut de Linguistique de Louvain* 30/1-3, 7-25.
- Mertens, P. (2004). Un outil pour la transcription de la prosodie dans les corpus oraux. In *Traitement automatique des langues* 45/2, 109-130.
- Mertens, P. (2009). Prosodie, syntaxe et discours : autour d'une approche prédictive. In H.-Y. Yoo & E. Delais-Roussarie (éds), *Actes du colloque IDP 2009* (pp. 19-32). Paris.
- Miller, J. (2007). *Swiss French Prosody. Intonation, rate and speaking style in the Vaud Canton*. Thèse de doctorat, University of Illinois at Champaign-Urbana.
- Nolan, F. (2003). Intonational equivalence : an experimental evaluation of pitch scales. In *Proceedings of ICPhSc* (pp. 661-774). Barcelone.
- Simon, A. C. (2004). Le domaine de la variation prosodique régionale. Aspects phonologiques et phonétiques du français parlé à Liège. Communication aux *Journées PFC*, Paris, 9-11 décembre 2004.
- Simon, A. C., Avanzi M. & Goldman J.-P. (2008). La détection automatique des proéminences syllabiques. Un aller-retour entre manuelle et le traitement automatique. In J. Durand, B. Habert & B. Laks (éds), *Congrès Mondial de Linguistique Française* (pp. 1685-1698). Paris : Institut de Linguistique Française.