

La synchronisation des profils temporel et mélodique en français spontané

François Poiré* et Henrietta J. Cedergren**

*The University of Western Ontario, Département de français, London, On., Canada, N6A 3K7, fpoire@uwo.ca

**Université du Québec à Montréal, Département de linguistique et de didactique des langues, C.P. 8888, Succ. A Montréal, Qc, Canada, H3C 3P8, cedergren.henrietta@uqam.ca

ABSTRACT

This study is concerned with the timing of two properties of the Intonational Phrase (IP) in spontaneous speech: its durational profile (DP) and its tune. High or low continuation or finality, or high-low continuation are distinguished in a corpus of 2395 IPs from 16 speakers of Montreal French. DP is derived from normalized variation of syllable duration provided by a Z-score analysis. DP is characterized by the passage from negative normalized values to positive ones in the last two syllables of the IP. Results show that most of the time the tune doesn't influence the evolution of DP. Only two types of IP with high-low continuation show a different behaviour: when a discourse marker is introduced at the end of the IP and when hesitation occurs at the same position. In both cases, the DP will be aligned with the penult of the IP.

1. INTRODUCTION

Toute recherche portant sur la parole spontanée doit résoudre en premier lieu le choix de l'unité de référence à l'intérieur de laquelle les phénomènes sont étudiés. Le domaine du contour intonatif intègre est souvent privilégié en raison de ses liens tant à l'organisation interne du référent linguistique (quasi isomorphie aux groupes syntaxiques, aux groupes de sens, [Sel96]), qu'aux propriétés physiques de l'énoncé (correspondance aux groupes de souffle, à la distribution des pauses, [Pie80]). Il a d'ailleurs été démontré [Ced94] qu'il était possible de modéliser les relations entre la variation de la durée syllabique et l'organisation prosodique de surface de l'énoncé si le domaine du contour intonatif (le «Syntagme Intonatif», SI) servait d'unité de référence majeure.

La présente étude s'intéresse à la synchronisation des profils temporel et mélodique de l'énoncé. Nous cherchons à savoir s'ils constituent les deux faces d'une même unité de production de la parole spontanée qui se développent parallèlement à une portion particulière du texte. Nous prenons comme point de départ l'étude d'un corpus de 16 locuteurs du français de Montréal pour lequel les propriétés relatives à la variation de la durée syllabique ont été mises à jour à l'aide d'une analyse par régressions multiples ([Ced95], [Ced94]).

2. MÉTHODOLOGIE

2.1 La codification des contours intonatifs

Notre corpus comprend trois minutes de parole pour 16 locuteurs du corpus Sankoff-Cedergren du français parlé de Montréal. Ces extraits ont été segmentés et codifiés selon certaines propriétés prosodiques dont le découpage en SI. Ces propriétés prosodiques (proéminences, découpage en différentes unités) ont été identifiées à l'origine sur une base impressionniste selon une procédure d'accord inter-juge [Poi90]. Pour la présente étude, à l'aide du logiciel PitchWorks, nous avons catégorisé «objectivement» le contour de chaque SI selon la nature du dernier mouvement de fréquence fondamentale (F0). Dans les faits, une catégorisation binaire s'est avérée suffisante pour décrire l'ensemble des contours, à savoir une montée de F0 sur la dernière syllabe du domaine (un accent haut) ou bien une descente de F0 en fin de domaine dont l'origine est un sommet fréquentiel associé à une syllabe autre que finale. En bref, il s'agit de distinguer les SI selon la position du dernier accent haut relativement à la syllabe finale. Cette distinction découle des propriétés accentuelles du français [DiC00],[DiC99] et des liens qui associent intonation et accentuation dans cette langue [DiC98], [Ros93].

Dans un second temps, un critère fonctionnel a été introduit afin de distinguer les profils mélodiques. En effet, de par leur nature, les entrevues de type sociolinguistique présentent peu de variation dans le choix de l'expression. On y retrouve essentiellement des réponses plus ou moins longues aux questions posées. Ces réponses prennent la forme d'une suite de SI qui selon leur position dans le propos du locuteur signalent la continuité ou la finalité. Ces SI peuvent être articulés prosodiquement par une montée ou descente mélodique sur la dernière syllabe. Les SI ont alors été étiquetés selon les quatre types majeurs qui suivent :

La Continuité Haute (CH)

La Finalité Haute (FH)

La Finalité Basse (FB)

La Continuité Haute-Basse (CHB)

Dans un SI à finalité basse, le dernier accent haut du contour est réalisé sur une syllabe autre que la dernière du domaine. Ce type de contour marque habituellement la fin d'un tour de parole et est généralement suivi d'une pause. Dans exactement 80% des cas, la position

du dernier accent haut est déterminée par la longueur (nb de syllabes) du dernier constituant du texte (un SN ou un SP, par exemple). Ainsi, le SI « nous parlions comme des paysans de France » voit le dernier accent haut réalisé sur la dernière syllabe du mot « paysans » (syllabe soulignée). Dans les autres cas, le dernier accent haut survient au début du dernier mot ou à la frontière d'une dislocation syntaxique du texte (« on aimait ça / les vacances »).

La quatrième catégorie (CHB) se distingue des autres de trois façons. Premièrement, tout comme la catégorie CH, elle exprime la continuité mais avec un accent haut porté par l'avant-dernière syllabe du domaine. Deuxièmement, elle se distingue de la catégorie FB par le fait qu'elle exprime la continuité même si le dernier accent haut n'est pas portée par la syllabe finale. Troisièmement, elle diffère des trois autres en raison de propriétés particulières du texte. En effet, tant CH, FH que FB montrent un texte « complet », c'est-à-dire que le contenu propositionnel du SI correspond minimalement à un constituant majeur de l'énoncé. Dans le cas des CHB, le texte comporte une entorse à cette notion de texte complet. Dans tous les cas, la pénultième de domaine porte le dernier accent haut.

CHB1 : le SI se termine par une voyelle d'hésitation

exemple : [mais je pense euh]

CHB2 : le SI se termine par un mot de fonction (article, préposition, conjonction)

exemples : [on avait l'impression que]

[présenté à la manière de]

CHB3 : le SI se termine par le marqueur discursif « là » typique de l'oral montréalais (qui diffère du démonstratif « là »)

exemple : [et après ça là]

L'exemple qui suit, tiré du locuteur 081, montre une suite de SI étiquetée selon ces critères :

Mais]CH

On avait l'impression que]CHB2

L'université]CH

À l'université]CH

Les professeurs et les étudiants]CH

Étaient au service des administrateurs]FH

2.2 Le profil temporel

Les profils temporels sont établis sur la base d'un indice de variation de la durée relative à la moyenne d'une classe de segments donnée [Ced94]. Cet indice est représenté par une cote-z pour chaque occurrence des segments calculée pour chacun des locuteurs du corpus. Par exemple, pour un locuteur donné, la

moyenne des durées de la voyelle [o] est établie. Chaque occurrence de cette voyelle se voit ensuite attribué un indice de variation relatif à cette moyenne par le calcul de la cote-z. La somme des indices segmentaux d'une syllabe donne l'indice de variation pour cette syllabe. Nous avons établi un profil temporel pour chaque SI du corpus. Nous avons ensuite calculé pour chaque locuteur la moyenne des indices de variation pour chacune des positions dans le SI (syllabe finale, syllabe finale -1, -2, -3, -n) et ce pour chaque catégorie de SI (CH, FH, FB, CHB1, CHB2, CHB3). Cette procédure permet d'évaluer l'influence relative sur la durée segmentale (et donc syllabique) de la position du segment dans le SI. Il en ressort un profil temporel canonique – une signature temporelle – propre à ce domaine prosodique qui prend la forme d'un passage de valeurs négatives vers des valeurs positives de l'indice de variation lors de la réalisation des deux dernières syllabes du domaine (Figure 1).

Après analyse des profils temporels et intonatifs individuels, il s'est avéré que tous les locuteurs présentaient les mêmes différences ou similarités de synchronisation selon les différentes catégories de SI. Nous présentons donc dans la section suivante les résultats en tenant compte des données moyennes de tout le corpus.

3. RÉSULTATS

La première table donne la distribution globale des SI du corpus selon le type de contour. On y voit que la majorité se termine par un accent haut (CH +FH = 77.45%). Par contre, 13.82% des SI montrent un contour descendant sur la ou les dernière syllabes du domaine (FB). Cette différence dans la réalisation du dernier accent haut n'a par contre aucune incidence sur la réalisation du profil temporel (Figure 1)

Table 1: Distribution des SI par type de contour (16 locuteurs)

	N =	En %
CH	1418	59.20
FH	437	18.25
FB	331	13.82
CHB1	56	2.34
CHB2	92	3.84
CHB3	61	2.55
Total	2395	100

Les trois premiers types de contour montrent un même profil temporel qui s'exprime principalement par le passage de valeurs négatives à des valeurs positives de la cote-z lors de la réalisation des deux dernières syllabes du SI. Dans les trois cas, la variation positive culmine sur la syllabe finale du domaine, marquant une synchronisation parfaite avec le profil intonatif. La finalité haute se distingue en ce point culminant de la

continuité haute par un indice positif plus élevé, ce qui pourrait s'expliquer par un allongement marquant la finalité. Par contre, la finalité basse montre un indice moins élevé que la continuité haute pour cette même syllabe de fin de domaine.

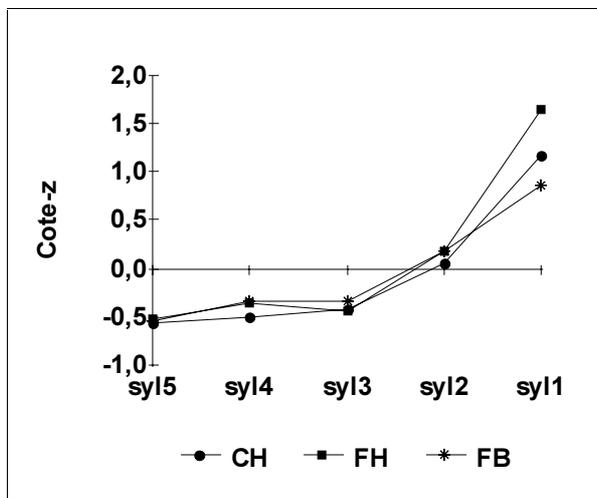


Figure 1. Profil temporel moyen des SI de type CH, FH et FB (16 locuteurs).

Dans tous les cas, la tendance générale exprimée par le profil temporel est clairement marquée. Elle peut s'expliquer par une décélération graduelle de la vitesse de production qui s'accroît dans la portion finale du domaine et culmine sur la dernière syllabe à laquelle est aussi associée la limite droite du contour mélodique.

Ces trois types de contour ont en commun d'être associés à du texte « complet », tel que défini dans la section 2.1. Qu'advient-il des SI de type CHB qui ne comptent que pour 9.73% du corpus (Table 1) ?

Les CHB se distinguent des CH, FH et FB par leur contour ainsi que par le texte auquel ils sont associés. Dans les trois cas, soit qu'un ajout au contenu propositionnel survient (hésitation, « là » discursif), soit que le texte est découpé sans respecter les frontières de syntagmes syntaxiques. Ils se distinguent aussi par de plus grandes valeurs positives de cote-z réalisées sur la syllabe où culminent ces valeurs (Figure 2). Ces valeurs positives supérieures ne sont pas surprenantes si l'on considère que CHB1 et CHB2 correspondent à des moments de l'énonciation où le locuteur interrompt ou ralentit le flot normal de la parole (avec ou sans pause) pour mieux reprendre le fil de son discours. Il s'agirait alors d'une décélération plus marquée. Nous pourrions peut-être étendre cette interprétation au CHB3 si nous disposions d'une analyse de l'insertion des marqueurs discursifs à la parole continue.

Mais les CHB ne forment pas une classe de contours homogène. D'une part, CHB2 ([on avait l'impression que]) présente le même profil temporel que CH, FH et FB avec une plus grande variation des valeurs positives de cote-z sur les deux dernières syllabes du domaine. On peut donc croire qu'ici les profils temporel et intonatif sont synchronisés. D'autre part, CHB1 et

CHB3 montrent un comportement différent. Les valeurs positives de l'indice de variation culminent sur la pénultième du domaine qui est en fait la dernière syllabe du texte normal (avant la syllabe d'hésitation, avant le « là » discursif). Ces valeurs positives redescendent ensuite sur la syllabe finale à des niveaux comparables à ceux de la syllabe finale de la Figure 1. Le profil temporel des CHB1 et CHB3 est donc désynchronisé relativement au profil intonatif qui inclut par définition la syllabe finale du SI

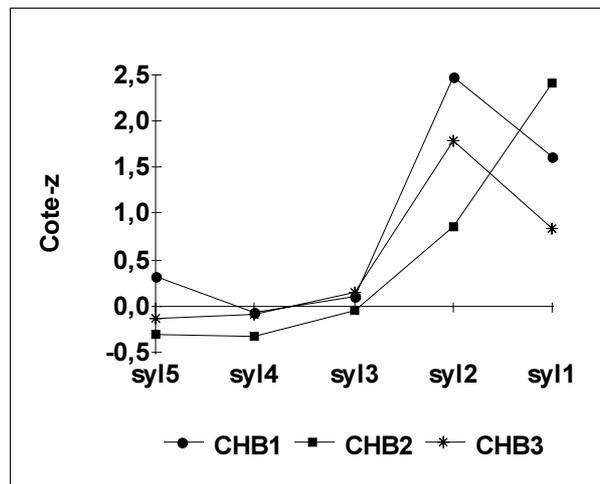


Figure 2. Profil temporel moyen des SI de type CHB1, CHB2 et CHB3 (16 locuteurs).

4. DISCUSSION

La Figure 1 montre clairement que les deux caractéristiques servant à la définition du profil prosodique, soit le critère fonctionnel et celui lié à la distribution de l'accent haut final, n'influencent pas la réalisation du profil temporel dans le cas des SI de type CH, FH et FB. Cela illustre une synchronisation parfaite des profils mélodique et temporel lorsque le texte est lui-même « complet » tel que défini à la section 2.1.

Dans le cas des contours à Finalité Basse (FB), il est intéressant de noter que l'absence d'accent haut porté par la dernière syllabe n'influence en rien la réalisation du profil temporel. On peut y voir une confirmation ne serait-ce que partielle de la théorie métrique/ autosegmentale qui prédit ce type d'indépendance entre les deux aspects de la prosodie de l'énoncé. L'analyse des FB présente deux autres points d'intérêt. D'une part, le positionnement du dernier accent haut semble dépendre principalement d'une contrainte d'ordre pragmatique (signaler la fin d'un tour de parole); d'autre part, elle vient ajouter un chapitre à la longue discussion sur la nature de l'accentuation française en finale de mot et sur ses corrélats physiques. L'analyse des FB montre que la proéminence propre à cette position en français résulte de plusieurs facteurs qui n'incluent pas nécessairement la présence d'un accent haut associé à la dernière syllabe.

Les CHB, même si minoritaires dans ce corpus, jettent quant à eux un éclairage intéressant sur la « planification » des unités de production que sont les SI. Le moins surprenant est le CHB1 (hésitation) dans la mesure où il semble logique qu'une hésitation ne soit normalement pas prévue dans la production du texte. Si le profil temporel montre cette disjonction entre le texte et la syllabe d'hésitation, le contour mélodique doit quant à lui jouer le rôle « d'intégrateur » à ce matériel segmental inattendu. Les CHB2 nous intéressent pour deux raisons. Premièrement, et contrairement à plusieurs théories syntaxiques contemporaines, il semble que le mot de fonction qui termine ces SI soit facilement détachable du constituant majeur auquel il appartient. Ainsi, un article devrait accompagner le nom qu'il détermine et non clore le domaine qui précède ce nom. Le profil temporel identique à celui des CH, FH et FB montre que l'unité est bel et bien voulue comme se terminant par ce mot « faible ». Deuxièmement, Le fait que le dernier accent haut de ce type de contour soit associé au mot lexical qui précède ce mot de fonction va dans le sens de Dell [Del84] qui argue qu'il existe des propriétés accentuelles en français qui dépendent non pas de la position dans le domaine prosodique (dernière syllabe) mais bien de la propriété abstraite « stress » associée à une syllabe dont le noyau n'est pas un « e muet ». Ce dernier point va à l'encontre de Beckman [Bec86] qui voit l'accentuation française comme un procédé purement démarcatif. Enfin, le CHB3 (« là », marqueur discursif) ouvre la porte à une étude psycholinguistique sur le moment où est rattaché ce type de marqueurs au corps principal de l'énoncé. On pourrait croire que les éléments appartenant à cette catégorie relevant de la grammaire du discours sont insérés tardivement dans le flot de la parole. Leur exclusion du profil temporel propre aux autres SI nous porte à penser ainsi. Et encore une fois, le contour intonatif vient intégrer cette syllabe à l'énoncé.

D'autres études seraient nécessaires afin de confirmer le comportement des propriétés tant temporelles qu'intonatives du SI, en particulier dans d'autres variétés du Français. Cette remarque tient aussi pour des types fonctionnels de contours autres que finalité et continuité que la nature même du présent corpus nous empêche d'explorer plus à fond.

REMERCIEMENTS

Cette recherche a été réalisée avec l'aide du Conseil de la Recherche en Sciences Humaines du Canada, subvention no 410-96-0950.

BIBLIOGRAPHIE

- [Bec86] Beckman M. (1986) Stress and non-stress accent, Foris.
- [Ced95] Cedergren, H. et H. Perreault (1995) "On the analysis of syllable timing in everyday speech", *JCPHS* 95, pp.232-235.
- [Ced94] Cedergren, H. et H. Perreault (1994) "Speech rate and syllable timing in spontaneous speech", *Proceedings of the International Conference on Spoken Language Processing*, pp. 1087-1090
- [Del84] Dell, F. (1984) "L'accentuation dans les phrases en français", in F. Dell, J.-R. Vergnaud, and D. Hirst (eds), *Les représentations en phonologie*, Herman.
- [DiC00] Di Cristo, A. (2000) "Vers la modélisation de l'accentuation du français (seconde partie)", *French Language Studies*, 10, pp. 27-44.
- [DiC99] Di Cristo, A. (1999) "Vers la modélisation de l'accentuation du français (première partie)", *French Language Studies*, 9, pp 143-179.
- [DiC98] Di Cristo, A. (1998) "Intonation in French", in D. Hirst and A. Di Cristo (eds), *Intonation systems. A survey of twenty languages*, Cambridge University Press.
- [Pie80] Pierrehumbert, J. (1980) *The phonology and phonetics of English intonation*, PhD thesis, MIT.
- [Poi90] Poiré, F., J.M. Sosa, H. Perreault and H. J. Cedergren (1990), "Le syntagme intonatif en langage spontané: étude préliminaire", *Revue Québécoise de linguistique*, vol. 19, 2, pp. 93-109.
- [Ros93] Rossi, M. (1993) "A model for predicting the prosody of spontaneous speech", *Speech Communication*, 13, pp. 87-107
- [Sel96] Selkirk, E. (1996) "Sentence prosody: intonation, stress, and phrasing", in J. A. Goldsmith (ed), *The Handbook of Phonological Theory*, Blackwell.