

# Valeurs pragmatiques de la proéminence prosodique lexicale : de l'outil vers l'analyse

Geneviève Caelen-Haumont

Laboratoire Parole et Langage  
Université de Provence, 29 av. R. Schuman – 13621 Aix-en-Provence Cedex 1, France  
Tél.: ++33 (0)4 42 95 36 39 - Fax: ++33 (0)4 42 59 50 96  
Mél: gcaelen@lpl.univ-aix.fr - <http://www.lpl.univ-aix.fr/jep-taln04/>

## ABSTRACT

Based on previous studies, this paper aims 1° at checking a new model of superficial phonology annotation in tonal sequences dedicated to the analysis of the F0 prominence structure in words (INTSMEL), whatever the language type, tonal or not, 2° at using it in a semantic and a pragmatic approach of spontaneous dialogue. INTSINT and INTSMEL are both integrated in the Praat speech analysis software, the former dedicated to the annotation of intonation, the latter to that of lexical prominence.

## 1. INTRODUCTION

Le contenu de cet article est dans le prolongement des études antérieures. Ainsi dans Caelen-Haumont & Bel [5], nous avons distingué deux espaces prosodiques : celui de l'intonation en relation avec l'organisation prosodique de l'énoncé et de ses groupes, et par ailleurs du « mélisme », en relation avec l'item lexical. Dans Caelen-Haumont [2], à propos d'un dialogue improvisé (et également de chants improvisés marathi), nous avons réalisé une analyse perceptive des valeurs et fonctions pragmatiques des mélismes. Dans Caelen-Haumont & Auran [4], nous avons défini un outil d'analyse phonologique des mélismes, qui permet de coder automatiquement sous Praat les structures phonologiques superficielles des mélismes. L'objectif présent de cet article est de tester la validité de cet outil phonologique pour l'analyse pragmatique du même corpus. Pour une étude plus approfondie, se reporter à Caelen-Haumont [3].

## 2. INTONATION ET MÉLISME

Caelen-Haumont & Bel [5] proposent de distinguer dans la prosodie deux strates, l'une qui serait la partie linguistique et conventionnelle, correspondant à l'intonation (en relation avec la structure linguistique), et l'autre qui serait la partie exprimant l'affectivité (« l'émotion-racine » à la source de l'individu), exprimée par le *mélisme*, d'expression locale souvent mono-lexicale. Le mélisme serait caractérisé par une amplitude de F0 importante, ce qui implique des valeurs maximales de F0, et serait ainsi souvent lié à une rupture mélodique ou prosodique de la trame linguistique, rupture intervenant donc en dehors des frontières de syntagmes.

Etablie sur des données prosodiques, cette partition est par ailleurs appuyée par le contexte des recherches actuel en psycholinguistique et neuronal (aires cérébrales). Tout d'abord, comme on le sait, Scherer [10] a testé deux

modèles, *covariance* et *configuration*, le premier stipulant que l'information sur l'émotion et l'information sur le contenu linguistique fonctionnent de manière indépendante, ce qui implique que les traitements linguistiques et paralinguistiques puissent s'effectuer en parallèle. Les indices de F0 associés à cette approche sont le F0 global et l'amplitude de F0. Le second pose que le type de contour de F0 est un élément linguistique, ce qui induit deux types de variations de F0, linguistiques et paralinguistiques, et sous-entend que l'information liée au contour et à l'émotion n'est accessible que via les traits linguistiques de l'énoncé. L'indice de F0 correspondant est le type de contour. Au terme d'expérimentations perceptives, il apparaît que les deux approches sont pertinentes, la covariance semblant plus adéquate pour décrire la parole affectée par les facteurs biologiques (i.e. émotionnelles), la configuration, pour la parole affectée par les conventions linguistiques et socioculturelles.

Cette dichotomie des fonctions est encore reprise par le modèle de Scherer "pull" et "push" [9], 2 processus qui rendent compte des expressions de la parole émotionnelle sous contrôle cognitif. Il y aurait ainsi 2 types d'effets au niveau des expressions de la parole affective, les effets "push" produits par des "déterminants internes au sujet" causant des modifications physiologiques en réponse à une activation émotionnelle, et les effets "pull" produits par des "déterminants externes" en relation avec le contexte environnemental social où se trouve le locuteur.

Par ailleurs au niveau des aires cérébrales, il semble attesté dans Pell [8] d'une part que l'hémisphère gauche soit plutôt spécialisé dans le traitement de la linguistique et de la prosodie linguistique (niveau représentationnel), alors que l'hémisphère droit aurait plus d'habileté à traiter la prosodie émotionnelle, avec les indices de F0, véhicules attestés des intentions liées aux émotions.

Ainsi selon nous, ces différentes études valident à propos de la parole affective, la distinction intonation / mélisme, établie à partir de l'analyse des données, comme deux espaces spécialisés de l'information, en interaction, l'un lié plus spécifiquement à l'expression de la prosodie linguistique, l'autre à celle de la prosodie affective.

## 3. CARACTÉRISATION PROSODIQUE DU MÉLISME

En l'absence de terme approprié, nous avons emprunté au domaine musical le terme de *mélisme*. Ce terme s'applique dans le cadre d'un item lexical, ou au plus d'une suite de mots, à la *forme* acoustique et mélodique, désignant une structure phonologique de surface, avec une granularité adaptée à ses objectifs [2][3].

Concernant cette granularité, la division du registre du locuteur selon 4 niveaux, usuelle depuis Delattre (1966), nous a paru insuffisante pour nos objectifs. Notre codage étant effectué de manière automatique sous Praat par la procédure INTSMEL, en accord avec la procédure INTSINT, nous avons en fait divisé chaque niveau en 3, le « cœur » représentant 50% du niveau, chacune des marges inférieure et supérieure, 25%, ce qui, hormis les marges extrêmes (25%) du registre du locuteur, correspond à une succession de 7 plages de 50% (échelle logarithmique). Nous arrivons ainsi à 9 niveaux (Table 1 ci-dessous) : aigu (a), supérieur (s), haut (h), élevé (e), moyen (m), centré (c), bas (b), inférieur (i), grave (g). Pour une justification de ces niveaux, se reporter à [3].

**Table 1 :** Matrice des séquences tonales pour la description des configurations mélodiques des mots, en particulier les mélismes : en gras les séquences mono- et bitonales du mélisme, en italiques, les plateaux.

ton	aigu	supérieur	haut	élevé	moyen	centré	bas	inférieur	grave
	a	s	h	e	m	c	b	i	g
a	<b>aa</b>	<b>as</b>	<b>ah</b>	<b>ae</b>	<b>am</b>	<b>ac</b>	<b>ab</b>	<b>ai</b>	<b>ag</b>
s	<b>sa</b>	<b>ss</b>	<b>sh</b>	<b>se</b>	<b>sm</b>	<b>sc</b>	<b>sb</b>	<b>si</b>	<b>sg</b>
h	<b>ha</b>	<b>hs</b>	<b>hh</b>	<b>he</b>	<b>hm</b>	<b>hc</b>	<b>hb</b>	<b>hi</b>	<b>hg</b>
e	<b>ea</b>	<b>es</b>	<b>eh</b>	<i>ee</i>	<i>em</i>	<i>ec</i>	<i>eb</i>	<i>ei</i>	<i>eg</i>
m	<b>ma</b>	<b>ms</b>	<b>mh</b>	<i>me</i>	<i>mm</i>	<i>mc</i>	<i>mb</i>	<i>mi</i>	<i>mg</i>
c	<b>ca</b>	<b>cs</b>	<b>ch</b>	<i>ce</i>	<i>cm</i>	<i>cc</i>	<i>cb</i>	<i>ci</i>	<i>cg</i>
b	<b>ba</b>	<b>bs</b>	<b>bh</b>	<i>be</i>	<i>bm</i>	<i>bc</i>	<i>bb</i>	<i>bi</i>	<i>bg</i>
i	<b>ia</b>	<b>is</b>	<b>ih</b>	<i>ie</i>	<i>im</i>	<i>ic</i>	<i>ib</i>	<i>ii</i>	<i>ig</i>
g	<b>ga</b>	<b>gs</b>	<b>gh</b>	<i>ge</i>	<i>gm</i>	<i>gc</i>	<i>gb</i>	<i>gi</i>	<i>gg</i>

Ces 9 niveaux en combinaison créent 81 séquences mono- et bitonales. En fait ces 81 *tons* rendent compte de toute configuration mélodique du mot ou du groupe de mots, pour les langues tonales ou non tonales, mais pour notre analyse des mélismes, seulement 45 tons (correspondant aux niveaux a, s, h, les plus aigus) sont retenus.

Par ailleurs, pour une meilleure précision, nous avons fait une distinction entre les variations intra-niveau et les plateaux. Par opposition aux plateaux qui par définition sont non orientés (ex. : *aa*), les variations intra-niveau sont toujours annotées avec le symbole +/- (ex. : *mm+* ; *ii-*, etc.), qui code simplement l'existence d'une pente ascendante vs. descendante, déterminée par un seuil. Les niveaux sont naturellement gradués ( $a > g$ ), mais dans une suite, deux ou plusieurs cibles identiques de tons peuvent coexister, et alors on ne sait pas, quand elles ne sont pas adjacentes, où se trouve la F0 max. Une lettre capitale permet d'éviter toute ambiguïté : ainsi <as sA As sS> indiquerait que 1° sA, As > as, 2° S > s, (et bien sûr que A > S).

Nous pouvons maintenant préciser que les corrélats prosodiques du mélisme, justifiés par l'expérience, sont 1° l'implication des niveaux les plus aigus, soit les niveaux *a* et *s* ou 2° une large excursion de F0 (interne ou externe, si une rupture de F0 intervient entre le mot courant et le mot précédent / suivant), atteignant au minimum le niveau *h*, à condition que l'amplitude soit suffisante (excluant les niveaux *e* et *m*), l'un ou l'autre

corrélait impliquant généralement un ralentissement sensible du débit et éventuellement une augmentation importante de l'énergie.

#### 4. INTSINT ET INTSMEL

INTSINT et INTSMEL sont 2 procédures automatiques [4], fonctionnant sous PRAAT. Elles se développent l'une et l'autre à partir de la procédure MOMEL définie par Hirst & Espesser [7], qui permet de calculer et même de reconstruire la valeur de F0 sous-jacente aux segments non voisés et pauses, réalisant alors une courbe de F0 continue, proche de la perception auditive. Ainsi indépendamment de la structure linguistique, MOMEL repère dans le continuum les valeurs min / max relatives.

INTSINT et INTSMEL proposent tous deux une annotation mélodique automatique. La différence entre ces deux systèmes repose sur leur portée, et leur type de calcul des séquences tonales et sur la référence à la strate linguistique. Ainsi en bref, selon Hirst & Espesser [7], Hirst & Di Cristo [6], INTSINT est un codage de l'intonation (sans référence à la couche linguistique), qui sur la base de tons absolus (extrema du registre et milieu), analyse les valeurs relativement à la cible précédente. De manière complémentaire, INTSMEL code les séquences tonales de manière absolue dans le cadre de séquences courtes (mot ou suite de mots). En fait la procédure INTSMEL par elle-même est indépendante de la couche linguistique et fonctionne comme telle, mais l'usage que nous en faisons personnellement en l'appliquant sur les items segmentés et étiquetés, l'inscrit dans le contexte du mot ou d'une suite. Ainsi dans le champ du codage de F0, on retrouve cette même dichotomie évoquée paragraphe 2, le système INTSINT codant l'intonation, le système INTSMEL, les mélismes.

#### 5. ANALYSE PRAGMATIQUE

Cette seconde partie concerne l'application de l'outil d'annotation INTSMEL à notre corpus, ce qui facilite une analyse plus objective des valeurs et fonctions pragmatiques des mélismes.

##### 5.1. Corpus et objectifs du dialogue

Inspiré du Corpus HCRC Map Task de Brown & al. [1], le corpus étudié est l'un des 8 dialogues, d'une quinzaine de mn chacun, enregistrés dans de très bonnes conditions (chambre sourde), et avec pistes séparées pour chaque locuteur. Ce dialogue met en scène un jeu de rôles avec 2 locutrices, une « touriste » et une « employée de l'office de tourisme », avec des conflits d'objectifs, la touriste préférant les activités sportives, l'employée préconisant les activités culturelles, et des ambiguïtés spatiales (les plans n'ayant pas la même version de la ville). Il faut réactualiser le plan à la borne d'information, établir un programme, fixer l'itinéraire en fonction des nouveautés de la ville (travaux, sens circulatoire, noms des rues). L'étude porte comme précédemment sur les réalisations de la locutrice O4 (étudiante, 22 ans, français standard).

Nous formulons les hypothèses suivantes : 1° la locutrice s'implique subjectivement dans son rôle d'employée du

tourisme 2° les indices de cette implication se trouvent au niveau mélodique dans les mélismes 3° les items lexicaux « mélismés » correspondent aux objectifs du dialogue et aux motivations personnelles de la locutrice.

En fonction des objectifs généraux du dialogue exposés ci-dessus, nous posons que ces objectifs sont répertoriés grâce à 5 champs lexicaux des domaines sémantique et pragmatique :

- champ A : les coordonnées spatiales de la ville, des objets touristiques (noms et directions)
- champ B : les qualités descriptives des objets culturels (spatiales, temporelles, architecturales, économiques etc.)
- champ C : les outils informatiques, manipulation, conseils, mise à jour du plan,
- champ D : l'appréciation subjective des objets touristiques, des tâches respectives, l'argumentation,
- champ E : l'interaction verbale, particulièrement les verbes et adverbess de l'interaction, les particules phatiques.

### 5.2. Mélismes et expérimentation

Pour cette expérimentation, nous n'avons retenu dans un premier temps que les mélismes des items lexicaux atteignant exclusivement les niveaux les plus aigus (a, s).

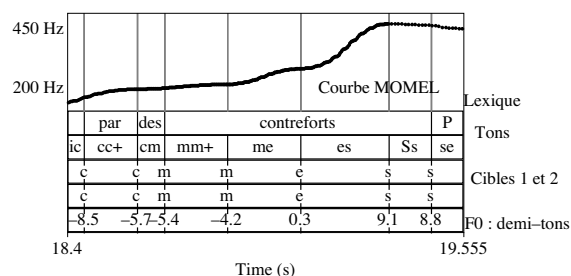
**Table 2 :** Pourcentages des mélismes en fonction des champs sémantiques / pragmatiques des mots lexicaux, A : coordonnées spatiales, B : qualité des objets, C : tâche technique, D : évaluation subjective et argumentation, E : interaction verbale.

	A	B	C	D	E
72	14 / 72	16 / 72	1 / 72	26 / 72	15 / 72
100%	19%	22%	1%	36%	21%
	57 / 72		79%	15 / 72	
				21%	

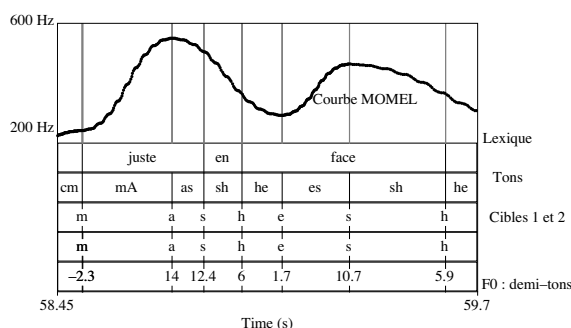
Sur les 1860 mots du corpus O4, 72 ont été l'objet de ce type de mélisme, se distribuant en 35% à la finale d'un groupe prosodique (groupe syntaxique avec pause), et 65% ailleurs. Après segmentation et étiquetage, les cibles ont été recalculées au sein des bornes des mots lexicaux pour définir les séquences tonales.

La distribution des mélismes en fonction des champs lexicaux est présentée table 2 ci-dessus. On constate que 79% des mots lexicaux mélismés (a et s) correspondent aux champs lexicaux liés aux objectifs du dialogue (A, B, C, D), soit les 3/4. Le quart restant correspond au lexique de l'échange discursif (E, 21%). Ces résultats sont proches de ceux de l'étude perceptive (Caelen-Haumont, [2]) : en effet sur les 151 mélismes sélectionnés auditivement (soit le double), 71% correspondaient aux champs lexicaux des objectifs du dialogue, 28 % au lexique des échanges discursifs. Notons que nous n'avons étudié présentement que les mélismes a et s. Les exemples ci-dessus (Figures 1 et 2) présentent des extraits de parole avec mélisme, le premier extrait « ... protégé par des contreforts ... », de forme ascendante puis plateau,

avec 14.5 demi-tons d'amplitude, ainsi que les séquences tonales qui le décrivent.



**Figure 1 :** exemple de mélisme sur le mot *contreforts*, avec de bas en haut, les valeurs de F0 en demi-tons, les cibles MOMEL (cibles 1 : valeurs directes, cibles 2 : recalculées), les séquences tonales (structure phonologique superficielle), le lexique, et enfin la courbe MOMEL (Hz).



**Figure 2 :** Séquence de mots « momélisée » *juste en face*. Même légende que Figure 1.

L'exemple 2 (« *juste en face* [...] y'a l'marché ») est une suite de 3 mots mélismés, présentant une série de pentes montantes et descendantes, avec une amplitude très importante sur *juste* (16.3 demi-tons), parvenant dans la zone maximale du registre de la locutrice (près de 600 Hz). Comme on le voit la procédure INTSMEL décrit avec précision la courbe de F0, ses formes et ses valeurs, distinguant bien notamment entre les variations inter-niveau (Figure 1, ex : *mm+*) et les plateaux (ex : *sS*). Dans la figure 2, inversement, les plateaux sont absents. Pour une étude plus approfondie, cf [3].

### 5.3. Valeurs, croyances, buts perlocutoires, fonctions prosodiques

Ces résultats nous permettent d'une part, par le compromis satisfaisant qu'il offre entre généralité et précision, de valider notre outil de sélection et d'annotation des mélismes, et d'autre part nous autorisent à poursuivre l'analyse dans le domaine pragmatique. Nous présentons ici une analyse préliminaire. Pour une version plus étendue, cf [3].

L'analyse retient tout d'abord que la locutrice s'est parfaitement investie dans sa tâche d'employée du tourisme. En effet si la majorité du vocabulaire mélismé se trouvait en dehors des champs lexicaux liés aux objectifs du dialogue, nous aurions pu en conclure que le

locuteur n'avait pas été motivé par sa tâche, ce qui manifestement n'est pas le cas.

Mais sur quoi repose cet investissement ? Autrement dit quelles sont les valeurs qui le fondent ? En fait dans les mots mélismés, on trouve 2 types de valeurs, celles qui sont transmises par les consignes du dialogue, à savoir informer, établir un programme de visites culturelles, élaborer un itinéraire, mettre à jour le plan, et par ailleurs celles qui sont propres à la locutrice O4 : l'importance qu'elle accorde aux éléments naturels (les arbres du parc, la roseraie, les oiseaux ...), traditionnels (marché, artisanat, costumes ...), aux éléments esthétiques des monuments anciens, conviviaux (accueil et atmosphère des restaurants, amabilité des commerçants ...). En réalité ces différentes valeurs, d'ailleurs compatibles entre elles, sont plus profondément des croyances, véritables sources d'énergie (cf l'état *d'excitation*, ou *arousal* en anglais, bien décrit dans la littérature), pour la parole et donc la prosodie. De ce fait ces croyances sont également de 2 lieux culturels, de l'aider dans ses choix, l'informer sur le plan touristique, dans sa tâche technique, et de le faire sur le registre de l'amabilité et de la disponibilité. Relativement à la locutrice, c'est croire (et faire croire) qu'un objet touristique donné (château, église, restaurant traditionnel ...) nécessite un déplacement, que les éléments et les repères de la ville (ronds-points, sens uniques ...) sont une connaissance indispensable pour atteindre mieux et plus vite le point de visite, qu'une caractéristique de cet objet (historique, esthétique, exotique, caractère exceptionnel ...) est susceptible de la séduire. En fait la croyance repose sur le sentiment que ces valeurs sont partageables et motivantes pour autrui. types, celles qui sont issues de la tâche (et dans ce cas, tous les locuteurs des différents dialogues qui se sont investis, les partagent), et celles qui sont propres à la locutrice.

## BIBLIOGRAPHIE

- [1] G. Brown, A. Anderson, R. Shillcock, and G. Yule. *Teaching talk: Strategies for production and assessment*. Cambridge University Press, 1984.
- [2] G. Caelen-Haumont. Perlocutory Values and Functions of Melisms in Spontaneous Dialogue. In *Proceedings of the 1<sup>st</sup> International Conference on Speech Prosody*, SP 2002, Aix-en-Provence,, 195-198, on line <<http://www.lpl.univ-aix.fr/sp2002/papers.htm>>, 2002.
- [3] G. Caelen-Haumont. Melodic lexical prominence and melisms: an automatic tonal annotation by INTSMEL and its application to pragmatic analysis (submitted).
- [4] G. Caelen-Haumont, C. Auran. The Phonology of Melodic Prominence: the Structure of melisms. In *the 2nd International Conference on Speech Prosody Proceedings*, SP 2004, (to be published).
- [5] G. Caelen-Haumont, B. Bel. Le caractère spontané dans la parole et le chant improvisés : de la structure intonative au mélisme, *Revue Parole*, 15/16, 251-302, 2000.
- [6] D. Hirst, A. Di Cristo, R. Espesser. Levels of Representation and Levels of Analysis for the Description

Relativement à la tâche, c'est croire (et faire croire) à la nécessité de convaincre la touriste de se rendre dans les De ce fait, les mélismes qui mettent en relief par le registre aigu (niveaux *a* et *s* large excursion mélodi-que) tout ce vocabulaire attaché aux valeurs de la tâche et de la locutrice, exercent, indépendamment de la strate linguistique, et comme elle, une fonction perlocutoire par l'action sur autrui, volontaire, consciente ou pas. C'est sans doute par les notes aiguës du mélisme que l'auditeur identifie, attachée à ces valeurs exprimées lexicalement, l'expression d'une vérité du locuteur, d'une subjectivité ou « émotion-racine », qu'il peut y adhérer, ou au contraire la réfuter. L'espace lexical du mélisme est à ce titre un lieu de rencontre avec l'autre, une communication privilégiée de l'intersubjectivité et du passionnel.

Dans cette perspective, pour éviter la circularité du raisonnement, une étude complémentaire devra être menée sur tous les mots lexicaux du corpus qui entrent dans les 5 champs, afin de savoir quelle est la proportion des mots mélismés (*a*, *s*, *h*), et quelles sont les caractéristiques acoustiques des mots qui ne le sont pas.

## 6. Conclusion

Nous avons défini dans ces pages un outil de description phonologique (de surface) des prééminences mélodiques des unités lexicales, ou mélisme, et son application à l'analyse sémantique et pragmatique de ces contours. Cet outil est adéquat dans la mesure où il simplifie la courbe mélodique naturelle de F0, et la structure sur les plans quantitatif et qualificatif, et rend l'analyse linguistique et pragmatique plus objective. L'analyse phonologique doublée d'analyses sémantique (champs lexicaux) et pragmatique, permet de caractériser les objectifs du locuteur et ses valeurs. La prochaine étape de l'analyse portera sur l'analyse des contours des mélismes.

- of Intonation Systems. In Horne (ed.): *Prosody: Theory and Experiment Text, Speech and Language. Technology*, 14. Kluwer Academic Publishers, 51-87, 2000.
- [7] D. Hirst, R. Espesser. Automatic modelling of fundamental frequency using a quadratic spline function. In *Travaux de l'Institut de Phonétique d'Aix*, 15, 71-85, 1993.
- [8] M. D. Pell. Surveying Emotional Prosody in the Brain. In B. Bel, I. Marlien (eds.) *Proceedings of the 1<sup>st</sup> International Conference on Speech Prosody*, SP 2002, Aix-en-Provence, on line, <<http://www.lpl.univ-aix.fr/sp2002/papers.htm>>, 2002.
- [9] K. R. Scherer. Vocal correlates of emotion. In Wagner and Manstead Eds., *Handbook of psychophysiology: Emotion and social behavior*, London, Ed. Wiley, 1989.
- [10] K. R., Scherer, D. R. Ladd, K. E. A. Silverman. Vocal cues to speaker affect: Testing two models, *J.A.S.A.*, 76 (5), 1346-1356, 1984.