

Hauteurs mélodiques en français : variations continues ou catégorielles ?

David Le Gac¹ Hiyon Yoo² Katarina Bartkova³

(1) Université de Rouen, DySoLa, Rue Thomas Becket, Mont Saint Aignan

(2) Université Paris Diderot, Laboratoire de Linguistique Formelle, Case 7003, 75205 Paris cedex

(3) Université de Lorraine, ATILF, 44 avenue de la libération, B.P. 30687, 54063 Nancy cedex7

david.le.gac@gmail.com, yoo@linguist.jussieu.fr,

katarina.bartkova@atilf.fr

RESUME

Dans cet article, nous présentons les résultats d'une expérience visant à déterminer si les variations de f_0 à la fin du groupe intonatif terminal du français sont catégorielles ou continues. Nous avons demandé à sept locuteurs natifs d'imiter des stimuli où la f_0 a été resynthétisée en un continuum de 26 stimuli séparés d'un demi-ton. Les résultats sont mitigés : si certains locuteurs sont capables de reproduire presque parfaitement le continuum de stimuli, les performances d'autres locuteurs sont catégorielles, suggérant l'existence de trois voire quatre catégories de hauteurs en français.

ABSTRACT

Pitch height in French: Gradual or categorical variations?

This paper reports the results of an experiment on the question of whether the realizations of f_0 variations at the end of the final IP are categorical or gradient. We conducted an imitation task with resynthesized stimuli where the final pitch height was varied in steps of one semi-tone. Results are ambivalent, since both strategies are possible. However, we argue that there is enough evidence for establishing at least three pitch categories.

MOTS-CLES : intonation, tâche d'imitation, français, variations catégorielles ou continues, hauteurs mélodiques.

KEYWORDS: intonation, imitation task, French, continuous, categorical, pitch level, range.

1 Introduction

Une des questions centrales des études sur l'intonation est de déterminer le nombre de catégories phonologiques qui sous-tendent les variations de hauteurs mélodiques. Ainsi, la théorie métrique et autosegmentale (MA) « standard », qui domine à l'heure actuelle le champ des études sur l'intonation, postule-t-elle deux catégories tonales – « ton bas » ~ « ton haut ». C'est en général l'analyse retenue pour décrire l'intonation de l'anglais et de bien d'autres langues. Cependant, l'hypothèse des deux catégories tonales n'est pas définitive. Même pour l'anglais, des auteurs avaient proposé plus de deux niveaux (cf. Ladd, 2008 pour une revue) ; plus récemment, (Ladd & Morton, 1997) ont élaboré une expérience en perception pour savoir s'il existait un ton extra haut dans cette langue.

En français, le nombre de niveaux mélodiques phonologiquement distincts est encore loin d'être clair ; cela se reflète dans les nombreuses théories existantes. S'il existe un consensus sur une opposition minimale entre un contour descendant et un contour

montant, les approches diffèrent sur le nombre de contours montants distinctifs, et particulier sur la hauteur atteinte par le contour terminal.

Les travaux d'inspiration MA (Di Cristo, 1998 ; Jun & Fougeron, 2000) n'utilisent effectivement que deux tons ; la grammaire tonale de (Post, 2000), cependant, génère trois hauteurs distinctives en fin d'IP, L% ~ Ø% ~ H%. Au *Conclusif Majeur* atteignant le registre *infra-grave*, les approches de (Delattre, 1966) ou de (Rossi, 1999) opposent le *Continuatif Majeur*, réalisé dans *l'aigu*, et le *Contour Interrogatif*, avec un glissement vers le *sur-aigu*. Notons que Delattre (*ibid.*) propose également le trait « + », lequel est interprétable comme un niveau supplémentaire. Quant à (Mertens, 1990), il suppose quatre niveaux tonals en fin de groupe intonatif (B-, B, H, H+).

Un des moyens pour déterminer le nombre de catégories intonatives est d'aborder le problème du point de vue expérimental et plus précisément sous l'angle de la perception, associée ou non à la production : il s'agit alors de savoir si les variations de hauteurs sont perçues (et réalisées) de façon catégorielle ou continue.

Dans le cadre MA, on suppose que les variations de hauteurs des deux catégories H et B sont des variations continues du « pitch range » (empan mélodique), paralinguistiques et implémentées au niveau phonétique. Cela a été notamment étayé par (Liberman & Pierrehumbert, 1984), lors d'une expérience dans laquelle 10 sujets ont pu produire 10 registres globaux de f0 distincts sans les regrouper en catégories plus petites.

Afin de savoir si des sujets percevaient de façon catégorielle la différence entre un ton haut et un ton extra-haut marquant l'emphase, (Ladd & Morton, 1997) ont mis en œuvre une série de tests d'identification et de discrimination. Face à des résultats ambivalents, les auteurs concluent que la distinction entre les pics accentuels « normal » et « emphatique » en anglais peut être considérée comme catégorielle en *production* mais pas en *perception*. (Vanrell Bosch, 2006) applique des tests similaires et montre toutefois qu'en catalan, les variations de hauteur accentuelles sont perçues de façon catégorielle et permettent notamment de distinguer les questions polaires et les questions partielles.

Pour étudier le caractère catégoriel ou continu de l'alignement des tons du contour « rise-fall-rise » en anglais, (Pierrehumbert & Steele, 1989) ont construit une expérience où des sujets devaient imiter des stimuli de resynthèse (tâche d'imitation). Leurs résultats montrent clairement une distinction binaire entre deux catégories d'alignement tonal. Dans deux autres travaux utilisant aussi la tâche d'imitation, (Dilley & Brown, 2007) et (Redi, 2003) ont mis en évidence une répartition catégorielle des variations continues de pitch range et d'alignement tonal des stimuli présentés aux sujets.

Cependant, dans une autre étude utilisant également une tâche d'imitation, (Dilley, 2007) a obtenu des résultats inverses, lesquels montrent que les locuteurs étaient capables de reproduire les variations continues de pitch range des stimuli plutôt que de réaliser les catégories prévues par la théorie MA. Néanmoins, pour cette auteure, ce n'est pas parce que les locuteurs arrivent à imiter le continuum que cela remet en cause l'existence de catégories ; elle suggère que les stimuli correspondent à des variations graduelles à l'intérieur d'une unique catégorie linguistique.

En ce qui concerne le français, une contribution importante a été faite par (Post, 2000), laquelle a employé des tests d'identification et de discrimination pour étudier la

perception de différents contours terminaux en français. Les résultats de l'expérience concernant la distinction entre H*H% et H*Ø sont ambigus : si la courbe de la tâche d'identification a bien une forme en « S », typique d'une distinction catégorielle, le pic de discrimination est, quant à lui, clairement décalé par rapport au centre de la courbe ; aussi, les résultats globaux de (Post, 2000) ne livrent-ils pas de preuve solide de l'existence des deux catégories H% et Ø% prédites par sa grammaire tonale.

Dans cet article, nous étudions les variations de hauteur mélodique à la fin du groupe intonatif terminal en français en utilisant la démarche expérimentale de la tâche d'imitation. Notre but est double : il s'agit de tester tout d'abord si les imitations des variations de f₀ sont continues ou catégorielles ; ensuite, si ces dernières s'avèrent catégorielles, l'objectif est de définir le nombre de niveaux mélodiques distinctifs.

2 Protocole expérimental

2.1 La tâche d'imitation

Les résultats de (Post, 2000) et de (Ladd & Morton, 1997) amènent ces auteurs à considérer que les tâches d'identification et de discrimination, ne sont peut-être pas les tâches les plus appropriées pour mettre en évidence le caractère catégoriel ou continu de la perception des catégories intonatives. (Pierrehumbert & Steele, 1989) discutent aussi ce problème et signalent que les tests d'identification forcent les sujets à répondre selon des catégories prédéfinies ; de telles tâches servent avant tout à déterminer où se situe la *frontière* entre deux catégories, mais pas de mettre au jour l'existence de catégories ; les tâches d'identification doivent donc être réalisées lorsque l'existence des catégories phonologiques étudiées est indiscutable. Aussi, pour la présente étude, nous avons choisi la méthode de la tâche d'imitation, utilisée par (Dilley, 2006; Dilley & Brown, 2007; Pierrehumbert & Steele, 1989; Redi, 2003), et qui semble être la tâche la plus appropriée pour tester l'existence de catégories intonatives présumées selon (Gussenhoven, 2006).

2.2 Stimuli et sujets

Nous avons enregistré un locuteur de 37 ans de la région parisienne qui a prononcé la phrase "Elle est là" plusieurs fois dans différents contextes afin d'obtenir des contours finaux avec les différents niveaux mélodiques possibles. Nous avons ensuite sélectionné la phrase dont les valeurs de durée et de f₀ de la syllabe pénultième et finale étaient les plus proches des valeurs moyennes pour l'ensemble des phrases enregistrées. La f₀ de cette phrase « naturelle » a ensuite été resynthétisée (algorithme PSOLA de PRAAT) en faisant varier systématiquement d'un demi-ton la hauteur mélodique finale, de l'infra-grave au suraigu. Une mélodie plate, proche des valeurs moyennes mesurées dans les énoncés naturels, a été appliquée sur les deux premières syllabes. Nous avons ainsi obtenu 26 stimuli, qui ont été ensuite mélangés aléatoirement par ensembles.

Sept locuteurs natifs du français (4 femmes et 3 hommes), étudiant ou travaillant dans une université française, ont passé l'expérience. L'un d'entre eux, MRG, est doctorant en phonétique et fait du théâtre ; les six autres locuteurs n'ont jamais passé d'expérience en linguistique ; aucun des participants n'a été informé du but de l'expérience.

2.3 Passations et mesures

Chaque sujet a passé l'expérience en deux sessions, constituées de sept ensembles (« blocs ») de stimuli chacune. Chaque session débutait par une session d'entraînement de neuf stimuli. Nous avons recueilli 2548 imitations en tout (26 stimuli *14 blocs = 364 stimuli/sujet) pour l'analyse. Nous avons utilisé le logiciel « CorpusRecorder » développé par le groupe « Parole » du LORIA à Nancy, qui a été modifié pour les besoins de l'expérience afin que les sujets n'aient aucune restriction de temps et qu'ils puissent réécouter les stimuli et réenregistrer leur production autant de fois qu'ils le souhaitent ; la consigne donnée aux sujets était de répéter aussi fidèlement que possible le stimulus qu'ils entendaient (cf. Pierrehumbert & Steele, 1989).

Les valeurs de f_0 en Hertz ont été calculées automatiquement toutes les 10 millisecondes et vérifiées manuellement. Dans cette étude, nous n'avons pris en compte que la valeur minimale ou maximale de la voyelle finale, qui sont considérées comme les valeurs pertinentes dans les théories actuelles de l'intonation. Ces valeurs ont été ensuite converties en demi-tons (« DT », sur la base de la valeur moyenne des valeurs minimales et maximales) afin de normaliser les voix masculines et féminines et les variations de hauteur chez un même locuteur. L'utilisation des demi-tons est également en adéquation avec les stimuli, obtenus en variant d'un demi-ton la hauteur finale.

3 Résultats de la tâche d'imitation

Les résultats de l'expérience sont présentés ci-dessous, des figures 1 à 3, sous la forme de graphiques représentant les valeurs médianes de f_0 en demi-tons pour chaque stimulus, ainsi que le premier et troisième quartiles, et d'histogrammes de la distribution des hauteurs mélodiques. Les résultats montrent qu'on peut regrouper les locuteurs en trois groupes que nous détaillons ci-dessous.

1. Les réalisations du locuteur MRG sont remarquablement proches des stimuli initiaux, (cf. Figure 1). Les résultats montrent que la reproduction du continuum est possible. Certains facteurs peuvent cependant expliquer cette performance. Contrairement aux autres locuteurs, MRG est un doctorant en phonétique et a une expérience d'acteur ; il a accompli en outre l'expérience en un temps double par rapport aux autres locuteurs (8 min 17 par bloc pour MRG, contre 4 min 50 pour les autres).

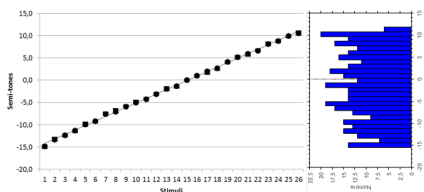


FIGURE 1 – Locuteur MRG : valeurs médianes de f_0 (en demi-tons) avec premier et troisième quartiles et histogramme de la distribution des hauteurs mélodiques.

2. A l'opposé, des résultats des locutrices FZJ et FDA (Figure 2) suggèrent l'existence d'au moins trois catégories distinctes. L'histogramme de FZJ (Figure 2a)

présente trois modes distincts autour de -8, 0 et 3 DT ; les valeurs médianes montrent trois voire quatre groupes de points, formant des « paliers » : le premier regroupe les stimuli 1 à 5, le second les stimuli 9 et 15, le troisième les stimuli 16 à 19 et un quatrième regrouperait les stimuli 22 à 26. Entre ces groupes, on constate des zones de réponses transitoires avec des quartiles indiquant notamment une variation plus importante. Les données de FDA (figure 2b) corroborent l'existence de catégories distinctives, sous la forme de trois paliers, regroupant les réponses aux stimuli 1 à 5, 7 à 16 et 23 à 26. L'histogramme de FDA présente trois modes autour de -15, -3 et 8 DT, le mode préminent étant celui autour de -3 DT, ce qui correspond au large plateau du deuxième groupe de points. Pour ces deux locutrices, les stimuli 6 à 8, caractérisés par d'avantage de variation, peuvent être considérés comme transitoires, et marqueraient une frontière entre une catégorie de hauteur grave et les catégories plus aiguës.

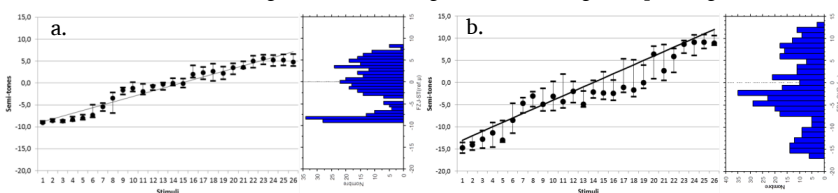


FIGURE 2 – Locutrices a. FZJ et b. FDA.

3. Les résultats pour les quatre autres locuteurs (FBG, FLM, MBB et MSJ, Figure 3) sont moins faciles à interpréter. Les valeurs médianes ressemblent à celles de MRG, sans formation claire de paliers comme pour FZJ et FDA ; elles suggèrent que ces locuteurs arrivent à imiter de près les stimuli, sans beaucoup de variation. Cependant, leurs histogrammes correspondent davantage à ceux de FZJ et FDA ; ils présentent des modes distincts, ce qui suggère au contraire l'existence de catégories de hauteurs mélodiques.

Par ailleurs, un examen plus fin des graphiques montre que des « décrochages » viennent perturber la progression linéaire des valeurs médianes. Ces décrochages consistent soit en des sauts entre deux valeurs médianes immédiates de 2 DT au minimum et sans chevauchement entre le troisième quartile de la première valeur et le premier quartile de la deuxième valeur, soit en une ou deux valeurs médianes transitoires marquées par d'avantage de variation. Ces décrochages plaideraient pour des catégories mélodiques, et ce, d'autant plus que là où ils apparaissent, on constate un « creux » dans les histogrammes. Nous proposons donc d'interpréter ces décrochages, associés aux « creux » des histogrammes, comme l'indice de frontières de catégorie mélodique. Ainsi, pour les quatre locuteurs, tout comme pour FZJ et FDA, on constate un premier mode autour de -10DT. Ce mode est suivi d'un creux aux alentours de -6/8 DT, correspondant à un décrochage dans les valeurs médianes. Ce décrochage est particulièrement clair chez MBB et FBG (Figure 3 a/c) et s'exprime par des valeurs de transitions chez MSJ et FLM (Figure 3 b/d). Ce mode et ces valeurs médianes permettent de rendre compte de la catégorie tonale « bas », associée à l'assertion. Pour les valeurs à partir de -6DT, le nombre de modes et décrochages dépendent des locuteurs. Ainsi, l'histogramme de FBG (Figure 3a) suggère une distribution bimodale dans ces valeurs plus aiguës avec des pics autour de -10DT et 4DT. Pour MBB, l'histogramme suggère également deux pics, autour de -10DT et 7DT ; par contre, pour

ce locuteur, l'association des décrochages (entre les réponses aux stimuli 7/8 et 13/14) et des creux des histogrammes (entre -8DT et 0DT) nous amène à proposer une troisième catégorie dont les frontières sont situées autour des stimuli 7/8 et 13/14. L'histogramme de FLM (Figure 3b) révèle quatre modes autour de -11DT, -4DT, 2DT et 7DT associé à des décrochages de valeurs médianes (stimuli 6/7, 13 et 19). Enfin, pour le locuteur MSJ (Figure 3d), les graphiques sont difficiles à interpréter sauf pour les valeurs basses.

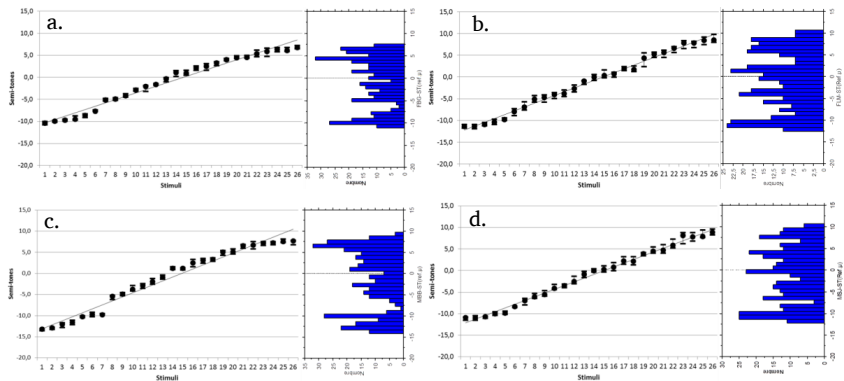


FIGURE 3 – Locuteurs a. FBG, b. FLM, c. MBB et d. MSJ

4 Discussion

L'expérience que nous avons menée auprès de sept locuteurs a produit des résultats ambivalents. D'un côté, les « paliers » formés par les valeurs médianes des imitations de FZJ et FDA, associés aux modes des histogrammes ne peuvent s'expliquer que si l'on postule l'existence de catégories de hauteur ; (Pierrehumbert & Steele, 1989), qui ont utilisé le même paradigme expérimental, arrivent à la conclusion que l'alignement tonal est catégoriel en anglais à partir de graphiques et d'histogrammes, dont les « formes » sont très proches des nôtres. D'un autre côté, cinq locuteurs sur sept ont pu reproduire un continuum relativement proche de celui des stimuli ; trois interprétations peuvent être avancées pour expliquer ce continuum.

1. les variations de hauteurs mélodiques sont graduelles ; il n'existe pas de catégories mélodiques distinctes, au moins pour certains locuteurs. Cela va dans le sens de la théorie MA « standard » et des résultats expérimentaux de (Lieberman & Pierrehumbert, 1984) et (Dilley, 2007) entre autres, selon lesquels ces variations résultent d'une modification graduelle du « pitch range », implémentée au niveau phonétique. Cependant, puisque les stimuli s'étendent de l'infra-grave au suraigu, cette explication remettrait en cause l'existence même des deux catégories de base « ton bas » ~ « ton haut », catégories indépendamment bien établies. Cette explication ne nous semble donc pas satisfaisante : la reproduction d'un certain continuum ne prouve pas en soi l'absence de catégories. De plus, les histogrammes de trois de ces cinq locuteurs

présentent clairement trois à quatre modes et nous avons remarqué que les « décrochages » des valeurs médianes, associés à des minima dans les histogrammes pouvaient raisonnablement être interprétés comme des indices de frontières catégorielles ; cela est d'autant plus plausible qu'on observe dans les graphiques et histogrammes des locuteurs FBG, MBB et FLM une distinction nette entre les hauteurs graves et les hauteurs plus aiguës, distinction qu'on retrouve chez les locutrices FZJ et FDA, dont les productions sont indéniablement catégorielles. L'observation d'autres décrochages dans les valeurs médianes ainsi que la présence de modes dans les histogrammes des locuteurs FBG, MBB et FLM plaident en faveur de l'existence de trois, voire quatre catégories de hauteurs mélodiques.

2. On peut dès lors proposer que des catégories tonales existent effectivement mais que celles-ci constituent chacune une dimension graduelle avec peut-être une valeur préférée ou prototypique ; une idée similaire a été avancée par (Gussenhoven 2006).

3. Enfin, une troisième interprétation plus plausible est que les productions graduelles s'expliquent par la tâche même d'imitation : les locuteurs qui reproduisent le continuum de stimuli ont une compétence dans l'imitation de détails phonétiques fins d'autres locuteurs¹. Cependant, comme nous l'avons argumenté plus haut, être capable de reproduire un continuum ne veut pas dire que les catégories n'existent pas, d'autant plus que nos données montrent que la performance exceptionnelle réalisée par MRG est liée au temps dévoué à l'expérience : MRG a passé en moyenne 8 min 17 par bloc, la moyenne par bloc étant de 4 min 50 pour les autres.

Les données de FZJ et FDA corroborent de manière indirecte l'idée que les réalisations graduelles des autres locuteurs s'expliqueraient par la tâche d'imitation elle-même. Ce que nos résultats suggèrent c'est que FZJ et FDA ne sont pas des « imitatrices » aussi performantes que les autres locuteurs. Autrement dit, FZJ et FDA ramènent spontanément leurs productions à leur système de catégories tonales, et le fait qu'elles se caractérisent par plus de variation montre que l'imitation des hauteurs mélodiques d'une autre voix est une tâche difficile et que leurs propres catégories prévalent sur les hauteurs cibles à imiter.

5 Conclusion

Si les valeurs médianes de certains locuteurs laissent supposer que les hauteurs mélodiques varient de manière graduelle et non catégorielle, nous avons avancé des arguments en faveur de l'existence d'au moins trois catégories de hauteur mélodique. Notre expérience a suggéré que la tâche d'imitation elle-même était probablement à l'origine de la reproduction du continuum par certains locuteurs, certainement plus à mêmes que d'autres d'imiter d'autres voix. Au niveau de la théorie phonologique, si l'on admet l'existence d'au moins trois catégories de hauteur, alors le modèle MA « standard » avec ses deux tons « Haut » et « Bas » se révèle insuffisant, et il serait nécessaire de reconsidérer les modèles phonologiques à plusieurs niveaux.

¹ Dans une perspective des changements sonores (*sound change*) (Ohala 1981) avait avancé l'idée selon laquelle la perception est bimodale, et que l'auditeur a la capacité de se concentrer plus sur les caractéristiques du signal acoustique (comment le message est transmis) que le contenu linguistique du message même.

Références

- DELATTRE, P. (1966). Les dix intonations de base du français. *The French Review* 40, pages 1-14.
- DI CRISTO, A. (1998). Intonation in French. In DI CRISTO, A. & HIRST, D. (éditeurs), *Intonation Systems: a Survey of Twenty Languages*. CUP, pages 195-218.
- DILLEY, L. et BROWN, M. (2007). Effects of pitch range variation on F0 extrema in an imitation task. In *Journal of Phonetics*, 35, pages 523-551.
- DILLEY, L. (2007). Pitch range variation in English tonal contrasts: Continuous or categorical? In *Proceedings of the XVIth ICPhS*, Saarbrücken, Allemagne, pages 1153-1157.
- GUSSENHOVEN, C. (2006). Experimental approaches to establishing discreteness of intonational contrasts. In S. SUDHOFF & al. (éditeurs), *Methods in Empirical Prosody Research*, Mouton de Gruyter, pages 321-334.
- JUN, S.-A. et FOUGERON, C. (2000). A Phonological model of French intonation. In A. BOTINIS (ed.) *Intonation: Analysis, Modeling and Technology*. Kluwer, pages 209-242.
- LADD, D.R. et MORTON, R. (1997). The perception of intonational emphasis: Continuous or categorical? *Journal of Phonetics* 25, pages 313-342.
- LADD D. R. (2008) *Intonational Phonology*, second edition. Cambridge University Press.
- LIBERMAN, M. et PIERREHUMBERT, J. (1984). Intonational invariance under changes in pitch range and length. In: M. Aronoffand R.T. Oehrle (eds.): *Language sound structure: Studies in phonology presented to Morris Halle*. Cambridge, MIT Press, pages 157-233.
- MERTENS, P. (1990). L'intonation. In BLANCHE-BENVENISTE, C. et al (éditeurs), *Le français parlé*, Paris: Éditions du CNRS, pages 159-176.
- OHALA, J. J. (1981). The listener as a source of sound change. dans C. S. Masek, R. A. Hendrick, & M. F. Miller (eds.), *Papers from the Parasession on Language and Behavior*. Chicago: Chicago Ling. Soc. pages 178-203.
- PIERREHUMBERT, J. et BECKMAN, M. (1988). *Japanese Tone Structure*. Linguistic Inquiry Monograph 15, MIT Press, Cambridge.
- PIERREHUMBERT, J. et STEELE S.A., (1989). Categories of tonal alignment in English. *Phonetica* 46, pages 181-196.
- POST, B. (2000). *Tonal and phrasal structures in French intonation*. Holland Ac. Graphics.
- REDI, L. (2003). Categorical effects in the production of pitch contours in English. In *Proceedings of the 15th ICPhS*, Barcelona, pages 2921-2924.
- ROSSI, M. (1999). *L'intonation, le système du français : description et modélisation*. Ophrys, Paris.
- VANRELL BOSCH, M.M. (2006). A scaling contrast in Majorcan Catalan interrogatives. In *Proceedings of Speech Prosody 2006*, Dresdes, Allemagne.