

Approches prosodiques

Du signal à la modélisation linguistique



Volume en l'honneur de Philippe Martin

Édité par

Elisabeth Delais-Roussarie, CNRS – Université Paris-Diderot

& Hiyon Yoo, Université Paris-Diderot

<http://sites.google.com/site/phmartin2012/>

PRÉFACE

Les 28 et 29 juin 2012 ont eu lieu à Paris des Journées en l'honneur de Philippe Martin. Cette manifestation avait pour but de regrouper autour de notre collègue et ami – qui achevait sa dernière année universitaire avant de partir à la retraite – d'anciens collègues, des amis ou des étudiants. Durant cette manifestation, des communications scientifiques portant sur des thématiques chères à Philippe Martin ont été présentées. Ces journées étaient également un moyen de retracer certaines étapes importantes de sa carrière, de lui témoigner notre sympathie.

Ce volume se veut à l'image de ces journées. Il se composera de deux parties distinctes. Dans une première partie, le parcours académique et les travaux de Philippe Martin sont rappelés. Cela se fera par la publication de lettre ou discours rédigés par d'anciens collègues. La seconde partie du volume regroupe neuf communications scientifiques :

- Trois d'entre elles proposent des analyses prosodiques de parlars ou de phono-styles particuliers. La communication d'Antoine AUCHLIN, Jean-Philippe GOLDMAN et Anne-Catherine SIMON propose une description prosodique du parler de Philippe Martin. Katarina BARTKOVA, Denis JOUVET et Natalia SEGAL présentent un algorithme permettant de segmenter automatiquement en groupes prosodiques la parole journalistique, en s'appuyant sur la théorie prosodique de Philippe Martin. Quant à Fernand CARTON, il propose une analyse des caractéristiques prosodiques du parler ouvrier de la région de Tourcoing à la fin du XIX^{ème} siècle.
- Trois communications prennent la forme d'analyses théoriques qui s'inspirent des travaux de Philippe Martin. La communication d'Elisabeth DELAIS-ROUSSARIE et Brechtje POST tente de montrer que le contraste de pente, avant d'être une forme particulière du principe du contour obligatoire (OCP), relève davantage d'une forme d'opacité phonologique, dans la mesure où plusieurs phénomènes sont regroupés sous ce mécanisme du contraste de pente. Tomas DUBEDA propose dans sa communication d'aborder la question de l'intonotactique et de montrer son importance pour rendre compte des profils mélodiques observés en français et en tchèque. Pour finir, David LE GAC essaie de montrer que le contraste de pente et le principe du contour obligatoire (OCP) font souvent de pair pour contraindre l'apparition de certaines formes mélodiques.

- Une communication de Jean-Sylvain LIENARD revient sur la fonction Peigne, développé par Philippe Martin. A partir d'une présentation historique, il montre les différentes évolutions et les apports de la fonction peigne dans la détection de la fréquence fondamentale.
- La communication de Massimo PETTORINO, Anna DE MEO et Marilisa VITALE présente les résultats d'un test de perception visant à évaluer dans quelle mesure des modifications prosodiques, effectuées à l'aide de Winpitch, nuisent à l'intelligibilité d'un message en italien. L'impact des modifications apportées sur l'intelligibilité est évalué aussi bien auprès d'une population native qu'avec des locuteurs dont l'italien est une L2.
- La communication de Willy SERNICLAES porte sur la perception de la parole. Il tente d'esquisser un parallèle entre les modèles de perception des traits segmentaux, basés essentiellement sur des traits hiérarchiquement organisés, et le modèle proposé par Philippe Martin pour le traitement de la prosodie.

SOMMAIRE

PRÉFACE	3
SOMMAIRE	5
PREMIÈRE PARTIE : Parcours et bibliographie de Philippe Martin, Philippe et ses collègues	7
Avant-propos	9
Parcours académique et productions scientifiques de Philippe Martin	11
Résumé synoptique de la carrière académique de Philippe Martin	13
Publications scientifiques de Philippe Martin	17
Philippe MARTIN vu par des collègues	29
Philippe Intime, <i>Pierre Léon</i>	31
Philippe Martin, son œuvre et sa carrière en quelques mots, <i>Elisabeth Delais-Roussarie</i>	35
SECONDE PARTIE : Communications scientifiques	37
Description prosodique de styles ou de parlers particuliers	39
Philippe Martin, tel qu'en lui-même, enfin la prosodie le révèle, <i>Antoine Auchlin, Jean-Philippe Goldman et Anne-Catherine Simon</i>	41
Organisation prosodique de la parole dans le style radiophonique, <i>Katarina Bartkova, Denis Jouvét et Natalia Segal</i>	51
Etude prosodique du parler ouvrier de Tourcoing (Nord) à la fin du XIXème siècle, <i>Fernand Carton</i>	59
Structure Prosodique et Théorie de l'intonation	77
Le contraste pente peut-il s'analyser comme une manifestation d'OCP, <i>Elisabeth Delais-Roussarie et Brechtje Post</i>	79
Pour une « intonotactique » des événements mélodiques, <i>Tomáš Duběda</i>	97
Contraste de pente et principe du contour obligatoire : « l'union fait la force », <i>David Le Gac</i>	113
Phonétique Appliquée et Traitement du signal	131
Variations sur un peigne, <i>Jean-Sylvain Liénard</i>	133
The effect of rhythm and pitch contour manipulation on intelligibility in L2 Italian informative speech, <i>Massimo Pettorino, Anna De Meo et Marilisa Vitale</i>	145
La perception des traits phonologiques: des logatomes à la parole spontanée, <i>Willy Serniclaes</i>	151

PREMIÈRE PARTIE

PARCOURS ET BIBLIOGRAPHIE DE
PHILIPPE MARTIN

PHILIPPE ET SES COLLÈGUES

AVANT-PROPOS

A la fin d'une carrière académique, on peut toujours regarder en arrière, mais les seules choses dont on est sûr, c'est :

- Qu'on ne s'imaginait sans doute pas ce qu'on est devenu...
- Que la vie est faite de hauts et de bas...
- Que le parcours n'est pas linéaire...



En somme, la vie est très comparable à la parole, à sa mélodie, et à la façon dont tu les as appréhendées et modélisées. De fait, on peut retenir de tes travaux deux choses essentielles :

- Les hauts et les bas sont essentiels à la structuration de la mélodie ;
- Le signal de parole est fait de nœuds : il faut brosser, peigner, et parfois en utilisant différentes méthodes....

Les travaux sur l'intonation et sur la détection de pitch de Philippe MARTIN sont des apports considérables à la recherche en linguistique et en traitement du signal. Mais ce recueil ne veut pas se limiter à cela... Philippe est aussi un chercheur avec qui les uns et les autres ont apprécié travailler, discuter dans des colloques, rire autour d'un verre.

PARCOURS ACADÉMIQUE ET
PRODUCTIONS SCIENTIFIQUES

RÉSUMÉ SYNOPTIQUE DE LA CARRIÈRE ACADÉMIQUE DE PHILIPPE MARTIN

Formation

Doctorat ès Sciences (Université Libre de Bruxelles), 1973

Doctorat en Linguistique Appliquée (Université Nancy 2), 1972

Sciences Appliquées (spécialité acoustique), École Polytechnique (Université Libre de Bruxelles), 1967

Parcours professionnel

Professeur (CE), Université Paris 7 Denis Diderot (2009-2012)

Professeur (PR1), Université Paris 7 Denis Diderot (2002-2009)

Conseiller scientifique, Centre National d'Étude des Télécommunications CNET Lannion- France (2007-08)

Professeur titulaire (PR1), Department of French, University of Toronto, Canada (1986-2002)

Directeur, Experimental Phonetics Laboratory, Department of French, University of Toronto, Canada (1991-96)

Conseiller scientifique, Centre National d'Étude des Télécommunications CNET Lannion- France (1987-89)

Professeur (PR2), Institut de Phonétique, Université de Provence, France (1978-86)

Professeur agrégé (PR2), Department of French, University of Toronto, Canada (1975-78)

Professeur assistant (MC), Department of French, University of Toronto, Canada (1972-75)

Chargé de cours (MC), Department of French, University of Toronto, Canada (1970-72)

Chargé de recherches, Institut de Phonétique, Université Libre de Bruxelles, Belgique (1974-1981)

Collaborateur scientifique, Section de préhistoire, Musée Royal de l'Afrique Centrale, Belgique (1971-1979)

Ingénieur de recherche, Experimental Phonetics Laboratory, University of Toronto, Canada (1967-1981)

Ayant bénéficié d'une double formation (scientifique, Doctorat en Acoustique et sciences humaines, Doctorat en Linguistique), Philippe Martin a mené une carrière scientifique et d'enseignement reflétant ces deux orientations. D'abord chercheur en phonétique expérimentale au

Canada (Université de Toronto) et en Belgique (Institut de Phonétique, ULB), il est devenu enseignant en phonétique et linguistique française à l'Université de Toronto. En 1978, il a été nommé professeur à l'Université de Provence (Institut de Phonétique), puis il a été détaché à l'Université de Toronto de 1986 à 2002 (Département d'Études Françaises). En 2002, il a été nommé professeur à l'UFR de Linguistique de l'Université Paris Diderot, responsable de l'EA de phonétique expérimentale et directeur de cette UFR en 2004. Il a publié 7 livres (seul ou en collaboration), plus de 200 articles (dans les domaines de la phonétique acoustique, du traitement du signal, et surtout en phonologie de l'intonation dans ses rapports avec la syntaxe). Il a participé à près de 180 colloques et congrès, et a été invité à donner plus de 170 conférences dans 15 pays. Au cours de sa carrière, Philippe Martin a également conçu plusieurs appareillages d'analyse acoustique de la voix (dont certains ont été commercialisés), ainsi que des logiciels de conception de circuits imprimés (placement et routage automatique), de télémétrie pour la recherche archéologique, de conception assistée d'échantillons de tissus, de gestion d'horaires de travail en entreprises, de placement automatique de l'accent lexical en italien, et d'analyse acoustique de la voix (WinPitch).

Thèmes de recherche

- Etudes expérimentales et théoriques des rapports prosodie-syntaxe en français, anglais, italien, portugais, espagnol, roumain, catalan, mandarin ;
- Traitement du signal de parole (en particulier mesure et visualisation de la fréquence fondamentale)
- Méthodes contrastives d'enseignement de l'oral (prosodie anglais, français, portugais, mandarin) ;
- Développement de logiciels pour la transcription et l'alignement de grands corpus oraux.

Activités administratives

Directeur, UFR Linguistique, Université Paris Diderot (2004-2012)

Responsable EA333 (Atelier de Recherche sur la Parole), UFRL, Université Paris Diderot (2004-2009)

Responsable du programme de réforme du nouveau Magistère de l'UFRL Paris 7 (2004).

Membre du conseil d'administration de l'UFRL Paris Diderot.

Membre du conseil scientifique de l'UFRL Paris Diderot.

Directeur, Experimental Phonetics Laboratory, Department of French, University of Toronto, Canada (1991-96)

Autres activités

Développement d'un programme d'analyse temps réel de la fréquence fondamentale et d'édition de paramètres prosodiques pour la recherche phonétique et l'enseignement ("WinPitch") (1996).

Conception et réalisation d'une carte PC d'analyse de la fréquence fondamentale pour l'enseignement (Projet DOD, Gouv. du Canada), ("PM1000") (1988).

Conception et réalisation d'un spectrographe temps réel couleur pour l'exposition "Les Immatériaux", Centre Pompidou, Paris, et le Musée des Sciences de la Villette, Paris (1985).

Conception et réalisation d'un analyseur de mélodie temps réel multifonctions, commercialisé par "Voice Identification Inc.", USA, et destiné à l'enseignement (PM100), la recherche phonétique (PM200) et aux mesures en phoniatry (PM300) (1981).

Conception et réalisation d'un analyseur de mélodie temps réel, commercialisé par F-J Electronics, Danemark ("Pitch Computer") (1977).

Directions de thèse

- Hélène Canto (1991)
- Isabelle Guaitela (1991)
- Alexandre Sévigny (2000)
- Ivan Chow (2001)
- Iraj Bazgari (2006)
- Catherine Mathon (2007)
- Virginia Gripon (2009)
- Géraldine Vercherand (2010)
- Li Jun (2011)
- Sylva Novakova (2011)
- Natalia Segal (2011)
- Inyoung Kim (2012)

PUBLICATIONS SCIENTIFIQUES

Ouvrages ou monographies

Intonation du français, mesures, théories, modèles. Armand Colin, Paris, 2009, 256 p.

Phonétique Acoustique. Armand Colin, Paris, 2008, 160 p.

Classification formelle automatique et industries lithiques (D. Cahen co-auteur), Annales du Musée Royal de l'Afrique Centrale - Série Sciences humaines, n. 76, 1972, 112 p.

Prolégomènes à l'étude des structures intonatives (P.R. Léon co-auteur). Didier, Montréal, 1969, 226 p.

Édition d'ouvrages

Promenades Phonétiques (co-rédacteur avec H. Gezundhajt). Éd. Mélodie-Toronto, 1997, 152 p.

Accent, intonation et modèles phonologiques (B. Ferguson et H. Gezundhajt co-rédacteurs). Éd. Mélodie-Toronto, 1994, 202 p.

Mélanges Léon (Rédacteur). Éd. Mélodie-Toronto, 1992, 587 p.

Toronto English - Studies in Phonetics, (P.R. Léon et al. co-auteurs). Didier, Montréal, 1979, 168

Édition de revue

Information-Communication, EPL, University of Toronto, Volume 11, June 1990.

Information-Communication, EPL, University of Toronto, Volume 9, Aug 1988.

Articles dans des revues à comité de lecture

Ponctuation et structure prosodique. *Langue Française* 172 : 99-114, 2011.

La prosodie du français: Une approche pas très syntaxique. *Journal of French Language Studies* 21/1 : 9-52, 2011.

Lexique et mélodie: un rendez-vous manqué ? *Cahiers de lexicologie* 96 : 151-165, 2010.

Structure prosodique, structure de contrastes. *Revue TRANEL* 47 : 103-116, déc. 2007.

Les formants vocaliques et le barrissement de l'éléphant. *Histoire des Théories Linguistiques X*: 9-27, 2007.

WinPitch LTL, a Multimodal Pronunciation Software. *Revue de l'ALSIC* 8/2, mai 2005.

L'intonation de la phrase dans les langues romanes : l'exception du français. *Langue française* (mars 2004): 36-55.

L'intonation en parole spontanée. *Revue Française de Linguistique Appliquée* IV-2: 57-76, 2000.

Intégration des technologies de la parole pour l'apprentissage des langues. *Revue de l'ALSIC*, déc. 2000 (en coll. avec Mercier, G. et Guyomard, M.).

L'intonation du français et du portugais: phonétique et phonologie. *Revista de l'ANPOLL* 5/7, 1999.

Prosodie des langues romanes: Analyse phonétique et phonologie. *Recherches sur le français parlé* 15: 233-253, 1999.

Essai d'étude différentielle intonative. *IRAL XXVIII* 2: 135-151, 1990, (en coll. Avec Daniel Lepeit).

Prosodic and Rhythmic Structures in French. *Linguistics* 25-5: 925-949, 1987.

Spectral Comb Gives Real-Time Pitch Analysis. *Speech Technology* 1983: 96-98.

Phonetic Realizations of Prosodic Contours in French, *Speech Communication* 1: 283-294, 1982.

Mesure tridimensionnelle par ultrasons pour la recherche archéologique. *Notae Prehistoricae* 2, 1982.

Questions de phonosyntaxe et de phonosémantique en français. *Linguisticae Investigationes* II: 93-126, 1978.

L'analyse des indices et la détection automatique des traits (en coll. avec J-P. Haton), *Revue d'Acoustique* 97: 209-211, 1978.

Analyse phonologique de la phrase française. *Linguistics* 146: 35-68, 1975.

A propos de l'accentuation des pronoms personnels en français. *Le français moderne* (4): 348-350, 1975.

Les problèmes de l'intonation: recherches et méthodes. *Langue française* 19: 4-42, 1973.

Applied Linguistic and the Teaching of Intonation. *Modern Language Journal* 56 (3): 139-144, 1972. (en coll. avec P.R. Léon).

Classification formelle automatique et industries lithiques. *Bulletin de la Société Royale Belge d'Anthropologie et de Préhistoire* (83): 19-21, 1972. (en coll. avec D. Cahen).

Linguistique appliquée et enseignement de l'intonation. *Etudes de linguistique appliquée* 1971 (3): 36-45. (en coll. avec P.R. Léon).

Articles dans des Actes de colloques à comité de lecture

La structure prosodique : Ratures et ajout dans l'oral spontané. *Actes des journées d'Hommage à Claire Blanche-Benveniste*, Paris ENS, 2-4 décembre 2010. (à paraître).

La structure prosodique en français et ses réalisations variées. *Journée du français: Colloque « Comme on nous parle : culture et média francophones »*, 10 décembre 2010, University of Oxford, in *Les Voix des Français*, Abecassis and Ledegen ed. Peter Lang (à paraître)

Les parenthèses: analyse macrosyntaxique et prosodique sur corpus. *Actes du colloque La Parataxe*, Université de Neuchâtel, Suisse/ (avec J-M. Debaisieux), (sous presse).

The role of Fo in Mongolian stress. *Proc. Speech Prosody* 2012, (avec Sang Yumei).

The Autosegmental-Metrical Prosodic Structure: not fit for French? *Proc. Speech Prosody* 2012.

Multi methods pitch tracking. *Proc. Speech Prosody* 2012.

Automatic detection of voice creak. *Proc. Speech Prosody* 2012.

Prosodic similarities in French spoken in the Mascareignes. *Proc. Speech Prosody* 2012.

Traits nécessaires et suffisant pour l'indication de la structure prosodique. *Actes du Colloque IDP09*, Paris 9-11 septembre 2009, 2011.

Prosodic characteristics of read and spontaneous speech in French. *Proc. of the The 17th International Congress of Phonetic Sciences (ICPhS XVII)*, Hong Kong, August 17-21, 2011.

Souveraineté-Association en linguistique: L'exemple de l'intonation et de la (macro)syntaxe. *Actes du colloque « À l'école de l'oral »*, Univ degli Studi di Torino, 6 mai 2011.

Segmentation prosodique et production orale. *Actes du colloque (Dés)organisation de l'oral*, Rennes 24-25 mars 2011.

Time selected multiple algorithms for reliable Fo tracking in difficult recording conditions. *Proc. ExLing 2011*, Paris May 26-27, 2011.

WinPitch, a multipurpose multimodal tool for speech analysis of very large scale corpora. *Proc. New Tools and Methods for Very-Large-Scale Phonetics Research*, University of Pennsylvania, January 28-31, 2011

A dynamic view of the prosodic structure: The example of French. *Proc. ExLing 2010*, Athens, August 25-27, 2010.

Détection semi-automatique des syllabes proéminentes avec une segmentation automatique en pseudo-syllabes, *Actes des XXV JEP*, Mons (Belgique), 25-28 mai 2010, 185-188.

Suffixes complexes : quand c'est fini ça recommence. *Actes des XXV JEP*, Mons (Belgique), 25-28 mai 2010, 241-244.

Prosodic structure revisited: a cognitive approach. *Proc. Speech Prosody 2010*, Chicago, May 11-14, 2010.

Learning the prosodic structure of a foreign language with a pitch visualizer. *Proc. Speech Prosody 2010*, Chicago, Ill, May 11-14, 2010.

Prominence detection without syllabic segmentation, Prosodic Prominence: Perceptual and Automatic Identification. *Speech Prosody 2010 Satellite Workshop*, Chicago, Ill, May 11-14, 2010.

Spontaneous speech intonation in Italian, An experimental analysis with a macrosyntactic point of view. *Proceeding Terzo Convegno Internazionale sulla Comunicazione Parlata / 3rd International Conference on Spoken Communicatio*, to appear (2010).

Spontaneous speech intonation in Italian, Spontaneous speech to appear intonation in Italian, An experimental analysis with a macrosyntactic view. *Proc. GSCP 2009*.

Contours mélodiques de continuation majeure à La Réunion et à Maurice et aux Seychelles. *Actes du Colloque International sur La variation du français dans les espaces créolophones et francophones*, Saint Denis de la Réunion, 3-5 juin 2009.

Iconicity of melodic contours in French. *Actes Prosico 2008*.

Postfixes et suffixes interrogatifs : un cas d'ambiguïté prosodique ? *Actes de la conférence de la section tchéco-slovaque de l'ISPhS 2008*, 19 janvier 2008, 111-119.

Crosscorrelation of adjacent spectra enhances fundamental frequency estimation. *Proc. Interspeech*, Brisbane, 22 – 26 September 2008

Implementing an Intonation Model in a Speech Recognition System. *Proceedings Speech Prosody 2008*, Campinas, Brésil, 6-9 mai 2008 (with K. Bartkova et N. Segal).

A fundamental frequency estimator by crosscorrelation of adjacent spectra. *Proceedings of Speech Prosody 2008*, Campinas, Brésil, 6-9 mai 2008.

Contraintes rythmiques et syntaxiques dans la relation prosodie-syntaxe. *Actes du colloque CerLiCo « Grammaire et Prosodie »*, 1-2 juin 2007, 13-26.

Acoustic analysis of the neutral tone in Mandarin. *Proc. InterSpeech 2007*, Antwerp August 2007 (avec Li Jun).

L'intonème conclusif: une fin (de phrase) en soi ? *Actes IDP07*, Genève, 12-14 septembre 2007. (avec M. Avanzi).

Experimental Analysis of Prosody-Syntax Interaction in Spontaneous Speech. *Proc. Comunicazione parlata*, Napoli, 2006.

Prosody, Syntax, Macrosyntax. *Proc. ISCA Tutorial and research workshop on experimental linguistics*, Athens (Greece), 28-30 August 2006.

Phonologies and Phonetics of French Prosody. *Proc. Speech Prosody 2006*, Dresden.

Enseigner la prosodie avec le logiciel WinPitch LTL: Possibilités technologiques et pédagogiques *Actes du 2ème Colloque International Nouvelles Technologies et Éducation en Milieu Formel et Informel*, Casablanca, Maroc, 6-8 avril 2006. (avec I. Benali, R. Chiarelli, et H. Yoo).

Models of analysis and prosodic labeling systems. *Proc. AISV 2005*, Salerno, Italie.

Pour une syntaxe sans ellipse. *Actes du colloque sur les nouvelles syntaxes*, Université de St Étienne, St Étienne, France, 29 octobre 2005. (avec J-M. Debaisieux et H-J. Deulofeu).

Phonetic Variations of Prosodic Contours in French. *Actes du colloque "Phonological Variation: The Case of French"* (Bulletin PFC), University of Tromsø, Norway, August 24-27, 2005.

Identification des langues à partir des indices suprasegmentaux: cas du français et du grec. *Proc. MIDL Workshop*, Paris, 29-30 novembre 2004 (avec N. Dimou).

Tones and Intonation in Declarative and Interrogative Sentences in Mandarin. *The International Symposium on Tonal Aspect of Languages: Emphasis on Tone*

Languages, March 28-30, 2004 Beijing, China, (avec G. Boulakia et Zeng XiaoLi).

WinPitchPro, a Tool for Text to Speech Alignment and Prosodic Analysis. *Proc. Speech Prosody 2004*, Nara, Japon, 23-26 mars 2004.

WinPitchPro, a text to speech alignment and acoustic analysis for large corpora. *Proc. LREC*, Lisbonne, 26-28 mai 2004.

The C-ORAL-ROM CORPUS. A Multilingual Resource of Spontaneous Speech for Romance Languages. *Proc. LREC 2004*, Lisboa, Portugal, May 26-28, 2004 (avec Emanuela Cresti, Fernanda Bacelar do Nascimento, Antonio Moreno Sandoval, Jean Veronis, Philippe Martin, Khalid Choukri)

WinPitch LTL II, a multimodal pronunciation software NLP and Speech Technologies in Advanced language Learning Systems, *Proc. Of Instill/CALL2004*, Venice Italy, 17-19 June 2004, 177-182.

WinPitch Corpus, a Text to Speech Alignment Tool for Large Multimodal Corpora. *Workshop on Compiling and processing Spoken language Corpora*, Belém, Portugal, May 24th, 2004.

Un logiciel d'enseignement de la prosodie multimedia. *Actes des Journée ATALA "TAL et Apprentissage des Langues"*, Grenoble 22 octobre 2004, 71-82.

Élaboration d'un dictionnaire multimédia de la langue Parkatêjê, une langue Timbira de l'Amazonie Brésilienne », *Proc. 11th EURALEX*, Lorient, France, 6-10 juillet 2004 (avec L. Araújo).

Alignement de grands corpus, macrosyntaxe et intonation du français spontané. *Colloque DYALANG*, Université de Rouen, 29 avril 2003.

Accent de mot et intonation en Parkatêjê, une langue timbira de l'Amazonie Brésilienne. *Actes du colloque Interfaces Prosodiques*, Nantes 27-29 mars 2003, (avec L. Araújo).

Sans contraste, pas d'interface. *Actes du colloque Interfaces Prosodiques*, Nantes 27-29 mars 2003, (avec G. Boulakia et J. Deulofeu).

Teaching and Learning Second Language Oral production Skills in a Distance Education Setting. *Proc. of WorldCALL 2003*, Banff, Canada, 7-10 mai 2003 (avec A. Germain).

WinPitch Corpus, a Software Tool for Alignment and Analysis of Large Corpora. *Proc of the 3rd E-MELD Workshop*, LSA Institute, Lansing, July 11th-13th.

Regional Variations of Sentence Intonation in French: The Continuous Contour in Parisian French. *Proc. of Speech Prosody 2002*, Aix-en-Provence, April 11-14, 2002.

Ralentisseur du signal de parole par autocorrelation. *Actes des Journées d'Etudes sur la Parole 2002*, Nancy, 24-27 juin 2002.

Prosodic features finish off ill-formed utterances, don't they? *Proc. Congreso de Fonetica Experimental*, Universidad de Sevilla, España, 5-7 mars 2001, (avec J. Deulofeu et G. Boulakia).

ToBi : l'illusion scientifique ? *Actes des Journées "Prosodie" 2001*, Grenoble, 10-11 octobre 2001.

Quand la prosodie bien-forme des énoncés mal formés. *Actes des Journées "Prosodie" 2001*, Grenoble, 10-11 octobre 2001 (avec J. Deulofeu et G. Boulakia).

Peigne et brosse pour Fo : Mesure de la fréquence fondamentale par alignement de spectres séquentiels. *Actes des 23èmes JEP*, Aussois, France, juin 2000, 245-248.

WinPitch 2000: a tool for experimental phonology and intonation research. *Proceedings of the Prosody 2000 Workshop*, Kraków, Pologne, 2-5 October 2000.

Intonation of Spontaneous Speech in French. *Proc. of the ICPH99*, San Francisco, Aug 1999.

Intonation : a Case for Experimental Phonology. *Actes du Colloque de Royaumont*, ICP, Grenoble, 1999.

WinPitch Language Teaching and Learning: Ecouter, voir et manipuler la production orale pour l'apprentissage en langue seconde » (avec A. Germain), *Actes de JILA99* (Journées Internationales de Linguistique Appliquée), Nice 24-25 juin 1999. 110-113.

Association prosodie-syntaxe: validation par synthèse. *Actes des XXV Journées d'Etude sur la Parole*, Martigny, juin 1998, 224-227.

Accessibilité et intelligibilité des corpus linguistiques sur les réseaux: Analyse et perspectives. *Actes du XXIIe Colloque international de linguistique fonctionnelle: Les langues à l'aube du XXIème siècle*, (Claude Tatillon ed.), Paris. (with D. Sheffel-Dunand & Françoise Mougeon). Evora, Portugal, Mai 1998.

WinPitch: un logiciel d'analyse temps réel de la fréquence fondamentale fonctionnant sous Windows. *Actes des XXI Journées d'Etude sur la Parole*, Avignon, mai 1996, 224-227.

Prosodic Transcription Systems for French. *Prosodic Labeling Workshop, ICPHS 95 Congress*, Stockholm, August 1995.

The Prosody of Mauritian Creole: Some Experimental Aspects. *Proceedings of the XIIIrd International Congress of Phonetic Sciences*, Stockholm August 1995.

Congruent Strategies in Prosody Generation for Text-to-Speech Synthesis. *Proceedings of the 14th Int. Congress on Acoustics*, Sept. 1992, Beijing.

Positionnement automatique de l'accent lexical de l'Italien. *Actes des XVIIIèmes Journées d'Études sur la parole*, Montréal 1990, pp.149-152.

Automatic Generation of Prosody for Speech Synthesis in Italian. *Proceedings of the First International Conference on Speech Synthesis*, Autrans, France, Sept 1990, 149-152.

Automatic Assignment of Lexical Stress in Italian. *Proc. Eurospeech 89*, Paris, sept 1989, 222-225.

The Acquisition of Prosodic Patterns through Interactive Visual Display. *Proceedings of the 8th Annual Microcomputer in Education Conference*, Arizona State University, Phoenix, March 1988. (en coll. avec M. Adriaen), 1-11.

Structure rythmique et structure prosodique en français. *Actes des 14èmes JEP*, Hammamet, 1987.

Antonio mangia la zuppa inglese. *Proceedings of the XIII International Congress of Phonetic Sciences*, Tallin Aug 1987. (en collaboration avec O. Profili).

A Logarithmic Spectral Comb method for Fundamental Frequency Extraction. *Proceedings of the XIII International Congress of Phonetic Sciences*, Tallin, Aug 1987.

Structures prosodiques et structures rythmiques, *Actes des 13ème JEP*, Aix-en-Provence, 1986.

A Fast Spectral Comb Algorithm for Fo Detection. *Proceedings from the 12 International Congress of Acoustics*, Toronto 1986.

Semantics, Syntax and Intonation. *Proceedings of the Xth Int. Congress of Phonetic Sciences*, Foris, 1983.

Real Time Fundamental Frequency Analysis using the Spectral Comb Method. *Proceedings of the Xth Congress of Phonetic Sciences*, Foris, 1983.

Phonetics and Phonology: The Example of Intonation. *Proceedings of the XIIIth Int. Congress of Linguists*, Tokyo, 1982, 71-83.

Prosodic Structures in French. *Preprints of the Working Group on Intonation, XIIIth Int. Congress of Linguists*, Tokyo, 1982, 71-83.

Mesure de la fréquence fondamentale par intercorrélation avec une fonction peigne. *Actes des XIIèmes JEP*, Montréal, 1981.

Variations prosodiques inter et intralocuteurs. *Actes des XIèmes JEP*, Strasbourg, 1980.

Automatic Location of Stressed Syllables in French. *Proceedings of the International Congress of Phonetics*, Miami, 1091-1094, 1979.

Un analyseur syntaxique pour la synthèse du texte. *Actes des Xèmes JEP*, Grenoble, 1979, 227-236.

Perception des séquences de contours prosodiques des phrases synthétisées. *Actes des 9èmes Journées d'Etude sur la Parole*, GALF, Lannion, 1978.

Combinatory Aspects of the Intonation-Syntax Relationship. *Actes de l'école d'été en linguistique mathématique*, CNUCE, Pisa, 1977.

L'intonation: Aspects linguistiques et reconnaissance de la parole. *Actes des 8èmes JEP*, Aix-en-Provence, mai 1977, vol. II.

Problèmes de neutralisation des marques prosodiques - Application à la reconnaissance automatique. *Actes des 8èmes Journées d'Etude sur la Parole*, GALF, Aix-en-Provence, 1977, 306-311.

Un analyseur-visualiseur de mélodie à microprocesseur. *Actes des 8èmes Journées d'Etude sur la Parole*, GALF, Aix-en-Provence, 1977, 211-215.

Modèles en intonation et syntaxe. *Actes des 8èmes Journées d'Etude sur la Parole*, GALF, Aix-en-Provence, 1977, 84-87.

Utilisation d'un modèle prosodique pour la synthèse par concaténation de mots en français et en anglais. *Actes des 8èmes Journées d'Etude sur la Parole*, GALF, Aix-en-Provence, 1977, 247-254.

About a theory of intonation *Phonologica* 19, *Akten der 3 Internationalen Phonologie-Tagung*, Wien, 1976.

Synthèse par règles de l'intonation de la phrase. *Actes des 7èmes Journées d'Etude sur la Parole*, GALF, Nancy 1976, 207-213.

Intonation et reconnaissance automatique de la structure syntaxique. *Actes des 6èmes Journées d'Etude sur la Parole*, GALF, Toulouse 1975.

Phonologie de l'intonation de la phrase. *Actes du XIVèmes Congrès International de linguistique et de philologie romane*, Naples, 1974, 213-227.

Classification formelle automatique et industries lithiques (1973). *Actes du VIIIe Congrès International des Sciences Préhistoriques et Protohistoriques*, Belgrade, Sept. 1971, 33-39. (en coll. avec D. Cahen).

Reconnaissance automatique de patrons intonatifs. *Acta Universitatis Carolinae Philologica* 1. *Phonetica Pragensia* III - Prague 1972 - 77-81.

L'analyseur de mélodie du laboratoire de phonétique expérimentale de l'Université de Toronto, *Actes du VIIème Congrès International des Sciences Phonétiques*, Mouton, La Haye 1972, 1272-1274

Vers une description fonctionnelle et structurale automatique d'objets linguistiques. *Actes du VIIème Congrès International des Sciences Phonétiques*, Mouton, La Haye 1972, 1154-1160.

Chapitres dans des ouvrages scientifiques

La structure prosodique dynamique: Rature et insertion de texte dans l'oral spontané, *Hommages à Claire Blanche-Benveniste*, Univ. Degli Studi di Torino, 2012 (à paraître).

The Representation of Intonation, in *Spoken Communication between Symbolics and Deixis*, In F. Albano Leoni, I. Chiari, F. Dovetto, A. Giannini, M. Pettorino (eds.): Cambridge Scholars Publishing, 2010. 241-257.

Le français, usages de la langue parlée, *Claire Blanche-Benveniste*, Louvain, Peeters (1 chapitre), 2010, 33-48.

A propos de la perception et la transcription des unités prosodiques. In M. Bilger (ed): *Les enjeux de la transcription de la langue parlée*, Presses Universitaires de Perpignan, 2008.

WinPitch Corpus: Presentation and user's manual. In E. Cresti and M. Moneglia (eds): *C-ORAL-ROM, Integrated Reference Corpora for Spoken Romance Languages*, Benjamins, London, 2005.

Prosodie et technologie. In E. Guimbretère (ed.): *La prosodie au cœur du débat*, Collection Dyalang, Université de Rouen. 135-150. (en coll. avec P. Léon),

L'apport des technologies de visualisation et de synthèse dans l'enseignement des langues. In Patricia Raymond et Claudette Cornaire (eds.): *Regards sur la didactique des langues*, Editions Logiques, Montréal, pp. 105-131. (en coll. avec A. Germain).

Teaching Oral Production with Computers, In Holland, M. and Delcloque, Ph. (eds.): *Speech Technology in Language Teaching*, Swets & Zeitlinger, Zurich. (with A. Germain).

Modelling F0 in Various Romance Languages: Implementation in Some TTS Systems. In E. Keller & al. (eds): *Improvements in Speech Synthesis*, Wiley & Sons, New-York, 2001.

La Musique de la Phrase. In H. Gezundhajt & P. Martin (eds): *Promenades en Phonétique*, Editions Mélodie-Toronto, 1997, 97-112.

Sentence Intonation in 4 Romance Languages. In Botinis et al. (eds): *Intonation : Theory, Models and Applications*, ESCA, Athens, 1997, 227-230.

Interaction prosodie-syntaxe en français: cas des adverbes en –ment. In C. Sorin and al. (eds): *Levels in Speech Communication: Relations and Interactions*, Elsevier, Amsterdam, 1995, 171-182.

Intonation de la phrase en français: cas des adverbes en –ment. In Ferguson, Gezundhajt, Martin (eds): *Accent, intonation et modèles phonologiques*, Ed. Mélodie-Toronto, 1994, 1-14.

A propos du statut théorique de la structure prosodique de la phrase. In *Mélanges Fernand Carton, Verbum*, 1991, Univ. de Nancy, XIV, 307-310.

Il était deux fois l'intonation... In *Mélanges Léon*, Editions Mélodie-Toronto, 1992, 293-304.

Prosodie: Parole et Langue. In *Calliope: La parole et son traitement automatique*, Dunod, 1989, 131-146.

Détection de la fréquence fondamentale. In *Calliope: La parole et son traitement automatique*, Dunod, 1989, 301-309.

Sur les principes d'une théorie syntaxique de l'intonation. In P. Léon et M. Rossi (eds): *Hommage à G. Faure*, Didier, Montréal, 1981.

Vers une théorie syntaxique de l'intonation. In M. Rossi et al. (eds): *Intonation: de l'acoustique à la sémantique*, Klincksieck, 1981.

Une théorie syntaxique de l'accentuation en français. In P. Léon et M. Rossi (eds): *L'accent en français contemporain*, Didier Montréal, 1980, 1-12.

La reconnaissance de patrons intonatifs. In Léon, Faure, Rigault (eds): *Prosodic Features Analysis / Analyse des faits prosodiques*, Didier, Montréal, 1970, 175-191.

Autres contributions

Petite histoire de l'analyse de la fréquence fondamentale, In L-J. Boë et C-E. Vilain (éd.): *Colloque « Un siècle de phonétique expérimentale : de Théodore Rosset à John Ohala »*, 2010, Grenoble, ENS édition, 317-329.

Intonation theory: Autosegmental vs. Phonosyntactic. Part 3: *Some intonation facts in Portuguese*. Cyber Semiotic Institute.

Intonation in Spontaneous Speech – Part 2: *Intonation and Macrosyntax*. Cyber Semiotic Institute.

The Prosodic Structure in French: Properties and Constraints. Cyber Semiotic Institute.

Intonation and Syntax – Another Point of View, Cyber Semiotic Institute.

Une méthode de calcul rapide du peigne spectral pour la mesure de la fréquence fondamentale, *Travaux de l'Institut de Phonétique d'Aix-en-Provence*, 1986, 359-369.

Sur la non-congruence des structures syntaxiques et prosodiques. *Trav. de l'Institut de Phonétique d'Aix-en-Provence*, 1981.

Réalisation de filtres actifs sans éléments de précision. *Trav. de l'Institut de Phonétique d'Aix-en-Provence*, 1981.

Une nouvelle méthode de mesure de la fréquence fondamentale par intercorrélation avec une fonction peigne. *Rapport d'Activité de l'Institut de Phonétique ULB* 15: 117-128, 1981.

An Experimental Study of Toronto English Sentence Intonation. *Trav. de l'Institut de Phonétique d'Aix*, 1980.

L'intonation des phrases à structures non connexes. *Rapport d'Activité de l'Institut de Phonétique ULB*, 12/1: 95-106, 1978.

Résumé d'une théorie de l'intonation. *Bulletin de l'Institut de Phonétique de Grenoble*, vol. VI: 57-87.

Questions de dominance des faits prosodiques sur les marques syntaxiques. *Studi di Grammatica Italiana, Acc. della Crusca Firenze*, vol. VI, 1977, 23-31.

Une théorie pour l'intonation de l'italien. *Rapport d'Activités de l'Institut de Phonétique* 11/2, mai-octobre 1977, 95-113.

A Theory for English Intonation. *Rapport d'Activités de l'Institut de Phonétique* 11/1, Avril 1977, 83-96.

L'accentuation en français: Théorie présuppositionnelle. *Rapport d'Activités de l'Institut de Phonétique*, Bruxelles, No.10/1, 1976, 75-82.

Eléments pour une théorie de l'intonation. *Rapport d'activités de l'Institut de Phonétique*, Bruxelles, No. 9/1, mars 1975, 97-126.

Une grammaire de l'intonation de la phrase française. *Rapport d'Activité de l'Institut de Phonétique*, Bruxelles, No. 9/2, 1975, 77-96.

PHILIPPE MARTIN VU PAR DES COLLEGUES

Philippe Intime

*Pierre Léon*¹

leon048@rogers.com

Université de Toronto

Philippe assure que je l'ai jeté, tout petit, dans un bain de fréquence fondamentale. Et puis, comme dans *Astérix*, il est sorti de la marmite avec une force herculéenne, même pas mesurable en décibels, renversant toutes les théories sur son passage. Je voudrais en raconter brièvement ici la genèse, à laquelle j'ai modestement contribué.

J'étais alors enseignant de phonétique et de linguistique française à l'université de Toronto depuis 1965. En 1966, j'avais créé un mini laboratoire de phonétique expérimentale avec un spectrographe acheté grâce à une bourse de recherche après un court passage chez Pierre Delattre, à New York. Mon expérience en phonétique expérimentale avait commencé à Paris, à l'Institut de Phonétique, au labo de Marguerite Durand, en 1948. Époque héroïque où on enfumait le cylindre du kymographe de Passy, pour obtenir un tracé dont il fallait compter le nombre de vibrations pour extraire la note fondamentale. C'était une occupation absorbante et pas toujours sûre. Le spectrographe n'allait pas se révéler un instrument tellement plus facile à manier, lui à cause de la richesse d'information qu'il produisait. J'avais bien besoin d'un ingénieur pour me guider !

On parlait beaucoup alors de la géniale invention d'Harlan Lane, à Michigan. Il avait mis au point ce dont je rêvais pour mes étudiants de phonétique corrective, un analyseur de mélodie qui donnait la courbe intonative d'un enregistrement de parole, sur écran de télévision. L'étudiant apprenant devait tenter de reproduire la même courbe. Si l'imitation visuelle était bonne, sa correspondance sonore devait l'être également. Élémentaire!

Je fais un saut à Michigan où Harlan m'accueille avec sa gentillesse américaine, dans son splendide laboratoire, haut lieu de technologie moderne. Il faut dire qu'à l'époque les ordinateurs étaient de la taille des armoires moyenâgeuses. Impressionnant! Harlan me fit la démonstration espérée, me montrant tout de suite qu'il fallait trouver pourquoi les courbes intonatives présentaient des déviations inattendues, empêchant le bon usage de l'instrument. Je suis donc reparti à Toronto, avec un projet copiant celui de Lane et soumis aussitôt au Conseil de recherches canadiennes. Il fut accepté et Harlan vint plus tard retrouver Philippe qui était arrivé entre temps à Toronto.

J'avais une douzaine d'étudiants de thèse, avec lesquels j'avais entrepris des recherches sur le français canadien, laissant prudemment aux Québécois le soin de s'occuper de leur propre parler. Nous nous intéressions particulièrement à la prosodie pour laquelle il y avait peu de recherches publiées en dehors des pays nordiques et anglo-saxons. On s'est tous mis à établir une bibliographie, tout en multipliant les spectrogrammes des phrases où l'analyseur d'Harlan Lane devenait fantaisiste.

¹ [Note des éditeurs] : Lors des journées en l'honneur de Philippe Martin, Pierre Léon était encore parmi nous. Il n'avait pas pu faire le déplacement jusqu'à Paris, mais nous avait envoyé ce texte pour qu'il soit lu devant l'assemblée. C'est à titre posthume que nous le publions dans ces actes.

Là, Philippe est apparu comme le Saint-Esprit, venant d'en haut – Montréal – à la recherche d'âmes à illuminer. Modeste, avec une tête d'adolescent qu'il a toujours gardée, il a dit en regardant les courbes intonatives : « Ya effectivement un problème. Mais avec un petit programme d'informatique, ça devrait s'arranger. » Et il a ajouté : « Continuez vos spectrogrammes! » On lui en a pondu des centaines. Pendant ce temps-là, il apprenait la phonétique et la linguistique. Six mois plus tard, il en savait plus que tout le monde.

J'avais déjà publié avec mes étudiants un premier volume : *Recherches sur la structure phonique du français canadien*, dans la nouvelle collection, *STUDIA PHONETICA*, dont Didier me confiait la direction. Avec Philippe, maître ingénieur ingénieux, on n'allait pas s'arrêter en si bon chemin. Toutes nos recherches allaient constituer le second volume de *STUDIA PHONETICA* : *Prolégomènes à l'étude des structures intonatives*. C'était une grosse bibliographie, suivie de l'étude de Philippe sur la naissance des technologies nouvelles et les problèmes posés par la mise au point d'un nouvel analyseur de mélodie. Ce volume fut suivi d'autres recherches auxquelles Philippe a beaucoup contribué et que j'ai exposées au Congrès International des Sciences Phonétiques à Montréal, sous le titre suggéré par André Rigault :

Où en sont les recherches sur l'intonation? Et dont parlera, plus tard, élogieusement, notre bon collègue aixois, Mario Rossi.

J'avais invité à Toronto Georges Faure, seul phonéticien français à cette époque à s'être intéressé à l'intonation. Il nous a été précieux dans les séminaires et comme catalyseur d'un rassemblement des grands chercheurs des études prosodiques (14 et 15 novembre 1969). On a réuni leurs contributions dans le troisième volume de *STUDIA PHONETICA*, *Prosodic Feature Analysis/Analyse des faits prosodiques*.

Une *Table Ronde* réunissait, le 15 novembre, pour clore ce colloque, les phonéticiens britanniques, nord-américains, canadiens et français, impliqués dans la recherche d'un analyseur de mélodie et d'un synthétiseur de parole avec : Peter B. Denes, de la *Bell Telephone*, qui lança la discussion après son rapport : *Speech Analysis and Speech Synthesis in Speech Training*. À l'interrogation de John Firth « Puisque vous avez un laboratoire, pourquoi ne nous donnez-vous pas un analyseur de mélodie? » Denes répliqua : « Il est impossible de construire un instrument physique qui mesure un phénomène linguistique ». Néanmoins, Denes expliqua les trois grandes méthodes employées alors pour résoudre les problèmes de l'analyse du spectre vocal, concluant avec ses propres recherches et confirmant que l'analyse en temps réel n'était pas encore possible avec les moyens dont on disposait alors. Après un long débat, Philippe exposa alors comment il avait réussi ce que les laboratoires de la Bell n'avaient pas encore pu réaliser. Il en fit la démonstration convaincante. Il eut un beau succès enthousiasmant Bolinger qui nous demanda un article pour son *Intonation*.

Pour être modeste, disons que Philippe a beaucoup amélioré, par la suite ce qui devait devenir son merveilleux *WinPitch*. Philippe volera bientôt de ses propres ailes, quittant Toronto pour Paris où il continuera une brillante carrière. Toronto ne s'en remettra pas.

On pourrait croire que j'ai résumé la carrière d'un ingénieur, auteur d'une *Phonétique acoustique*, alors que l'on sait combien Philippe a apporté de nouveau dans la manière d'aborder les problèmes de la prosodie, brillamment exposés dans son autre ouvrage, *L'Intonation*, également publié chez Armand Colin à Paris.

Philippe m'a beaucoup aidé de son expertise acoustique pour ma *Phonostylistique* et la dernière édition de mon *Phonétisme et prononciations du français*,

Philippe m'a fait l'honneur et le grand plaisir de m'offrir un livre d'*Hommages* quand j'ai pris ma retraite. J'aurais aimé contribuer à ce volume d'*Hommages* en lui rendant la pareille par un article de phonétique expérimentale, domaine où il est orfèvre et non seulement par des réminiscences historiques. Mais je suis de plus en plus effrayé par la somme de connaissances informatiques, mathématiques et autres -tiques, que je laisse aux jeunes chercheurs le soin de célébrer comme il faut la gloire de Philippe.

Je voudrais seulement terminer par quelques souvenirs de bon compagnonnage du temps du lab de Toronto quand pour tromper l'hiver canadien nous étions allés sur le lac gelé, dans une cabane de pêche au trou sur la glace. Philippe se rendit célèbre, dès le premier soir, par la prise d'un vairon de cinq centimètres. On en remit à plus tard la pesée. Personne d'autre n'avait pris de poisson mais on s'était bien réchauffés de scotch et de plantureuses grillades.

Le souvenir le plus reconnaissant que je dois à Philippe est de m'avoir évité de finir au goulag soviétique, durant la Guerre Froide. Nous étions partis en équipée à la recherche du synthétiseur de parole idéal. Gunnar Fant nous avait reçus à Stockholm et de là nous devions nous rendre au Tata Institute à Bombay, via Genève où notre avion faisait escale. Là, j'ai des problèmes avec le détecteur de métal que je fais sonner sans veste, sans chaussures, sans pantalon et finalement en costume d'Adam dans une cabine spéciale. Finalement arrive affolée une hôtesse : « C'est pas lui, c'est la machine. Mais rhabillez-le vite, prenez son bagage et courez le mettre dans l'avion. On l'appelle pour la dernière fois ! » On me traîne en courant jusqu'à l'échelle qui monte à la carlingue. On me pousse dans l'appareil, sans ménagement. On me flanque sur un siège, dans l'hostilité générale. J'entends du russe. Bizarre, bizarre. Puis un haut-parleur demande, en anglais si Mr. Pierre Léon est dans l'avion. Je lève la main et on me flanque dehors, sans autre ménagement. Philippe, qui m'attendait dans l'avion de Bombay, avait tout vu de son hublot et m'avait fait appeler par le commandant de bord. Arrivé à Moscou sans visa ni billet, durant la guerre froide, c'était la Sibérie, où je serais encore à casser des cailloux dans la neige.

J'ai beaucoup admiré la dextérité de Philippe à conduire sa jeep au milieu des vaches sacrées des rues de Bombay et son intrépidité au milieu des troupeaux de sacrés éléphants, de même que son stoïcisme à recevoir des paquets de sac de poudre de riz lorsque nous devions longer les innombrables fêtes des mariages que nous trouvions sur la route du Taj Mahal. La suite plus froide fut la mise en train du fameux synthétiseur qui nous avait fait courir le monde.

Merci aussi, Philippe, et bravo pour ta brillante carrière et la somme de réflexions que tu as apportée à la phonologie du français et à la phonétique expérimentale.

Références

Bolinger D. (1972) (ed.) *Intonation, Selected Readings*,

Harmondworth, Penguin Books.

Léon P. (dir.1969) *Recherches sur la structure phonique du français canadien*, Studia Phonetica 1, Montréal,Paris,Bruxelles, Didier.

- Léon, P. et Martin, Ph. *Prolégomènes à l'étude des structures intonatives* (1970) *Studia Phonetica* 2, Montréal, Paris, Bruxelles, Didier.
- Léon P. Faure, G. et Rigault, A. (dir.) *Prosodic Feature Analysis/Analyse des faits prosodiques* (1970) *Studia Phonetica* 3, Montréal, Paris, Bruxelles, Didier,
- Léon, P. « Où en sont les études sur l'intonation », *Actes du 7^{ème} Congrès des Sciences Phonétiques*, La Haye, Mouton, 113-156.
- Léon, P. (1992), *Phonétisme et Prononciations du français*, Paris Nathan. 6^{ème} édition 2011, Paris, Armand Colin,.
- Léon, P. (1993) *Précis de Phonostylistique*, Paris, Nathan. 2^{ème} édition 2000, Paris, Armand Colin.
- Martin, Ph. *Mélanges Léon*, Toronto, (1992) Editions Mélodie,
- Martin Ph. *Phonétique acoustique, Introduction à l'analyse acoustique de la parole* (2008) Paris, Armand Colin.
- Martin Ph. *Intonation du français*, (2009) Paris, Armand Colin.

Philippe Martin, son œuvre et sa carrière en quelques mots

Discours d'ouverture des Journées

Elisabeth Delais-Roussarie

Elisabeth.roussarie@wanadoo.fr

CNRS-UMR 7110/ Laboratoire de Linguistique Formelle, Université Paris-Diderot

Abstract:

Ce texte reprend l'intégralité des propos prononcés par Elisabeth Delais-Roussarie lors de l'ouverture des journées en l'honneur de Philippe Martin, à Paris, les 28 et 29 juin 2012.

Bonjour à tous, et un grand merci d'être venu pour rendre hommage à notre collègue et ami, Philippe Martin. Comme il achève cette année sa dernière année universitaire, Mathieu Avanzi, Hiyon Yoo et moi-même avons décidé d'organiser ces journées un peu comme un "au revoir". Nous avons donc contacté des amis et collègues de Philippe, et nous devons dire que l'accueil que nous avons reçu des uns et des autres, collègues, amis, anciens collègues, étudiants, a été des plus favorable. Nous espérons maintenant que ces journées, qui comprendront des communications scientifiques, mais aussi des démonstrations, des intermèdes festifs et des tables rondes, seront un succès.

Avant de lancer à proprement parler ces journées, il nous semble important de dire quelques mots sur la formation et la carrière de Philippe Martin. J'ai décidé de m'arrêter sur quelques points forts de sa formation qui ont sans doute fortement contribué à donner un caractère singulier à sa démarche pluridisciplinaire originale dans le domaine de la parole et de la prosodie.

Philippe Martin a une double compétence. D'une part, il a reçu une solide formation scientifique en sciences appliquées à l'Ecole Polytechnique de l'Université de Bruxelles. Cela lui a même permis d'obtenir en 1973 le titre de Docteur es Sciences de l'Université de Bruxelles. Parallèlement, il a reçu une formation en linguistique qui s'est achevée par l'obtention d'un doctorat en linguistique appliquée en 1972 à l'Université de Nancy.

A la suite de cette double formation, la carrière académique de Philippe Martin s'est construite autour de trois pôles universitaires:

- A Toronto, où il a été ingénieur de recherche de 1967 à 1981, puis Maître de Conférences de 1972 à 1975, puis Professeur de 1986 à 2002;
- A Aix en Provence, où il a été recruté comme Professeur des Universités de 1978 à 1986;
- A l'université Paris-Diderot, où il a été professeur de 2002 à nos jours.

Dans ces différents postes, Philippe Martin a fréquemment rempli des fonctions de direction. Il a par exemple été Directeur du laboratoire de Phonétique expérimentale de l'Université de Toronto de 1991 à 1996, et Directeur de l'UFR de Linguistique de l'Université Paris-Diderot de 2005 à 2012.

Dans ces différents lieux, et durant ces années, Philippe Martin a développé une approche originale pour appréhender les faits prosodiques, approche que sa double formation a rendu

possible.

Parmi les faits marquants de sa carrière, on peut noter les travaux qu'il a accomplis dans le domaine de l'extraction de pitch et de la visualisation de la mélodie. On lui doit par exemple la fonction peigne qui joue un rôle essentielle, encore de nos jours, dans l'extraction du pitch en vue de sa visualisation. On lui doit aussi le développement du logiciel WINPITCH qui permet, dans un environnement convivial, d'analyser et de modifier la mélodie de la parole dans des fichiers sonores de durée importante. Ces éléments ont ouvert la voie à des recherches sur corpus, comme dans le cadre du projet européen C-ORAL-ROM, où Philippe Martin a joué un rôle important.

Sur le plan de la modélisation théorique des faits prosodiques, Philippe Martin a développé une théorie de l'intonation qui repose à la fois sur de bonnes compétences formelles et sur une observation minutieuse des données, y compris à grande échelle. On retiendra surtout dans les travaux sur ce thème :

- l'analyse qu'il a proposé des relations entre structure syntaxique et structure prosodique,
- le rôle de mécanisme tonal comme le contraste de pente dans le marquage de la structuration prosodique, et donc de la structure syntaxique.

Les communications qui se suivront durant ces journées, comme celles qui composent ce volume, portent sur différents thèmes développés par Philippe dans sa carrière: la phonostylistique, la structuration prosodique, les mécanismes intonatifs à distance, la phonétique expérimentale et le traitement de signal.

SECONDE PARTIE

COMMUNICATIONS
SCIENTIFIQUES

DESCRIPTION PROSODIQUE
DE STYLES OU DE PARLERS PARTICULIERS

Philippe Martin, tel qu'en lui-même enfin la prosodie le révèle

Antoine Auchlin¹, Jean-Philippe Goldman¹, Anne Catherine Simon²

antoine.auchlin@unige.ch, jean-philippe.goldman@unige.ch, anne-catherine.simon@uclouvain.be

¹Université de Genève, Département de Linguistique

²Université catholique de Louvain, Institut Langage & Communication, Centre VALIBEL

Abstract:

This contribution provides a phonostylistic description of Philippe Martin's speaking style. The sample of speech under scrutiny was recorded when Philippe Martin (PM) provided a talk at a University Conference. This speech sample is first analyzed in its main prosodic dimensions, both informally and through semi-automatic tools (ProsoReport; ProsoDyn). It is then compared with samples from C-PROM corpus, including comparable speaking style (University Conference; same context of enunciation) as well as six other speaking styles. Second, the sample is submitted to manipulation, through prosocopy, in order to give PM different French regional accents (Meridional, Bruxellois, Liégeois and Vaudois). Transplanting prosody procedure and requirements are detailed, while the oral presentation will provide for audio output.

1. Introduction

Cette contribution vise d'une part à décrire le phonostyle de Philippe Martin (dorénavant PM), et, d'autre part, à manipuler des échantillons de sa parole afin de lui faire adopter d'autres caractéristiques prosodiques, en particulier marquées par des accents régionaux. Ces deux objectifs mettent en œuvre des outils automatiques d'annotation, d'analyse prosodique et de (re-)synthèse de la parole développés par les auteurs dans leurs travaux antérieurs.

À partir d'un enregistrement de PM dans une activité de présentation scientifique (durée 6'24''), on analyse différents paramètres macroprosodiques (débit, moyenne et étendue de la f_0 , densité accentuelle, etc. voir Goldman et al. (2007), Simon et al. (2010) afin de comparer cet échantillon à d'autres locuteurs dans le même type d'activité communicative (issus du corpus C-PROM, voir Avanzi et al. 2010]) et d'en décrire les similitudes et les différences.

Ensuite, on procède par prosocopie (Morel & Bazinger 2004, Roekhaut et al. 2010) pour resynthétiser la voix de PM en lui appliquant différents styles prosodiques et on analyse les effets produits (Suciu et al. 2007).

2. Données et méthodes

2.1. Échantillon et corpus de comparaison

L'échantillon retenu pour cette analyse consiste en la production de PM lors d'une conférence où il présentait, en compagnie de son co-auteur, une étude sur les contours de finalité (Avanzi & Martin 2007). Il s'agit d'une parole semi-préparée, non lue, publique (dans le sens où elle est adressée à un

auditoire en situation professionnelle) et relativement monologique. Cet extrait a été comparé aux 24 échantillons du corpus C-PROM (Avanzi et al. 2010), représentant 7 situations de parole (lecture, discours politique, journal parlé, conférence scientifique, demande d'itinéraire, interview radiophonique et narration conversationnelle). Dans la comparaison prosodique, le phonostyle de PM est contrasté simultanément avec l'ensemble des phonostyles de C-PROM, et avec les échantillons de parole produits dans la même situation (la conférence scientifique décrite ci-dessus) par trois autres locuteurs.

2.2. Annotation des données

Dans la ligne des annotations constituées pour le corpus C-PROM, l'échantillon de parole de PM a fait l'objet d'une transcription orthographique du texte segmenté en « unités séparées par des pauses » (qui ne correspondent pas nécessairement à un domaine prosodique univoque ; Gendrot et al. (2012) et d'une phonétisation. Sur cette base, on a réalisé un alignement du signal au niveau du phone, de la syllabe et du mot graphique à l'aide de l'outil EasyAlign (Goldman 2011). Une fois vérifié manuellement, cet alignement a permis divers traitements :

- une détection automatique des syllabes proéminentes (Simon et al. 2008, Goldman et al. 2012) associée à une annotation manuelle des mots clitiques vs. non clitiques, afin de distinguer les syllabes proéminentes en position d'accent final de celles en position d'accent initial ;
- un rapport prosodique global (Goldman et al. 2007) mesurant des paramètres macroprosodiques tels que : taux d'articulation, débit de parole et débit d'articulation (Grosjean & Deschamps 1972), registre moyen et étendue du registre ; mélodicité ; proportion de syllabes proéminentes, et distribution selon la position accentuable ; etc. ;
- une annotation manuelle en unités de rection syntaxique (selon la méthode présentée dans (Simon & Degand, 2011) afin d'analyser la corrélation entre unités syntaxiques et unités séparées par des pauses.

3. Le phonostyle de PM

Que révèle la prosodie de PM de son phonostyle personnel dans une situation de parole professionnelle ?

3.1. Description intuitive informelle

Le phonostyle, c'est, avant tout, une « impression », un « sentiment », que la façon de parler dans la situation déterminée laisse à l'auditeur, plus ou moins intensément, et de façon plus ou moins diffuse, ou plus ou moins nette et précise. Comme le rappelle le concept « d'émaillage » de Bally (cité par Léon 1993), il peut suffire de parsemer la parole ici ou là de telle particularité pour que le discours entier s'en trouve qualifié. On pourrait penser que cela ne vaut que des genres de discours très normés et formels ; mais un subjonctif, ou une articulation rétroflexe sonnante snob ou précieux dans une parole informelle ou familière ont le même effet. Tout émaillage se pose sur un fond, plus ou moins constant et homogène, qui contribue, également, à l'émergence de cette impression.

Dans l'extrait analysé - limitons d'emblée toute prétention à généraliser - différentes particularités tissent, les unes avec les autres, un sentiment un peu paradoxal, que celui qui parle dit ce qu'il dit sans y croire complètement ou sans y accorder de réelle importance.

Une première particularité, trait phonostylistique objet d'émaillage, pourrait, à lui seul, donner dans cette circonstance une impression de bonhomie, de convivialité, de proximité: le paraverbal, qui se manifeste sous forme de: petit gloussement, amorces de rire, expirations bruyantes par la bouche mi-ouverte formant soupir (soupir d'aise, non de dépit ou d'exaspération, par exemple).

Ce trait se trouve en cooccurrence avec une autre particularité, les importantes variations de débit, d'autant plus marquantes qu'elles surviennent de façon abrupte (vs progressive) et inopinées, et entraînent une baisse sensible de la précision articulatoire.

Certaines de ces variations abruptes de débit sont, en outre, accompagnées de chutes d'intensité sur plusieurs syllabes. Le « code d'effort » (Gussenhoven 2002) fait son travail, et fournit la traduction selon laquelle celui qui parle n'accorde pas beaucoup d'importance à ce qu'il dit.

Ce sentiment peut être renforcé par le « relief dynamique syllabique » relativement plat. Autrement dit, le taux de proéminences syllabiques qui paraît assez faible - notamment en raison des passages accélérés comportant peu de proéminences.

Au plan mélodique, si la parole n'est pas monotone, l'espace tonal dans lequel elle évolue donne une impression de régularité, de constance. Ni les données de f_0 (non présentées ici), ni la conduite très différenciée de PM concernant son marquage des proéminences accentuelles initiales vs. finales (Fig. 4a & b infra) ne soutiennent cette impression, dont on se demande si des paramètres prosodiques non explorés jusqu'à présent pourraient rendre compte.

Enfin, un nombre élevé d'allongements syllabiques de type hésitation et de « euh » signe cette parole comme familière, contribuant de façon convergente à l'élaboration de l'impression d'ensemble.

Bonhomie, convivialité, détachement voire désinvestissement, en plus court : une sorte de désinvolture. Mais si, dans l'élaboration progressive d'une impression phonostylistique chez l'auditeur intervient sa connaissance des paramètres de la situation de parole, alors intervient aussi le fait que l'auditeur dispose, ou non, d'informations encyclopédiques concernant la personne qui parle (son « ethos préalable », cf. Kerbrat-Orrechionni & De Chanay, 2006). En ce qui nous concerne, impossible de revenir en arrière : notre interprétation du phonostyle de PM dans cet extrait se retourne, se complexifie: le « détachement » se transforme en ironie, et la désinvolture en « deuxième degré de modestie ». Hahhh...

3.2. Description outillée

L'outil ProsoDyn (Goldman 2012) permet de visualiser les variations synchrones de quatre paramètres macroprosodiques : le débit d'articulation (en syll./sec.), la hauteur moyenne du registre et son étendue (en ST), et la densité accentuelle, définie comme la proportion de syllabes proéminentes par rapport aux syllabes non proéminentes. L'affichage dynamique de ces paramètres, dissociés ou associés, permet de détecter les variations et d'analyser si une variation de l'organisation temporelle du discours est corrélée, par exemple, à une variation du registre mélodique, et de quelle manière.

ProsoDyn permet en outre d'afficher ces paramètres conjointement à une annotation textuelle du

discours étudié, dans notre cas une annotation en unités syntaxiques maximales. Dans la section qui suit, nous analysons la manière dont PM insère des pauses ou réalise des variations importantes de débit ou de registre en lien avec les unités syntaxiques qu'il produit.

3.2.1. Débit. Ce qui est frappant c'est que, avec une très grande régularité, on observe une accélération brusque et importante du débit à la fin des unités maximales de rection syntaxique suivie d'un ralentissement marqué au début de l'unité suivante, une fois qu'elle a été amorcée. Toutes les macro-unités ne sont pas construites sur ce schéma, mais on pourrait dire que c'est une manière prototypique de les construire (voir Fig. 1). À chaque fin d'unité de rection, on observe une accélération du débit et, à chaque début, un fort ralentissement associé à des pauses (planification) et à des marques du travail de formulation (voir aussi Fig. 2 et section 3.2.2).

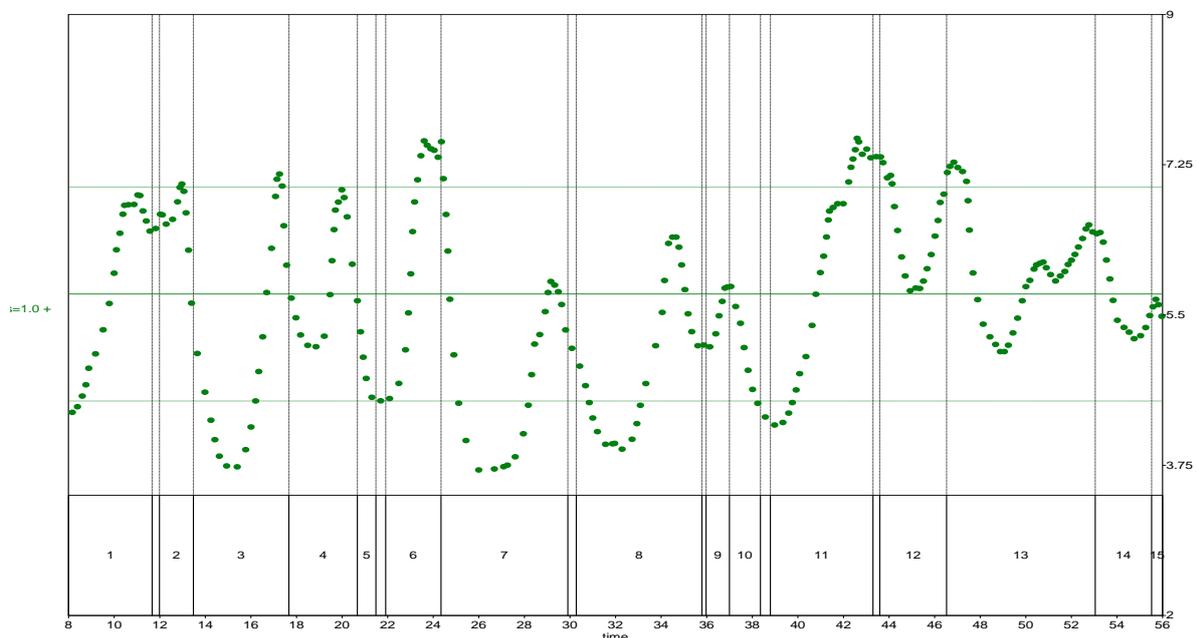


Figure 1 : ProsoDyn illustrant les variations de débit d'articulation relativement aux unités syntaxiques maximales : 1:voilà alors nous on pf on a beaucoup parlé ce matin de jesais pas aspects morphologiques ou idées morphologiques de de la prosodie # 2:et ici c'est c'est un petit peu différent 3:c'est # c'est basé sur euh peut-être le concept de base de la phonologie 4:c'est-à-dire que # euh c'qu'est en phonologie y a des contrastes 5:y a des différences # 6:et euh notre avis dans la structure prosodique aussi 7:alors je prends un exemple euh # avec deux deux mots prosodiques ou deux groupes accentuels # 8:et là je parle pas de continuation euh plutôt de tonème conclusif d'intonème conclusif # 9:mais de ce qui vient avant 10:voyez sur éléphanteau # 11:et euh c'est souvent décrit comme une continuation majeure etcetera avec une montée # 12:en fait ce quand on regarde un petit peu ça d'un point de vue purement phonologique 13:on se dit ben # y a un ensemble de traits qui peuvent fonctionner pour assurer la différence avec # justement le contour terminal 14:ça peut être n'importe lequel # de ceux qu'à j'ai mis 15:mais y en a bien d'autres possibles bien sûr # 16:euh en fait y a une sorte de neutralisation qui se passe # enfin neutralisation partielle 17:et un seul trait suffit.

On peut interpréter cet usage du débit comme une stratégie pour éviter une alternance de tour de parole : le locuteur accélère brusquement sa vitesse d'articulation en arrivant à une fin d'unité syntaxique – cette localisation correspondant à une place potentielle de transition où un passage de

tour au locuteur suivant serait envisageable (Selting 2000). Ayant « fermé » cette transition potentielle, le locuteur démarre l'unité suivante sur un débit toujours accéléré ; une fois qu'elle est entamée, il ralentit et peut planifier ce qui va suivre, ce qui transparaît dans l'apparition de marques d'hésitations et de pauses. Cette exploitation de la variation du débit s'apparente à une stratégie en contexte de concurrence pour la prise de parole, ce qui n'est pas le cas dans une situation (relativement monologique) de conférence scientifique...

Dans d'autres cas, la macro-unité se termine par un ralentissement du débit, par ex. quand elle est syntaxiquement inachevée ou qu'elle se termine par une hésitation. La section suivante décrit de manière plus détaillée la distribution des pauses dans quelques unités syntaxiques.

3.2.2. Pauses et marques du travail de formulation. Les pauses silencieuses peuvent tantôt être analysées comme des pauses structurantes, et considérées comme des marques associées aux frontières prosodiques majeures (Lacheret & Victorri 2002, Simon & Mertens 2009), tantôt comme des pauses d'hésitation, lorsqu'elles sont co-occurentes à des marques du travail de formulation (Candea 2000) (interruptions, répétitions, particules d'hésitation comme « euh », allongements vocaliques d'hésitation, etc.).

Dans le phonostyle de PM, on observe régulièrement l'insertion d'une pause longue (400 ms ou plus) à l'intérieur d'un constituant syntaxique, généralement juste après un mot de type « connecteur » ou « conjonction » qui projette une suite à venir. Dans l'exemple (1) ci-dessous, les pauses longues sont notées ##, les unités de rection syntaxiques sont annotées entre crochets droits [] et les marqueurs de discours ou marques du travail de formulation sont indiquées entre chevrons < > ; les accolades { } signalent une transcription incertaine :

Ex. 1. ## <euh> [c'est que <euh> en phonologie il y a des contrastes] [il y a des différences] ## <et> <euh> [{à} notre avis dans la structure prosodique aussi] <alors> [je prends un exemple euh ## avec deux deux mots prosodiques ou deux groupes accentuels]

Cette segmentation, appuyée par les accélérations et ralentissements de débit, mobilise la stratégie connue en versification comme enjambement. Les frontières d'une structure (unités séparées par des pauses, p.ex.) ne coïncident pas avec celles d'un autre niveau, unités de rection, ou « empan de formulation ». L'effet majeur de l'enjambement est un « lissage » des aspérités et irrégularités inhérentes au discours et sa structuration.

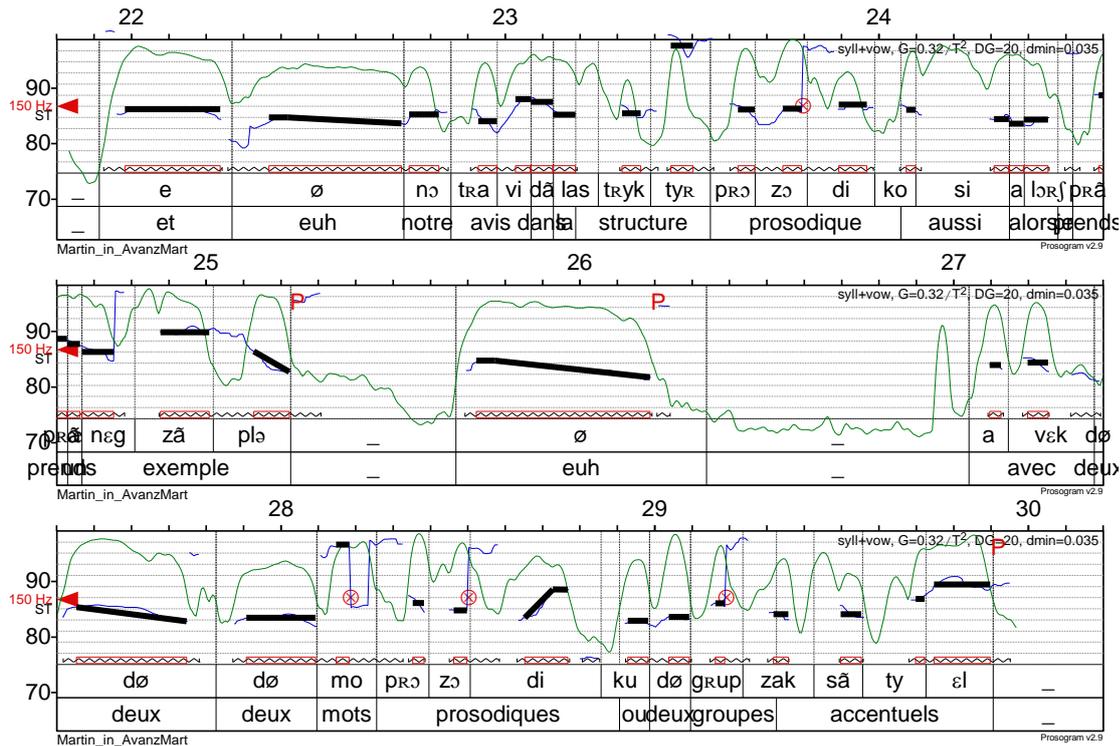


Figure 2. Prosogramme (Mertens 2004) d'un extrait de la parole de PM : chaque trait noir épais représente la mélodie stylisée du noyau syllabique ; le trait bleu représente la détection de la f_0 et le vert celle de l'intensité. La durée de chaque segment est mesurable au moyen de la graduation en dixièmes de seconde sur le bord supérieur du schéma.

4. Analyse prosodique comparée

Dans cette section, on compare les paramètres macro-prosodiques de la parole de PM à ceux d'autres locuteurs issus de situations de parole similaires ou non (voir § 2.1).

4.1. Paramètres temporels

Dans l'échantillon analysé, la vitesse d'articulation (hors pauses) de PM s'élève à 5.5 syl./sec. et sa vitesse de parole à 4.8 syl./sec. Par rapport à l'ensemble du corpus, ces valeurs sont relativement élevées, mais elles se situent dans la moyenne des trois autres échantillons de « conférence scientifique » (voir Fig. 3) (seule une des conférencières adopte un débit d'articulation plus rapide, s'élevant à 5.6 syl./sec.).

Du point de vue de ces deux paramètres, on peut dire que la conférence scientifique (cnf) se situe un peu en deçà d'un phonostyle très homogène, le journal parlé radio-phonique (jpa), texte lu avec un débit extrêmement rapide. La Figure 3 illustre clairement que d'autres styles sont très hétérogènes, comme le discours politique (pol) et, dans une moindre mesure, la lecture neutre (lec).

Le taux d'articulation de PM est également élevé (86.4% de son temps de parole est occupé par de la parole articulée), tout comme celui des trois autres conférenciers (variant de 83.7 à 88.6 %).

On peut remarquer que, par rapport à l'ensemble des échantillons du corpus C-PROM, le locuteur PM est parmi ceux qui réalisent les syllabes proéminentes dotées d'une f₀ relative élevée, tandis que les syllabes non proéminentes sont mélodiquement basses, tant pour les syllabes initiales que pour les syllabes finales. Du point de vue du registre tonal et de la mélodicité, la première impression auditive d'espace tonal régulier ou constant (§3.1) n'est donc pas confirmée par les mesures quantitatives.

5. Prosodie et prosodie régionale

La prosodie, ou transplantation prosodique, consiste à modifier un enregistrement de parole au moyen de l'algorithme OLA (Moulines & Charpentier, 1990) pour lui conférer des valeurs intonatives et temporelles provenant d'un autre locuteur. Autrement dit, le timbre de PM est maintenu (du fait des caractéristiques spectrales) mais les valeurs de la courbe mélodique et des durées syllabiques sont issues d'une relecture du même texte par un autre locuteur. Il s'agira alors de segmenter précisément les deux enregistrements (cf. Goldman 2011) et de s'assurer que les courbes mélodiques ne sont pas entachées d'erreurs de détection (Martin 2011). De plus il faut que les intervalles syllabiques coïncident exactement. Il a donc fallu faire relire le passage de PM en précisant dans la transcription les marques de formulations comme les hésitations, les interruptions et reprises syntaxiques, les épenthèses, les élisions, les liaisons facultatives, les pauses, les prises de souffle. Quatre locuteurs (un Marseillais, une Bruxelloise, un Liégeois et un Vaudois) se sont prêtés au jeu, sur quelques extraits : on leur a demandé de forcer leur accent régional, sachant que seuls les aspects prosodiques seraient pris en compte et non les spécificités segmentales ([R] vibré belge, ouvertures vocaliques vaudoises, nasales marseillaises).

Évalués informellement, les résultats audio sont plutôt décevants : la parole de PM peine à prendre l'accent. Plus scientifiquement, on conclut, à l'instar de (Boula & Brahim 2004), que la contribution des paramètres prosodiques à la constitution de l'accent régional (son identité – identifiabilité) est limitée, relativement à celles des facteurs segmentaux. L'inventaire des traits prosodiques pertinents pour chaque paire de parlers, cependant, reste à faire.

6. Conclusion

L'étude outillée de la parole de Philippe Martin la place comme relatif « outlier » par rapport aux autres paroles directement comparables (autres conférences) ; dans l'ensemble du corpus, cette parole, sur plusieurs paramètres, se place à proximité de paroles prononcées dans des situations moins publiques, et virtuellement plus improvisées (demandes d'itinéraires, i.a.). Son étude outillée était (contredit, parfois) l'analyse intuitive, et permet de la quantifier de façon précise. Mais les secrets de son phonostyle n'ont pas tous été percés

References

- Avanzi, M. et al. (2010). An annotated corpus for French prominence studies. *Proceedings of Prosodic Prominence: Perceptual and Automatic Identification, Speech Prosody 2010 Workshop*.
- Avanzi M. & P. Martin (2007). "L'intonème conclusif : une fin (de phrase) en soi ?". *Cahiers de linguistique française*, 28, pp. 247-258.

- Boula de Mareüil, P., & B. Brahimi (2004). Rôle du segmental et du suprasegmental dans la perception de l'accent maghrébin en français. *Actes des XXVes Journées d'Étude sur la Parole*.
- Cahen, D. & Ph. Martin (1972). *Classification formelle automatique et industries lithiques*. Annales du Musée Royal de l'Afrique Centrale, Série Sciences humaines, n°76.
- Candea, M. (2000). *Contribution à l'étude des pauses silencieuses et des phénomènes « d'hésitation » en français oral spontané. Etude sur un corpus de récits en classe de français*. Thèse de Doctorat, Université Paris III-La Sorbonne Nouvelle.
- Mertens, P. (2004). Un outil pour la transcription de la prosodie dans les corpus oraux ». *Traitement Automatique des Langues*, 45/2.
- Moulines, E. & F. Charpentier (1990). Pitch-synchronous waveform processing techniques for text-to-speech synthesis using diphones. *Speech Communication*, 9/5-6.
- Ferguson, B., Gezundhajt, H. & Ph. Martin (1994). *Accent, intonation et modèles phonologiques*. Mélodie-Toronto.
- Gendrot, C., Adda-Decker, M. C. & Schmid (2012). Comparaison de parole journalistique et de parole spontanée : analyses de séquences entre pauses. *Actes de la conférence conjointe JEP-TALN-RECITAL*.
- Gezundhajt, H. & P. (1997). *Promenades Phonétiques*. Mélodie-Toronto.
- Goldman, J.-P. et al. (2007). Phonostylographe: un outil de description prosodique. Comparaison du style radiophonique et lu. *Cahiers de linguistique française*, 28, pp. 219–237.
- Goldman, J.-P. (2011). “EasyAlign : an Automatic Phonetic Alignment Tool under Praat”. *Proceedings of Interspeech*. pp. 3233-3236.
- Goldman, J.-P. (2012) ProsoDyn: a graphical representation of macroprosody for phonostylistic ambiance change detection. *Proceedings of Speech Prosody*.
- Goldman, J.-P. et al. (2012). A Continuous Prominence Score Based on Acoustic Features, *Proceedings of Interspeech*.
- Grosjean, F., & Deschamps, A. (1972). « Analyse des variables temporelles du français spontané ». *Phonetica*, 26, pp. 129-156.
- Gussenhoven, C. (2002). Intonation and Interpretation: Phonetics and Phonology. *Proceedings of Speech Prosody*.
- Kerbrat-Orecchioni, C. & H. de Chanay (2006), 100 minutes pour convaincre: l'éthos en action de Nicolas Sarkozy, in Broth. M. Le français parlé dans les médias. Stockolm: Acta Universitatis Stokolmiensis, pp. 309-329.
- Lacheret-Dujour, A. & B. Victorri (2002). La période intonative comme unité d'analyse du français parlé: modélisation prosodique et enjeux linguistiques. *Verbum*, 24/1-2.
- Léon, P. (1993). *Précis de phonostylistique, Parole et expressivité*. Nathan Universités, Paris.
- Léon, P. & Ph. Martin (1969). *Prolégomènes à l'étude des structures intonatives*. Didier, Montréal.
- Léon, P. & Ph. Martin (1979). *Toronto English Studies in Phonetics*. Didier, Montréal, 1979.
- Martin Ph. (1992). *Mélanges Léon*. Éd. Mélodie-Toronto.

- Martin Ph. (2008). *Phonétique Acoustique*, Armand Colin, Paris.
- Martin Ph. (2009). *Intonation du français, mesures, théories, modèles*. Armand Colin, Paris.
- Martin, Ph. (2011). WinPitch, a multimodal tool for speech analysis of endangered languages. *Proceedings of Interspeech*, pp. 3273-3276.
- Morel, M., & Bänzinger, T. (2004). Le rôle de l'intonation dans la communication vocale des émotions: test par la synthèse. *Cahiers de l'Institut de Linguistique de Louvain*, 30/1-3.
- Roekhaut, S. et al. (2010). A Model for Varying Speaking Style in TTS systems. *Proceedings Speech Prosody*.
- Selting, M. (2000). The Construction of Units in Conversational Talk. *Language in Society*, 29/4, pp. 477-517.
- Simon, A.C. Avanzi, M. & J.-Ph. Goldman (2008). La détection des proéminences syllabiques. Un aller-retour entre l'annotation manuelle et le traitement automatique. *Actes du congrès Mondial de Linguistique Française*.
- Simon, A. C., & P. Mertens (2009). Automatic detection of prosodic boundaries in spoken French. A step towards the identification of basic discourse units. Unpublished Ms.
- Simon, A.C. et al. (2010). Les phonostyles: une description prosodique des styles de parole en français ». In Abecassi, M. & G. Ledegen, *Les voix des Français. En parlant, en écrivant*, Peter Lang, Berne, pp. 71-88.
- Simon, A.C.; Degand, L. (2011). L'analyse en unités discursives de base : pourquoi et comment ? *Langue Française*, 170/6, pp. 45-59.
- Suciu, I., Kannelos, I., & Moudenc, T. (2007). Expressivité et synthèse vocale. Isotopies expressives, cohérence discursive et structures prosodiques. *Cahiers de Linguistique française*, 28, pp. 199–206.

Organisation prosodique de la parole dans le style radiophonique

Katarina Bartkova¹, Denis Jouvét² & Natalia Segal³

Katarina.Bartkova@univ-lorraine.fr, Natalia.Segal@gmail.fr, Denis.Jouvet@loria.fr

¹ ATILF-UMR 7118, Université de Lorraine, France,

²LORIA-INRIA, Nancy (3) Reverso-Softissimo, Paris

Abstract:

An algorithm of prosodic trees allowing a hierarchical representation of the prosodic organization of speech is applied in this study on a corpus of broadcast news data in French. The developed algorithm contains the detection of prosodic boundaries and also the detection of the secondary (didactic) accents situated on the first syllable of lexical words. The detection of the prosodic boundaries is based on an approach that integrates linguistic knowledge and a hierarchical structuring using the inversion and the amplitude of the F0 slopes described in Martin (1987). The results provide an insight into the most frequently used prosodic structures in this broadcasting speech style.

1. Introduction

Notre étude s'inscrit dans le domaine du traitement automatique de la prosodie utilisant un cadre théorique linguistique. L'algorithme de construction d'arbres prosodiques revisité ici, a été présenté initialement en détails dans Segal & Bartkova (2007).

Notre approche automatique de segmentation prosodique repose sur une description théorique sous forme d'arbres prosodiques, cadre théorique mis en place pour de la parole préparée (Martin, 1987), puis adapté pour de la parole semi-spontanée (Segal & Bartkova, 2007).

2. Structure et mot prosodique

Notre travail se base sur l'existence d'une structure prosodique organisant hiérarchiquement les groupes prosodiques (groupes accentuels). Cette structure prosodique résulte de phénomènes de contraste des pentes mélodiques observées sur les syllabes accentuées. La structure prosodique est a priori indépendante, mais malgré tout associée à la structure syntaxique ; chaque structure ayant son propre ensemble de contraintes.

On s'accorde généralement pour se focaliser sur ou autour de la syllabe accentuée lors de la description des phénomènes prosodiques. Les unités prosodiques minimales - mots prosodiques - contiennent un accent lexical final et éventuellement un accent (didactique ou d'insistance) initial qui est facultatif. Une unité prosodique minimale contient un mot lexical (mots de catégories grammaticales ouvertes), et optionnellement des mots grammaticaux (mots des catégories fermées), et sa longueur ne dépasse généralement pas 8 syllabes.

Notre approche utilise des paramètres prosodiques mesurés sur les syllabes finales (durée

vocalique et F0) pour trouver les frontières prosodiques, ainsi que des contraintes rythmiques pour interdire les mots prosodiques trop longs (pas plus de 8 syllabes par mot prosodique). L'approche est basée sur la connaissance des frontières des mots obtenues par l'alignement forcé et prend en compte quelques contraintes lexicales sur l'organisation des mots prosodiques.

Les mots prosodiques identifiés sont organisés en une structure prosodique arborescente dont le nombre de niveaux n'est pas limité. Ici, les mots prosodiques n'ont pas de pattern mélodique standard, mais leurs paramètres prosodiques et les mouvements mélodiques sont imposés par la structure prosodique dont les deux règles principales sont

- l'Inversion de la Pente Mélodique (**IPM**)
- l'Amplitude de Variation Mélodique (**AVM**)

3. Paramètres prosodiques

Les valeurs de F0 en semi-tons et les valeurs de l'énergie ont été calculées toutes les 10 ms à partir du signal de parole en utilisant l'analyse acoustique Aurora. La phonétisation du texte pour réaliser l'alignement forcé est obtenue à partir du lexique Bdlex, et d'une phonétisation automatique pour les mots manquants.

Nous avons opté pour l'utilisation des valeurs observées uniquement sur les voyelles, ignorant ainsi les consonnes car elles introduisent majoritairement des perturbations micro mélodiques au niveau des paramètres étudiés. Par ailleurs, la prise en compte de la durée vocalique permet d'éviter tout problème lié aux formes des syllabes (syllabe fermée vs. ouverte, syllabe avec attaque ou coda complexes, etc.). De fait, les durées vocaliques peuvent être considérées comme plus homogènes, car moins contraintes par la structure interne des syllabes que les consonnes

L'énergie de chaque voyelle correspond à la valeur moyenne calculée sur l'ensemble des trames de la voyelle, puis normalisée par la moyenne des énergies des voyelles du groupe de souffle ou des 5 voyelles précédant la voyelle courante (si le groupe de souffle est trop court).

La durée des voyelles a été normalisée par la moyenne des durées des voyelles non-accentuées (voyelles en position interne des mots pluri-syllabiques) se trouvant dans le même groupe de souffle que la voyelle courante. Quand le nombre de voyelles était inférieur à 5, le calcul de la durée moyenne a été élargi sur les groupes de souffle adjacents.

Pour chaque voyelle la pente du F0 a été calculée par régression linéaire. Cette approche nous a permis de lisser des valeurs inappropriées de F0, comme, par exemple, les premières valeurs de F0 après une consonne plosive non voisée. En complément de la pente, nous avons calculé également, pour chaque voyelle, le delta de mouvement de F0 par rapport à la voyelle précédente.

4. Méthodologie

Le découpage de signal de parole par des paramètres prosodiques est précédé par un découpage de texte en mots prosodiques accentuables, regroupant les mots cliques avec les mots lexicaux. Ce découpage permet par la suite de considérer les paramètres prosodiques

uniquement sur les dernières syllabes (voyelles) des groupes accentuables dont l'accentuation était confirmée ou infirmée par la valeur des paramètres prosodiques.

Deux paramètres principaux : la pente de F0 et la durée normalisée de la voyelle (autre que la voyelle [ə] en position finale quand le mot est pluri-syllabique) ont été utilisées pour la détecter les frontières prosodiques. Pour définir le seuil de la durée qui sépare une voyelle accentuée d'une voyelle hors accent un histogramme a été calculé représentant la distribution des durées normalisées des voyelles en position non-accentuée (syllabes autres que dernières syllabes des unités lexicales) et en position accentuée (syllabes suivies de pause). La frontière pertinente entre les deux distributions des durées se situait entre la valeur 130% et 150%, par conséquent une voyelle dont la durée dépassait une fois et demi la valeur étalon, a été considérée comme accentuée.

La même démarche a été appliquée pour rechercher le seuil séparant les valeurs des pentes de F0 sur des frontières prosodiques et hors frontières prosodiques. Le calcul de la distribution des valeurs des pentes entre voyelles non-accentuées (syllabes internes des mots pluri-syllabiques) et voyelles accentuées (syllabes suivies d'une pause) a situé le seuil séparant ces deux groupes de pente vers la valeur correspondant au seuil de glissando de 0.32 obtenu sur la parole ('t Hart et al. 1990).

Pour la détection de l'accent secondaire (didactique), l'énergie de la voyelle est utilisée en complément de la pente de F0.

5. Corpus

Dans cette étude nous avons utilisé des extraits (environ 1h30) du corpus Ester constitué d'enregistrements radiophoniques, où les passages semi-spontanés (interviews) alternent avec des passages de parole préparée (bulletins d'information) et qui représente un style que l'on peut qualifier de phonostyle radiophonique ou journalistique.

La description prosodique des différents phonostyles a suscité beaucoup d'intérêt dans la littérature scientifique. Ces études présentent souvent la variation des paramètres prosodiques d'une façon globale (Goldman et al., 2007 ; vitesse d'articulation ou de débit accéléré, F0 moyenne plus haute ...) par rapport à la parole considérée comme non-marquée (neutre). Le phonostyle des journalistes présente un caractère partiellement standardisé bien qu'il laisse un espace à un style personnel. Des études consacrées à ce phonostyle cherchaient non seulement à cerner le style vocal des journalistes lorsqu'ils passent à l'antenne, mais également à définir un style spécifique par chaîne de radio en vérifiant si des chaînes radiophoniques concurrentes comportaient une distinction prosodique entre elles (Hupin & Simon 2009).

L'utilisation de notre algorithme de découpage du signal nous permet de situer notre étude du style radiophonique sur un niveau plus linguistique, et d'étudier les structures prosodiques dans ce style en les confrontant avec le cadre théorique développé pour la parole lue préparée (Martin 87) et déjà testé sur la parole semi-spontanée (Segal & Bartkova 2007).

Le signal de parole de notre corpus a été segmenté par un alignement forcé réalisé avec le module du décodage acoustique du logiciel Sphynx (logiciel de reconnaissance automatique de la parole développé au CMU). Cet alignement forcé permet d'obtenir les durées des sons, ainsi que l'emplacement et la durée des pauses. Comme le signal de parole était plutôt de

bonne qualité, nous pouvons supposer que la segmentation s'est déroulée sans problèmes majeurs. Néanmoins un décalage entre le signal et la transcription phonétique peut se produire quand la transcription orthographique s'éloigne du contenu acoustique du signal, c'est-à-dire, quand elle contient des erreurs ou se révèle trop « normée » et ne capte pas des écarts de prononciation qui sont difficilement dérivables à partir d'une transcription canonique.

6. Découpage prosodique

La première étape du découpage prosodique consiste à déterminer le degré de la profondeur des frontières prosodiques en utilisant séparément, d'une part la durée des voyelles des dernières syllabes des unités accentuables, et d'autre part la valeur de la pente de F0. Les deux découpages sont alors combinés afin de garder le rang le plus élevé des deux paramètres dans le découpage (hiérarchie) final. Ainsi, l'indice d'une frontière définie à partir d'une durée vocalique longue mais ayant une pente de F0 neutralisée (CN) sera corrigé afin d'indiquer une frontière à un niveau plus élevé dans l'arbre prosodique du constituant prosodique considéré.

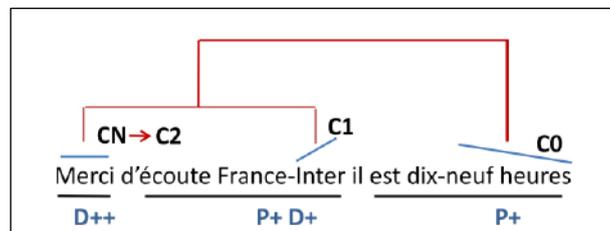


Figure 1 : Exemple d'arbre prosodique marqué par la pente (P+) et/ou la durée (D+).

L'étude de la distribution des groupes prosodiques en fonction de leurs durées exprimées en nombre de syllabes montre que le nombre de syllabes par groupe prosodique (Fig. 2, couleur bleue) diminue fortement à partir de 4 syllabes. Un rythme secondaire est introduit dans les groupes prosodiques par le biais des accents didactiques secondaires se trouvant sur les premières syllabes des mots pluri syllabiques, créant des arcs accentuels (Fónagy, 1980) et diminuant la longueur de la chaîne entre deux syllabes accentuées (Fig. 2, couleur rouge).

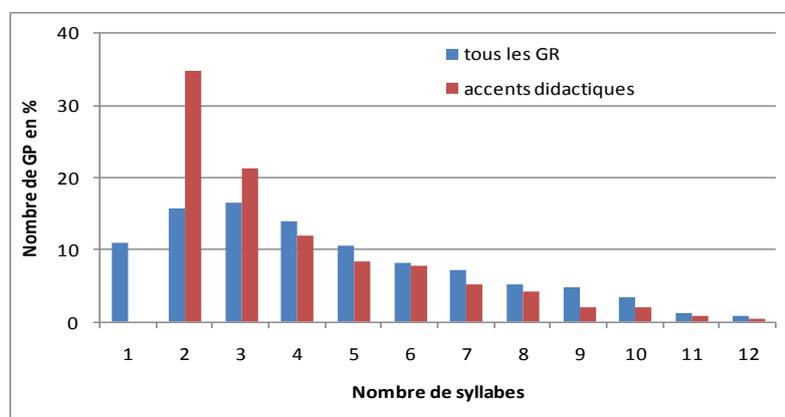


Figure 2 : Fréquence d'occurrence des groupes prosodiques

Comme indiqué plus haut, des contraintes rythmiques ont été implémentés dans le logiciel

afin d'interdire, quand c'était possible, les mots prosodiques trop longs (pas plus de 8 syllabes par mot prosodique) et de regrouper des mots prosodiques séparés par de courtes pauses (<100ms) et ayant une frontière prosodique neutralisée, ne provoquant pas de reset des paramètres prosodiques. Quand un groupe prosodique dépassait 8 syllabes, une recherche de frontière a été initiée en utilisant des seuils de décision abaissés.

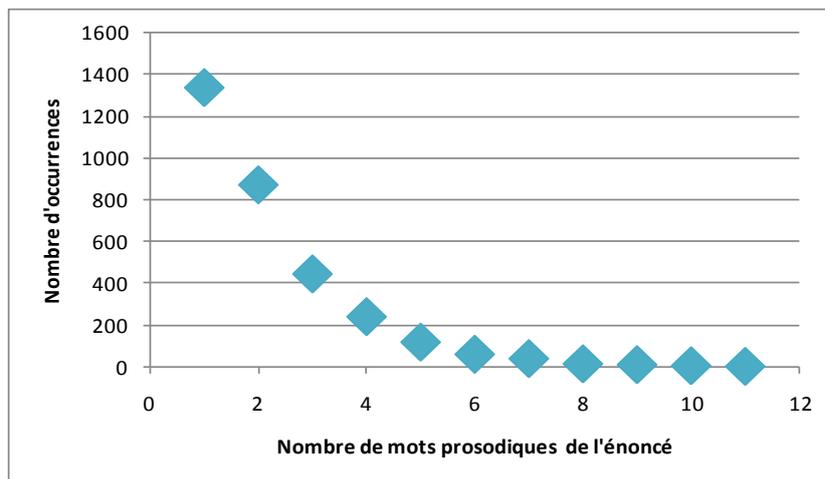


Figure 3 : Nombre de mots prosodiques dans l'énoncé

7. Arbres prosodiques

Les arbres prosodiques les plus fréquents étaient ceux contenant deux branches (deux mots prosodiques) ayant majoritairement des mouvements montants de la pente sur leurs deux constituants.

Les pentes montantes de F0 étaient légèrement plus fréquentes sur les frontières prosodiques (38%) que les pentes descendantes (26%). Dans 12% des cas la frontière était considérée comme neutralisée par les deux paramètres et dans 24% de cas c'est le paramètre de la durée vocalique qui déterminait la frontière prosodique.

Pour des énoncés constitués de trois mots prosodiques, la structure prosodique la plus fréquemment utilisée était celle de l'énumération avec des pentes montantes ayant approximativement la même amplitude (cf. Fig. 4). Ces données corroborent celles observées dans notre étude précédente (Segal & Bartkova 2007), où nous avons trouvé des structures prosodiques avec essentiellement des pentes montantes, avec peu d'oppositions basées sur l'inversion des pentes.

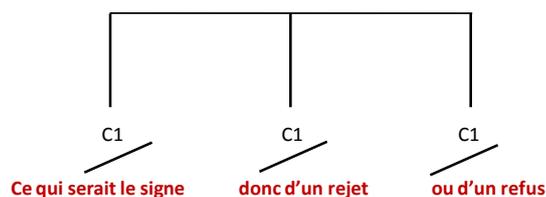


Figure 4 : Arbre prosodique d'énumération

Or, le style radiophonique peut être considéré comme préparé ou en tout cas semi-spontané, par conséquent il s'apparente également avec le style préparé. En effet, la structuration prosodique s'exprime ici également par l'inversion des pentes de F0 ainsi que par la variation de leurs amplitudes. En dehors de l'énumération, les arbres prosodiques privilégiés pour hiérarchiser des énoncés de 3 constituants prosodiques sont ceux utilisant l'inversion et l'amplitude des pentes dans une dépendance à droite (entre C2 et C0) ou à gauche (entre C2 et C1).

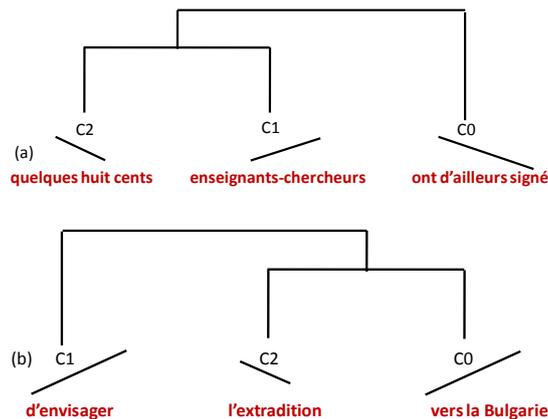


Figure 5 : Arbres prosodiques avec branchement à droite (a) et à gauche (b)

Quand l'énoncé contient 4 mots prosodiques, la structure prosodique que l'on observe le plus fréquemment et celle d'une inversion de pente au début de l'énoncé basculant vers une structure d'énumération sur les constituants restants tout en privilégiant la pente montante du F0.

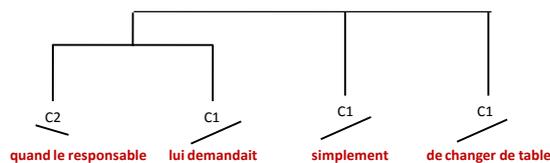


Figure 6 : Arbre avec inversion puis répétition des pentes

Nous avons également observé des arbres non prévus par le cadre théorique dans lequel nous avons situé notre étude. En effet sur plusieurs arbres prosodiques des répétitions de la même direction des pentes ont été observés alors même que leurs amplitudes ne permettaient pas d'effectuer la structuration attendue.

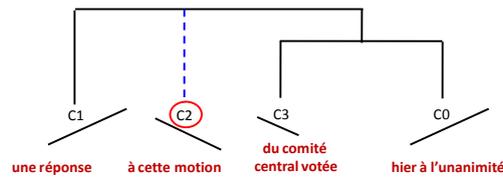


Figure 7 : Arbre prosodique non conforme au cadre théorique utilisé

8. Conclusion

Cette étude nous a permis d'appliquer l'algorithme de structuration prosodique et de détecter automatiquement des frontières prosodiques sur une base de données contenant de la parole journalistique (préparée et semi-spontanée). L'approche utilisée nous a permis d'étudier les structures prosodiques le plus fréquemment utilisées dans ce style de parole ainsi qu'observer son organisation rythmique.

Pour compléter cette étude, il serait utile de découper le corpus étudié afin d'introduire une distinction entre le style des bulletins d'information et des interviews. Il serait également intéressant de comparer les arbres prosodiques observés avec les arbres syntaxiques obtenus par un traitement automatique, afin de permettre l'étude de la congruence entre ces deux structurations.

Références

- Fónagy, I. (1980). L'accent français: accent probilitaire. *Studia Phonetica*, 15, pp. 123-233.
- Goldman J.-Ph., Auchlin A., Simon A. C., & M. Avanzi (2007), "Phonostylographe : un outil de description prosodique. Comparaison du style radiophonique et lu". *Nouveaux cahiers de linguistique française*, 28, pp. 219-237.
- Hupin, B. & A.C. Simon (2009). Analyse phonostylistique du discours radiophonique. Expériences sur la mise en fonction professionnelle du phonostyle et sur le lien entre mélodicité et proximité du discours radiophonique. *Recherches en communication*, 28, pp. 103-121
- Martin Ph. (1987). Prosodic and rhythmic structures in French, *Linguistics*, 25, pp. 925-949.
- Segal, N. & Bartkova, K. (2007), "Prosodic structure representation for boundary detection in spontaneous French", *Proceedings of ICPHS*, pp. 1197-1200.
- 't Hart J., Collier R. & A. Cohen (1990), *A Perceptual Study of Intonation*. Cambridge U.P., London.

Etude prosodique du parler ouvrier de Tourcoing (Nord) à la fin du XIX^{ème} siècle

Fernand Carton

carton.fernand@free.fr

Université de Lorraine

Abstract:

This contribution is the instrumental analysis of two old corpora, characteristic of the speech of the Picardy type in a city of the North of France, at the linguistic border between French and Dutch. The speakers are workmen who worked all their life in textile factories, in full expansion at the end of the XIXth century and at the beginning of XXth. Starting from old recordings, we study the prosody of familiar autobiographical accounts: temporal, stress and intonative structures. We comment on the statistical data, the melody curves and the spectrograms. Our study show prosodic features that are different from those of current standardized French, and also from features influenced by West-Vlaams languages especially for one of the speakers.

1. But de la recherche

Connaissant depuis plus de quarante ans¹ le goût très vif du dédicataire pour l'humour, j'avais envie de le taquiner en intitulant ma contribution: *Prosodie franco-belge...* Le parler qui fait l'objet de cette étude est transfrontalier comme le dédicataire, mais le titre est plus précis. A Tourcoing, dans la triangulaire français-West-Vlaams²-picard, c'est le système segmental du picard qui s'imposait chez deux ouvriers du textile nés en 1874 et 1877 (Carton, 2012) : c'était la langue courante des ouvriers de l'industrie textile du Nord à l'époque de sa pleine expansion. L'analyse instrumentale devrait mettre en évidence des différences prosodiques entre langues de même souche, français et picard.

2. Locuteurs

Deux locuteurs ont été choisis en raison de critères socioprofessionnels, de leur âge et de leur histoire.

EM est né en 1874 à Reckem (Belgique, Figure 1), village frontière à 6 km à l'est d'Halluin (France) dans une famille de dix enfants. Son père, ouvrier frontalier, a parlé le Wvl dans sa jeunesse, mais sa langue courante était le picard de Tourcoing, comme sa mère, ouvrière elle aussi. Ils sont venus travailler dans les tissages de Tourcoing, alors en pleine expansion industrielle, alors qu'EM avait 6 ans. Il a été scolarisé dans le primaire en France, et il dit n'avoir jamais parlé le Wvl mais le comprendre un peu. EM s'est marié à une française, comme de nombreux immigrants dont l'industrie lainière française avait grand besoin. Le ménage a habité La Marlière, hameau frontalier, à 10 km au sud de Reckem. Ouvrier dans des usines de tissage de Tourcoing, il ne parlait que le picard local avec ses familiers.

¹ Depuis une certaine thèse soutenue à Nancy...

² Flamand occidental. En abrégé Wvl.

L'enregistrement a été réalisé par nous, à la suite d'une enquête pour l'Atlas linguistique picard (Carton, 1989 et Carton, 1998) en 1962, au domicile du locuteur.

JBM est né en 1877 dans le quartier du Pont Rompu au nord de Tourcoing, près de Neuville-en-Ferrain, dans une famille de six enfants. Ses parents étaient ouvriers du textile et natifs de Tourcoing. Dès la sortie de l'école primaire, il a été manœuvre dans diverses filatures de laine à Tourcoing jusqu'à 78 ans. Marié à une tourquennoise, ouvrière comme lui, il passait ses loisirs à cultiver un *courti*³ et à jouer aux boules dans les *bourloires* de cabaret. L'enregistrement a été effectué par nous en 1966 à l'Hospice de Tourcoing, où il résidait, en présence de retraités, pour le mettre plus à l'aise.

La prononciation d'EM a reçu la qualification d'« accent flamand » par tous les natifs du Nord auxquels nous avons fait entendre les corpus, contrairement à celle de JBM.

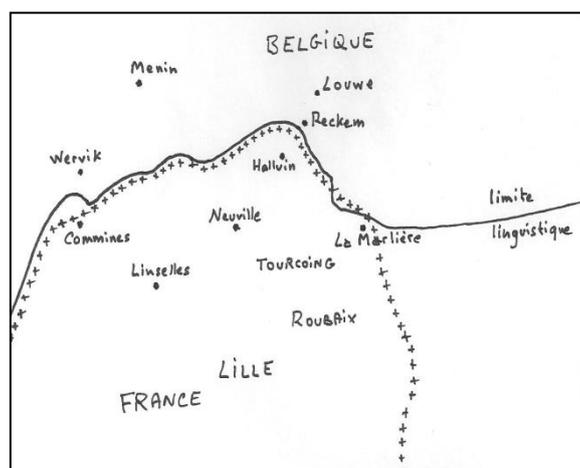


Figure 1 : Carte schématique des lieux cités

3. Corpus

Le corpus EM comporte trois anecdotes autobiographiques⁴ :

- Le cerf-volant (01:35⁵). Un cerf-volant lumineux manié par EM est pris pour un phénomène apocalyptique. Les faits se passent à La Marlière vers 1880.
- Incident au tissage (02:01). EM provoque sans le vouloir une réaction chimique. Les faits se passent à Tourcoing en 1895.
- Une farce de conscrits (02:02). Dans un cabaret du Pont-de-Neuville à Tourcoing, des conscrits, fêtant leur incorporation, font ingurgiter un purgatif à un intrus. Les faits se passent en 1894.

Le corpus JBM comporte trois souvenirs autobiographiques :

- Le travail d'un ouvrier de filature (01:20). JBM a été «homme de peine» et remplaçant

³ Petit jardin ouvrier.

⁴ Ces enregistrements figurent dans les données du Centre de ressources en documentation orale (CRDO) en cours de transfert au Speech and Language Data Repertory (SLDR). On peut les entendre en ligne, sur <http://carton.fernand.free.fr>.

⁵ Durée de l'enregistrement.

du veilleur de nuit de son usine à Tourcoing, vers 1900.

- La maison et la nourriture (01:26). Description d'une mesure de la fin du XIXème siècle. Frugalité des repas.
- Jeux et fêtes d'un enfant d'ouvrier vers 1900 (01:15). La guise (bâtonnet), la toupie flamande, les œufs de Pâques.

4. Protocole expérimental

Les corpus ont fait l'objet d'une analyse acoustique afin de les segmenter plus sûrement. L'opération était difficile du fait des problèmes d'élocution et de l'ancienneté du matériel enregistreur⁶. Nous avons bénéficié de l'aide de Katarina Bartkova⁷. Pour pouvoir comparer les durées, elle les a normalisées. La durée moyenne des voyelles non accentuées (dans des syllabes non suivies de pause) a été préférée à la durée syllabique, car les structures syllabiques sont très diverses. Cette durée moyenne a été calculée pour chaque séquence de parole délimitée par des pauses. Cette durée reflète mieux la vitesse d'articulation de la parole que la durée des syllabes, dont la structure est très variable. K. Bartkova a ensuite calculé une durée moyenne et la valeur de l'écart-type pour les 8 macro-classes homogènes des unités (occlusives sourdes, occlusives sonores, fricatives sourdes, fricatives sonores, nasales, liquides, semi-voyelles, voyelles). Elle a utilisé la durée normalisée des sons qui n'étaient pas immédiatement suivis de pause (Figure 2).

Sons	Moyenne	EcTyp	NB
ptk	112	35	190
vzʒ	88	31	96
bdg	100	48	104
mn	91	30	111
fsʃ	120	48	127
rl	76	79	224
SV	107	40	38
VOY	125	58	725

Figure 2 : Durées moyennes et écarts-types des macro-classes chez EM

En effet, dans une position immédiatement suivie de pause, les sons s'allongent de façon significative, souvent indépendamment de la vitesse d'articulation. Pour décider si la durée d'un son est allongée ou non, sa durée a été comparée à la somme de la durée moyenne et de l'écart-type de sa macro-classe. Si la durée normalisée du son est plus grande que cette valeur

⁶ Magnétophone à bande magnétique Philips modèle 1959, coupant à 4000 Hz.

⁷ Maître de conférences de phonétique à l'Université de Lorraine. Nous la remercions pour son aide précieuse.

(moyenne + écart-type), la durée est considérée comme allongée.

L'écart-type important pour [R] et [L] chez EM peut s'expliquer par le fait que les réalisations de ces phonèmes varient chez lui en fonction de leur position.

Sons	Moyenne	EcTyp	NB
ptk	109	35	94
vzɔ	85	34	37
bdg	98	32	54
mn	86	26	50
fsʃ	128	50	76
rl	65	37	65
SV	101	39	27
VOY	137	104	339

Figure 3 : Durées moyennes et écarts-types des macro-classes chez JMB

L'écart-type pour les voyelles est plus grand chez JBM que chez RM, sans doute à cause du débit plus irrégulier.

Nous avons éliminé des corpus les bégaiements, les syllabes en *creaky voice* avec trémulations, caractéristiques manifestement liées à l'âge des locuteurs (Carton, 1982). Nous avons travaillé à partir des spectrogrammes et des variations du fondamental: les tracés ont été réalisés par K. Bartkova avec WaveSurfer. Après segmentation, les signes phonétiques figurant sur les tracés sont générés par le codage SAMPA.

5. Structuration temporelle

5.1. Durées vocaliques

La comparaison avec la durée moyenne calculée montre qu'il y a dans les deux corpus des allongements en syllabe pénultième, accompagnés d'une chute mélodique : c'est un trait caractéristique du picard du Nord (Carton 1979), identifié comme un « accent traînant » Ex:

JBM [dɛz e vjɛl me:'zɔ e#⁸] « dans une vieille maison hein».

La durée du [e] est 247 ms, celle de [ɔ] est 260ms.

Durée vocalique moyenne + écart-type pour ce locuteur: 241: la voyelle pénultième est longue.

EM [j eto ma:'lat#] « il était malade » (Figure 12).

La durée du premier [a] est 185 ms, celle du second: 126 ms; la durée vocalique

⁸ Le signe # indique une pause silencieuse; le trait oblique sépare les groupes accentuels.

moyenne + écart-type pour ce locuteur est 183: la voyelle pénultième est longue. Dans ces exemples, l'accent lexical est final, marqué par une forte prééminence mélodique.

		micherry_1_duree.txt	
60,3251269	60,5564365	l	231
60,5564365	61,0190556	#	463
61,0190556	61,1002169	oe	81
61,1002169	61,1915233	t	91
61,1915233	61,2463071	r	55
61,2463071	61,3579038	oe	112
61,3579038	61,4552973	m	97
61,4552973	61,9341487	an	479
61,9341487	62,0660357	l	132
62,0660357	62,2547356	ei	189
62,2547356	62,4474936	s	193
62,4474936	62,5895258	ai	142
62,5895258	62,6645999	l	75
62,6645999	62,7701096	i	106
62,7701096	62,9080837	tch	138
62,9080837	62,9892449	y	81
62,9892449	63,0359126	l	47
63,0359126	63,1556255	b	120
63,1556255	63,2935996	y	138
63,2935996	63,5675188	t	274
63,5675188	64,1498508	#	582
64,1498508	64,2715233	.	133

Figure 4 : Extrait de la segmentation (à droite, durées en millisecondes): JBM eutremant les selles i tchulbut'# « sinon, les chaises culbutent »

5.2. Traitement des semi-voyelles

La diérèse n'est pas générale en picard, mais elle est presque constante chez nos locuteurs. Parce qu'EM et JBM analysent la séquence [tʃi] comme une suite de deux phonèmes : ils prononcent [yi] et intercalent parfois une semi-voyelle de transition [lytʃi], comme cela se fait dans une prononciation possible du mot *piano* [pijano] en français et en néerlandais. La séquence du français *soir* [war] est analysée par EM comme une séquence de deux voyelles: [so'ar] contient alors deux syllabes au lieu d'une. De même [ka'Ri'o] « chariot », [edo'ar] « Edouard » sont trisyllabiques dans nos corpus. Ce traitement est moins constant avec yod: *tchurieux* « curieux » est dissyllabique comme en français.

5.3. Traitement des consonnes

5.3.1. Réductions. Les deux locuteurs ont tendance à réduire les groupes consonantiques (*clusters*). Ce relâchement articulaire semble en partie lié à l'âge, car le picard du Nord se caractérise par une forte énergie articulaire (Carton, 1979). Ex :

JBM juste in morceau > jus' in morceau⁹
 [t] disparaît.

Les constrictives sonores intervocaliques ont tendance à s'amuir.

⁹ Les syllabes accentuées sont soulignées.

EM j'de'o « je devais »
[v] disparaît.

5.3.2 Assimilations consonantiques. Les assimilations totales sont très nombreuses chez les deux locuteurs, comme il est attendu dans un parler familial de personnes âgées. (Figure 5) Elles sont toutes anticipantes¹⁰. Elles sont plus nombreuses chez JBM que chez EM.

	EM	JBM
voisement	4	7
dévoisement	6	9
mode articuloire	4	3
lieu d'articulation	10	16
total	24	35

Figure 5 : Nombre d'assimilations consonantiques

Ex. de voisement et de dévoisement :

- homorganique : *i savo po faire s'journée* « il ne pouvait pas travailler »¹ avec [s₃>z₃]
- hétéro-organique : *i-a nin d'carreu* « il n'y a pas de carreaux¹¹ »: [dk>tk]

Ex. assimilation de mode articuloire.

EM : *chand qu'j'ai tiré au r'sort* « quand j'ai tiré au sort »¹² : [k₃>g₃>d₃]

Le [g] devient fricatif sous l'influence de la fricative [ʒ].

L'assimilation de lieu articuloire est totale et constante, dans nos corpus, pour l'article défini, proclitique épïcène en picard, ainsi que pour les adjectifs possessifs atones. *Le, me, te* et *se* sont des formes à la fois masculines et féminines: le *schwa* s'élide et leur consonne devient implosive. La consonne absorbante, initiale du mot suivant, est généralement plus longue que la moyenne normalisée en syllabe non accentuée. Ex :

JBM *ch'est de (l') terre* #: [ʃ e dø t:ɛ:R] « c'est de la terre (battue) »

Durée de [t]: 207 ms au lieu de 109, soit 52 % d'allongement car il est initial de syllabe accentuée.

EM *e(ll') dans'minm'* [ɛl 'da:s mem] : le [l] implosif subit l'assimilation totale, le [d] est long car il est à l'initial de syllabe accentuée.

¹⁰ Cette terminologie est empruntée à Georges Straka.

¹¹ Carrelage.

¹² Conscription en militaire.

EM *aussitôt que (j') su révéyi mi* #[osi'to kø s:y eve'i mI] «aussitôt que je suis réveillé, moi »

La comparaison des durées montre que la sifflante est la consonne [t] la plus longue (131 ms)

Mais la particularité de ces parlers est que, dans 26,92% des assimilations totales de *le* et *me, te se*, il n'y a pas d'allongement de la consonne absorbante, quand elle est initiale de syllabe accentuée, donc forte, ni quand c'est une occlusive, ni quand les consonnes en contact sont homorganiques. On peut distinguer trois degrés d'assimilation, selon que la consonne absorbante est:

- initiale de *cluster* accentué : EM *e(m') mère è c'minche à rire* [e'kmẽ:ʃ]. Assimilation totale sans allongement du possessif par la consonne initiale du verbe; [k] dure 107 ms alors que la moyenne des [k] non accentués est de 109 ms; le [m] dure 680 ms (Figure 6) ;
- initiale de syllabe non accentuée: JBM *les sell' i vont su (l') côtéé* [sy ko'te#] « les chaises vont sur le côté¹³ ». Durées consonantiques en ms : [s] 120, [k] 122, [t] 140 : c'est une assimilation sans allongement compensatoire ;
- homorganique et non accentuée : JBM *èn bonne assi(t') de potache* [asi dø po'taʃ] : le [d] dure 95 ms, moins que moyenne des non accentués (98 ms). Absence d'allongement compensatoire.

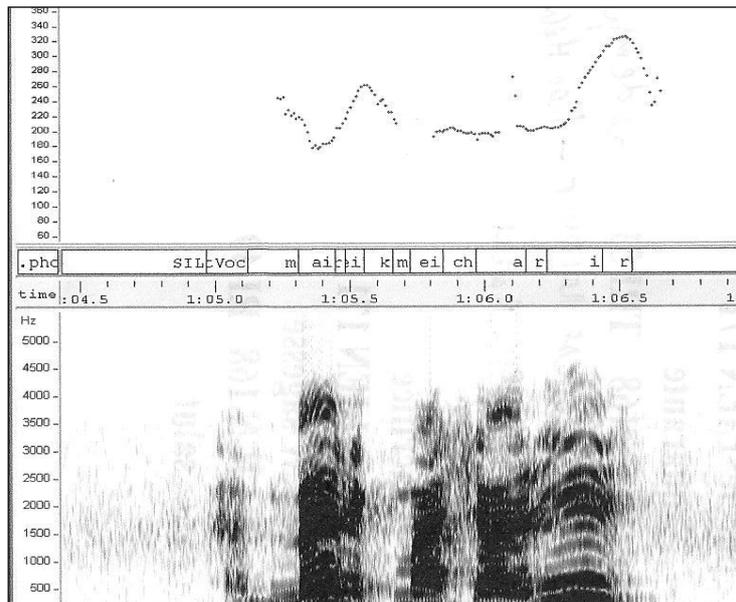


Figure 6 : *e(m') mère è c'minche à rire*. De haut en bas: courbe mélodique, segmentation (notation SAMPA), échelle des durées, spectrogramme.

Cet affaiblissement articulaire est lié, non pas à la rapidité du débit, mais à l'âge des locuteurs et au caractère dialectal du corpus.

¹³ Les chaises sont inclinées car le sol est inégal.

6. Structuration accentuelle

6.1 Accent à fonction démarcative

6.1.1. Groupe accentuel. Un découpage par « groupe de souffle », délimité par des pauses silencieuses, n'est pas congruent avec l'analyse syntaxique dans une grande partie des corpus. Des pauses dues à l'hésitation, à la recherche d'un mot etc. interrompent parfois l'énoncé. Le groupe accentuel est une unité moins aléatoire : il correspond, en général, à une unité syntaxique. Mais ce n'est pas toujours le cas, surtout chez JBM. Nous avons délimité té groupes accentuels en nous fondant sur le fait que l'accent est final de groupe. L'analyse instrumentale montre que celui-ci est réalisé par une proéminence mélodique accompagnée, sauf exception signalée, d'une durée et d'une intensité accrue. Ex:

EM *i n'avo nin d'leumir' par deurir' là t# à ç cabaret cha fait#* « il n'y avait pas de lumière derrière¹⁴ là à ce cabaret alors ».

La courbe mélodique (Figure 7) montre les proéminences de Fo, les allongements des deux [i] accentués et l'absence de proéminence accentuelle sur *là*, adverbe final de groupe, qui porte le contour de finalité. Les proéminences mélodiques permettent d'identifier 7 groupes accentuels qui ont respectivement 3+3+4+4+2+3+2 syllabes.

6.1.2 Enclise. Des morphèmes enclitiques s'intègrent fréquemment aux groupes accentuels (cf. Carton, 1979) – des postfixes en termes macrosegmentaux (Martin, 2009). Ils sont mono- ou dissyllabiques :

- pronoms personnels : *mi* « moi », *ti* « toi », *li* « lui », *eusses* « eux », *nous-autes* « nous autres », *ti-s-autes* « vous autres »,
- pronom démonstratif: *cha* « ça »,
- interjection *hé* (non-nasalisation de *hein*),
- adverbe à valeur d'insistance: *même*,
- particules à fonction discursive: *quo*, « quoi », *là*.

Autre cas d'enclise : [t:] suivi de *schwa* est un morphème verbal usuel en picard du Nord. Il oppose la 6^{ème} personne à la 3^{ème} de l'indicatif présent, imparfait et au conditionnel présent. Il a été généralisé à toutes les conjugaisons (cf. Carton 1967). Ex:

JBM *i-étottent* [e'tɔt:ə] «ils étaient»; EM *i crittent* [i 'krit:ə] «ils crient»; *i dittent* [i 'dit:ə] «ils disent ».

¹⁴ Dans la cour, derrière la salle du cabaret.

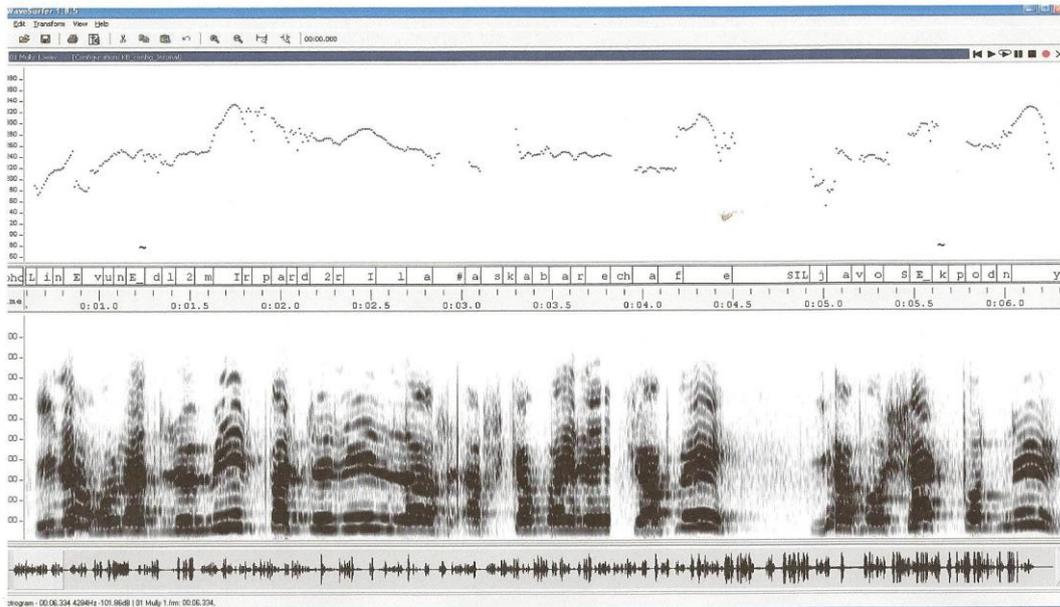


Figure 7 : (par d)erir'là # à ç cabaret cha fait#i-avo chinqu'pots d'nu

Ces enclises ont pour effet de mettre en relief la syllabe accentuée par une rupture mélodique bilatérale.

	EM	JBM
pronoms	6,27	0,38
adverbes	6,69	5,82
morphème verbal en -te	2,6	0,83
Total	15,66	7,03

Figure 8 : Pourcentage de groupes avec enclise

Les différences sont liées au type de discours: les récits d'EM mettent en scène plusieurs acteurs qu'il identifie au fur et à mesure et dont il rapporte les propos. Le discours rapporté en style direct contient de nombreuses incises (*i dit, j'dis* etc.), des marqueurs qui sont autant de groupes accentuels. Il y a aussi l'influence du flamand, mais elle est difficile à établir. Les souvenirs de JBM sont plus descriptifs et moins animés.

6.2 Accent à fonction expressive

La fonction expressive est le plus souvent assurée par les ruptures mélodiques bilatérales qui renforcent les effets de contraste. Les accents expressifs de type dialectal (Carton, 1979) sont aussi réalisés par des allongements vocaliques:

JBM *t'in fais pas #* « ne t'en fais pas ! »

Les voyelles sont toutes plus longues que la moyenne (Figure 3)

EJBM *i-a des grant' punitions qui pnttent d'zeur nou têtes #* » il y a de grandes Punitions¹ qui sont suspendus au-dessus de nos têtes»

Nos corpus ne contiennent que trois accents d'insistance de type français. La répétition des accents initiaux produit un effet d'accumulation. Un accent « émotif » (232 ms) frappe le [k]'initial du terme d'injure dans le troisième, où l'accent initial de lexème est réalisé par la durée consonantique accrue et une forte montée mélodique:

EM *les pinchonneux #:les coulonneux#*« les amateurs de pinsons et de pigeons² (de concours)

EM *cha coulo cha boulo#* « ça coulait, ça bouillait! ».

EM *cré capenouls* « sacrés garnements! »

6.3. Structure rythmique

Le rythme du récit en picard, fondé sur l'accentuation, est plus prégnant que l'organisation syntaxique dans la structuration des groupes accentuels. Le nombre de syllabes par groupe accentuel est un indicateur de la structuration rythmique et un élément de comparaison entre les corpus.

Après application des critères, nous obtenons les données suivantes : EM 239, JBM 223, soit un total de 462 groupes. Après calcul des pourcentages aux fins de comparaison entre les locuteurs, nous reportons ces données sur le graphique (Figure 9): en losanges les réalisations de EM, en carrés celles de JBM. En abscisse figure le nombre de syllabes par groupe, en ordonnée le nombre d'occurrences de chaque type syllabique. Le tableau présente un pic très net pour EM, indice d'une régularité rythmique (ternaire et binaire) qui caractérise un récit souvent raconté. Chez JBM le rythme est parfois syncopé, avec des pauses à l'intérieur des syntagmes, des bégaiements et es hésitations.

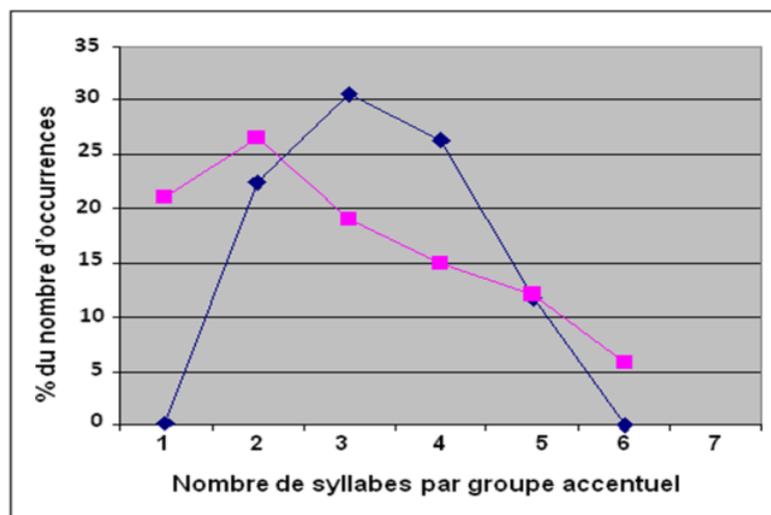


Figure 9 : Comparaison du nombre de syllabes par groupe accentuel

¹ Phénomènes apocalyptiques.

² Concours de pigeons voyageurs et de pinsons siffleurs

Le corpus EM présente trois cas d'alternance faible (*schwa*) et forte, probablement imputables à l'influence flamande: [bʏl'ʒik] « Belgique », [dʏ'zU:r] « dessous », [dʏ'rI:r] « derrière ». La première syllabe de ces mots est relâchée, la seconde est tendue (opposition *tense/lax*).

L'analyse de Martinet du *schwa* comme « lubrifiant phonique » (Martinet, 1960) convient bien pour analyser nos corpus. Le fait qu'il soit fermé par exemple dans [ørpe'rir] « revenir » montre qu'il n'obéit pas à la « loi de position ». C'est un indice sérieux pour le considérer comme un non-phonème.

Groupes accentuels	Traduction
<i>cha fait (l')directeur</i>	alors le directeur
<i>i vnot cachi</i>	venait chercher
<i>après mi#</i>	après moi (venait me chercher).
<i>Batiss' i dit</i>	« dBaptiste,, il dit
<i>va deurm<u>ir</u></i>	va dormir (passer la nuit, remplacer le veilleur de nuit)
<i>én pair' d'heur'</i>	quelques heures
<i>au soir#</i>	le soir »
<i>v'la que [kø:#]</i>	(mais) voilà que...
<i>j'dépasso (l') plafond#</i>	je dépassais le plafond (de rémunération) !
<i>j'dépasso (l') plafond#</i>	je dépassais le plafond !
<i>j'allos trouver (l') directeur</i>	Je suis allé voir le directeur:
<i>vèt' euj' dis</i>	« Regarde, je dis,
<i>j'ai cor èrchu èn let'#</i>	J'ai encore reçu une lettre!
<i>t'in fais pas</i>	- Ne t'en fais pas,
<i>i dit Batisse</i>	il dit, Baptiste!
<i>ch'est moi qui est maîte i dit#</i>	C'est moi le maître, d il dit.

Figure 10: Exemple de segmentation des groupes accentuels chez JBM.

Parmi les groupes de dissyllabiques, les marqueurs à fonction discursive (*cha fait* « cela fait (que) », *v'là que* « voilà que » etc.) sont particulièrement nombreux. Dans la séquence analysée Figure10, on remarque que 6 groupes accentuels sur 15, soit 40%, ne correspondent pas à des syntagmes.

6.4. Effets phonétiques de l'accentuation

6.4.1. Crase. La nasale accentuée absorbe la non-accentuée. Ex:

EM [va 't ε n fo 'vi:r] « va-t-en un'fois³ voir ».

Le [ε] accentué dure 140 ms, seulement 50 ms de plus que [a] et [o] : absence d'allongement.

6.4.2. Insertion d'un segment vocalique transitoire en finale libre accentuée (Straka, 1959 : 298) : 9 occurrences de ce phénomène, 3 chez JBM, 6 chez EM :

EM : *non?* ['nɔ] : question expressive

Montée du Fo de 160 à 360 Hz (Figure 15);

JBM *spuper* [su'p^øe] « souper »

EM [ko'd^øɔ̃] « *Caudron* », nom d'une ruelle (Figure 11).

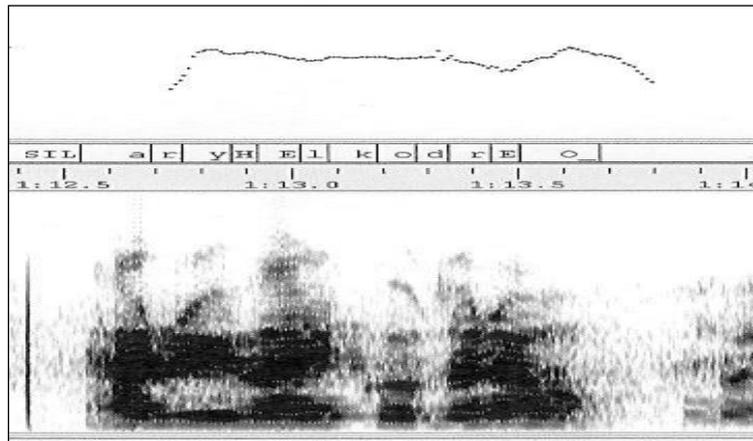


Figure 11 : Diphthongaison de *Caudron* sous accent lexical

6.4.3. Basculement accentuel La prééminence qui affecte le second élément de la diphtongue a tendance à basculer sur le premier élément. Le lexème est alors accentué sur la pénultième, en conformité avec les règles d'accentuation de mot en néerlandais et en West-Vlaams (flamand occidental), dont les diphtongues sont descendantes (cf. Booij, 1995). Ex:

EM [po't^o>po'te^o] « poteau, copain ».

Le basculement peut aboutir à l'amuissement du second élément. Ce phénomène est sporadique chez EM, alors qu'il est constant à Linselles⁴. Ex¹ [r^eɛ̃] · [ʔre] « rien » (Figure 15).

³ Belgicisme, (calque du néerlandais).

⁴ Point 15 de l'*Atlas linguistique picard* (Carton & Lebègue, 1989, 1998), au nord de Tourcoing.

7. Structures intonatives

Nous avons travaillé sur les « mots prosodiques » (Martin, 2009 : 95-111). Nous n'avons pas pu appliquer la représentation par niveaux (Léon, 1996), à cause de l'instabilité du niveau de base (correspondant au fondamental « usuel »). Elle est liée au grand âge des locuteurs et à la nature des corpus.

7.1. Fonctions discursives

7.1.1. Continuation. C'est le contour de loin le plus fréquent: chez EM 83,2%, chez JBM 90,1%. Chez EM il y a davantage de dialogues rapportés en style direct; le récit stéréotypé est plus vivant et plus drôle. Il rit lui-même de ses propos scatologiques. La distinction majeure/mineure permet de hiérarchiser les séquences de type paratactique (Carton, 1972). L'accent lexical sur la dernière syllabe comporte souvent une variation mélodique concave caractéristique, comme celle qui figure parmi les intonations régionales notées dans (Martin, 2009)⁵. Dans le narratif, EM monte régulièrement à la fin des groupes, pour entretenir l'intérêt, et la courbe mélodique a la forme d'un « S couché », ce qui est réalisation typique de la demande d'approbation en français standardisé (Martin, 2009). Ex:

EM [j eto ma:'lat] « il était malade » (Figure 12).

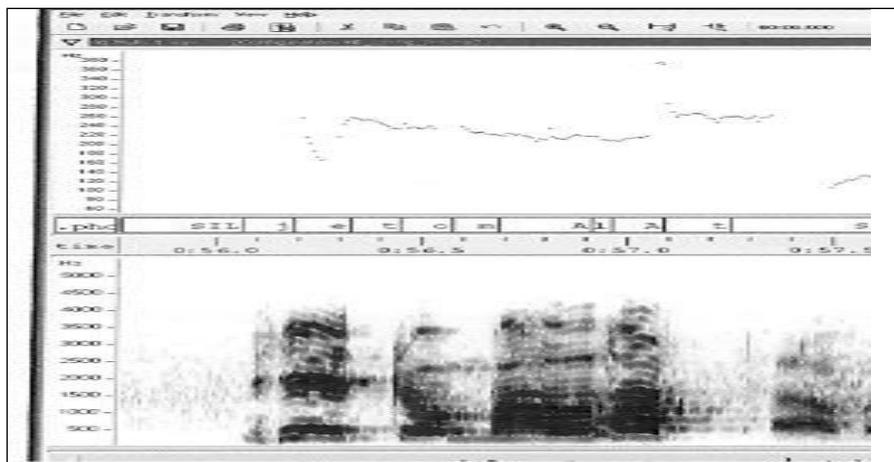


Figure 12 : *i-éto malate*: allongement pénultième et rupture mélodique. (cf. 5.1).

Même concavité mélodique dans EM *du jalap*⁶, favorisée par une durée accrue de la pénultième (cf. 5.1.).

Un des corpus contient une séquence de 6 groupes qui se termine par une continuation majeure alors qu'on attend une finalité. La cause est l'ellipse du verbe : EM *j'éto tchurieux pour aller vir el*

⁵ Page 224: Picardie. Il s'agit en fait de l'enregistrement d'un ouvrier retraité du textile à Roubaix, ville limitrophe de Tourcoing, recueilli par nous (Carton, Rossi, Léon, Auteserre, 1983).

⁶ Produit purgatif de cette époque.

cabarétir' tcheul'grimace # « j'étais curieux d'aller voir la cabaretière quelle grimace⁷ (elle faisait) ». Montée de 190 à 280 Hz sur la dernière syllabe (Fig.16).

JBM *chinqu'siz-afnats* « cinq six enfants »: allongement et descente pré-accentuelles (durée de [a] 179 ms, de [ã] 148 ms).

7.1.2. Finalité. Le contour descendant est rare dans les corpus : 4,91% des contours chez JBM, 13,82% chez EM. La pente mélodique est nettement plus faible que celle de la continuation (Ex Figure 7). Le contour de finalité ressemble à celui de la parenthèse basse. Ex:

JBM *j'condujo l'ouv^urach' à les incaisseus'* #« je conduisais l'ouvrage⁸ aux encaisseuses »

Le contour descendant ou plat de finalité ne se trouve chez EM que dans les enclises Ex:

EM *cha fait i-eur^uèn't eusses* « alors ils reviennent, eux »

7.1.3. Parenthèse. Les incises n'ont pas un contour spécifique. Dans les dialogues rapportés au style direct, es incises *i dit comm'cha, ej'dis, elle dit à mi* etc. (cf. 6.1.3) identifient le locuteur ou permettent de préciser les modalités et/ou les circonstances de l'élocution. Contrairement à l'apostrophe (7.2.3), Les parenthèses sont « basses », elles peuvent comporter plusieurs groupes et ont un contour légèrement descendant. Ex:

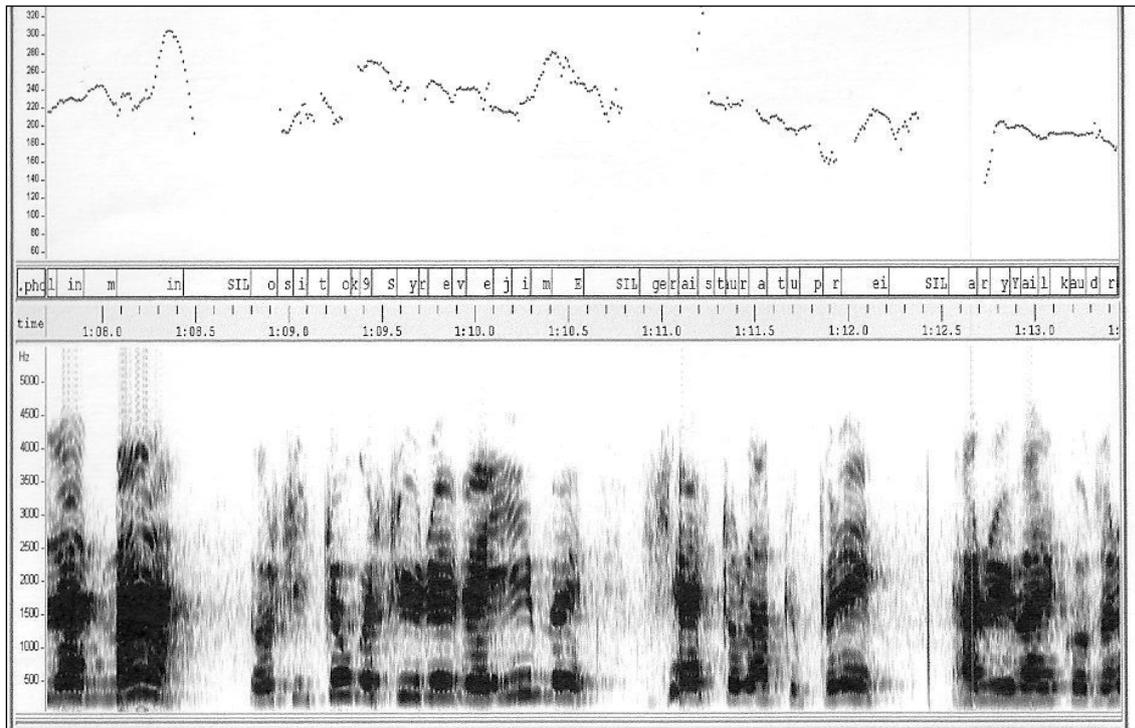


Figure 13 : aussitôt que j'su révéyi mi – j'restos là tout près...

⁷ « Quelle tête elle faisait ».

⁸ Les grosses bobines de fil. *Ouvrier signifie* « travailler ».

EM (*el lin*)'min'# *ussitôt que j'su révéyi mi # j'restos là tout près # à (ruelle Cudron « le lendemain, aussitôt que je suis réveillé, moi – j'habitais là tout près, dans la ruelle Caudron »* La parenthèse est en pente descendante de 220 à 180 Hz (Figure13).

7.2. Fonctions modales

7.2.1. Interrogation. Le contour est toujours fortement montant du fait de l'expressivité inhérente aux dialogues en style direct. De ce fait, nous n'observons pas de différence entre question totale et partielle (cf. Delattre, 1966).

- question « partielle » avec pronom:

EM *uo qu'i a Marie #* « qu'est-ce qu'il y a, Marie? »

La montée va de 200 à 300 Hz et le mot mis en apostrophe redescend légèrement à partir du sommet.

- question « totale » :

EM *t'as nin acor' vu c'té. c'l' 'étwal'rouch' là #* « tu n'as pas encore vu cette étoile rouge là? ».

EM a d'abord employé le démonstratif français puis le picard (*ceie*). Les proéminences mélodiques sont de plus en plus élevées et marquent une croissance de l'émotion: EM passe de 180 (attaque haute) à 360 Hz (suraigu). L'adverbe est légèrement descendant (Figure 14).

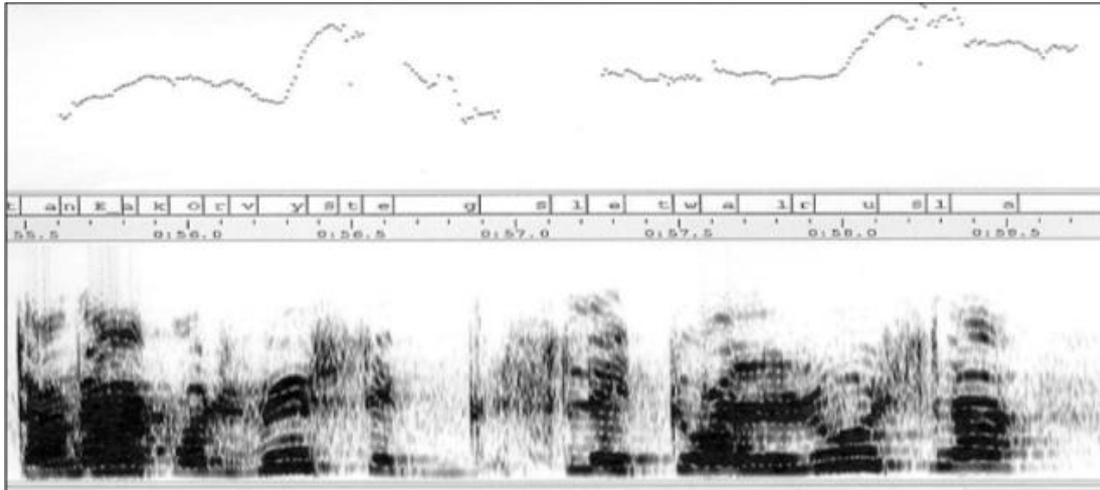


Figure 14 : question totale avec enclise

7.2.2. Impératif. L'ordre est toujours réalisé par un contour montant, dont l'amplitude est fonction du degré d'expressivité.

JBM *vèt ej dis #* « regarde! dis-je »: contour montant de 50 Hz du [ε]

EM *va-t-in 'n'fo vir tin garchon là#* « va-t-en un' fois voir ton garçon là! »: montée en escalier : les voyelles accentuées sont de plus en plus hautes (220-300-340 Hz pour un fondamental usuel de 140 Hz).

7.2.3. Apostrophe. La fonction d'appel est réalisée, comme en français, par une brève et importante montée mélodique de la voyelle, affectée d'une forte intensité. Ex;

EM *Louiss' Louiss'#* « Louise, Louise! ». Les [i] montent à 340 Hz.

Le nom propre mis en apostrophe est légèrement descendante à partir de la proéminence de [a], cf. 7.2.1. Ex:

EM *quo qu'i a Marie #* « qu'est-ce qu'il y a, Marie? »

7.3 .Fonction expressive

La fin des récits est très expressive Ex:

EM *qu_o sais ré du tout mi # neon#* « quoi? (je ne) sais rien du tout, moi!-Non? »(Figure 15)

EM *quo! J'sais ré du tout, mi* « je ne sais rien du tout, moi ».

Dans l'expression de surprise, l'enclitique est à la même hauteur que la voyelle accentuée (380 Hz). EM mime la réaction incrédule de la cabaretière : *Non?* atteint aussi le suraigu, avec diptongaison de l'interjection (Figure 15).

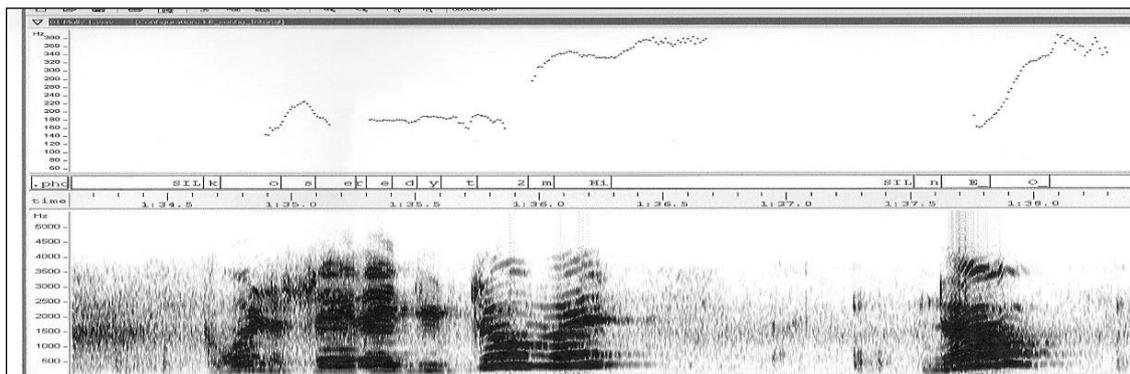


Figure 15 : *quo! J'sais ré du tout, mi* : exclamation et Enclise

La figure 16 et le tableau 17 sont des exemples de l'analyse intonative: segmentation, interprétation.

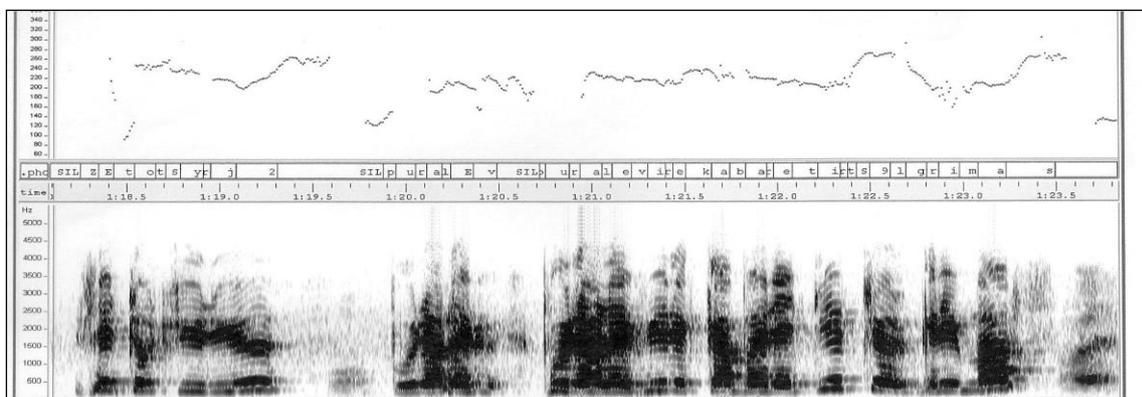


Figure 16 : *j'éto's tchurieux pour aller ...vir el cabarétir' tcheul'grimace*

Mots prosodiques	Traduction	Intonation
<i>el'lind'main#</i>	le lendemain	continuation majeure
<i>premi ouv<u>r</u>ach'#</i>	premier travail	continuation majeure
<i>j'éto<u>s</u> tchurieu#¹</i>	j'étais curieux	continuation mineure
<i>pour aller</i>	d'aller	[bégalement]
<i>pour aller <u>vir</u></i>	d'aller voir	continuation mineure
<i>eu(l')cabar<u>é</u>tire</i>	la cabaretère	continuation majeure
<i>tcheul' grim<u>a</u>ce#²</i>	quelle grimace	continuation majeure
<i>V'là in rin<u>tr</u>ant</i>	voilà qu'à mon arrivée	continuation majeure
<i>cré cap<u>é</u>noul'</i>	sacrés garnements	exclamation
<i>e ll' dit à <u>mi</u>#</i>	Me dit-elle	parenthèse basse
<i>quo qu'i <u>a</u></i>	qu'est-ce qu'il y a	question partielle
<i>Marie#</i>	Marie	apostrophe
<i><u>ouais</u>#</i>	oui	question totale
<i>te n'<u>dos</u> nin d'<u>mand</u>er</i>	tu ne dois pas demander	continuation mineure
<i>quo qu'i <u>a</u></i>	ce qu'il y a	ordre
<i>te sais bin <u>quo</u>#</i>	tu sais bien ce qui s'est passé	exclamation

Figure 17: Exemple d'analyse intonative

8. Bilan

Nous avons relevé dans les corpus 9 traits picards ou flamands qui ne se trouvent pas dans le français standardisé:

- allongements vocaliques en pénultième de groupe accentuel,
- diérèse quasi constante,
- assimilations anticipantes totales fréquentes, parfois sans allongement de la consonne absorbante,
- rareté des contours de finalité,
- enclises très nombreuses,
- pas de coïncidence régulière entre groupe accentuel et syntagme,
- diphtongaisons sporadiques de voyelles accentuées
- réductions et amuïssements de voyelles non accentuées,
- variation mélodique concave fréquente sur la voyelle accentuée.

¹ Début du spectrogramme.

² Fin du spectrogramme.

Références

- Booij G. (1995). *The Phonology of Dutch*. Oxford, Claredon Press.
- Carton F. (1977). Insistance dialectale: l'accent d'insistance dans les dialectes d'oïl. *Studia Phonetica*, 12. pp.59-92.
- Carton F. (2012). Un parler picard au contact du flamand occidental: étude de phonologie fonctionnelle. In: Le foctionnalisme en France, SIL/Université de Leon (Espagne).
- Carton F. & M. Lebègue (1989,1998). *Atlas linguistique et ethnographique picard*, Vol I et II, Coll. Atlas linguistiques de France par régions, Paris, Editions de CNRS.
- Carton F., Rossi M., Auteserre D. & P. Léon (1983). *Les accents des français*, Paris, Hachette.
- Carton F., (1977). L'identification régionale par l'intonation", *Bulletin d'audiophonologie*, Besançon, 100, pp. 185-196.
- Carton F (1979). L'accentuation dans le français dialectal du Nord de la France. *Studia Phonetica*, 15, pp. 66-92.
- Carton F., et coll. (1982). Appréciation de la voix des personnes âgées. Sénescence de la relation/Aging and Communication, *Bulletin d'audiophonologie*, pp. 149-163.
- Delattre P. (1966). *Studies in French and comparative Phonetics*, La Haye.
- Leon P. (1971). *Essais de phonostylistique*. *Studia Phonetica* 4.
- Leon P. (1996). *Phonétisme et prononciations du français. Avec des travaux pratiques 'application et leurs corrigés*. Paris, Nathan Université.
- Martin Ph. (2009). *Intonation du français*. Paris, Armand Colin.
- Martinet A. (1960). *Eléments de linguistique générale*. Paris, Armand Colin.
- Straka G. (1959). Durée et timbre vocalique. *Zeitschrift fürPhonetik und allgemeine Sprachwissenschaft*, 12, pp. 276-300.

STRUCTURE PROSODIQUE
ET THÉORIE DE L'INTONATION

Le contraste pente peut-il s'analyser comme une manifestation d'OCP

Elisabeth Delais-Roussarie et Brechtje Post

elisabeth.roussarie@wanadoo.fr, bmbp2@cam.ac.uk

UMR 7110-Laboratoire de Linguistique Formelle, Université Paris-Diderot

Department of Theoretical and Applied Linguistics, University of Cambridge

Abstract

In French, continuation is normally signaled by a rising tonal contour realized at the right edge of non-final prosodic phrases at all levels of structure (e.g. minor vs. major continuation according to Delattre's terminology, H* for accentual phrases AP, H*H- for intermediate phrase *ip*, and H*H% for intonational phrase IP¹). In some cases, however, continuation contours at the end of non-final prosodic constituents are realized as a fall. These falling realizations have been mentioned in the literature (see Delattre 1966; Martin 1975 among others). It is not clear, however, in which contexts they occur and what triggers them, especially when they appear to be obligatory. In this paper, we review the different contexts in which these falling realizations occur, before exploring to what extent the Obligatory Contour Principle (OCP) – which prohibits two tonal specifications of the same type consecutively in surface phonological forms – can be invoked to account for these falling realizations. A crucial question to address is whether the OCP could be held to account for French facts. Our analysis shows that there are at least three distinct cases of falling continuations which cannot be straightforwardly accounted for in a unified fashion with reference to a single mechanism such as the OCP.

1. Introduction

Dans de nombreux travaux consacrés à l'intonation du français, il est généralement admis que la frontière droite des constituants prosodiques non-terminaux est indiquée par la réalisation d'un contour mélodique montant (*cf.*, entre autres, Delattre 1966; Di Cristo 1998; Jun & Fougeron 2000; Post 2000, Delais-Roussarie *et al.* 2015.). L'ampleur de ces mouvements montants est souvent décrite comme liée à la force relative de la frontière prosodique (voir, par exemple, l'opposition entre les continuations mineures et majeures chez Delattre 1966). Ainsi, par exemple, dans la séquence sous (1), la distribution et l'ampleur relative des mouvements mélodiques montants sont cruciales pour interpréter le SN ajout circonstant *la semaine dernière*. Si la montée mélodique réalisée sur la syllabe finale de *voisine* est plus importante que celle réalisée sur la syllabe finale de *dernière*, le SN ajout est rattaché à la séquence qui suit, à savoir *il était en vacances* (1a). En revanche, si la syllabe /jɛʁ/ reçoit un mouvement mélodique montant plus ample que la syllabe /zin/, l'ajout est à interpréter avec la première partie de la séquence *Pierre a rencontré le fils de ma voisine* (1b).

¹ H*, H- et H% correspondent à des symboles utilisés dans le cadre métrique-autosegmental pour noter une montée mélodique se produisant respectivement sur la syllabe accentuée du syntagme accentuel (AP), à la frontière droite du syntagme intermédiaire, ou à la frontière droite du syntagme intonatif (*cf.*, pour plus de détails, Jun & Fougeron 2000 et Delais *et al.* 2015)

- (1) Pierre a rencontré le fils de ma voisine la semaine dernière il était en vacances.
 a. Pierre a rencontré le fils de ma voisine ; la semaine dernière, il était en vacances.

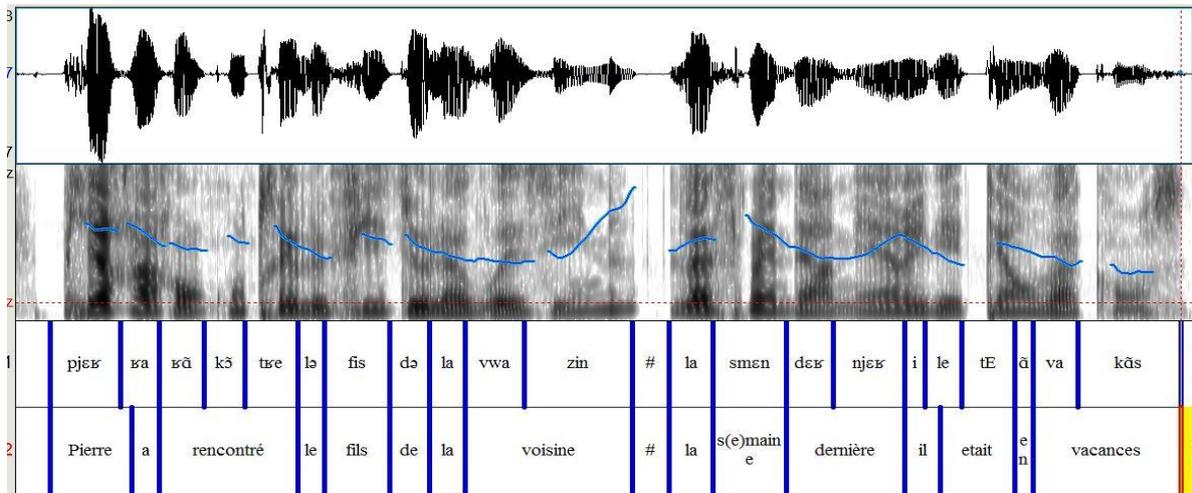


Fig. 1 : Signal, spectrogramme et courbe mélodique associés à l'énoncé (1a) *Pierre a rencontré le fils de ma voisine ; la semaine dernière, il était en vacances.*

- b. Pierre a rencontré le fils de ma voisine la semaine dernière ; il était en vacances.

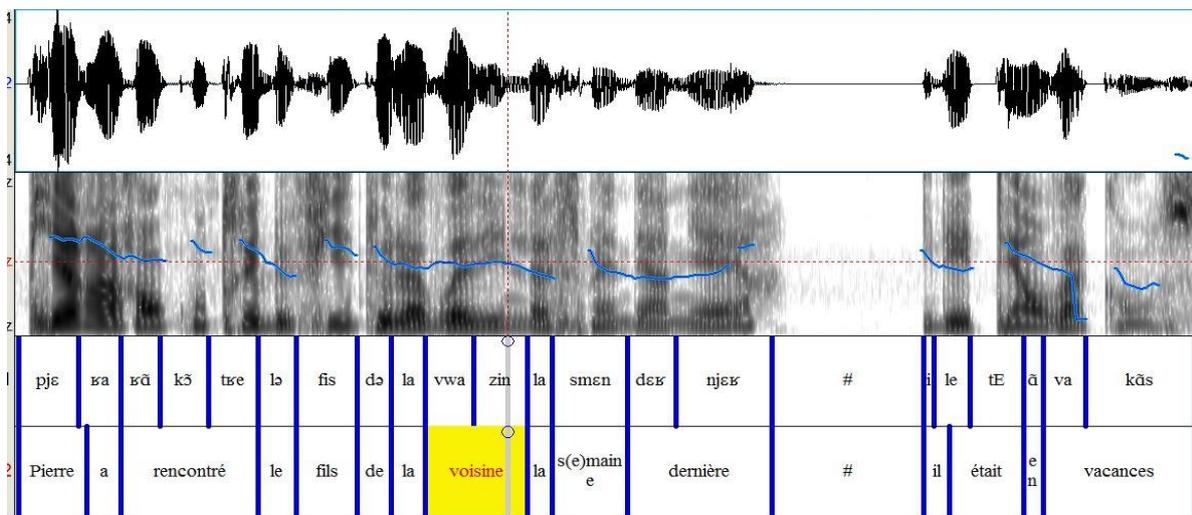


Fig. 2 : Signal, spectrogramme et courbe mélodique associés à l'énoncé (1b) *Pierre a rencontré le fils de ma voisine la semaine dernière ; il était en vacances.*

Malgré la fréquente réalisation de contours montants à la frontière droite des groupes intonatifs non-terminaux (groupe accentuel GA, syntagme intermédiaire et syntagme intonatif), plusieurs auteurs ont montré que des contours mélodiques non-montants sont également utilisés (cf. Delattre 1966 ; Martin 1975, 1981, 1987 *et seq.*). Ainsi, par exemple, Delattre (1966) insiste sur le fait que, dans certains contextes non discriminants, les contours de continuation mineure peuvent être descendants :

« De toutes les courbes d'intonation que nous analysons ici, celle de la continuation mineure est la seule qui n'ait pas une direction fixe. Elle peut descendre aussi bien que monter lorsqu'elle précède une courbe à niveau plus élevé; par exemple, lorsqu'elle précède la continuation majeure (4), la question (4+) ou l'implication (4-). (...) Cette descente se fait simplement, semble-t-il, pour briser la monotonie, pour donner de la variété à la ligne mélodique; elle n'a aucune fonction significative. » Delattre 1966 :10

Selon Delattre, le rôle des continuations montantes (en particulier majeures) est de regrouper des groupes prosodiques en un groupe plus large². Ainsi, pour (1), le contour montant permet de regrouper les groupes accentuels (Pierre) (a rencontré) (le fils de ma voisine) sous (1a), et (Pierre) (a rencontré) (le fils de ma voisine) et (la semaine dernière) sous (1b). Cette analyse vaut également pour les exemples sous (2) et (3), bien que les mouvements descendants apparaissent à différents niveaux dans la structure prosodique. Sous (2), le mouvement mélodique à la frontière droite de *déchargement* peut être réalisé descendant, l'important étant que celui à la fin de *casiers* permette de regrouper *le déchargement* et *des casiers*. De la même façon, sous (3), la subordonnée *quand je reste trop longtemps à Carcassonne* doit être regroupée avec *j'en ai ras le bol*, et une continuation majeure doit être réalisée à la fin de *bol*. La réalisation mélodique est en revanche descendante à la fin de la subordonnée temporelle, comme on le voit dans la figure 4.

- (2) Il réglait le déchargement des casiers sur les chariots des mareyeurs [corpus FR_EUROM 1, Chan et al. 1995]

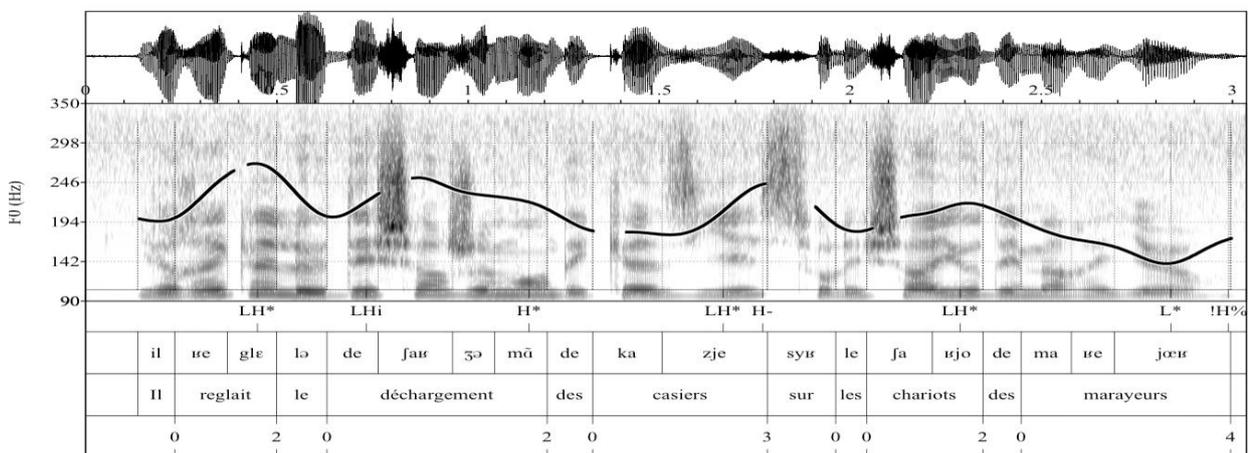


Fig. 3 : Signal, spectrogramme et courbe mélodique associés à l'énoncé (2) *Il réglait le déchargement des casiers sur les chariots des mareyeurs.*

- (3) disons que j'aime bien avoir et la ville et la campagne si j'avais **quand je quand je reste trop longtemps à Carcassonne j'en ai ras le bol** et quand je reste trop longtemps à Toulouse euh je j'en ai assez aussi [Acsynt, JASE 5]

² cf. Delattre 1966, p.10 : « Le rôle du niveau 4 est clairement de réunir de petites unités de sens en une grande unité de sens qui n'est pas la dernière de la phrase. D'où la terminologie de continuation mineure et continuation majeure. Ici, l'intonation de continuation majeure (4) fait sentir que toutes les unités mineures qui la précèdent, et qui ont l'intonation de continuation mineure (3), appartiennent à la grande unité de sens qui se termine au niveau 4 ».

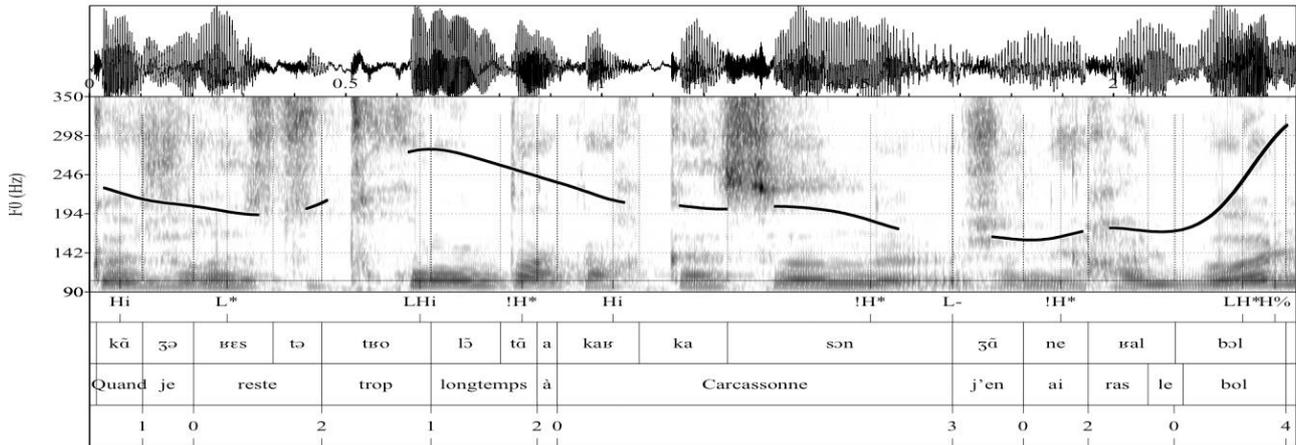


Fig. 4 : Signal, spectrogramme et courbe mélodique associés à la séquence (3) *quand je quand je reste trop longtemps à Carcassonne j'en ai ras le bol*.

Dans ses travaux, Philippe Martin a essayé de rendre compte de ces réalisations par le mécanisme de contraste de pente CDP (Martin 1981, 1987 *et seq*), insistant du même coup sur le fait que l'occurrence de ces formes est liée à l'intonotactique (*cf.* Duběda, ce volume). Il nous semble cependant que l'apparition de ces formes trouve leur explication dans des causes diverses, relevant en cela d'une certaine opacité. Notre objectif dans cet article sera donc :

- de dresser un inventaire des contextes où ces contours descendants non-terminaux apparaissent;
- de proposer des explications pour justifier ces réalisations;
- de voir si le *mécanisme de contraste de pente* (CDP) est suffisamment adéquat pour en rendre compte. Cela nous amènera à réfléchir à d'autres principes explicatifs, comme notamment le *Principe du contour obligatoire* (OCP) (*cf.*, entre autres, Leben 1973 et Yip 1988).

Notre contribution sera structurée comme suit. Dans un premier temps, nous expliquons en quoi consistent le mécanisme de contraste de pente (CDP) et le principe du contour obligatoire (OCP). Ensuite, nous dresserons, à partir d'exemples précis extraits de différents styles de parole, un inventaire des contextes prosodiques dans lesquels des contours descendants non-terminaux apparaissent. A partir de là, nous nous demanderons en discutant des résultats obtenus comment le mécanisme de contraste de pente et le principe d'OCP pourraient rendre compte des faits. Cela nous permettra de pondérer la portée du mécanisme de contraste de pente, notamment au vu de la multiplicité des contextes.

2. Cadres théoriques et mécanismes tonaux

Dans cette section seront présentés les mécanismes et principes qui peuvent être utilisés pour rendre compte des alternances de formes tonales entre contours montants et contours descendants. Nous nous limiterons essentiellement au mécanisme du contraste de pente et au principe du contour obligatoire. Comme la mise en œuvre de ces mécanismes est souvent liée à la structuration prosodique, nous aborderons ce point dans la section 2.2. Cela nous permettra d'ailleurs de mieux

distinguer les points de convergences et de divergences entre les deux mécanismes, et surtout cela fournira des indications sur le cadre prosodique que nous utiliserons pour décrire les faits dans la section 3.

2.1 Alternance et dissimilation de formes tonales

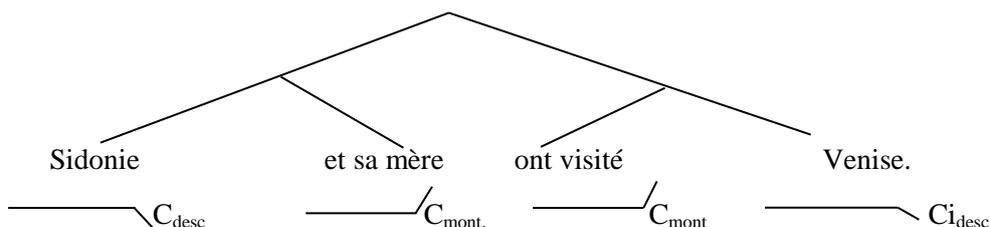
2.1.1 Martin et le contraste de pente. Le modèle de l'intonation du français développé par Philippe Martin prend comme unité de base le mot prosodique, et propose de générer le profil intonatif associé à un énoncé à partir de 3 éléments clefs :

- la forme intonative associée au dernier mot prosodique de l'énoncé est déterminée par la modalité illocutoire de l'énoncé. Un contour descendant $C_{i_{desc}}$ est ainsi associé aux assertions (4a), et un contour montant $C_{i_{mont}}$ aux questions (4c).
- la structure prosodique, c'est-à-dire le marquage des relations de dépendance et de sélection à droite entre les différents mots prosodiques de l'énoncé, découle de la structure syntaxique de surface comme indiqué sous (5).
- la forme des contours peut être vue comme un marqueur des relations de dépendance entre les mots prosodiques. Ainsi deux mots prosodiques entrant dans une relation de dépendance recevront généralement des contours de formes inverses (contraste de pente mélodique) :

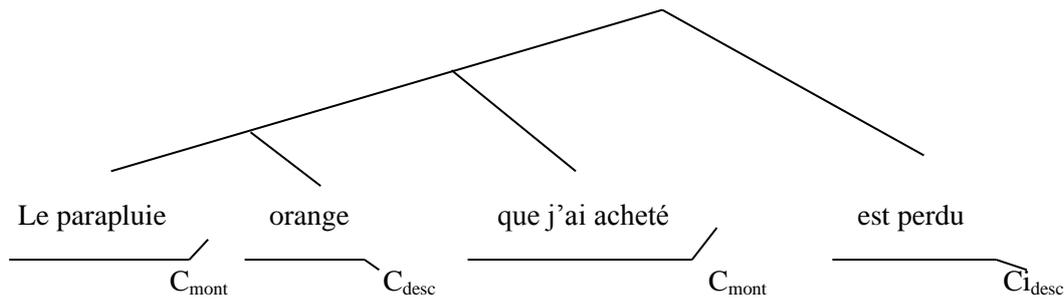
In French, syntactic relations are very often indicated by morphological markers such as the number and gender of the adjective, the number and person of the verb, etc. Similarly, if the existence of a prosodic structure is assumed, prosodic markers must exist to indicate the relations existing in the sentence between prosodic words. These markers can be found on the stressed syllable of the prosodic word and function merely through contrasts of melodic slope between the contours of the prosodic words involved in the relation of selection. Martin 1987:929

A partir de ces éléments, les énoncés sous (4)a-d seront réalisés intonativement à partir des profils donnés sous (5)a-c.

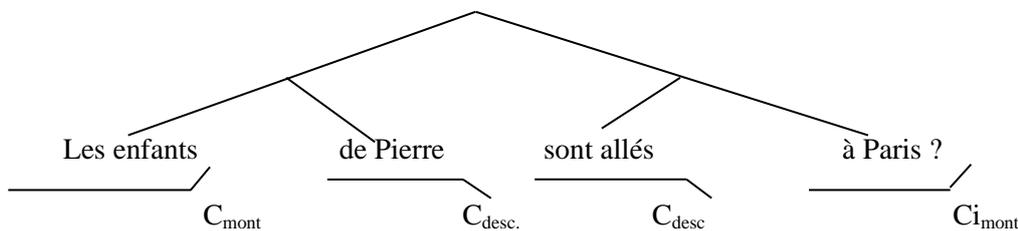
- (4) a. Sidonie et sa mère ont visité Venise (d'après Martin 1987 : 942)
 b. Le parapluie orange que j'ai acheté est perdu (d'après Martin 1987 : 943)
 c. Les enfants de Pierre sont allés à Paris ?
 d. Les enfants de François sont partis.
- (5) a. (Sidonie) (et sa mère) (ont visité) (Venise.)



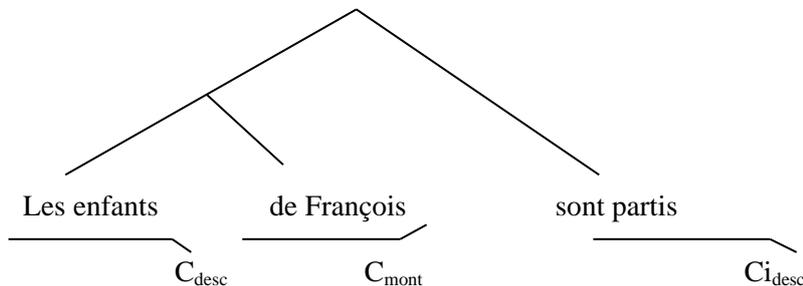
b. (Le parapluie) (orange) (que j'ai acheté) (est perdu.)



c. (Les enfants) (de Pierre) (sont allés) (à Paris) ?



d. (Les enfants) (de François) (sont partis)



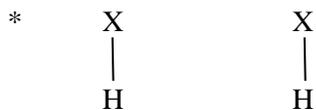
L'étude des exemples montre clairement qu'un contour intonatif non-montant n'apparaît en position non-terminale que dans des contextes où la séquence de même niveau à sa droite dans la structure prosodique s'achève par un contour montant. Ainsi, comme les mots prosodiques (à sa mère) (5a), (que j'ai acheté) (5b), (à Paris) (5c) et (de François) (5d) s'achèvent par des contours montants, les mots prosodiques qu'ils sélectionnent à leur gauche peuvent être réalisés avec un contour descendant, à savoir (Sidonie) (5a), (orange) (5b), (de Pierre) (5c) et (les enfants) (5d).

Le mécanisme de contraste de pente s'opère en tenant compte de la structuration prosodique, et non à partir de la seule suite linéaire des mots prosodiques. Aussi obtenons-nous des énoncés où les contours montants et descendants alternent toujours en surface (*cf.* 5b et 5d), mais cette réalisation découle entièrement de la structure syntactico-prosodique.

2.1.2 Le principe du contour obligatoire : OCP. A côté du mécanisme de contraste de pente qui rend compte du fait que deux unités entretenant une relation de sélection dans la structure prosodique doivent s'achever par un contour prosodique de forme contrastée (ou inverse), un autre principe a été développé en phonologie pour rendre compte de l'alternance entre formes tonales : le principe du contour obligatoire (OCP).

Ce principe, considéré comme universel, a initialement été formulé par Leben (1973) dans le cadre de la phonologie autosegmentale, et plus précisément en tonologie. Il rend compte du fait que, dans les langues à tons, deux tons identiques ne pouvaient être associés à deux syllabes/voyelles qui se suivent (l'un des deux tons étant alors effacé). Cette impossibilité est représentée sous (6).

(6) OCP : principe du contour obligatoire et représentation bannie



Ceci étant, ce principe a ensuite été étendu à d'autres unités associées aux positions squelettales, tels les traits articulatoires, mais aussi à l'accentuation, l'alternance rythmique pouvant être vue comme une instance d'OCP (*cf.*, pour des exemples précis, Yip 1988). Ainsi, on peut considérer qu'OCP est à l'œuvre dans les phénomènes suivants:

- le refus de gémination en japonais;
- la non-application du Rendaku en japonais si cela conduit à avoir plus de deux obstruents voisés dans un même morphème.

- (7) a. iro+ koni => irogoni (papier de couleur)
 b. yu+toofu => yudoofu (tofu bouilli)
 c. mais kami-kaze => kamikaze (vent divin)

De nombreuses règles phonologiques résultent du principe d'OCP. De fait, ce principe contraint l'application de règles qui permettent de corriger des formes intermédiaires ou de surface qui contiendraient deux unités semblables dans un même domaine. Ceci étant, le principe d'OCP et les règles qui en résultent s'appliquent selon les langues dans des domaines particuliers.

I have suggested that the OCP can trigger several different kinds of phonological rule in which the environment crucially includes two identical elements. I shall now make precise the way in which the OCP triggers phonological rules. The main contribution of the OCP is that it allows us to separate out condition and cure. The OCP is a trigger, a pressure for change, but the cure must be specified by the language (...). I assume that rules can only apply if four things are specified: a Domain (syllable, word); a Tier; a Trigger; and a Change. Yip 88 : 74

En faisant d'OCP un principe qui contraint les formes soit au niveau lexical, soit lors de la dérivation, on laisse l'espace pour une formulation rigoureuse des règles qui impose de savoir si ces dernières sont facultatives ou obligatoires, et quel est leur domaine d'application.

Selon nous, explorer les phénomènes traités comme relevant du contraste de pente peut apporter un éclairage nouveau. Pour ce faire, il est cependant essentiel de bien mettre en parallèle les domaines et structures prosodiques utilisés dans les deux cas : les travaux de Philippe Martin, et les travaux sur l'OCP en phonologie prosodique.

2.2 De la structure prosodique à l'application de mécanismes tonaux

Le mécanisme du contraste de pente tout comme les règles résultant du principe OCP s'appliquent en tenant compte de la structure prosodique. Dans le cadre d'OCP, les domaines d'application correspondent soit à des unités de la structure prosodique soit à des morphèmes. Lorsqu'il s'agit d'unités prosodiques, elles sont généralement définie dans le cadre de la théorie Prosodique (*cf.*, entre autres, Selkirk 1978 *et seq.*, Nespor & Vogel 1986, Frota 2012). En général, les niveaux retenus pour la structure prosodique correspondent à ceux sous (8), les équivalences entre modèles étant prises en compte (*cf.* aussi Frota 2012).

(8) *Les constituants de la structure prosodique*

Syntagme intonatif

Syntagme phonologique / Syntagme mineur ou syntagme intermédiaire

Syntagme phonologique / Syntagme mineur ou syntagme accentuel (AP)

Mot prosodique

Pied

Syllabe

En outre, la structure prosodique dans le cadre de la Théorie Prosodique (*cf.* Selkirk 1978, 1986 entre autres) possède plusieurs caractéristiques:

- une fois que les constituants prosodiques sont construits, généralement à partir d'informations morpho-syntaxiques, l'application des règles phonologiques se fait dans ces domaines, sans tenir compte de leur structuration prosodique interne ou de la structure syntaxique. Ainsi, par exemple, en Bengali, le principe OCP impose en surface au sein des syntagmes intonatifs (IP) une alternance de tons H et L (c'est-à-dire de contours montants/ hauts et descendants/ bas), que ceux-ci soient associés à des syllabes et correspondent à des accents mélodiques ou qu'ils soient réalisés à la fin de groupes prosodiques de niveau inférieur. En aucun cas, en tous cas, l'application du principe d'OCP ne tient compte de la structure interne du domaine dans lequel il s'applique, donc le syntagme intonatif dans le cas du Bengali (*cf.* Hayes & Lahiri 1991).
- la structure prosodique est généralement plus plate que la structure syntaxique, notamment du fait du nombre limité de constituants de niveau supérieur au mot prosodique, à savoir le syntagme phonologique mineur (qui correspond d'ailleurs en gros au mot prosodique chez Martin (1975 *et seq.*)), le syntagme phonologique majeur

et le syntagme intonatif.

- Dans certains travaux, la structure prosodique doit aussi respecter le principe de l'étagement strict si bien que des structures comme (5b) et (5d) pourraient poser problème car on ne voit pas toujours à quoi correspondent, en termes d'unités prosodiques, les séquences comme (le parapluie orange), (que j'ai acheté), (le parapluie orange que j'ai acheté), etc.

Pour formuler le mécanisme CDP, en revanche, Martin (1975 *et seq*) ne fait pas appel à des constituants particuliers, dont la formation serait dérivée de la structure syntaxique. Il prend plutôt comme approximation de la structure prosodique une structure syntaxique simplifiée dont les nœuds terminaux correspondraient à des mots lexicaux précédés des mots grammaticaux qui en dépendent (pour une structure prosodique du même ordre, *cf.* Dell 1984). La formulation du mécanisme de contraste dépend alors des dépendances dans l'arbre, tout comme l'assignation des degrés d'accent chez Dell (1984).

Pour tenter d'appréhender si le contraste est une instance d'OCP, il est nécessaire d'étudier son application dans des constituants de la structure prosodique. C'est ce que nous ferons dans la section suivante. Pour cela, nous utiliserons des unités de la théorie prosodique (Selkirk, 1978 *et seq.*), dont la définition s'appuie sur les modalités d'appariement entre les structures prosodique et syntaxique suivantes :

- le syntagme phonologique (ou groupe accentuel AP) résulte de l'application conjointe de contraintes métriques (sur la taille minimale et maximale) et de la contrainte d'alignement (X_{head} , right), ces dernières étant cependant plus importantes (*cf.* Delais-Roussarie, 1996)
- le syntagme intonatif résulte de l'alignement avec des constituants syntaxiques apparaissant dans des positions linéaires particulières (dislocation, incidents, fin de clauses, etc.) (*cf.* Delais & Post, 2008, en particulier).

En outre, lors de la présentation des faits, nous nous poserons au moins deux questions fondamentales:

- A quel niveau, l'inversion de pente s'applique-t-elle ? Est-elle analysable dans un domaine/ constituant prosodique particulier ?
- A un niveau donné, le mécanisme d'inversion de pente est-il obligatoire ou non ?

3. Les faits et leur analyse

Le contraste de pente tel que formulé par Martin (1975 *et seq*) est un mécanisme puissant et élégant qui rend compte de certains aspects du phénomène de dissimilation tonale, mais qui, selon nous, ne permet pas de faire des distinctions entre les différents cas observés. De fait, les formes de surface alternantes obtenues relèvent, d'après nous, d'une certaine opacité, et ne peuvent donc pas être analysées comme issues d'un unique mécanisme. Aussi, pour présenter les faits allons-nous faire une distinction entre trois catégories. En outre, nous nous demanderons à chaque fois :

- Si le mécanisme de contraste s'applique de façon obligatoire ou optionnel ;

- Si les faits sont analysables comme relevant d'OCP à un niveau donné ;
- Si le phénomène est phonologique, ou simplement lié à l'implémentation phonétique ;
- Si l'application du contraste ou sa non-application a des conséquences sur l'interprétation discusivo-pragmatique de la séquence.

3.1 Dissimilation tonale, contraste de pente et syntagme phonologique

Dans ses écrits, Philippe Martin propose plusieurs cas où le mécanisme CDP s'applique entre deux syntagmes phonologiques (noté PhP dans les exemples). Dans les exemples sous (9), un contour descendant est ainsi réalisé à la fin d'un syntagme phonologique non terminal comme en témoignent les réalisations de *la sœur* (9a), *nos voisins* (9b), *le déchargement* (9c) et *avec l'introduction* (9d).

(9) Contraste de pente et contour L à la fin des syntagmes phonologiques non-terminaux

a. La sœur de Paul s'en va → [(la sœur)_{PhP} L* (de Paul)_{PhP} H* (s'en va)_{PhP} L%]

b. Nos voisins chinois repartent demain →
[(Nos voisins)_{PhP} L* (chinois)_{PhP} H* (repartent demain)_{PhP} L%]

c. Il réglait le déchargement des casiers sur les chariots des Mareyeurs [Eurom 1] (cf. (2), fig.3)→

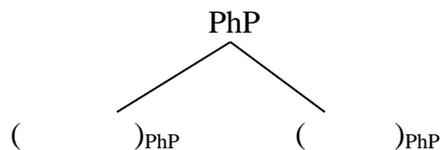
[(il réglait)_{PhP} H* (le déchargement)_{PhP} L* (des casiers)_{PhP} H* (sur les chariots)_{PhP} H* (des mareyeurs)_{PhP} L%]

d. Avec l'introduction de l'euro, nous nous trouvons au cœur d'une évolution de même ampleur [*Discours de Jacques Chirac, bicentenaire de la Banque de France*] →

[(Avec l'introduction)_{PhP} L* (de l'euro)_{PhP} H-], {(nous nous trouvons)_{PhP} H* (au cœur d'une évolution)_{PhP} H* (de même ampleur)_{PhP} } L%]

Dans ces exemples, les syntagmes phonologiques recevant un contour non-montant L à leur frontière droite partagent tous une même caractéristique : ils sont intimement liés au syntagme prosodique qui les suit immédiatement à droite. En ce sens, on peut dire qu'ils forment avec eux une sorte de syntagme phonologique composé dont la structure est donnée sous (10).

(10) Syntagme phonologique composé



Dans les structures de ce type, résultant souvent d'une très forte relation morpho-syntaxique entre les deux syntagmes liés (nom et complément du nom en (9a), (9c) et (9d), nom et adjectif le modifiant à sa droite en (9b), etc.), le premier élément s'achève fréquemment par un mouvement mélodique non-montant, le mouvement montant étant à la fin du second conjoint.

En plus de cette caractéristique structurale, il est important de noter que le mouvement mélodique associé au premier conjoint doit simplement être d'ampleur moindre que celui du second conjoint, cela permettant de bien marquer le rapprochement entre les deux unités (*cf.* Delattre 1966 et Martin 1987). La réalisation intonative varie cependant en fonction de la taille des éléments en présence et du débit. En outre, à un débit très lent, elle peut ne pas être observée. Dès lors, ce phénomène est davantage facultatif.

Comme cette restructuration et ce changement tonal ne sont pas sans rappeler les phénomènes de rétraction et/ou d'effacement d'accent qu'on peut observer dans les mots composés, on peut presque analyser ce cas comme une désaccentuation partielle ou totale, qui conduit à l'absence de réelle cible tonale à la fin du premier conjoint. Sur le plan de l'interprétation, elle favorise la cohésion du groupe, au même titre d'ailleurs que les phénomènes d'arc accentuels (*cf.* Fonagy 1979 et Di Cristo 1999 entre autres).

Ce mécanisme en revanche est difficilement analysable comme relevant du Principe du Contour Obligatoire. Au moins deux raisons distinctes le montrent :

- La réalisation non-montante n'est pas provoquée par la présence d'un autre ton haut qui pourrait par exemple être associé à un accent initial (*cf.* pour un exemple d'accent initial, la montée initiale sur *déchargement*, fig. 3). De fait, cette configuration tonale s'observe à la fin de syntagmes phonologiques courts qui ne contiennent pas d'autre mouvement mélodique (*la sœur* et *nos voisins* sous (9)). Le mécanisme CDP n'est pas borné au syntagme phonologique.
- On ne peut pas non plus considérer que ce mécanisme est automatique dès que plusieurs syntagmes phonologiques sont regroupés dans un constituant plus large, tel le syntagme intermédiaire. Il viserait alors à interdire des successions de mouvement montant H* au sein d'un constituant de niveau supérieur. Si on regarde attentivement le second syntagme intermédiaire de l'énoncé (9d), on s'aperçoit que deux groupes accentuels qui se suivent au sein d'un même *ip* s'achèvent par un contour montant : (nous nous trouvons) et (au cœur d'une évolution). Dès lors, le principe d'OCP ne peut pas être invoqué pour expliquer le contraste observé à la fin de certains syntagmes phonologiques.

3.2 Dissimilation tonale, contraste de pente et syntagme intonatif IP

A la fin des syntagmes intermédiaires et des syntagmes intonatifs en position non-finale d'énoncé, on observe généralement un contour de continuation montant (H*H- ou H*H%), comme en témoignent l'exemple sous (11).

(11) [(Quand on habitait)_{PhP} H* (à Marseille)_{PhP} H-], {(mon fils)_{PhP} H* (faisait de l'accordéon)_{PhP} L%}.

Dans certains cas, cependant, il arrive qu'un contour non-montant soit réalisé à la fin des syntagmes intermédiaires. Ainsi, sous (12), un contour plat ou un contour descendant est souvent

réalisé à la fin du premier syntagme intermédiaire { *quand les enfants sont en vacances* }, { *chaque fois que les enfants sont à Toulouse* } ou { *quand je quand je reste trop longtemps à Carcassonne* }.

(12) a. Quand les enfants sont en vacances, vous vous reposez ?

→ [{(Quand les enfants)_{PhP} H* (sont en vacances)_{PhP} L- } { (vous vous reposez ?)_{PhP} } H%]

b. Chaque fois que les enfants sont à Toulouse, leurs parents leur manquent, mais s'ils sont à Bordeaux, ils s'ennuient aussi.

→ [{(chaque fois)_{PhP} !H* (que les enfants) H* (sont à Toulouse) L- } { (leurs parents)_{PhP} H* (leur manquent)_{PhP} H- } {(mais s'ils sont à Bordeaux)_{PhP} H* } (ils s'ennuient aussi)_{PhP} } L%]

c. quand je quand je reste trop longtemps à Carcassonne j'en ai ras le bol et quand je reste trop longtemps à Toulouse euh je j'en ai assez aussi (cf. fig. 4 et exemple (3))

→ [{(quand je quand je reste)_{PhP} L* (trop longtemps)_{PhP} !H* (à Carcassonne)_{PhP} L- } {(j'en ai ras le bol)_{PhP} H- } {(et quand je reste)_{PhP} H* (trop longtemps)_{PhP} H* (à Toulouse) H- } {(euh je j'en ai assez aussi)_{PhP} } L%]

A première vue, ces exemples pourraient être analysés comme des instances d'OCP. Le principe du contour obligatoire interdirait la succession de contours montants de continuation à la fin de syntagmes intermédiaires regroupés dans une même IP. Ainsi, sous (12a), le contour réalisé à la fin de l'*ip* initial serait non-montant puisque le second et dernier *ip* de l'énoncé serait de la forme H*H%.

Ceci étant, une telle analyse est loin d'être aussi valable pour les exemples (12b) et (12c). Des contours non-montants sont observés à la fin du premier *ip* de l'énoncé (*Chaque fois que les enfants sont à Toulouse* et *quand je quand je reste trop longtemps à Carcassonne*), et des contours montants à la fin du second (*leurs parents leur manquent* et *j'en ai ras le bol*), mais ces deux *ip* sont suivis par des *ip* qui s'achèvent également par un contour montant (*mais s'ils sont à Bordeaux* et *et quand je reste trop longtemps à Toulouse* respectivement). Pour avoir en surface à une succession de contour montant et descendant, il faudrait que les premiers *ip* des énoncés se terminent par un contour montant H*H-, et les seconds par un contour descendant T*L-, les troisièmes étant bornés par des contours montants. Pour rendre compte de la réalisation observée, l'explication la plus valable réside dans une analyse comparable à celle proposée au niveau des syntagmes phonologiques : si deux *ip* forment un *ip* composé, essentiellement pour des raisons de dépendances syntaxiques, le premier *ip* doit s'achever par un contour mélodique de forme inverse à celui réalisé à la fin du second *ip*.

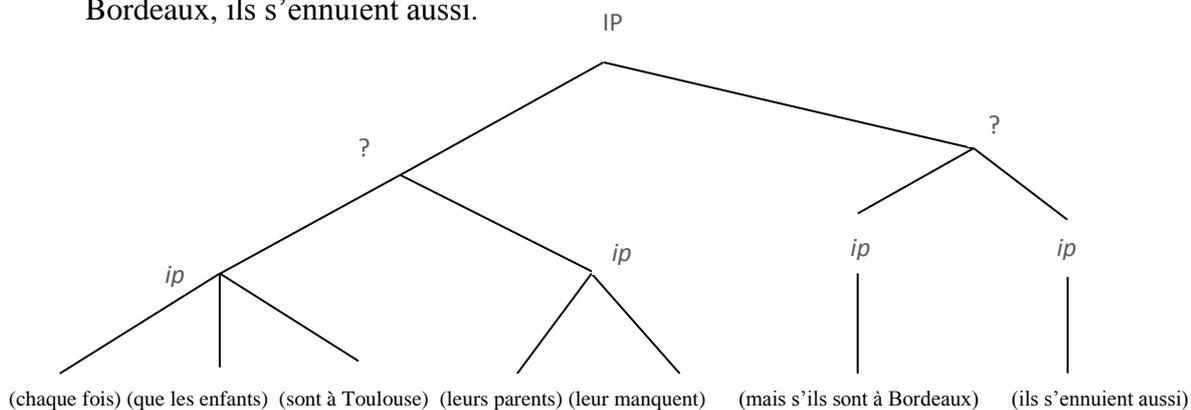
(13) Syntagme intermédiaire composé et règle de réalisation



L'application de cette règle, étrangement proche du mécanisme CDP, permet d'expliquer pourquoi sous (12b) et (12c) le syntagme intermédiaire initial s'achève par un contour descendant. De fait, les syntagmes intermédiaires *chaque fois que les enfants sont à Toulouse* et *quand je quand je reste trop longtemps à Carcassonne* forment un syntagme intermédiaire composé avec les *ip* qui se trouvent à leur droite : $\{(Chaque\ fois\ que\ les\ enfants\ sont\ à\ Toulouse)_{ip}\ (leurs\ parents\ leur\ manquent)_{ip}\}$ et $\{(quand\ je\ quand\ je\ reste\ trop\ longtemps\ à\ Carcassonne)_{ip}\ (j'en\ ai\ ras\ le\ bol)_{ip}\}$. Comme ces syntagmes intermédiaires (composés) se situent en position non-finale, ils s'achèvent par un contour montant H*H- si bien que le premier conjoint se termine par un contour mélodique descendant !H*L-.

Le recours à ces syntagmes composés offre la possibilité de créer localement des niveaux de structuration supplémentaires, voire des structures récursives. Si on reprend les exemples sous (12b) et (12c), on voit que les constituants analysables comme des syntagmes intermédiaires composés ne correspondent à aucun niveau de constituance prosodique dans le cadre de la théorie prosodique (cf. (8)).

(14) Chaque fois que les enfants sont à Toulouse, leurs parents leur manquent, mais s'ils sont à Bordeaux, ils s'ennuient aussi.



D'une manière générale, les réalisations non-montantes observées à la fin des syntagmes intermédiaires non-terminaux relèvent de deux cas distincts et ne peuvent donc être considérées comme relevant d'un unique phénomène/ mécanisme tel le principe du contour obligatoire :

- Lorsque l'énoncé s'achève par un contour montant H*H%, le syntagme intermédiaire en position initiale qui précède peut être réalisé avec un contour non-montant, cela permettant de mieux marquer la distinction entre les deux contours (cf. exemple (12a)). Ceci dit, cette réalisation n'est pas forcément obligatoire, et résulte d'une volonté de renforcement de la dissimilation.
- Lorsqu'un syntagme intermédiaire forme avec le syntagme adjacent un syntagme intermédiaire composé, le mouvement mélodique associé à chacun des conjoints est de forme inverse³ (cf. exemples (12b) et (12c), règle (13)).

³ Cette analyse suppose qu'au moins trois niveaux de structuration prosodique soient considérés comme pertinents pour rendre compte des faits du français : le syntagme phonologique (ou groupe accentuel), le syntagme intermédiaire et le syntagme intonatif. Notons néanmoins que, même si l'existence du syntagme phonologique et du syntagme intonatif est communément admise, celle du syntagme intermédiaire est plus discutée (cf. Michelas 2011 ; Michelas & D'Imperio 2012)

3.3 Contraste de pente et énumérations

Il existe une autre catégorie de constructions syntactico-prosodiques dans lesquelles des contours terminaux non-montants sont fréquemment observés : les énumérations ou listes. Dans la littérature, il n'est pas rare d'assigner aux termes non-terminaux de l'énumération un contour montant, que chacun des termes correspondent à un syntagme phonologique (c'est-à-dire à une unité syntactico-prosodique minimale) comme sous (15a) ou à un syntagme intonatif (ou intermédiaire) comme sous (15b).

(15) a. Marie, François, Frédéric et Annie sont venus.

(Marie) (François) (Frédéric) (et Annie) (sont venus)

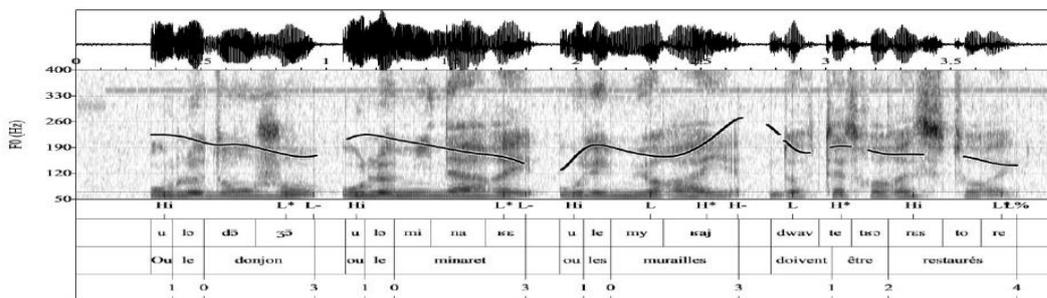
H* H* H* H* L %

b. les voisins sont arrivés, ils ont salué tout le monde, et ils sont rentrés.

[{(les voisins)_{PhP} H* (sont arrivés)_{PhP} H-} {(ils ont salué)_{PhP} H*(tout le monde)_{PhP} H-} {(et ils sont rentrés)_{PhP} }L%]

Cependant, dans certains cas, les conjoints non-terminaux de l'énumération peuvent s'achever par un contour non-montant comme sous (16)

(16) Ou le donjon ou le minaret ou les murailles doivent être restaurés.



D'une manière générale, l'utilisation de contours descendants à la fin des conjoints non-terminaux renforce l'idée de liste, en donnant l'impression que chaque élément est ajouté. Même si des recherches sont nécessaires pour vérifier jusqu'à quel point l'effet de liste est renforcé par l'utilisation de contours non-montants, on ne peut pas comparer ces cas avec les précédents. En aucun cas il ne s'agit d'un phénomène tonal guidé par le mécanisme de CDP ou par le principe d'OCP, mais d'un choix de contour délibéré à des fins particulières.

Dans certaines énumérations, en revanche, lorsque les conjoints sont construits à partir de deux syntagmes intermédiaires pour en former un troisième coordonné, le premier conjoint peut recevoir un contour non-descendant, comme dans les exemples (12b) et (12c). Dans ce cas, on a affaire à l'application de la règle sous (13). L'exemple sous (17) renvoie à un tel cas.

- (17) [{{{(je me levais)_{PhP} (le matin)_{PhP} L- }_{ip} {(j'étais avec des clients)_{PhP}}_{ip} H- }_{ip-coordonné} {{{(je mangeais)_{PhP} (à midi)_{PhP} L- }_{ip} {(j'étais avec des clients)_{PhP}}_{ip} H- }_{ip-coordonné} {{{(et je me couchais)_{PhP} (le soir)_{PhP} H- }_{ip} {(j'étais avec des clients)_{PhP}}_{ip} L- }_{ip-coordonné} L%].

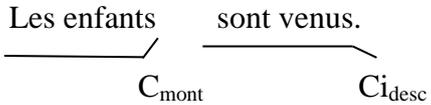
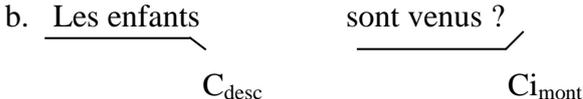
4. Discussion et conclusion

L'étude des différents cas dans lesquels des contours non-montants peuvent être réalisés à la fin des groupes prosodiques non-terminaux permet de distinguer au minimum trois cas, et force donc à ne pas traiter les faits comme relevant d'un unique mécanisme.

Dans le cas des énumérations (*cf.* § 3.3), réaliser un contour non-montant à la fin des conjoints non-terminaux comme sous (16) vise à renforcer le sentiment de liste. Ce type de réalisation peut s'observer aussi bien au niveau du syntagme phonologique que du syntagme intermédiaire (ou intonatif). Selon nous, des études complémentaires seraient intéressantes car elles permettraient d'évaluer plus finement l'impact des différentes réalisations.

Dans les autres contextes, il semble important de distinguer les cas où on a affaire à une réelle conjonction de constituants prosodiques (*cf.* par exemples (5a-d) et (9a-d) pour des conjonctions de syntagmes phonologiques et (17) pour des conjonctions de syntagmes intermédiaires) des cas comme (12a).

Philippe Martin utilise le mécanisme du contraste de pente aussi bien pour rendre compte des cas de conjonctions que pour les différences de réalisation tonale entre un SN sujet (montant) et un SV (descendant) dans une assertion (18a) ou une question (18b).

- (18) a. Les enfants sont venus.

 b. Les enfants sont venus ?


Selon nous, cependant, les réalisations sous (18) relèvent des règles générales de l'intonation du français, et ne peuvent donc pas être vues comme découlant d'un mécanisme tonal particulier. En (18b), où le SN sujet peut s'achever par un contour non-montant (comparable à ce qui est observé en (12a) au niveau du syntagme intermédiaire), la présence de réalisations non-montantes s'explique par une volonté de bien distinguer le contour non-terminal du contour terminal montant. D'après nos observations, même si des études plus systématiques seraient bienvenues, ces réalisations sont davantage liées à l'implémentation phonétique et seraient sensibles à des éléments comme la taille des constituants, le débit, etc. Néanmoins, il n'est pas possible de les considérer comme relevant d'un mécanisme phonologique tonal comme OCP, ou même le mécanisme CDP.

Reste alors les cas de compositions de constituants prosodiques. Ils constituent pour nous les seules situations où un mécanisme tonal peut être invoqué. Des différences existent cependant selon que la composition se situe au niveau du syntagme phonologique ou du syntagme intermédiaire. Dans les cas de syntagmes phonologiques composés, l'absence de contour mélodique montant à la fin du premier conjoint peut en effet s'analyser comme résultant d'une désaccentuation partielle,

voire d'un déplacement d'accent avec renforcement de l'accent initial (*cf.* exemple (2)). Il semble d'ailleurs que ce phénomène soit facultatif et sensible à des phénomènes tels le débit. Par contraste, au niveau du syntagme intermédiaire, la formation d'un groupe prosodique composé conduit presque systématiquement à la réalisation d'un contour descendant à la fin du premier conjoint. De plus, comme ce dernier s'accompagne d'un allongement de la syllabe finale du groupe, il ne peut pas être analysé comme une désaccentuation partielle. Des études plus systématiques seraient cependant nécessaires pour mieux comprendre les modalités d'implémentation phonétique de ce contour :

- L'importance du mouvement descendant est-elle déterminée par la taille du conjoint ?
- La forme des contours environnants joue-t-elle un rôle dans le calcul de la forme mélodique des contours associés aux frontières droites de chacun des conjoints ?
- La structure interne des syntagmes qui entrent dans la formation des syntagmes conjoints intervient-elle dans la réalisation du contour non-montant à la frontière droite du premier conjoint ?

Bien que des études plus systématiques soient nécessaires, les observations de données que nous avons faites laissent penser que la réalisation d'un contour non-montant à la fin du premier conjoint soit un indicateur de la coordination. En ce sens, l'analyse proposée par Philippe Martin (1987) va dans le bon sens, mais gagne à être restreinte à la simple coordination, et non à un marquage généralisé de la dépendance (*cf.* Martin 1987 et § 2.1.1). En tous cas, une analyse en ayant recours au Principe du Contour Obligatoire n'irait pas vers plus de transparence, mais impliquerait de mieux classer les différents contextes.

L'étude proposée ici ouvre des perspectives très intéressantes. De fait, elle impose de mener une réelle réflexion sur la structure prosodique et sur le nombre de niveaux de structuration. Comme l'indiquent les exemples (14) et (17), la création d'un syntagme coordonné, en particulier au niveau du syntagme intermédiaire, va à l'encontre de l'hypothèse de l'étagement stricte (*cf.*, entre autres, Nespor & Vogel 1986 et Selkirk 1986) puisqu'elle conduit à créer des structures récursives ou à ne plus limiter le nombre de niveaux. Des études supplémentaires sont nécessaires pour comparer ces cas de coordinations avec des cas analysés comment engendrant parfois des structures récursives (constructions disloquées, etc.). Cela permettra en effet de mieux appréhender comment ils doivent être analysés en termes de structuration prosodique. S'il s'avère que la création de niveau de structuration supplémentaire est justifiée, il faudra alors s'interroger sur les modalités de ses ajouts :

- S'agit-il d'un phénomène global ou local ?
- Les restructurations s'accompagnent-elles de neutralisation dans le marquage intonatif des niveaux de frontière prosodique ?

Nous envisageons donc de mener des travaux sur ce point aussi bien à partir de l'analyse de corpus que d'expériences de production ou de perception. De plus, il semble important de voir si un tel mécanisme se retrouve dans d'autres langues, et comment il se manifeste dans l'intonation.

References

- Chan, Dominic; Fourcin, Adrian; Gibbon, Dafydd; Granström, Bjorn; Hucvale, Mark; Kokkinakis, George; Kvale, Knut; Lamel, Lori; Lindberg, Borge; Moreno, Asuncion; Mouropoulos, Jiannis; Senia, Franco; Trancoso, Isabel; Veld, Cor in't; and Zeiliger, Jerome (1995), EUROM: A Spoken Language Resource for the EU, in *Proceedings of the 4th European Conference on Speech Communication and Speech Technology*, Eurospeech 1995. Madrid. vol. 1, pp. 867-870.
- Delais-Roussarie, Elisabeth (1996), Phonological Phrasing and Accentuation in French, in: Marina Nespor & Norval Smith (eds.), *Dam Phonology (HIL Phonology Papers II)*, Den Haag: Holland Academic Graphics, pp. 1–38.
- Delais-Roussarie, Elisabeth & Post, Brechtje (2008), Unités prosodiques et grammaire de l'intonation : vers une nouvelle approche, in *Actes des XXVIIème Journées d'Etudes sur la Parole (JEPTALN 2008)*, Avignon, Juin 2008
- Delais-Roussarie, Elisabeth ; Post, Brechtje ; Avanzi, Mathieu ; Buthke, Carolin ; Di Cristo, Albert ; Feldhausen, Ingo ; Jun, Sun-Ah ; Martin, Philippe ; Meisenburg, Trudel ; Rialland, Annie ; Sichel-Bazin, Rafèu et Yoo, Hiyon (2015), Intonational Phonology of French. Developing a ToBi system for French, in: Sonia Frota & Pilar Prieto (eds.), *Intonation in Romance*, Oxford University Press, pp. 63-100.
- Delattre, Pierre (1966), *Les dix intonations de base en français*, French Review 40, pp. 1–14.
- Dell, François (1984), L'accentuation dans les phrases en français, in : François Dell, Daniel Hirst & Jean-Roger Vergnaud (eds), *Formes sonores du langage*, Hermann, Paris, pp. 65 122.
- Di Cristo, Albert (1998), Intonation in French, in: Daniel Hirst & Albert di Cristo (eds.), *Intonation Systems. A Survey of Twenty Languages*, Cambridge: Cambridge University Press, pp. 88–103.
- Di Cristo, Albert (1999), Le cadre accentuel du français contemporain, Première partie dans *Langues 2-3*: 184-205. Deuxième partie dans *Langues 2-4*: 258-267.
- Duběda, Tomáš (ce volume), Pour une « intonotactique » des événements mélodiques.
- Fonagy, Ivan (1979), L'accent français : accent probabilitaire, l'accent du français contemporain, in *Studia Phonetica* 15 : 123-233.
- Frota, Sónia (2012), Prosodic structure, constituents and their representations, in: Abigail Cohn, Cécile Fougeron & Mary Huffman (eds), *The Oxford Handbook of Laboratory Phonology*, Oxford, Oxford University Press, pp. 255–265.
- Hayes, Bruce & Lahiri, Aditi (1991), Bengali Intonational Phonology, *Natural Language and Linguistic Theory* 9:1, pp. 47-96.
- Jun, Sun-Ah & Fougeron, Cécile (2000), A phonological model of French intonation, in: Antonis Botinis (ed.), *Intonation: Analysis, Modelling and Technology*, Dordrecht: Kluwer, 209–242.
- Leben, William (1973), *Suprasegmental Phonology*, PhD Thesis, MIT
- Martin, Philippe (1975), Analyse phonologique de la phrase française, *Linguistics* 146: 35 68.
- Martin, Philippe (1981), Vers une théorie syntaxique de l'intonation, in : Mario Rossi et al. (eds): *Intonation: de l'acoustique à la sémantique*, Klincksieck, 1981
- Martin, Philippe (1987), Prosodic and Rhythmic structures in French, *Linguistics* 25-5: 925-949.
- Michelas, Amandine (2011), *Caractérisation phonétique et phonologique du syntagme intermédiaire en français: de la production à la perception*, PhD thesis, Université d'Aix-Marseille I.

- Michelas, Amandine, & D'Imperio, Mariapaola (2012), When syntax meets prosody: Tonal and duration variability in French Accentual Phrases, in *Journal of Phonetics* 40(6): 816-829.
- Nespor, Marina & Vogel, Irene (1986), *Prosodic phonology*, Dordrecht: Foris.
- Post, Brechtje (2000), *Tonal and phrasal structures in French intonation*. Den Haag: Holland Academic Graphics.
- Selkirk, Elisabeth (1978/1981), On prosodic structure and its relation to syntactic structure, in: Thorstein Fretheim (ed.), *Nordic Prosody II: Papers from a symposium*, Trondheim, TAPIR, 111–140.
- Selkirk, Elisabeth (1986), *On derived domains in sentence phonology*, in *Phonology Yearbook* 3, 371–405.
- Yip, M. (1988), The Obligatory Contour Principle and Phonological Rules: A Loss of Identity, in *Linguistic Inquiry* 19:1, pp. 65 100.

Pour une « intonotactique » des événements mélodiques

Tomáš Duběda

dubeda@ff.cuni.cz

Université Charles de Prague, Institut de traductologie

Abstract

Some intonation grammars do not impose any constraints on the linear arrangement of melodic events, while others (including that of Ph. Martin) subject sentence intonation to a system of rules or trends which can be termed as “intonotactic”. We perform a contrastive analysis of intonation in French and Czech with the aim to detect and quantify some of these regularities. The aspects examined are: distribution of pitch accents in the nucleus and in the prenuclear part; effect of position within the intonation phrase; effect of phrase length; tendency to make the prenuclear accents identical; effect of nuclear contour type; effect of syntactic tree structure.

1. Introduction

L'un des paramètres qui permettent de caractériser les modèles intonatifs existants et de les comparer entre eux est le rôle que ceux-ci confèrent aux contraintes syntagmatiques. En appliquant ce critère, on peut dégager, *grosso modo*, les cinq types suivants :

- **Modèles non-contraints** : Pierrehumbert (1980:29) propose une grammaire de l'intonation anglaise où toutes les combinaisons linéaires d'accents mélodiques sont possibles et où le répertoire des accents prénucléaires est identique à celui des accents nucléaires.
- **Modèles « à deux inventaires »** : Gussenhoven (2004:275) propose une description de l'intonation anglaise à travers un inventaire de 5 accents prénucléaires et 8 accents nucléaires. La division de l'unité intonative (UI) en deux parties – nucléaire et prénucléaire –, chacune faisant appel à un inventaire particulier d'événements tonaux, représente un degré élémentaire (et largement accepté) d'application des contraintes syntagmatiques.
- **Modèles « assimilatoires »** : Ladd (1996:211) avance l'hypothèse que les accents prénucléaires tendent à être identiques dans une UI anglaise, et Gussenhoven (2004:266) propose la même règle pour le français. Cette hypothèse ajoute à la division nucléaire/prénucléaire une contrainte syntagmatique supplémentaire, réduisant fortement les possibilités combinatoires des accents prénucléaires.
- **Modèles « probabilistes »** : Dainora (2006) étudie la probabilité de transition entre les différents événements mélodiques dans l'UI. Dans un esprit similaire, Duběda (2010a) analyse sur un échantillon tchèque la distribution des accents prénucléaires en fonction de plusieurs variables, dont la position dans l'UI et la longueur de l'UI. Ces conceptions, opérant au niveau formel sans tenir compte de la syntaxe, tâchent de caractériser la combinatoire des événements tonaux à l'aide de méthodes statistiques.
- **Modèles à orientation syntaxique** : Martin (1975) propose pour le français un modèle où la sélection des contours dans l'énoncé dépend du contour final ; l'assignation des mouvements mélodiques se fait par le jeu du contraste de la pente, appliqué à la structure prosodique (qui

entretient des relations logiques avec la syntaxe). Il s'agit à notre connaissance de la tentative la plus poussée d'intégrer la composante syntagmatique dans un modèle prosodique ; elle a trouvé un écho chez d'autres auteurs, p. ex. Le Gac & Yoo (2007).

La liste embrasse ainsi cinq visions de l'intonation, qui vont du degré zéro de la syntagmaticité jusqu'à une incorporation profonde de la combinatoire des événements tonaux. S'il est considéré comme utile, depuis F. de Saussure, d'étudier les faits linguistiques conjointement sur deux axes – paradigmatique et syntagmatique –, il est étonnant qu'au niveau intonatif, l'approche paradigmatique reste toujours fortement dominante. En effet, à la différence des termes « phonotactique », « morphotactique » et « syntaxe », qui ont déjà été forgés, l'expression « intonotactique » (tout comme son équivalent anglais « intonotactics ») ne retourne, au moment de la rédaction du présent texte, aucune occurrence sur Google (précisons cependant que l'on peut trouver des termes allant dans le même sens, mais plus fortement liés à la syntaxe, comme « phonosyntaxe » chez G. Boulakia (1974) ou « intonosyntaxe » chez A. Di Cristo (1985).

Le présent article a pour objectif d'exploiter, dans une perspective comparée, l'organisation syntagmatique des accents mélodiques (AM) en français et en tchèque. Notre approche présuppose un système de transcription commun ou, tout au moins, assez comparable. Nous adoptons, à l'instar de la plupart des systèmes tonaux, une stylisation basée sur le contraste entre deux cibles – H et L –, auxquelles nous ajoutons trois tons de frontière en français et quatre en tchèque. À chaque fois que plusieurs possibilités de stylisation se présentent, nous écartons des solutions trop abstraites, basées sur des règles d'implémentation anti-intuitives (voir la critique de l'abstractionnisme tonal dans Martin (2009:78,208]). Sans remettre en question l'intérêt théorique de tels modèles (la simplicité – souvent assimilé à l'« élégance » (Pierrehumbert 1980:16 ; Gussenhoven 2004:272) – étant un principe scientifique valable), nous sommes d'avis que le réalisme tonal est une approche plus viable lorsqu'il s'agit de comparer plusieurs langues entre elles.

En ce qui concerne le tchèque, un système de stylisation basé sur de tels postulats a déjà été esquissé (Duběda 2011), et pour le français, il peut être facilement élaboré sur la base de systèmes existants (voir plus bas).

Le tchèque, langue slave occidentale, se caractérise par un accent fixe, frappant la première syllabe des mots lexicaux. Le principal paramètre responsable de la perception de l'accent est la fréquence fondamentale, alors que la durée et l'intensité restent largement inopérantes. Comme en français, le noyau prosodique se trouve généralement à la fin de l'unité intonative ; sa configuration tonale permet de signaler la modalité et la finalité/non-finalité de l'UI. Les contours nucléaires, notamment là où le dernier mot de l'UI comporte plus de deux syllabes, sont plus complexes qu'en français (Daneš 1957).

Nos hypothèses sont les suivantes :

- La distribution des AM est **différente dans le noyau et dans la partie prénucléaire** de l'unité intonative.
- La distribution des AM prénucléaires **varie selon la position dans l'UI**. Par exemple, Jun & Fougeron (2000:238) font remarquer que les accents initiaux sont plus fréquents au début de la phrase qu'à l'intérieur de celle-ci. Le tchèque, qui dispose d'un plus grand nombre d'AM différents que le français, devrait adopter un comportement plus complexe à cet égard.

- La distribution des AM prénucléaires **varie selon la longueur de l'UI**. Nous pouvons nous attendre aux effets décrits par Ph. Martin (2009:109), qui fait l'observation que dans les phrases longues, les contrastes prosodiques sont plus prononcés, alors que dans les phrases courtes, ceux-ci peuvent être neutralisés.
- Les accents prénucléaires **ne tendent pas à être identiques** à l'intérieur d'une même UI.
- La distribution des AM prénucléaires **interagit avec la nature du noyau prosodique**.
- Plus spécifiquement, la distribution des AM **est régie par la constituance syntaxique**.

À cet égard, nous aspirons à une confirmation partielle de la théorie de l'inversion de la pente pour le français, alors que nous sommes bien plus sceptique en ce qui concerne le tchèque.

2. Matériau analysé

Pour les deux langues, nous avons utilisé un texte journalistique (197 mots en français et 163 en tchèque). Huit locuteurs français (quatre hommes et quatre femmes entre 25 et 40 ans, tous parlant un français standard bien qu'une personne soit originaire de Suisse romande et qu'une autre soit bilingue français-grec) et huit locuteurs tchèques (quatre hommes et quatre femmes entre 19 et 35 ans, tous parlant un tchèque standard, avec des traces marginales d'un accent régional chez deux d'entre eux) ont été enregistrés. Puisque nous travaillons sur la lecture oralisée, non-interactive et plus cohérente au niveau prosodique que la parole spontanée, notre analyse peut rester largement formelle : les critères sémantiques et communicatifs sont délibérément laissés de côté, ce qui permet un meilleur traitement statistique. En revanche, l'échantillon utilisé ne nous permet pas d'étudier les contours finaux interrogatifs. En tout, 988 accents mélodiques ont été analysés pour le français et 1149 pour le tchèque.

3. Stylisation intonative

La grammaire tonale retenue est schématisée dans le Tableau 1.

	Tons de frontière initiaux	Accents prénucléaires	Accents nucléaires	Tons de frontière finaux
Français	%H %L	H* LH* Hi* LHi* L*	H* LH* L* HL*	L% H% 0%
Tchèque	%H %L	L*H L* S* H* H*L	L*H L* H*	L% H% M% 0%

Tableau 1 : Grammaire tonale

Pour le français, le système des tons de frontière est le même que celui de Post (2000) ; nous sommes d'avis que l'opposition entre trois tons de frontière finaux (un ton haut, un ton bas et un ton « passif », qui pourrait également être interprété comme une absence de ton) est nécessaire pour rendre compte de certains contrastes, comme p. ex. la montée de continuation (LH* 0%) et le contour interrogatif (LH* H%). Quant aux accents prénucléaires, nous adoptons le ton H* (présent chez Post 2000 et Jun & Fougeron 2000), nous réinterprétons le ton H* avec un ton L précédent comme un accent mélodique complexe LH* (solution que Post (2000:153) rejette, tout en admettant que le ton bas est généralement aligné avec la syllabe pré-accentuelle), nous distinguons, conformément à Jun & Fougeron 2000, entre accents finaux (H*, LH*) et initiaux (Hi*, LHi*), mais nous n'hésitons pas, à l'instar de Post (2000), d'attribuer aux proéminences initiales le statut d'accents mélodiques. La différenciation formelle entre accents initiaux et finaux est conforme au principe du parenthésage strict et permet de prédire d'autres caractéristiques prosodiques (les accents finaux s'accompagnent, à la différence des accents initiaux, d'un allongement). Un cinquième ton prénucléaire, L*, absent de Post (2000), mais accepté dans certaines circonstances par Jun & Fougeron (2000), a été ajouté pour rendre compte des configurations discutées par Martin (2009:108). Le répertoire des tons nucléaires est le même, à deux choses près : les accents initiaux ne sont jamais nucléaires, et un accent supplémentaire, HL*, permet de décrire le contraste entre le contour déclaratif non-marqué (L* L%) et le contour traditionnellement appelé « évidence » (HL* L%). Chez Post (2000), qui n'accepte pas de tons L*, ce même accent porte l'étiquette H+H*. La compatibilité séquentielle des événements tonaux est libre, sauf la combinaison HL* H% ; quatre combinaisons d'accents nucléaires et de tons de frontière sont possibles, mais non attestées dans nos données (LH* H%, H* H%, H* L%, L* H%).

En ce qui concerne le tchèque, le choix des tons de frontière initiaux est le même qu'en français. Le répertoire des accents prénucléaires comporte cinq éléments : L*H (« montée post-accentuelle »), L* (« catathèse »), S* (« réduit »), H* (« anathèse ») et H*L (« sommet accentuel »). L'accent S* correspond à une proéminence perçue mais très faiblement caractérisée au niveau intonatif (Duběda 2010b) ; une solution alternative consisterait à ne pas tenir compte de ce type d'accent dans la stylisation, ou bien de le considérer comme une forme réduite d'un autre type d'accent. Pour les contours nucléaires, nous adoptons le système proposé par Daneš (1957). Le contour déclaratif le plus fréquent est L* L%, le contour continuatif, L*H 0%. Un système à quatre tons de frontière s'est avéré nécessaire pour rendre compte de certains contours, notamment des mouvements montants-descendants, qui caractérisent les interrogations totales (absentes de notre corpus) et les déclaratives « marquées » (selon la terminologie de Daneš). L'opposition entre les tons de frontière H%, L%, M% et 0% peut être démontrée à l'aide des exemples suivants :

	L*H H%	continuation marquée
	L*H L%	déclarative marquée
	L*H M%	interrogation totale
	L*H 0%	continuation

Le trait le plus insolite de l'accentuation tchèque est sa forte propension aux tons L* : environ trois quarts des accents recensés sont soit L*H (montée post-accentuelle), soit L* (catathèse). La

compatibilité séquentielle des événements mélodiques est soumise à certaines restrictions : les accents prénucléaires S^* et L^* n'apparaissent jamais à l'initiale de l'UI, et les combinaisons $L^*M\%$ et $H^*0\%$ ne sont pas prévues par le système. Les contours $L^*H H\%$, $L^*H M^*$ et $L^* H\%$, quant à eux, sont possibles mais absents de nos données.

4. Vérification des hypothèses

4.1. La distribution des AM est-elle différente dans le noyau et dans la partie prénucléaire de l'UI ?

La fonction des AM nucléaires est diamétralement opposée à celle des AM prénucléaires : dans le noyau, l'AM fait partie d'un contour dans lequel se concentrent des oppositions phonologiques (modalité, finalité/ non-finalité) et dont la décomposition en accent mélodique et ton de frontière est plutôt formelle, alors que les accents prénucléaires sont des événements relativement autonomes, qui ont avant tout une fonction segmentative. C'est cette différence fonctionnelle qui permet d'expliquer en grande partie les données présentées dans les Figures 1 et 2 :

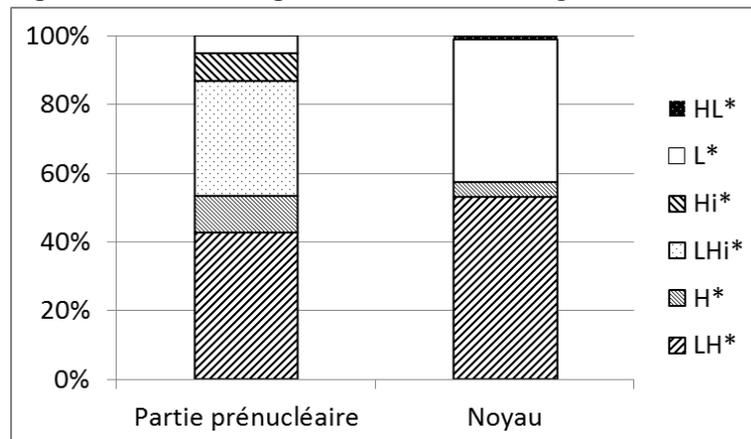


Figure 1 : Fréquence des AM en français : partie prénucléaire et noyau.

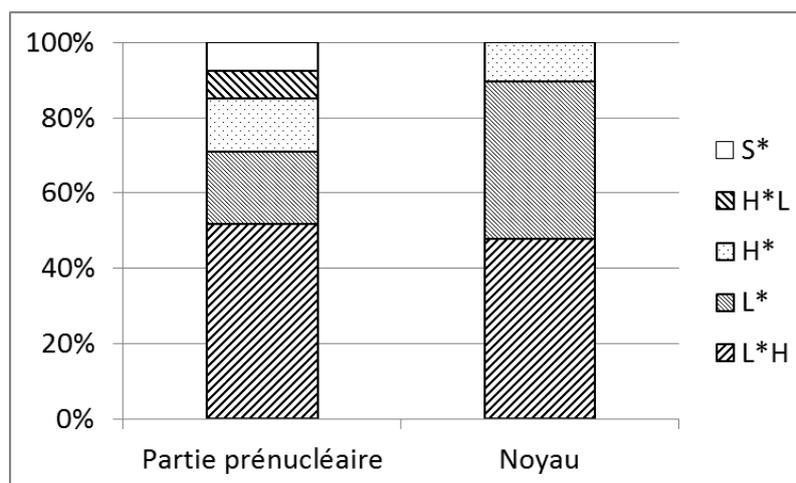


Figure 2 : Fréquence des AM en tchèque : partie prénucléaire et noyau.

En français, les accents prénucléaires se structurent en deux grandes classes dont l'importance est à peu près comparable : les accents finaux LH* et les accents initiaux LHi* ; chacune de ces classes est complétée par un nombre beaucoup moins important d'accents sans ton bas associé à gauche (H* et Hi*). À cela s'ajoute un certain nombre d'accents prénucléaires L* (5 %), dont la distribution semble être beaucoup plus libre que ne le font penser Jun & Fougeron (2000:216), qui estiment que ces tons apparaissent uniquement entre deux accents initiaux. Les accents nucléaires forment également deux classes importantes, qui reflètent la nature syntaxique de l'UI (non-finalité/finalité) : LH*/H* et L*. L'accent HL*, exclusivement nucléaire, apparaît très peu.

Le tchèque, de son côté, préfère les accents L*H et L*, tant dans la partie prénucléaire que dans le noyau. À cela s'ajoute un pourcentage bien moins important d'accents H*, et, uniquement dans la partie prénucléaire, les accents H*L et S*.

On constate donc, pour les deux langues, que la distribution des AM n'est pas identique dans le noyau et dans la partie prénucléaire : certains accents n'apparaissent que dans un de ces contextes, et d'autres apparaissent dans les deux, mais avec une fréquence différente.

4.2. La distribution des AM prénucléaires varie-t-elle selon la position dans l'UI ?

Nous avons comparé la distribution des accents prénucléaires en fonction de leur position à l'intérieur de l'UI. Seules les UI à 2–5 accents mélodiques ont été analysées, afin de garantir des effectifs permettant un traitement statistique. Nous avons comparé la distribution des AM de manière suivante :

- l'AM initial vs. les autres AM prénucléaires ;
- le dernier accent prénucléaire (précédant immédiatement l'accent nucléaire) vs. les autres AM prénucléaires ;
- l'évolution des effectifs pour les différents AM du début à la fin de la partie prénucléaire.

Pour le français, aucune tendance systématique n'a été identifiée : il semble que la mise en place des différents types d'AM dépende plutôt de facteurs locaux tels que la longueur du mot (plus la distance entre deux accents H* est petite, plus la tendance à omettre le ton bas est forte ; plus le mot est long, plus il attire les accents initiaux) et la structure syntaxique locale (tendance à créer des arcs accentuels). L'hypothèse de Jun & Fougeron (2000:238) sur la plus forte incidence des accents initiaux au début de l'UI n'a pas été confirmée.

En revanche, plusieurs tendances ont été mises en évidence pour le tchèque :

- L'UI ne commence jamais ni par L* ni S*, ces deux accents étant définis par rapport au contexte gauche (cette observation est bien évidemment liée à la méthode de stylisation). Par contre, les accents H* sont plus fréquents en position initiale : en comparant la position initiale avec toutes les autres positions prénucléaires, on obtient une différence significative (χ^2 ; $p < 0,05$) pour cinq différences sur six.
- La proportion d'accents L*H décroît vers la fin de l'UI (en comparant leur fréquence pour la n^e et pour la n+1^{ère} position dans tous les contextes où cette différence s'applique, le pourcentage est toujours inférieur pour la n+1^{ère} position, et trois différences sur six sont statistiquement significatives (χ^2 ; $p < 0,05$). Ce phénomène peut s'expliquer par le rétrécissement du couloir intonatif, une des conséquences de la déclinaison.

- Par contre, ni la fréquence des accents « catathétiques » L* ni celle des accents « réduits » S* n'augmentent vers la fin de l'UI, comme on aurait pu le supposer en vertu de ce même principe.

Nos données indiquent donc que la position à l'intérieur de l'UI n'est pas un facteur pertinent pour le choix des AM prénucléaires en français ; l'hypothèse de Jun & Fougeron (2000:238) selon laquelle les accents initiaux seraient plus fréquents au début de la phrase n'a pas été confirmée. Dans le cas du tchèque, quatre accents sur cinq sont soumis à certaines contraintes statistiquement significatives, qui régissent leur choix en fonction de la position au sein de l'UI.

4.3. *La distribution des AM prénucléaires varie-t-elle selon la longueur de l'UI ?*

Pour vérifier cette hypothèse, nous avons divisé les UI étudiées en deux groupes : les UI « brèves » (2 ou 3 AM) et les UI « longues » (4 ou 5 AM). En comparant la fréquence des différents AM prénucléaires entre ces deux groupes, nous n'avons dépisté aucune tendance systématique pour le français (les tests χ^2 , appliqués aux pourcentages enregistrés, ne retournent jamais de différence significative). Ici encore, il semble que le choix des accents est régi par des facteurs locaux.

En ce qui concerne le tchèque, nous constatons des différences significatives pour chacun des cinq accents prénucléaires : les accents H* et H*L sont plus fréquents dans les unités brèves (χ^2 ; $p < 0,001$, $p = 0,003$, respectivement), alors que les accents L*H, L* et S* sont plus fréquents dans les unités longues (χ^2 ; $p = 0,045$, $p = 0,012$, $p = 0,014$, respectivement). Pour tous les accents excepté L*H, cette situation est le résultat des contraintes positionnelles (H* et H*L sont surtout attestés en position initiale, et leur poids statistique est donc plus important dans les unités brèves ; L* et S* n'apparaissent jamais en position initiale, ce qui augmente leur fréquence relative dans les unités longues). Seul l'accent L*H mérite donc notre attention : sa plus grande fréquence dans les UI longues peut s'expliquer par la planification de l'énoncé, car il s'agit du seul accent à grande fréquence qui contient un ton haut, nécessaire pour freiner la catathèse qui aurait été trop rapide avec les accents L* ou S*.

Les données obtenues pour le français ne permettent donc pas de confirmer l'hypothèse de l'amplification des contrastes mélodiques dans des phrases longues (Martin 2009:109). Rappelons toutefois que pour la mettre réellement à l'épreuve, il faudrait adopter une méthodologie différente de la nôtre : en effet, notre analyse porte sur les unités intonatives, et non pas sur les phrases entières, et elle tient compte uniquement des types d'AM, et non de leur ampleur. En revanche, le tchèque semble être conforme à cette hypothèse, du moins en ce qui concerne le choix des accents L*H.

4.4. *Y a-t-il une tendance à l'uniformisation des accents prénucléaires à l'intérieur d'une même UI ?*

Comme les accents prénucléaires assument, dans les deux langues, un rôle principalement démarcatif, on pourrait avancer l'hypothèse qu'un seul type d'accent serait suffisant pour satisfaire à ce besoin. Or, les données empiriques indiquent tout l'inverse : la variabilité séquentielle des accents peut résulter de contraintes sémantiques et attitudinales (qui sont délibérément laissées de côté dans la présente étude), positionnelles (dont certaines ont été démontrées plus haut pour le

tchèque), syntaxiques (qui seront abordées plus bas), dissimilatoires ou idiolectales, sans oublier l'effet du simple hasard.

Les Figures 3 et 4 font état de la variabilité des accents prénucléaires dans les UI contenant 2 à 4 accents prénucléaires. La structure de chaque colonne indique l'uniformité accentuelle (1 à 4 accents différents attestés dans l'UI). Par exemple, dans l'hypothèse de l'uniformité absolue de la partie prénucléaire (un seul type d'accent permis), les trois colonnes porteraient sur toute leur hauteur la hachure indiquée sous « 1 ».

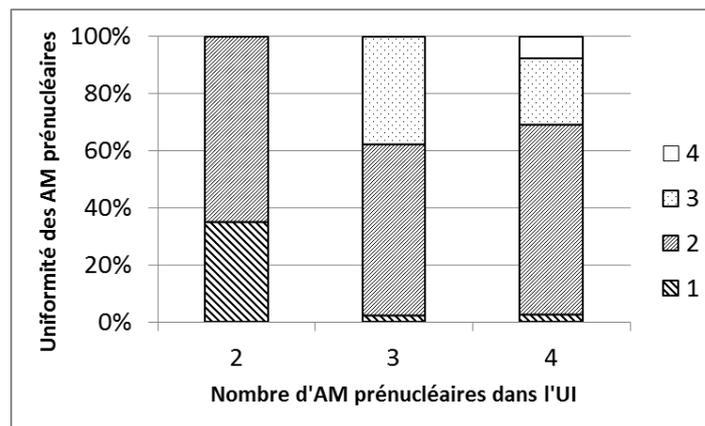


Figure 3 : Uniformité des AM prénucléaires en français.

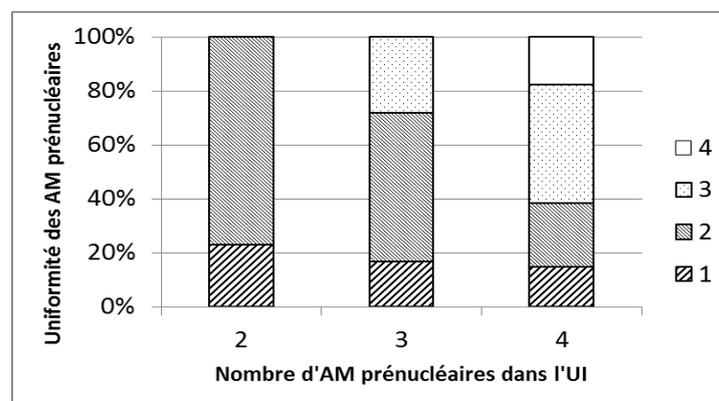


Figure 4 : Uniformité des AM prénucléaires en tchèque.

Les deux graphiques indiquent clairement que la tendance à limiter le choix des AM prénucléaires à un seul type dans chaque UI est extrêmement faible ; il en ressort au contraire que la variabilité des accents prénucléaires est la règle dans les deux langues.

4.5. La distribution des AM prénucléaires interagit-elle avec la nature du noyau prosodique ?

L'idée que la nature du noyau se reflète dans le choix des AM prénucléaires fait partie intégrante de certains modèles prosodiques (Martin 1975 ; Di Cristo 1998). Il semble logique de présumer une interaction unidirectionnelle plutôt que réciproque : la nature du noyau découle de conditions extérieures à l'UI (modalité, continuation/ finalité, expressivité), et elle est paradigmatique, alors que la séquence des accents prénucléaires est le résultat d'un arrangement syntagmatique à

l'intérieur de l'UI.

Les UI françaises ont été catégorisées en trois groupes :

- contour déclaratif ($L^* L\%$) ;
- contours continuatifs non-marqués ($LH^* 0\%$; $H^* 0\%$) ;
- contour continuatif marqué ($L^* 0\%$).

La catégorisation est basée sur la fonction du contour, qui se reflète également dans sa forme. Trois autres contours attestés ($HL^* L\%$; $HL^* 0\%$; $LH^* L\%$) n'ont pas été pris en considération du fait de leur très faible fréquence ($N = 4, 5$ et 8 , respectivement).

Ensuite, nous avons calculé la fréquence des AM prénucléaires pour chacun des trois types d'UI. Une différence significative a été trouvée dans le cas de l'accent prénucléaire L^* , qui apparaît dans les UI à contour continuatif non marqué, alors qu'il est extrêmement rare dans les UI à contour déclaratif (6% et 1% , respectivement ; χ^2 ; $p = 0,035$). Ce comportement, qui peut être qualifié de dissimilatoire, va donc dans le sens de la théorie de l'inversion de la pente.

Dans le cas du tchèque, nous avons catégorisé les UI de la manière suivante :

- contour déclaratif non-marqué ($L^* L\%$) ;
- contours déclaratifs marqués ($H^* L\%$; $L^*H L\%$) ;
- contour continuatif ($L^*H 0\%$).

Les contours $H^* H\%$, $H^* M\%$ et $L^* 0\%$, très peu fréquents ($N = 7, 3$ et 4 , respectivement), ont été éliminés.

Après avoir calculé la fréquence des AM prénucléaires séparément pour chacun des trois types d'UI, nous avons constaté une seule différence statistiquement significative entre les trois histogrammes : les accents L^* sont plus fréquents dans les UI à contour déclaratif marqué que dans les UI continuatives. À part cette observation, difficilement interprétable, la distribution des différents AM prénucléaires a été proche de celle qui est présentée sur la Figure 2. D'un point de vue global, la nature du noyau n'influerait donc pas sur le choix des AM prénucléaires.

Dans un deuxième temps, nous avons procédé à plusieurs analyses partielles, dont le dénominateur commun a été la réduction des degrés de liberté (c'est-à-dire un raffinement des conditions), ce afin de mettre au jour des régularités plus locales qui échappent à l'analyse globale. Pour éviter le danger des conclusions basées sur un trop petit nombre d'observations, nous considérons comme statistiquement significatives uniquement les différences qui satisfont à la fois au test χ^2 et à la condition qu'au moins un des deux ensembles comparés doit compter 10 éléments ou plus ; cette condition a d'ailleurs été respectée dans toutes les conclusions faites plus haut et elle le sera dans ce qui suit.

Premièrement, la reproduction de l'analyse sur un sous-ensemble d'UI limité à celles comportant seulement deux accents mélodiques (un prénucléaire et un nucléaire) a révélé des différences frappantes : l'accent prénucléaire L^*H est présent dans 58% des UI à contour déclaratif, alors qu'il est entièrement absent des UI continuatives, et la proportion d'accents prénucléaires H^* est de 26% dans les UI à contour déclaratif et de 84% dans des UI continuatives (le test χ^2 n'acceptant pas de valeur zéro, nous avons utilisé le chiffre $0,001$ à la place ; dans les deux cas, $p < 0,001$). Les deux observations s'expliquent par une tendance dissimilatoire : éviter une montée dans la partie prénucléaire si le contour final est montant. Les contours déclaratifs marqués n'ont pas été inclus dans l'analyse ($N = 3$).

Deuxièmement, nous avons étudié les accents précédant immédiatement le noyau (à l'exclusion des AM initiaux), toutes longueurs d'UI confondues. Les différences suivantes ont été recensées (toutes statistiquement significatives ; χ^2 ; $p < 0,05$) :

- différence entre contour déclaratif non-marqué et contour déclaratif marqué : H* plus fréquent et S* moins fréquent dans le premier cas (les deux tendances sont de nature dissimilatoire) ;
- différence entre contour déclaratif non-marqué et contour continuatif : L* plus fréquent, L*H et H*L moins fréquents dans le premier cas (le comportement de l'accent L* est assimilatoire, les deux autres tendances sont ambivalentes) ;
- différence entre contour déclaratif marqué et contour continuatif : L* plus fréquent, L*H, H* et H*L moins fréquents dans le premier cas (les tendances observées pour L* et H* sont de nature dissimilatoire, les deux autres sont ambivalentes).

Troisièmement, nous avons testé l'effet du noyau sur le premier accent prénucléaire dans l'UI (χ^2 ; $p < 0,05$) :

- différence entre contour déclaratif marqué et contour continuatif : aucun effet significatif ;
- différence entre contour déclaratif non-marqué et contour continuatif : H* plus fréquent dans le premier cas (tendance ambivalente) ;
- différence entre contour déclaratif marqué et contour continuatif : L*H moins fréquent dans le premier cas.

L'influence du noyau sur le choix des AM prénucléaires a donc été démontrée dans les deux langues. En français, la présence des accents L*, atypiques dans la partie prénucléaire, est corrélée avec les contours continuatifs. En tchèque, l'impact du contour nucléaire est plutôt local : il se manifeste surtout dans l'accent immédiatement prénucléaire. Dans les deux langues, le principe prédominant est celui d'une dissimilation tonale régressive, que l'on peut rapprocher du Principe du contour obligatoire (Leben 1973) et qui est également à la base de la théorie de l'inversion de la pente.

4.6. La distribution des AM est-elle régie par la constituance syntaxique ?

L'usage des AM prénucléaires pour marquer la structure syntaxique est un des éléments clés de la théorie prosodique de Ph. Martin (1975). Nous avons soumis ce principe à une vérification, en nous limitant à une seule position syntaxique : après avoir identifié dans chaque UI la frontière syntaxique la plus importante (entre ses deux constituants immédiats), nous avons testé l'interaction de l'accent mélodique situé devant cette frontière avec le noyau. L'analyse ressemble ainsi beaucoup à celle présentée dans la partie 7 (y compris la classification des noyaux), à ceci près que la variable dépendante est le type d'AM prénucléaire dans une position syntaxiquement spécifiée :

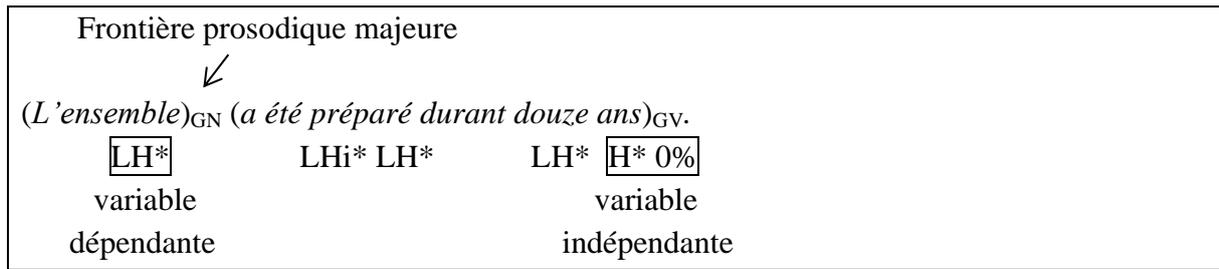


Figure 5 : Exemple d'analyse phonosyntaxique N° 1.

Seuls les accents finaux ont été pris en considération ; l'élimination des accents initiaux nous a paru nécessaire étant donné qu'ils ne sont pas obligatoires et que leur potentiel syntagmatique est selon toute vraisemblance local (frontière gauche des arcs accentuels). Nous partageons ainsi l'avis de Ph. Martin (1909:110), qui tient que les accents initiaux ne sont pas susceptibles d'indiquer des relations de dépendance. Les accents finaux, moins variables, sont systématiquement présents devant les discontinuités que nous tâchons d'analyser dans le présent cas. Là où un arc accentuel vient perturber la structure syntaxique (en vertu des principes de l'autonomie des groupements prosodiques (Grosjean & Dommergues 1983), nous respectons la primauté du parenthésage prosodique (Figure 6).

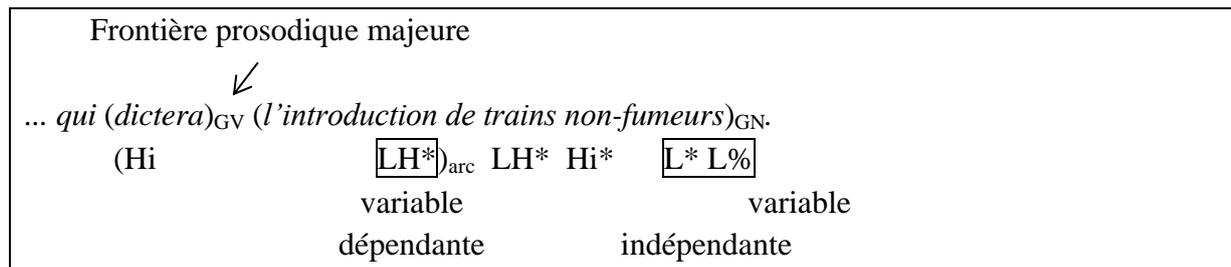


Figure 6 : Exemple d'analyse phonosyntaxique N° 2.

Comme les accents initiaux ne font pas partie de l'analyse et que les accents LH* et H* peuvent être considérés comme formant une seule catégorie logique, l'unique valeur réellement informative est le pourcentage d'accents L*, donné dans la Figure 7.

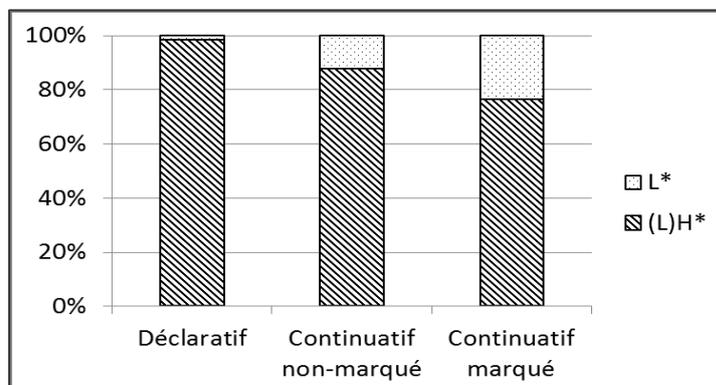


Figure 7 : Fréquence des AM prénucléaires français situés devant la frontière prosodique la plus forte, en fonction du type de noyau.

Pour les trois types d'UI, on voit que c'est l'accent « par défaut » H*/LH* qui prédomine largement dans le marquage de la frontière prosodique la plus importante. Cependant, les UI à contour continuatif non marqué ont une propension significativement plus forte à générer l'accent L* dans cette position (χ^2 ; $p = 0,002$). Nous avons donc confirmé le même effet que dans la partie 7, en le localisant sur une position syntaxique particulière : si le poids des accents L* dans les UI à contour continuatif non-marqué, toutes positions confondues, a été de 6 % (voir partie 7), ce même pourcentage passe à 12 % si on prend en compte uniquement la position devant la frontière majeure. Ceci nous permet de dire que l'inversion de la pente est un effet phonosyntaxique faible mais statistiquement significatif.

Dans le cas du tchèque, nous avons procédé de la même manière ; les résultats sont présentés dans la Figure 8.

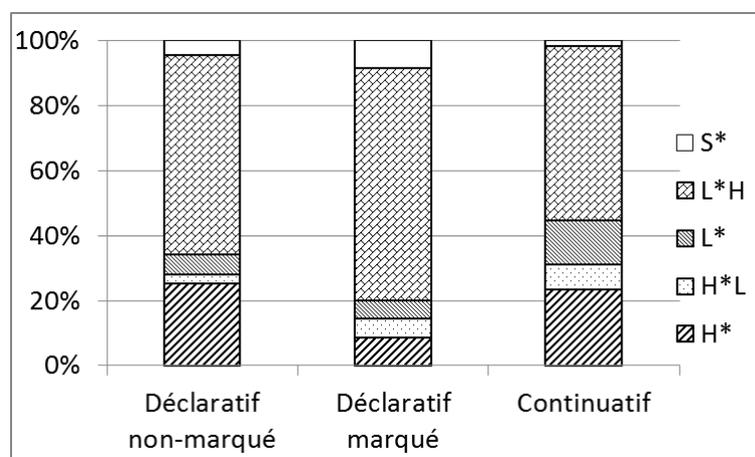


Figure 8 : Fréquence des AM prénucléaires tchèques situés devant la frontière prosodique la plus forte, en fonction du type de noyau.

Les différences suivantes ont été recensées (toutes statistiquement significatives ; χ^2 ; $p < 0,05$) :

- différence entre contour déclaratif non-marqué et contour déclaratif marqué : H* plus fréquent et L*H moins fréquent dans le premier cas (tendance difficilement interprétable ; l'effet peut être induit par le simple fait que la frontière majeure est située, dans un grand nombre d'UI, après le premier mot, ce qui augmente la fréquence des accents H*, typiquement initiaux) ;
- différence entre contour déclaratif non-marqué et contour continuatif : aucun effet statistiquement significatif ;
- différence entre contour déclaratif marqué et contour continuatif : L*H plus fréquent, L* et H* moins fréquents dans le premier cas (tendances difficilement interprétables).

En résumé, la position phonosyntaxique étudiée (l'AM à gauche de la frontière syntaxique la plus importante de l'UI) semble jouer un certain rôle phonologique en français en ce que le choix des accents mélodiques s'y fait différemment des autres positions, et ce dans le sens de la dissimilation tonale d'avec le contour final (inversion de la pente). Pour le tchèque, en revanche, aucun comportement systématique n'a été décelé dans cette position.

Nous sommes conscient du fait que d'autres positions phonosyntaxiques pourraient être testées dans le cas du tchèque où l'analyse effectuée n'a retourné aucun résultat intéressant ; faute de place, nous nous sommes limité ici à une seule variable que nous jugeons a priori importante pour les deux langues. Enfin, rappelons encore une fois que seul le type d'AM, et non pas son envergure, est pris en considération : il se peut qu'une analyse plus fine de la réalisation phonétique des AM confirme encore davantage le comportement dissimilatoire identifié.

5. Conclusion

Les principales conclusions, structurées selon les hypothèses formulées au début du présent article, sont résumées dans le Tableau 2.

Hypothèse	Français	Tchèque
1) La distribution des AM est différente dans le noyau et dans la partie prénucléaire de l'UI.	OUI	
	<ul style="list-style-type: none"> - Certains accents sont uniquement nucléaires ou uniquement prénucléaires. - La fréquence de certains accents est différente dans les deux contextes. 	
2) La distribution des AM prénucléaires varie selon la position dans l'UI.	NON	OUI
	<ul style="list-style-type: none"> - Pas d'effet. 	<ul style="list-style-type: none"> - Certains accents sont bloqués en début d'UI. - Effet significatif pour les accents H* et L*H (rétrécissement du couloir intonatif).
3) La distribution des AM prénucléaires varie-t-elle selon la longueur de l'UI ?	NON	OUI
	<ul style="list-style-type: none"> - Pas d'effet. 	<ul style="list-style-type: none"> - Effet pour l'accent L*H (planification de l'énoncé).
4) Les accents prénucléaires ne tendent pas à être identiques à l'intérieur d'une même UI.	OUI	
5) La distribution des AM prénucléaires interagit avec la nature du noyau prosodique.	OUI	OUI
	<ul style="list-style-type: none"> - Effet global pour les accents L* (dissimilation tonale). 	<ul style="list-style-type: none"> - Effet local (majoritairement de nature dissimilatoire).
6) La distribution des AM est régie par la constituance syntaxique.	OUI	NON
	<ul style="list-style-type: none"> - Effet pour l'accent L*. 	<ul style="list-style-type: none"> - Pas d'effet.

Tableau 2 : Hypothèses et conclusions.

Dans trois cas sur six, le comportement des deux langues va dans le même sens. Tout d'abord, il n'est pas surprenant de voir que la distribution des AM dans le noyau et dans la partie prénucléaire est partiellement différente : dans le premier cas, l'AM fait partie du contour final (plus complexe en tchèque qu'en français) et s'inscrit dans un paradigme phonologique ; dans le second cas, la dimension syntagmatique prend le dessus sur la fonction paradigmatisée. Le français comme le tchèque manifestent une tendance à faire varier les accents prénucléaires. Une interaction faible mais significative du noyau avec les AM prénucléaires a été observée dans les deux langues, qui s'explique principalement par une tendance à la dissimilation tonale.

Dans les trois autres cas, nos conclusions sont différentes pour le français et pour le tchèque. En ce qui concerne l'influence de la position et de la longueur de l'UI, le français ne manifeste aucune variation systématique des AM prénucléaires, dont le choix se fait plutôt en fonction de conditions locales. En revanche, le tchèque montre des effets syntagmatiques plus prononcés, qui peuvent s'expliquer par le rétrécissement du couloir intonatif et par la planification de l'énoncé. Finalement, l'impact de la structure syntaxique de l'UI sur le choix des AM a été observé uniquement en français, toujours dans le sens d'une dissimilation tonale.

Rappelons en passant que les conclusions tirées pour le tchèque confirment celles qui ont été publiées dans Duběda 2010a (cinq locuteurs, texte de longueur supérieure).

L'existence de contraintes syntagmatiques qui régulent l'agencement de l'UI est dès lors évidente, quoique la plupart des effets décrits soient de faible envergure ; c'est pour cela d'ailleurs que l'intérêt pour la dimension syntagmatique de l'intonation n'est que marginal. La question de savoir dans quelle mesure nos observations sont généralisables à d'autres phonostyles que la lecture oralisée reste ouverte. En revanche, le choix d'un texte authentique plutôt que des exemples construits augmente à notre avis la validité de nos conclusions.

Remerciements

Le présent article a été rédigé dans le cadre du projet GAČR P406/10/0101.

Références

- Boulakia, G. (1974). *Phonosyntaxe*. Thèse de doctorat, Paris.
- Dainora, A. (2006). Modeling intonation in English: A probabilistic approach to phonological competence. Goldstein, L. et al. (eds.), *Laboratory Phonology 8*, Walter de Gruyter, pp. 107–132.
- Daneš, F. (1957). *Intonace a věta ve spisovné češtině*. Prague: ČSAV [L'intonation de la phrase en tchèque standard].
- Di Cristo, A. (1985). *De la microprosodie à l'intonosyntaxe*. Université de Provence.
- Di Cristo, A. (1998). Intonation in French. Hirst, D. and Di Cristo, A. (eds.), *Intonation Systems. A Survey of Twenty Languages*. Cambridge University Press, pp. 195–218.
- Duběda, T. (2010a). Positional variability of pitch accents in Czech. Kobayashi, T. et al. (eds.), *Proceedings of Interspeech 2010*. pp. 1760–1763.
- Duběda, T. (2010b). Flat pitch accents in Czech. Kobayashi, T. et al. (eds.), *Proceedings of Interspeech 2010*, pp. 1756–1759.
- Duběda, T. (2011). Towards an inventory of pitch accents for read Czech. *Slovo a slovesnost 72/1*, pp. 3–12.

- Grosjean, F. & Dommergues, J.-Y. (1983). Les structures de performance en psycholinguistique. *L'année psychologique*, 83, pp. 513–536.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press.
- Jun, S.-A. & Fougeron, C. (2000). A phonological model of French intonation. Botinis, A. (ed.), *Intonation: Analysis, Modelling and Technology*. Dordrecht: Kluwer Academic Publishers, pp. 209–242.
- Ladd, D.R. (1996). *Intonational Phonology*. Cambridge University Press.
- Leben, W. (1973). *Suprasegmental Phonology*. Ph.D. dissertation, MIT.
- Le Gac, D. & Yoo, H.-Y. (2007). Indépendance et dépendances de l'intonation. *Nouveaux Cahiers de Linguistique Française n° 28*, Genève, pp. 187–198.
- Martin, Ph. (1975). Analyse phonologique de la phrase française. *Linguistics*, Vol. 146, pp. 35–68.
- Martin, Ph. (2009). *Intonation du français*. Armand Colin, Paris.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*. MIT.
- Post, B. (2000). *Tonal and phrasal structures in French intonation*. The Hague: Holland Academic Graphics.

Contraste de pente et principe du contour obligatoire : « l'union fait la force »

David Le Gac

david.legac@univ-rouen.fr

Normandie Univ, France - EA4701 : DySoLa

Abstract

This paper aims at establishing a relation between the Obligatory Contour Principle (“OCP”) and the mechanism of “Contraste De Pente” (pitch slope inversion) proposed by Philippe Martin. We show how the OCP may have an effect on intonation by investigating the intonation of French and Modern Greek, which both present challenging pitch alternations for OCP and “Contrast De Pente”. We propose that OCP and “Contrast De Pente” are both the instantiations of a more general constraint prohibiting identical elements in “direct” relationship on given phonological levels. The intonation of languages can then be described according to the activation of one of the two direct relationships, or both.

1. Introduction

Le mécanisme de Contraste De Pente (dorénavant « CDP ») est certainement la proposition la plus emblématique des travaux de Ph. Martin sur l'intonation. Dans son article de 1981, Ph. Martin propose une théorie de l'intonation du français dans laquelle les contours mélodiques associés aux syllabes accentuées s'organisent en une structure hiérarchisée ; la forme de ces contours est régie par le mécanisme du CDP, qui stipule que : « tout mot ou syntagme intonatif qui se joint à un mot ou syntagme situé à sa droite pour former un groupe plus grand présente un contour de sens de variation mélodique opposé à celui du syntagme auquel il se joint » (Martin 1981:165).

Nous nous proposons dans cet article d'expliquer pourquoi ce mécanisme revêt à nos yeux un intérêt et une source d'inspiration pour une théorie phonologique de l'intonation, notamment en le confrontant à l'un des principes fondamentaux de la théorie autosegmentale, le **Principe du Contour Obligatoire** (« PCO »), en anglais **Obligatory Contour Principle** (« OCP »).

Dans une première partie, nous définissons le PCO et montrons comment il peut s'articuler avec l'intonation. En nous basant sur la grammaire tonale de Beckman & Pierrehumbert (1986), nous illustrons cette articulation avec l'intonation du bengali telle qu'elle a été analysée par Hayes & Lahiri (1991).

Dans une deuxième partie, nous mettons en regard le PCO et le CDP. Bien qu'il semble possible et attrayant de prime abord de réinterpréter le CDP à l'aune du PCO, nous montrons que cette réinterprétation est illusoire car PCO et CDP reposent sur des présupposés théoriques différents. En outre, certains des présupposés du CDP se révèlent nécessaires pour rendre compte de faits intonatifs du français, tandis que le PCO s'avère insuffisant. Nous arguons qu'il en est de même dans une langue à accent lexical, le grec moderne, étayant l'idée que le CDP constitue un mécanisme grammatical général.

Enfin, dans la troisième partie, nous montrons qu'il est néanmoins possible d'unifier PCO et CDP, l'un et l'autre étant les manifestations d'un principe plus fondamental encore, excluant

« l'adjacence » de deux éléments identiques à un niveau structurel donné. Nous verrons qu'il est alors possible d'établir une typologie des langues basée sur ce principe.

2. Grammaires tonales et PCO

Dans cette première partie, nous définissons le PCO et expliquons comment il s'articule avec l'intonation dans le cadre de la théorie métrique et autosegmentale (« MA ») proposée par Pierrehumbert (1980), Beckman & Pierrehumbert (1986) et Pierrehumbert & Beckman (1988).

2.1. Grammaire tonale dans le cadre MA

Pierrehumbert (1980) puis Beckman & Pierrehumbert (1986) et Pierrehumbert & Beckman (1988) analysent l'intonation comme une concaténation de tons phonologiques de différents types – accents mélodiques (« pitch accents »), accents de groupe (« phrase accents ») et tons de frontière (« boundary tones ») – formant une grammaire à états finis.

Ce type de grammaire, qui constitue la base de la majorité des descriptions de l'intonation dans le cadre de la phonologie métrique et autosegmentale (voir Ladd (2008) et Gussenhoven (2004) pour une revue), implique deux postulats fondamentaux : 1) les tons n'entretiennent pas de relations entre eux, autrement dit, les contours n'ont pas de structure interne ; et 2) événements indépendants, les tons forment une chaîne où **toutes** les combinaisons sont autorisées, implication que Pierrehumbert (1980), Beckman & Pierrehumbert (1986) et Pierrehumbert & Beckman (1988) assument explicitement.

Ainsi, pour l'anglais, ces auteurs proposent la grammaire donnée *Figure 1*, où six accents mélodiques, suivis des deux accents de groupe B- et H- et des deux tons de frontière B% et H% se combinent librement pour générer les 24 contours possibles.

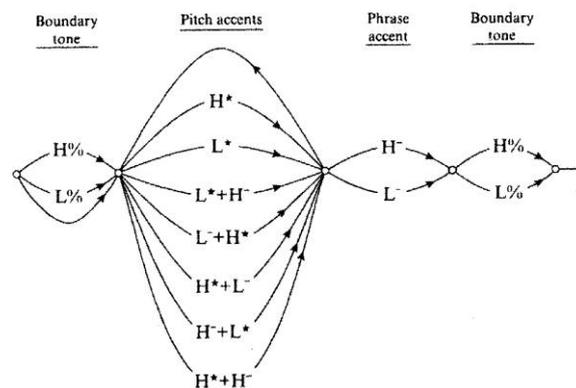


Figure 1 : Grammaire à états finis d'après Pierrehumbert (1980) pour générer les contours de l'anglais (repris de Ladd 2008:89).

Selon cette approche, les langues dont l'intonation ne réalise pas toutes ces possibilités sélectionnent simplement moins de tons (accents mélodiques et tons de frontière), et ce, de façon contingente. Au niveau de l'ensemble des langues, d'un point de vue statistique, aucune grammaire tonale ne devrait donc être privilégiée par rapport à une autre.

Même si une description d'un très grand nombre de langues est indispensable pour établir qu'une grammaire tonale doit être favorisée, nous adoptons l'hypothèse de travail provisoire qu'il existe au contraire des principes qui régulent les combinaisons tonales. L'article de Hayes & Lahiri (1991) illustre clairement et élégamment l'action sur l'intonation du bengali d'un principe bien connu des phonologues, le PCO.

2.2. Le PCO : généralités

Dans sa thèse de 1973 instaurant la phonologie autosegmentale, Leben (1973) propose que les tons lexicaux soient gouvernés par le **Principe du Contour Obligatoire** (« Obligatory Contour Principle », « OCP » en anglais) ; on le trouve formulé généralement de la façon suivante (1) :

(1) *Principe du Contour Obligatoire* (PCO)

Deux éléments identiques adjacents sont interdits

L'idée essentielle est que deux éléments adjacents doivent alterner et former ainsi un « contour », si bien que les séquences tonales en (2) ci-dessous sont prohibées :

(2) a. ⊕ H H b. ⊕ B B

Depuis Leben, il a été établi que le PCO ne concernait pas seulement les tons, mais toute séquence de traits situés sur un même niveau phonologique, ou « tier » dans la terminologie MA. Tout un débat a aussi eu lieu sur le caractère universel ou non du PCO et sur son statut exact au sein de la grammaire (voir à ce propos Myers 1997, Odden 1986, Yip 1988, Yip 2007). A l'heure actuelle, il semble admis que le PCO doit être considéré comme une contrainte pouvant être enfreinte et paramétrée différemment selon les langues (de ce point vue, le PCO s'adapte aisément dans le cadre de la théorie de l'optimalité comme une contrainte violable, cf. Myers 1997).

D'autre part, Yip (1988), Yip (2007) et Hayes & Lahiri (1991) proposent que le PCO s'applique dans des **domaines**, qui varient selon les langues. Le domaine peut être la syllabe, le morphème, le mot, et au-delà encore. Cela signifie que, dans une langue donnée, les séquences en (2) seraient interdites à l'intérieur d'un mot par exemple, mais que deux tons identiques appartenant à deux mots différents pourraient être adjacents. Le PCO serait aussi à l'œuvre entre les éléments de la grille métrique Yip (1988). Comme nous le verrons, selon Hayes & Lahiri (1991), le PCO s'applique systématiquement à l'ensemble des tons de la **phrase** en bengali.

Différentes stratégies peuvent être mises en œuvre pour éviter que le PCO ne soit violé. Le PCO peut **bloquer** certains processus qui auraient dû s'appliquer, ou au contraire, **déclencher** des mécanismes pour « réparer » cette violation, comme par exemple :

- la **polarité** tonale, règle très répandue dans les langues bantoues ou gur, qui consiste à inverser la valeur du ton d'un affixe par rapport à celui du radical ; le *Tableau 1* ci-dessous en donne un exemple tiré du konni, où le ton du suffixe du pluriel est inversé par rapport à celui du radical :

Sing.	Plur.	Ton du radical	Ton du suffixe du pluriel
ta ^B -ŋ ^H	ta ^B n-a ^H	B	H « pierre/s »
si ^H -ŋ ^H	si ^H -a ^B	H	B « poisson/s »

Tableau 1 : polarité tonale en konni provoquée par le PCO (Yip 2002:159)

- l'**insertion** tonale, selon laquelle, entre deux tons identiques va être inséré un troisième ton différent ; c'est par exemple le sandhi du ton 3 en mandarin, où la violation du PCO par la succession de deux tons 3 /B.B/ est évitée par l'insertion d'un ton H :
- (3) *Insertion tonale en mandarin* :
B.B → BH.B (Yip 2002:98)
- l'**effacement** ou la **dissimilation** d'un ton, illustré par une règle bien connue des bantouistes, la règle de Meeussen, qui stipule qu'un ton haut lexical soit disparaître, soit devient bas après un autre ton haut :
- (4) *Règle de Meeussen* :
H → {∅ ; B} / H __

Les exemples ci-dessus illustrent la pression que le PCO exerce sur les tons lexicaux. Dans leur article, Hayes & Lahiri (1991) montrent de manière convaincante comment le PCO entre en jeu au niveau supra-lexical et gouverne l'intonation du bengali, réduisant ainsi drastiquement le nombre de contours que peut générer une grammaire tonale à états finis comme celle de Pierrehumbert (1980).

2.3. PCO et intonation : le cas du bengali

La grammaire tonale de Hayes & Lahiri (1991) repose sur les hypothèses suivantes :

- la structure prosodique ne comporte que deux constituants au-dessus du pied métrique : le groupe phonologique (« phonological phrase », « P ») et le groupe intonatif (« intonation phrase », « I ») ;
- à ces constituants sont associés les tons H ou B, ou rien « ∅ » ; pour I, les bitons sont possibles, on ainsi peut avoir T_I ou T_IT_I ;
- le bengali, comme l'anglais, a des accents mélodiques simples (H*, B*) et un accent double (B+H*) ;
- l'intonation est structurée en un contour noyau, le « nucleus » et ce qui précède.

A partir de ces hypothèses, Hayes & Lahiri (1991:75) posent la formule générale en (5) pour le contour nucléaire. Rien que pour ce dernier, 54 combinaisons sont logiquement possibles. Or, sur ces 54 possibilités, seuls les 8 contours donnés

Tableau 2 sont attestés en bengali, dont une caractéristique remarquable est la stricte alternance entre tons hauts et bas :

(5) *Formule du contour nucléaire en bengali :*

$$T^* (T_P) (T_I) T_I$$

1.	$B^* H_I$	“Offering”
2.	$B^* H_I B_I$	Questions polaires
3.	$B^* H_P B_I$	Focus
4.	$B^* \quad H_P$ $B_I H_I$	Focus avec continuation
5.	$H^* B_I$	Déclaratives
6.	$H^* B_I H_I$	Décl. + continuation
7.	$B+H^* B_I$	Downstep
8.	$B+H^* B_I H_I$	Downstep + continuation

Tableau 2: Inventaire des contours nucléaires du bengali (Hayes & Lahiri 1991:72).

Outre l’absence d’un ton B_P en bengali, l’explication de cette sélection drastique et de l’alternance tonale provient naturellement de l’action du PCO, lequel agit comme une sorte de filtre : seuls les contours intonatifs générés par la grammaire 0 qui respectent le PCO sont retenus.

Le PCO permet aussi de comprendre une asymétrie tonale entre questions polaires et déclaratives. Les contours de questions sont **inversés** par rapports à ceux des déclaratives, avec alternance entre H et B due au PCO. Une exception apparaît cependant en ce qui concerne le focus étroit (cf.

Tableau 3) : celui-ci est marqué par la séquence / B^*H_P / dans les déclaratives mais cette séquence se neutralise au profit de B^* seulement dans les questions (i.e. il n’y a plus de H_P) :

a. Déclaratives	Neutres	H^*	$B_I(H_I)$
	Focus étroit	B^*H_P	$B_I(H_I)$
b. Questions	Neutres	B^*	$H_I(B_I)$
	Focus étroit	B^*	$H_I(B_I)$

Tableau 3: Contours nucléaires du bengali des phrases déclaratives et des questions polaires dans les énoncés « neutres » et ceux comportant un focus « étroit » (d’après Hayes & Lahiri 1991).

Comme le PCO est actif dans cette langue, Hayes & Lahiri (1991:76) proposent donc qu’il n’y ait qu’une seule marque de focalisation « étroite », la séquence / B^*H_P /, et qu’une règle générale, donnée en 0, efface le ton H_P devant tout autre ton haut, dont le ton H_I caractéristique des questions. L’existence d’une frontière – sous-jacente – de groupe phonologique P entre le B^* et le H_I est indépendamment confirmée par le blocage de processus d’assimilation. La règle (6) est en quelque sorte une règle de Meeussen, dont le domaine a été élargi à l’intonation et le sens d’application inversé, mais le but est toujours d’effacer un ton pour que la forme de surface respecte le PCO.

(6) *Effacement du H_P devant un ton H provoqué par le PCO (Hayes & Lahiri 1991:76) :*

$$H_P \rightarrow \emptyset / _ H$$

Hayes & Lahiri (1991) voient dans le PCO une contrainte probablement universelle, mais dont les effets varient d'une langue à l'autre. Ils remarquent en effet que, même en anglais, le PCO pourrait agir, mais à un niveau différent de celui du bengali : Beckman & Pierrehumbert (1986) suppriment l'accent mélodique H*+H proposé par Pierrehumbert (1980) ; ne restent que les accents H*, B*, H*+B, H+B*, B*+H et B+H*. L'anglais obéirait ainsi au PCO, mais au niveau de l'accent mélodique seulement. Le bengali, quant à lui, est un cas extrême où le PCO affecte tous les tons à l'intérieur de la phrase Hayes & Lahiri (1991:74).

Les diverses règles présentées en 2.2 ainsi que l'intonation du bengali, voire celle de l'anglais, illustrent la pression que peut exercer le PCO pour instaurer de l'alternance mélodique à des degrés et niveaux divers selon les langues. Or, le propre du CDP est précisément de créer de l'alternance mélodique, en enchaînant les contours montants et descendants au sein d'un domaine donné. On se posera alors tout naturellement la question suivante : le français serait-il un autre cas de langue où s'exercerait le PCO, sous la forme du CDP, restreignant ainsi les configurations tonales possibles générées par une grammaire à états finis ? C'est ce que nous allons évaluer dans la section suivante.

3. PCO vs. CDP

3.1. Le CDP comme règle réparatrice ?

En quoi plus précisément consisterait la relation entre PCO et CDP ? Le CDP étant un « mécanisme », on le considérera comme une règle de « réparation » s'appliquant à des éléments qui désobéissent au PCO à l'intérieur d'un domaine donné ; il s'agit alors de voir si les formes engendrées par le CDP peuvent être réinterprétées systématiquement à l'aune du PCO.

Supposons que dans un SN tel que [le frère de Max], en position **prénucléaire**, chaque syllabe accentuée soit associée à un ton H* (représentant le contour [+Montant]). Admettons que dans la grammaire de certains locuteurs du français, ce type de syntagme forme un **domaine prosodique** « DP » au sein duquel le PCO doit être respecté ; la séquence en 0 ci-dessous se trouve donc interdite pour ces locuteurs (les syllabes accentuées sont soulignées)¹.

(7) ☹ H* H*
 [le frère de Max]_{DP}

Le CDP serait la règle de « réparation » donnée en (8), énonçant qu'un ton H* est transformé en un ton B* (i.e. en un contour [-Montant]) devant un H* à l'intérieur du DP en question ; d'où la nouvelle séquence en (9), conforme au PCO. On aura remarqué la similitude de la règle du CDP en (8) avec la règle (6) d'effacement de H_P en bengali : la différence essentielle concerne la spécification d'un domaine d'application.

(8) *Réinterprétation du CDP*
H* → B* / [__ H*]_{DP}

¹ On objectera que cette séquence est attestée en français, c'est vrai, mais là n'est pas la question. Nous nous focalisons ici sur les locuteurs qui respectent le contraste de pente et tentons de déterminer leur grammaire tonale à l'aune du PCO.

- (9) ☺ B* H*
 [le frère de Max]_{DP}

Par contre, si on fait suivre le SN [le frère de Max] par le SV [a mangé] par exemple, la succession des deux tons H*, le premier associé à « Max », le second à « mangé », ne pose pas de problème car chaque syntagme forme son propre DP :

- (10) B* H* H*
 [le frère de Max]_{DP} [a mangé]_{DP}...

On se référera aux travaux de Philippe Martin et à tous les autres qui se sont occupés de l'interface syntaxe/prosodie en français pour définir plus rigoureusement le domaine prosodique dans lequel agirait le PCO. Les DP peuvent aussi être réajustés selon le débit, des contraintes de type rythmique, notamment afin d'équilibrer le nombre de syllabes dans les différents DP (Dell 1984, Martin 1989) ou éviter les rencontres accentuelles (« clash accentuel »). Par exemple, en (11)a, pour équilibrer le nombre de syllabes entre chaque domaine, le locuteur peut regrouper le SN sujet « Max » et le SV « a mangé » dans un seul domaine, donnant (11)b ; (11)b, enfreignant le PCO sous (8), s'applique pour aboutir à (11)c :

- (11) H* H* H*
 a. [Max]_{DP} [a mangé]_{DP} [les tartines]_{DP}...
 → ☹ **H*** **H*** **H***
 b. [Max a mangé]_{DP} [les tartines]_{DP}
 → ☺ **B*** **H*** **H***
 c. [Max a mangé]_{DP} [les tartines]_{DP}

Le CDP comme règle de réparation opère également dans le DP nucléaire. Considérons la phrase « C'est le frère de Max. », réponse à la question « Qui est venu ? » et admettons que « C'est le frère de Max » se termine par un ton B* nucléaire marquant le focus dans une assertion. Dans ce cas, le ton H* associé à « frère » reste tel quel car il n'enfreint pas le PCO :

- (12) ☺ H* B*_{NUCL}
 [c'est le frère de Max .]_{DP}

Par contre, dans le cas d'une question totale telle que « C'est le frère de Max ? », si on analyse la montée finale comme un H* nucléaire – dont la hauteur varie lors de l'implémentation phonétique, alors on se retrouve dans la situation donnée en (13)a, qui désobéit au PCO. le H* initial associé à « frère » est transformé en B* selon la règle (8), donnant la configuration tonale en (13)b, que l'on observe très souvent dans les questions.

- (13) a. ☹ H* H*_{NUCL}
 [c'est le frère de Max ?]_{DP}
 b. ☺ B* H*_{NUCL}
 [c'est le frère de Max ?]_{DP}

L'accent ou ton haut initial (« Hi ») semble aussi avoir une influence sur un H*. Jun & Fougeron (2002:155-7) discutent du patron [...B*][Hi...] dont on trouve un exemple *Figure 2* ci-dessous : le H* final de leur « Accentual Phrase » [le garçon] est remplacé ici par un B* (« L* ») ; les auteures expliquent cela en invoquant une contrainte empêchant la succession de trois tons hauts (« AVOIDHHH »). En fait, on peut tout-à-fait avoir un B* sans Hi qui précède. En remarquant que ce remplacement a tout l'air d'opérer à l'intérieur d'un syntagme – comme c'est le cas ici –, il suffit d'amender la règle (8) sous la forme plus générale donnée en (14), énonçant que H* devient B* devant **tout** H dans un DP.

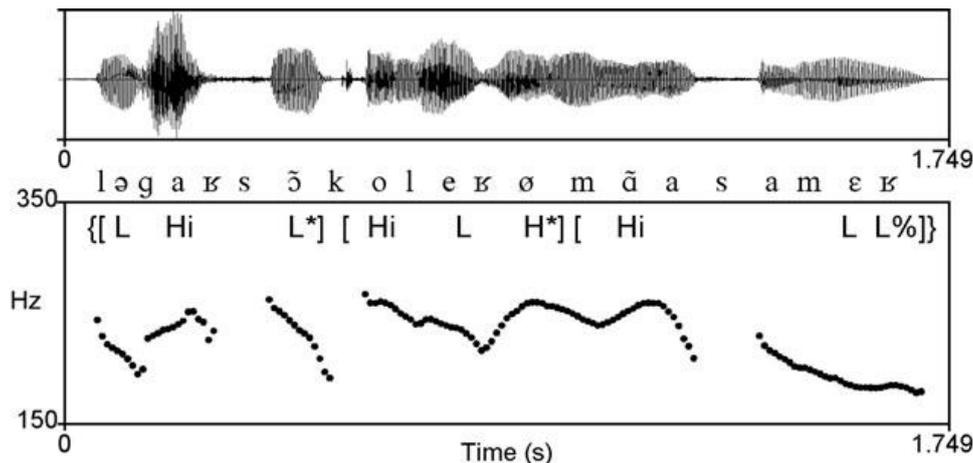


Figure 2: Ton L* (=B*) entre deux Hi dans le syntagme « le garçon coléreux ». (Jun & Fougeron (2002:155; figure 4.(d)).

- (14) Règle du CDP (2) :
 $H^* \rightarrow B^* / [_ H]_{DP}$

La règle (14), très similaire à la règle (6) du bengali, étaye ainsi l'idée selon laquelle le CDP fonctionne comme une règle motivée par un principe purement phonologique, le PCO. Cependant, en approfondissant quelque peu l'analyse, on s'aperçoit en réalité que le ton Hi met en difficulté cette hypothèse.

En effet, si on suppose, comme nous l'avons fait jusqu'à maintenant, que les tons bas autres qu'accentuels – i.e. les tons « L » sur « le », « -lé- » et « mer » – sont des tons purement phonétiques, alors le Hi et le H* qui suit sur « coléreux » sont adjacents au niveau phonologique. Pourquoi dans ce cas, le PCO n'est-il pas respecté ?

On peut arguer que le PCO (et partant la règle (14)) ne concerne pas les Hi, ce dernier ayant un statut particulier. Mais cela ne correspond pas à ce qui est attendu du PCO, lequel touche **tous** les tons d'une même catégorie : en bengali, H*, H_P ou H_I, qui sont tous des tons **hauts**, ne peuvent jamais être adjacents ; *idem* pour B* et B_I (il n'y a pas de B_P).

Une autre façon de résoudre le problème du Hi est de le placer sur un autre niveau (« tier ») phonologique que les tons T*. Cette hypothèse mettrait le Hi hors d'atteinte du PCO, et permettrait de recouvrir l'adjacence entre les deux tons H* pour que la règle (8) s'applique. La théorie MA ne prévoit cependant pas que les tons soient sur des tiers différents, et les motivations manquent pour cela, si ce n'est pour préserver l'adjacence entre les accents mélodiques T*...

Si le PCO fait pression en français, il agit donc de manière atypique, voire suspecte. L'alternative est finalement de rejeter l'idée selon laquelle le CDP est une instanciation du PCO, malgré les arguments avancés jusqu'à maintenant. En fait, le problème que pose le Hi n'est pas le seul ; faire du CDP une règle réparatrice entraîne d'autres difficultés comme nous allons le voir.

3.2. CDP, PCO et principe de localité

Le problème du Hi a mis en avant une condition nécessaire à l'application du PCO : il doit porter sur des éléments **adjacents** à un niveau donné. Tous les exemples de processus que nous avons évoqués en 2.2 s'appliquent à des éléments adjacents ; en bengali, la pression du PCO s'effectue de proche en proche sur tous les tons d'une phrase, mobilisant la règle 0 quand le contexte l'impose.

L'adjacence n'est pas une propriété anodine, elle est le corollaire d'un principe fondamental de la phonologie, le principe de **localité**, selon lequel tout processus phonologique opère sur des éléments « en contact » à un niveau donné. Ainsi les phénomènes d'assimilation et de dissimilation concernent typiquement des éléments qui se trouvent en contact, soit au niveau de leur représentation phonétique, soit au niveau de leur représentation sous-jacente (cf. De Lacy 2007, Alderete & Frisch 2007, Archangeli & Pulleyblank 2007, et Frisch et al. 2004). Le principe de localité sous-tend également la théorie de l'hypercorrection de Ohala (1981) sur l'évolution des sons (« sound change »), et constitue un des postulats fondamentaux de certaines théories phonologiques, comme la phonologie du gouvernement (cf. le « principe de localité stricte », Scheer 2004).

En interprétant le CDP comme une règle contextuelle telle que (8), le principe de localité est respecté : les tons affectés par cette règle sont adjacents au ton déclencheur, comme dans n'importe quel autre processus déclenché par le PCO. Cependant, si on décide de rejeter l'hypothèse selon laquelle le CDP s'applique sous la pression du PCO, la règle (8) et l'adjacence qu'elle implique seraient illusoire et pourraient même constituer un obstacle pour comprendre l'intonation du français, voire celle d'autres langues.

Dans sa définition originelle, le CDP est la manifestation des relations de dépendance qu'entretiennent les mots prosodiques au sein d'une structure hiérarchisée, le CDP ne concerne donc pas forcément des éléments adjacents. Dans la suite de cet article, nous allons présenter des faits intonatifs issus du français et du grec moderne qui ne peuvent être expliqués que par le CDP en tant qu'il exprime ce type de relations et de structure, et montrer que le PCO se révèle insuffisant.

3.3. CDP et structure de dépendance

3.3.1 Intonation du français. Soit le petit dialogue suivant :

- (15) a. Tiens au fait, les enfants de Paul, qu'est-ce qu'ils ont eu ?
 b. Et bien, Valérie, elle a eu la ferrari, Noémie, la librairie, et Amélie, la boulangerie.

Dans la réponse (15)b., les SN « la ferrari », « la librairie » et « la boulangerie » sont des focus, éléments de réponse à la question (15)a., auxquels sont joints les SN topiques « Valérie », « Noémie » et « Amélie » ; on a ainsi une succession de trois séquences Topique + Focus, dont un

exemple de courbe de f0 est donné *Figure 3* ci-dessous, prononcé par une locutrice (cf. Le Gac & Yoo 2007).

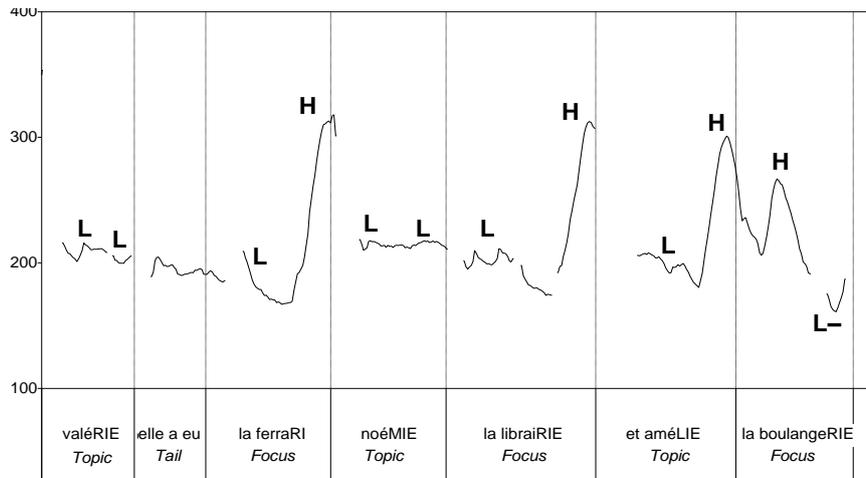


Figure 3: tracé de f0 prononcé par une locutrice d'un énoncé comportant une succession de trois groupes Topique+Focus.

Ce qu'il y a de remarquable chez les locuteurs qui réalisent ces patrons mélodiques, c'est le contraste tonal entre les deux premiers topiques et le dernier : celui-ci se termine par la montée typique des topiques (cf. entre autres, Rossi 1999, Di Cristo 1998), par contre, les deux premiers topiques se terminent par un ton bas. On s'aperçoit que les tons des topiques **dépendent** directement du ton focal subséquent : le focus final porte la chute terminale et le ton du topique qui le précède est haut ; les deux premiers focus portent un ton haut, les topiques sont bas.

Peut-on mettre cela sur le compte du PCO ? A savoir, les topiques 1 et 2 voient leur ton H* devenir B* selon la règle (8). La réponse est non, car cela signifierait que les topiques forment avec le focus subséquent un seul DP : cela contredit l'idée largement acceptée selon laquelle les topiques sont typiquement « détachés » par rapport au reste de la phrase, et va à l'encontre de la formation des DP que nous avons vue plus haut, où le SN sujet, le SV et le SN complément forment chacun leur DP. Le fait que les topiques et les focus constituent chacun un DP est en outre démontré par la possibilité d'avoir deux tons H successifs en fin de phrase, le ton H* du dernier topique suivi du ton Hi marquant la frontière gauche du focus. Topiques et focus sont donc bien dans deux DP différents ; l'explication par une relation de **dépendance** entre le ton T* des topiques et celui du focus est la plus adéquate.

Par ailleurs, le ton bas du premier topique est suivi dans notre exemple par le SV « elle a eu », lequel forme un DP se terminant par un autre ton bas, i.e. on a une séquence [B*]_{TOP} [B*]_{SV} [H*]_{FOC}. Cela signifie que le [B*] du topique n'est en fait même pas nécessaire du point de vue du PCO puisque un autre ton B* le sépare du H* focal, et met en évidence que la dépendance tonale du topique se fait « par-dessus » un autre ton. On fera une observation similaire pour le ton H* du dernier topique, lequel dépend du ton terminal nonobstant le ton haut initial du focus.

Bref, le PCO ne peut rendre compte de ce type d'énoncé. Les principes d'adjacence et de localité sont enfreints. Par contre, une structure de dépendance comme celle donnée *Figure 4* ci-dessous permet de dériver directement les tons observés.

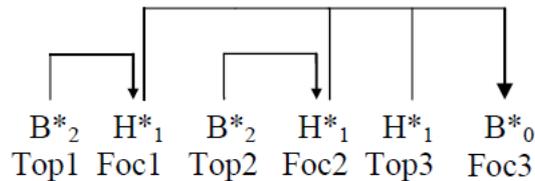


Figure 4 : relations de dépendance entre les tons finaux des topiques et des focus de la phrase de la Figure 3.

Dans cette structure, de la même manière que l'on fait dépendre le ton **haut** final du dernier topique (« Top3 ») du ton bas terminal (B^*_0), on fait également dépendre de ce dernier les tons hauts des deux premiers focus (« Foc1 et 2 ») ; on rend compte ainsi de l'homomorphie mélodique entre le dernier topique et les focus non terminaux. Ces montées mélodiques, traditionnellement analysées comme des montées de continuité, sont réinterprétées comme les manifestations d'une relation de dépendance au ton terminal B^*_0 .

Pour montrer que l'on a bien affaire à des relations de dépendance entre les tons, on transformera ce type de phrase en **question**, test proposé par Martin (1981). En effet, comme on le voit à la Figure 5, le focus final est associé maintenant à un ton haut final marquant la question totale ; le dernier topique ainsi que les deux premiers focus sont marqués par un ton bas, exactement ce que permettent de prédire le CDP et la structure Figure 4.

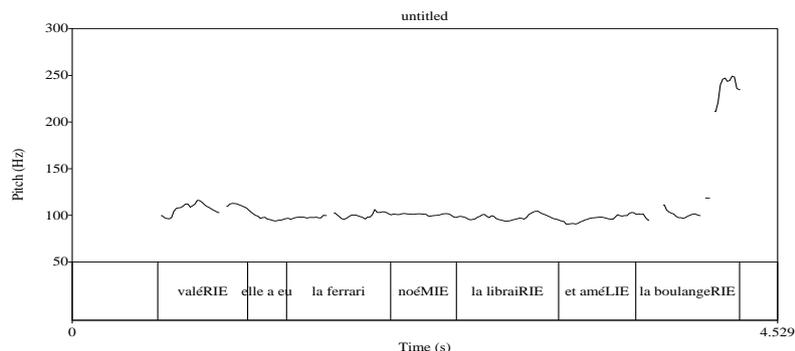


Figure 5 : tracé de f_0 de la phrase (15)b. prononcée comme une question totale. Le dernier topique « Amélie » ainsi que les deux premiers focus, « ferrari » et « librairie » sont maintenant marqués par un ton B^* .

On remarquera que, dans la question, les tons des deux autres topiques, de niveau 2 dans la hiérarchie Figure 4, restent bas alors que l'on s'attendrait à ce que ces tons soient hauts en vertu du CDP. Ici, il suffit de poser une règle stipulant que tout ton de niveau 2 ou inférieur ($T_3, T_4...$) dans la hiérarchie peut se réaliser bas (voir Le Gac & Yoo 2007 à ce propos), autrement dit le CDP n'opère plus à partir de ce niveau, ou tout du moins est facultatif.

(16) Tout ton de niveau 2 ou inférieur ($T_3, T_4...$) se réalise B^*

Cette configuration mélodique globale est typique des questions polaires chez certains locuteurs ; elles se caractérisent par une succession de tons bas adjacents que ne peut expliquer le PCO mais

dont rend compte le CDP comme marqueur relationnel dans une structure de dépendance, jouant de concert avec la règle (16).

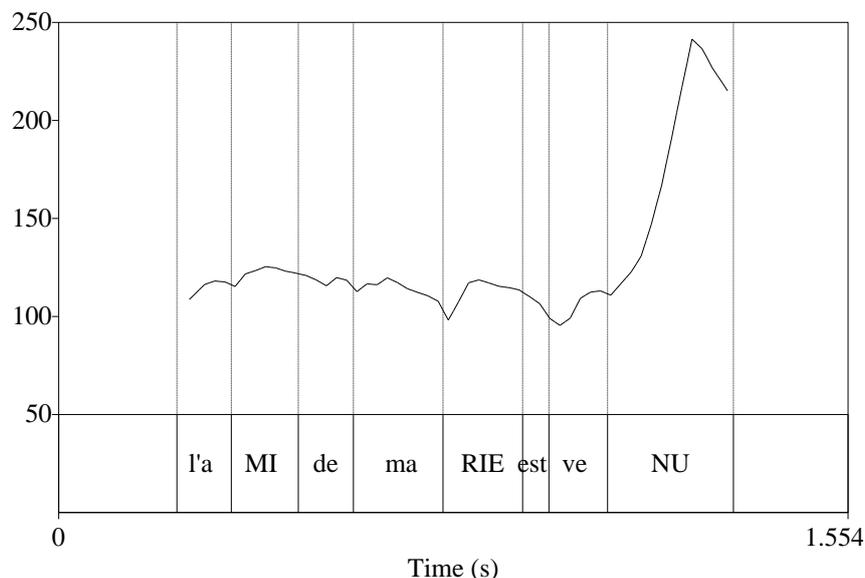


Figure 6: autre exemple de question totale avec une succession de deux tons B*. Si le ton B* sur « marie » est prévu en raison du H*_o final et du CDP, le B* sur « ami », dépendant de celui de « marie », ne l'est pas, on s'attend à un H* : la règle (16) s'applique.

On pourra objecter comme on l'a souvent fait pour le CDP, que ce dernier n'est pas la règle générale du français. Il n'existe néanmoins pas à notre connaissance d'étude empirique et systématique visant à déterminer la mesure dans laquelle le CDP s'applique en français, notamment en prenant en compte des facteurs sociolinguistiques et/ou le style de parole.

En français, certains processus phonologiques comme la « loi de position », l'antériorisation ou l'harmonie vocalique s'appliquent à des degrés très divers selon les variétés du français, les situations, voire les locuteurs au sein d'une même variété (cf. Detey et al. 2010).

Nous pensons qu'il en est de même pour le CDP : s'il ne concerne pas **toutes** les variétés de français, il ne fait aucun doute que le CDP appartient à la grammaire d'au moins certains locuteurs du français, et c'est cela qui importe le plus.

En effet, qu'il n'y ait que certains locuteurs qui réalisent le CDP, ne retire rien de l'intérêt de cette règle. Au contraire, pour traiter des faits dont le CDP fournit la meilleure explication, comme nous venons de le voir, c'est la grammaire intonative de ces locuteurs qui pose problème à des approches plus « conventionnelles » ; de ce point de vue, la charge de la preuve revient naturellement aux autres théories.

Qui plus est, en le transposant dans le cadre de la phonologie tonale, le CDP ne rend pas seulement compte de faits intonatifs du français mais aussi de ceux d'une autre langue, à accent lexical celle-ci, le grec moderne.

3.3.2 Grec moderne. Baltazani & Jun (1999), Le Gac & Yoo (2002) et Le Gac & Yoo (2011) ont montré que le grec moderne connaît une inversion tonale entre les phrases déclaratives et les questions, aussi bien sur le focus que sur les topiques ; cela est montré dans le *Tableau 4* et la *Figure 7* ci-dessous :

	Topiques	Focus	Post-Focus
Déclaratives	(B*+H) B* H%	H*	B- B%
Qu. Polaires	(B*+H) H* B%	B*	H- B%

Tableau 4: configurations tonales en grec moderne dans les déclaratives et les questions polaires

On observe que l'accent mélodique T* du focus, l'accent nucléaire, est l'inverse de celui du topique ; s'il y a plusieurs topiques, la séquence (B*+H) T* T% est répétée, avec le même schéma mélodique (cf. *Figure 7*).

Une structure où les mélodies accentuelles entretiennent des relations de dépendance comme celle que nous avons proposée pour le français rend immédiatement compte de ces patrons accentuels et de leur itération : chaque T*₁ de topique dépend directement du T*₀ focal, et est inversé par ce dernier.

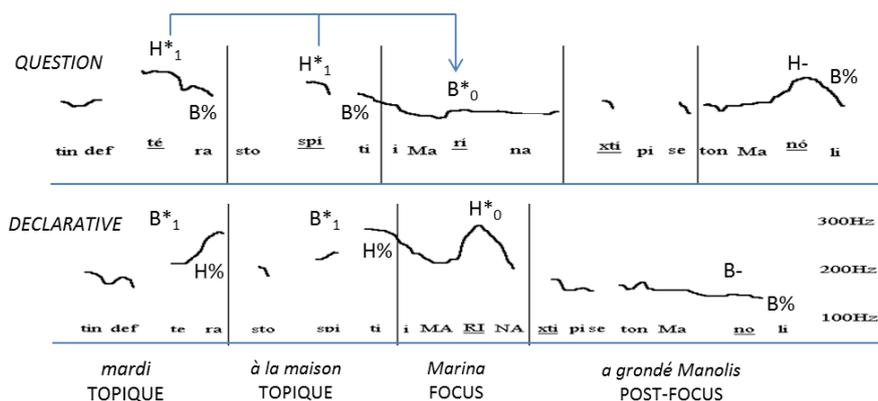


Figure 7 : tracés de f0 et structure tonale d'une question totale et d'une déclarative en grec moderne, comportant deux topiques, suivis du focus et d'un post-focus.

Quant à la séquence B*+H, il s'agit de l'accent pré-nucléaire « par défaut » en grec (Baltazani & Jun 1999, Arvaniti & Baltazani 2000) ; il est situé au niveau 2 dans la structure. A l'instar du français, il y a sélection à ce niveau d'un seul et même accent mélodique, le CDP n'opère plus.

Il est impossible en grec de rendre compte de ces inversions tonales avec le PCO : les tons de frontière T% des topiques jouent le rôle de barrière, ils rompent l'adjacence entre le T* focal et les T* des topiques.

Dans notre article (Le Gac & Yoo 2011), nous avons également voulu savoir si les contours des topiques ne dépendaient pas de facteurs sémantiques comme le degré d'activation, ou encore de l'emplacement de l'accent lexical du topique : les résultats ont confirmé que ces facteurs n'entraient pas en jeu et que les inversions ne pouvaient être pilotées que par un mécanisme phonologique, à savoir l'équivalent du CDP transposé dans un cadre tonal.

Le rôle du CDP est encore confirmé par les données présentées par Baltazani & Jun (1999) : les patrons mélodiques de parenthétiques et d'éléments insérés dans des listes exhibent les mêmes inversions quand on passe d'une déclarative à une question, et ce, de la même manière qu'en français.

3.3.3 En bref. Dans cette section, nous avons montré qu'un certain nombre d'inversions mélodiques du français et du grec ne peuvent être expliquées par le PCO ; une structure de dépendance, dans laquelle les tons ne sont pas forcément adjacents, se révèle nécessaire. Le CDP n'est pas la manifestation du PCO mais bien celle de relations de dépendance, insensibles au principe de localité.

Ce type de structure a mis au jour des similitudes entre français et grec : le CDP opère sur les tons T_1 qui dépendent directement du T_0 ; pour les tons T_2 , le CDP semble ne plus opérer dans les deux langues, et un accent tonal par défaut est sélectionné, B^* en français et B^*+H en grec.

Nous terminerons cet article en montrant qu'il est néanmoins possible, grâce à ce type de structure de dépendance, d'unifier PCO et CDP, et que ces derniers peuvent fonctionner de concert dans les langues.

4. PCO et CDP : l'union fait la force

D'une manière générale, les relations de dépendance hiérarchique et de précedence / succession linéaire peuvent être **médiates** ou **immédiates**. Dans la structure de dépendance *Figure 8*, T_1 dépend immédiatement de T_0 , T_2 de T_1 ; par contre, si T_2 dépend de T_0 par transitivité, cette relation est médiate, elle s'effectue **via** un autre élément, i.e. T_1 . De même, sur l'axe temporel, l'**adjacence** est la relation qui désigne la précedence (ou la succession) linéaire immédiate entre T_1 et T_0 et T_2 et T_1 ; ton T_2 , quant à lui, précède médiatement T_0 .

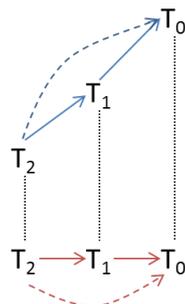


Figure 8 : Relations de dépendances hiérarchiques et de précedence linéaire immédiates (flèches pleines) et médiate (flèches en pointillées) entre les tons.

En considérant les choses d'un point de vue strictement relationnel, on pourra donc définir un principe d'alternance obligatoire, valable à la fois pour le PCO et le CDP :

(17) *Principe d'Alternance Obligatoire (PAO) :*

A un niveau donné, deux éléments identiques en relation immédiate sont interdits

Comme le PCO, le PAO peut être considéré comme un paramètre ou une contrainte de la grammaire universelle, qui « filtre » les configurations mélodiques générées par une grammaire tonale telle que celle de Beckman & Pierrehumbert (1986). Selon les langues, on spécifiera i) le **type** de relation directe qui entre en jeu, dépendance ou précédence/succession ; et ii) le **niveau** auquel le PAO agit.

Dans une langue comme le bengali, seule la relation de précédence est spécifiée pour le PAO, i.e. seul le PCO est actif, et ce, entre tous les tons de la phrase (le domaine est la phrase entière). En français, il semble que seule la relation directe de dépendance contraigne les sélections tonales : en effet, nous avons vu que deux tons H, le H_i et le H^* , peuvent apparaître ensemble dans un même DP, de même que deux B^* .

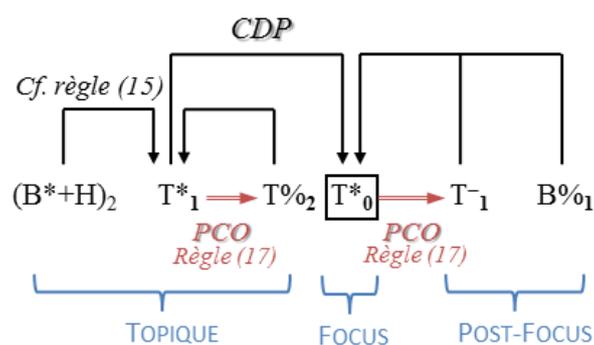
Y a-t-il des langues où sont activées les deux relations directes à la fois ? Le grec moderne pourrait être l'une de ces langues. Tout d'abord, selon l'inventaire des accents mélodiques dressé par Arvaniti & Baltazani (2000), le grec n'a pas d'accents mélodiques constitués de deux tons T^* identiques (H^*+H par exemple). En suivant la suggestion de Hayes & Lahiri (1991) à propos de l'anglais, on pourrait donc dire que le PCO élimine également ce type d'accents bi-tonaux en grec.

Ensuite, nous avons laissé en suspens les tons T^- et $T\%$. Or, en revenant au *Tableau 4*, on observe que ces tons sont systématiquement l'inverse du ton T^* qui les précède immédiatement, sauf en ce qui concerne le $B\%$ terminal : par exemple, le ton H^- final des questions polaires est précédé par le B^* focal, le $H\%$ des topiques est précédé par un B^* , etc.

En d'autres termes, on peut soutenir que les tons T^- et $T\%$ sont gérés par une règle de polarité tonale pour ne pas transgresser le PCO, comme celle que nous avons vue en 2.2 pour le konni (*Tableau 1*), avec pour ton déclencheur, le T^* précédent :

$$(18) \alpha T \rightarrow -\alpha T / \alpha T^* _$$

La *Figure 9* résume l'ensemble des relations et des règles que nous avons mises au jour pour les tons du grec moderne :



*Figure 9: les tons et leurs relations en grec moderne. Les flèches noires indiquent les relations de dépendance tonale. Les indices des tons marquent leur niveau dans la hiérarchie. On voit qu'en grec, le CDP ne s'applique qu'entre le T^*_0 et un T^*_1 situé à sa gauche ; au niveau 2, un accent B^*+H est inséré par défaut, de la même manière qu'en français (« règle 15 »). Entre un T^* et un $T\%$ ou T^- , c'est le PCO qui intervient avec la règle (18). Quant au $B\%$ terminal, il n'est concerné ni par le PCO ni par le CDP (cf. (iii) ci-dessous), et reste ainsi constant.*

On pourrait cependant avancer que ces inversions sont le produit d'une dépendance directe entre les tons de frontière T- et T% et les T* qui précèdent (cf. les flèches noires *Figure 9*), mais cela compliquerait les choses :

- i) les T% de topiques sont dans ce cas des tons de niveau 2 ; or à ce niveau, nous avons vu que le CDP est normalement inerte ;
- ii) la règle de CDP ne concerne que les T* et les relations de dépendance vers la **droite**, c'est ce qu'ont révélé les données du français et du grec jusqu'à maintenant ; il semble que ce soit une propriété intrinsèque de cette règle ; or, on a affaire à des tons de **frontière**, qui dépendent du T* à **gauche** ;
- iii) la présence constante du B% terminal démontre cette dernière propriété du CDP et le fait que les tons de frontière sont gérés par la règle (18) : si on suppose que ce ton dépend du T* focal qui **précède** (i.e. le B% final est un T₁), il ne peut, par son statut de ton de frontière et sa position, être inversé par le CDP ; la règle (18) ne le concerne pas non plus puisqu'il n'est pas adjacent au T* focal ; hors de portée à la fois du CDP et du PCO, le B% demeure constant, bas par défaut.

Les tons de frontières T% et T-, mis à part le B% terminal, sont donc inversés par la règle (18) sous la pression du PCO ; en grec, PCO et CDP sont donc bel et bien actifs. Quant au ton B%, Ladd (2008:284-5) en fait un ton « hors structure » en raison de son comportement particulier. En réalité, nous avons montré qu'il était structurellement intégré dans la phrase mais que son statut et sa position le mettaient hors de portée et du PCO et du CDP.

5. Conclusion

Le CDP et le PCO reposent tous les deux sur le principe selon lequel des éléments phonologiques doivent alterner dans un domaine donné. Nous nous sommes donc demandé dans cet article dans quelle mesure et comment on pouvait rapprocher le CDP et le PCO.

En nous appuyant sur un certain nombre d'alternances mélodiques observées en français et en grec moderne, nous avons montré que le PCO et le CDP étaient deux principes **irréductibles** l'un à l'autre mais complémentaires ; ils constituent deux facettes d'un principe plus général, le Principe d'Alternance Obligatoire (« PAO ») qui interdit que deux éléments en relation immédiate soient identiques à un niveau donné.

Ce principe est paramétré différemment selon les langues - et leurs variétés -, si bien qu'en bengali, seul le PCO est actif, et dans certaines variétés de français, le CDP ; en grec, cependant, les deux jouent de concert, et nous avons pu ainsi expliquer la constance problématique du ton B% final. Une fois paramétré, ce principe fonctionne comme une contrainte qui sélectionne les séquences tonales engendrées par une grammaire à états finis comme celle proposée par Pierrehumbert (1980) par exemple.

Nous espérons au terme de cet article avoir réussi à montrer tout l'intérêt que pouvait avoir le CDP et à travers lui, l'utilisation d'une structure de dépendance pour rendre de compte du fonctionnement de l'intonation. Pour ce faire, le CDP n'est toutefois pas seul, le PCO, principe directement issu de la théorie autosegmentale, peut lui apporter une aide précieuse. Ainsi réunis, il y

a fort à parier que l'intonation d'autres langues que le français ou le grec trouvera un éclairage nouveau.

Bibliographie

- Alderete, J. D. & S. Frisch (2007). Dissimilation in Grammar and the Lexicon. De Lacy, P. (dir.), *The Cambridge Handbook of Phonology*. Cambridge University Press, pp. 379-398.
- Archangeli, D. & D. Pulleyblank (2007). Harmony. De Lacy, P. (dir.), *The Cambridge Handbook of Phonology*. Cambridge University Press, pp. 353-378.
- Arvaniti, A. & M. Baltazani (2000). GREEK ToBI: a System for the Annotation of Greek Speech Corpora. *Proceedings of the Second International Conference on Language Resources and Evaluation (LREC 2000)*, vol. II, European Language Resources Association, pp. 555-562.
- Baltazani, M. & S-A. Jun (1999). Topic and Focus Intonation in Greek. *Proceedings of the 14th International Conference for the Phonetic Sciences*, vol. 2, San Fransisco, CA, pp. 1305-1308.
- Beckman, M. & J. Pierrehumbert (1986). Intonational Structure in Japanese and English. *Phonology* 3, pp. 255-309.
- Dell, F. (1984). L'accentuation dans les phrases en français. Dell F., D. Hirst & J.-R. Vergnaud, *Formes sonores du langage*, Hermann, Paris, pp. 65-122.
- Detey, S., J. Durand, B. Laks, B. & Ch. Lyche, dirs., (2010). *Les variétés du français parlé dans l'espace francophone*. Ophrys, Paris.
- Di Cristo, A. (1998). Intonation in French. Hirst D. & A. Di Cristo (dirs.), *Intonation Systems: A Survey of Twenty Languages*. Cambridge University Press, pp. 195-218.
- Frisch, S., J. Pierrehumbert, J. & M. Broe (2004). Similarity Avoidance and the OCP. *Natural Language & Linguistic Theory* 22:1, pp. 179-228.
- Gussenhoven, C. (2004). *The Phonology of Tone and Intonation*. Cambridge University Press.
- Hayes, B. & A. Lahiri (1991). Bengali Intonational Phonology. *Natural Language and Linguistic Theory* 9:1, pp. 47-96.
- Jun, S-A. & C. Fougeron (2002). Realizations of Accentual Phrase in French Intonation. *Probus* 14:1, pp. 147-172.
- De Lacy, P. dir. (2007). *The Cambridge Handbook of Phonology*. Cambridge University Press.
- Ladd, D. (2008). *Intonational Phonology*. 2nd ed. Cambridge University Press.
- Leben, W. (1973). *Suprasegmental Phonology*. PhD Thesis, MIT.
- Le Gac, D. & H-Y.Yoo (2002). Intonative structure of focalization in French and Greek. Beyssade C. et al. (dirs.), *Romance Languages and Linguistic Theory 2000*, J. Benjamins, pp. 213-231.
- Le Gac, D. & H-Y.Yoo (2007). For a dependency theory of intonation. *Proceedings of the 16th International Conference for the Phonetic Sciences*, Saarbrücken, Allemagne.
- Le Gac, D. & H-Y.Yoo (2011). Intonation of left dislocated topics in Modern Greek. *Proceedings of the Interspeech Conference*, Florence, Italie.
- Martin, Ph. (1981). Pour une théorie de l'intonation. Rossi M., A. Di Cristo, D. Hirst, Ph. Martin & N. Yukihiro (dirs.) *L'intonation : de l'acoustique à la sémantique*. Klincksieck, Paris, pp. 234-271.

- Martin, Ph. (1982). Phonetic Realisations of Prosodic Contours in French. In *Speech Communication I*. North-Holland Publishing Company, pp. 283-294.
- Martin, Ph. (1989). La prosodie. Calliope, *La parole et son traitement automatique*. Dunod, Masson, Paris, pp. 131-146.
- Martin, Ph. (1999). Intonation et syntaxe des langues romanes. *Atti macrosintassi e pragmatica : l'analisi linguistica del parlato*. Firenze, Italia.
- Martin, Ph. (2009). *L'intonation du français*. Armand Colin, Paris.
- Myers, S. (1997). OCP Effects in Optimality Theory. *Natural Language & Linguistic Theory* 15:4, pp. 847-892.
- Odden, D. (1986). On the Role of the Obligatory Contour Principle in Phonological Theory. *Language* 62:2, pp. 353-383.
- Ohala, J. (1981). The Listener as a Source of Sound Change. Masek C., R. Hendrick & M. Miller (dirs.), *Papers from the Parasession on Language and Behavior*. Chicago Linguistics Society, Chicago, pp. 178-203.
- Pierrehumbert, J. (1980). *The Phonology and Phonetics of English Intonation*. PhD Thesis, MIT.
- Pierrehumbert, B. & M. Beckman (1988). *Japanese Tone Structure*. Cambridge MIT Press.
- Post, B. (2000). *Tonal and Phrasal Structures in French Intonation*. Holland Academic Graphics, Nijmegen, The Hague.
- Rossi, M. (1999). *L'intonation, le système du français. Description et modélisation*. Ophrys, Paris.
- Scheer, T. (2004). *A Lateral Theory of Phonology: What Is CVCV, and Why Should It Be?* Mouton de Gruyter, Berlin.
- Yip, M. (1988). The Obligatory Contour Principle and Phonological Rules: A Loss of Identity. *Linguistic Inquiry* 19:1, pp. 65-100.
- Yip, M. (2002). *Tone*. Cambridge University Press.
- Yip, M. (2007). Tone. De Lacy, P. (dir.), *The Cambridge Handbook of Phonology*. Cambridge University Press, pp. 229-251.

PHONÉTIQUE APPLIQUÉE
ET TRAITEMENT DU SIGNAL

Variations sur un peigne

Jean-Sylvain Liénard

jean-sylvain.lienard@limsi.fr

LIMSI-CNRS, bp 133, 91403 Orsay Cedex

Abstract:

Philippe Martin designed the Spectral Comb method for the measurement of the speech fundamental frequency in the late 70's, for the purpose of his prosodic researches. We present some theoretical considerations showing that any F_0 measuring system has to deal with specific types of errors, and we explain how those errors have been contained in Martin's Spectral Comb. We also present a new development of the same technique, using several variant families - the missing teeth combs and the negative teeth combs - in order to specifically cancel the main types of F_0 errors.

1. Introduction

La mesure de la fréquence fondamentale de la parole est un problème très ancien de la phonétique acoustique. Jusqu'aux années 70 on utilisait une mesure directe sur le signal, qui avait le mérite de fonctionner en temps réel mais qui produisait des erreurs. Lorsque Philippe Martin, déjà passionné par l'analyse des contours mélodiques, a présenté sa méthode du peigne spectral, ce fut non seulement un progrès scientifique car la mesure était plus précise et plus robuste qu'auparavant, mais aussi un exploit technique car elle tournait en temps réel sur les microordinateurs des années 80.

Dans cet article nous revenons sur les fondements de la méthode. Celle-ci revient à prélever dans le spectre harmonique du signal une série de valeurs discrètes multiples d'une même valeur F_c , que l'on fait varier dans tout l'intervalle à explorer jusqu'à trouver la meilleure mise en correspondance, attribuée à F_0 . Nous montrons que cette idée, appliquée sans précautions, ne conduit pas à un système utilisable. Philippe Martin a trouvé d'emblée la solution qui permet d'en faire un véritable instrument de mesure, fiable et rapide. Nous présentons également des travaux menés dans le même esprit, qui nous ont conduits à développer des familles de peignes très particuliers dans le but de traiter des mélanges de signaux de parole.

2. Le peigne de Martin dans son contexte

La mesure de la fréquence fondamentale de la voix est un problème difficile, pour de multiples raisons détaillées dans l'ouvrage de référence de Wolfgang Hess (Hess 1983).

Jusqu'à la généralisation des méthodes numériques dans les années 70 on cherchait surtout à isoler le fondamental des harmoniques pour en mesurer la fréquence, avec deux problèmes majeurs: d'une part F_0 peut varier selon les locuteurs et les instants dans un intervalle très vaste (3 à 4 octaves), et d'autre part ce fondamental peut être d'intensité très faible ou franchement absent. Cette stratégie ne fonctionnait que si l'on connaissait d'avance l'ordre de grandeur de F_0 , et si sa variation restait limitée.

On savait pourtant, par l'observation des spectrogrammes en bande étroite et par des considérations auditives ou mathématiques (Schroeder 1968), que l'intervalle entre harmoniques, par définition égal à F_0 , pouvait constituer une meilleure base d'investigation. On savait aussi que l'autocorrélation était plus proche de la définition de la périodicité que la recherche du fondamental. Mais ces deux approches ne pouvaient pas donner naissance à de nouveaux systèmes d'estimation de F_0 avant que les technologies numériques ne soient accessibles pratiquement aux chercheurs en phonétique ou en traitement du signal.

Philippe Martin, alors ingénieur dans un laboratoire de phonétique, avait déjà mis au point un système de détection de F_0 , avec les technologies électroniques de l'époque (Martin 1973). Il en connaissait les défauts, mais il appréciait aussi l'avantage du temps réel, qui permet au phonéticien de voir immédiatement le résultat de la mesure et de travailler en interaction avec son instrument. Il était aussi attiré par les nouvelles techniques numériques et s'était initié à la programmation en langage assembleur, au plus près du système.

C'est ainsi que son travail l'amène à mettre au point, dans la fin des années 70, un système révolutionnaire à plusieurs égards. Ce système fonctionne sur des segments de signal de durée 32 ms prélevés toutes les 10 ms. Voici comment il en décrit le principe, le traitement étant appliqué à chaque prélèvement (ou trame) de signal (Martin 1981): "Le principe de la méthode du peigne réside dans la recherche de valeurs du spectre situées à des fréquences harmoniques et dont la somme soit maximale pour un intervalle fréquentiel donné. L'intercorrélation peigne-spectre revient au calcul de la somme des composantes spectrales correspondant à une structure harmonique donnée, dont la fondamentale varie dans une gamme prédéterminée (par exemple 70 à 1000 Hz)."

Dans ce même article Philippe Martin mentionne plusieurs points techniques, dont le plus important est d'assurer une décroissance des dents du peigne. Un autre point se trouve dans le prétraitement du spectre, visant à éliminer toute énergie autre que celle des harmoniques. Le résultat est parmi les meilleurs de l'époque: le système peut traiter la voix de divers locuteurs, dans des conditions difficiles (bruit, voix téléphonique, fondamental absent). Cette version rédigée en Fortran tourne sur un miniordinateur en 200 fois le temps réel.

L'année suivante, il présente une évolution du système, notamment en ce qui concerne la détection du voisement, et des résultats d'évaluation plus complets (Martin 1982), incluant une comparaison avec la méthode du cepstre, très en vogue à l'époque. Le peigne donne de meilleurs résultats sur tous les plans: moins d'erreurs sur divers signaux de parole, capacité à traiter des signaux réduits à une fréquence pure ou enregistrés via le canal téléphonique, résistance au bruit, rapidité.

Au fil des années qui suivent, il perfectionne sa méthode et en fait le coeur d'un logiciel d'analyse de la parole dénommé WinPitch, distribué par la société Pitch Instruments (Martin 1996). S'adressant à des phonéticiens, des chanteurs, des pédagogues de la voix, le fonctionnement en temps réel est indispensable, et c'est un tour de force qu'il accomplit, du point de vue de la programmation, pour faire tourner ce logiciel sur les micro-ordinateurs grand public, avec une interface graphique sophistiquée.

Depuis Philippe Martin a introduit dans les versions successives de WinPitch de nombreuses améliorations de son peigne de base, comme la brosse (Martin 2000), qui consiste à aligner plusieurs spectres successifs sur le premier d'entre eux, de façon à rendre plus robuste la détection locale de F_0 . Ce principe a été développé récemment pour déterminer une trajectoire locale optimale

reposant sur l'alignement de cinq spectres successifs (Martin 2008). Très récemment il a mis au point une approche multi-algorithmes (Martin 2012), qui vise à intégrer à WinPitch plusieurs algorithmes (en complément du Peigne et de la Brosse), chacun pouvant être meilleur que les autres dans des conditions particulières ou sur certains segments de signal entachés de diverses distorsions telles que bruit, écho, compression, sons simultanés.

3. Fondements du peigne spectral

Pour bien comprendre à la fois l'originalité et les développements du peigne de Martin, nous allons dans cette section tenter de montrer les difficultés liées à l'utilisation d'un peigne spectral, qui ressemblent à celles rencontrées par tout détecteur de F_0 , et les solutions qui permettent d'en tirer le meilleur parti.

3.1 La fonction de pitch (FP)

La plupart des détecteurs de pitch font aujourd'hui usage, souvent de manière implicite, d'une "fonction de pitch" (désignée FP dans la suite), prenant une certaine valeur à toute fréquence comprise dans l'intervalle de pitch exploré. De manière équivalente, la FP peut être exprimée en termes de durée de la période. Cette fonction porte des noms divers selon la manière de la calculer et selon les auteurs: fonction d'intercorrélation spectre-peigne chez Martin (Martin 1981), histogramme des périodes ou produit spectral chez Schroeder (Schroeder 1968), ou encore fonction d'autocorrélation dans les méthodes du même nom. La valeur F_0 cherchée est la valeur pour laquelle la FP est maximale (ou minimale dans certaines méthodes comme l'AMDF). Le problème que révèle la FP est que plusieurs maxima parasites peuvent accompagner, voire dépasser en amplitude, le maximum recherché correspondant à F_0 . Voyons ceci sur des exemples schématiques.

3.2 Le peigne uniforme infini

Considérons (figure 1) un spectre harmonique $|S|$ formé de bosses successives de même amplitude, placées aux fréquences pF_0 multiples de F_0 . Considérons également un peigne spectral infini C , formé de valeurs discrètes et unitaires (les dents du peigne), aux fréquences qF_c multiples de la fréquence du peigne F_c . Les indices p et q sont des entiers positifs.

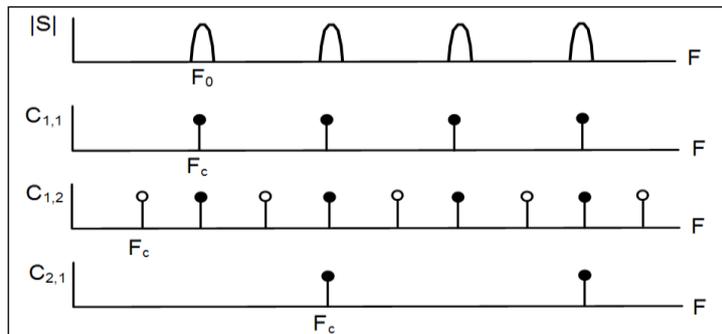


Figure 1: Représentation schématique d'un spectre harmonique uniforme de fréquence F_0 et d'un peigne uniforme infini de fréquence F_c , pour trois valeurs de F_c : $F_c=F_0$, $F_c=F_0/2$ et $F_c=2F_0$.

Lorsque F_c est calé sur F_0 (donc $p=q=1$, cas $C_{1,1}$) chaque dent du peigne est alignée avec une bosse spectrale. La somme des valeurs captées par le peigne est donc maximale.

Lorsque F_c est calé sur la sous-octave $F_0/2$ ($p=1$, $q=2$, cas $C_{1,2}$), seules les dents paires du peigne sont alignées sur des harmoniques. Mais comme chaque bosse spectrale est mise en correspondance avec une dent, la valeur totale captée par le peigne reste identique au cas précédent.

Lorsque F_c est calé sur l'octave $2F_0$ ($p=2$, $q=1$, cas $C_{2,1}$) les choses sont différentes: seules les bosses paires sont en correspondance avec des dents du peigne et la valeur totale captée diminue de moitié.

On voit là le problème posé par le peigne infini: le maximum correspondant à F_0 n'est pas unique. Détecter le maximum de la FP a toutes les chances de produire des erreurs sous-harmoniques à $F_0/2$, $F_0/3$ etc.

Ce point est illustré par la fonction de pitch de la figure 2, obtenue à partir d'un son harmonique réel (train d'impulsions de $F_0=250$ Hz, fenêtre de Hanning 50 ms) et d'un peigne infini. Le pic principal, pour lequel $F_c=F_0$, est étiqueté par $(p=1, q=1)$, en abrégé $(1,1)$. On voit apparaître des pics parasites sous-harmoniques $(1,2)$, $(1,3)\dots(1,q)$, des pics parasites harmoniques $(2,1)\dots(p,1)$, et des pics parasites fractionnaires $(2,3)$, $(4,3)$, $(3,2)\dots$ d'amplitude plus faible.

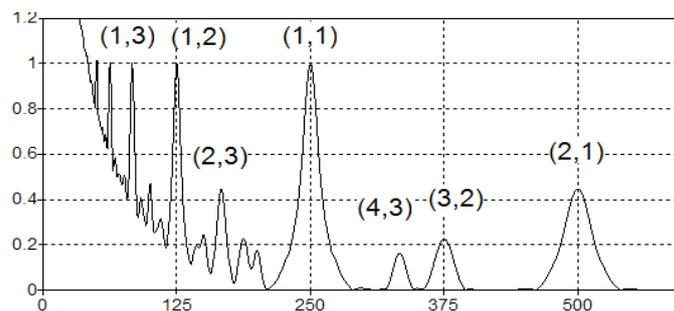


Figure 2: FP d'un son harmonique à $F_0=250$ Hz analysé par un peigne uniforme infini. Les pics sont indicés par (p,q) selon $F_c=(p/q)F_0$.

La remontée de la ligne de base de la fonction vers les fréquences basses correspond au cas où les dents sont tellement serrées qu'il s'en trouve plusieurs dans la largeur d'une bosse spectrale: le peigne ne se comporte plus comme un sélecteur de bosses mais comme un système intégrateur.

3.3 Le peigne limité

C'est le nombre de dents, supposé infini, qui cause l'ambiguïté exposée ci-dessus. Une solution évidente est de limiter ce nombre. Reprenons la figure 1 et imaginons que le peigne ne comporte que 3 dents. Quand $F_c=F_0$ le peigne capte les 3 premières bosses du spectre. La valeur de la FP est donc moins forte qu'avec le peigne infini, puisqu'on laisse de côté les harmoniques de rang supérieur à 3. Mais lorsque $F_c=F_0/2$ seule une bosse spectrale est captée, et le pic correspondant de la fonction de pitch se trouve 3 fois plus faible que le pic principal. Donc les erreurs sous-harmoniques sont beaucoup moins probables qu'avec le peigne infini.

Pourtant cette solution n'est pas la panacée: lorsque $F_c=2F_0$ le peigne à 3 dents capte les bosses 2, 4 et 6 du spectre. La fonction de pitch prend en ce point la même valeur que celle du pic principal, et ce sont alors les erreurs harmoniques qui deviennent les plus probables...

3.4 Le peigne à dents décroissantes, ou Peigne de Martin

On a compris d'après ce qui précède que le peigne infini produit des erreurs sous-harmoniques, et le peigne limité des erreurs harmoniques. La solution proposée par Philippe Martin se trouve à mi-chemin: elle consiste à faire décroître l'amplitude des dents en fonction de leur rang. La fonction de décroissance règle le compromis entre les deux types d'erreurs et tient compte du fait que le spectre de la parole voisée est lui-même décroissant en moyenne au-delà de 800 Hz. Philippe Martin a choisi une décroissance exponentielle avec un nombre de dents compris entre 5 et 10.

La figure 3 montre la fonction de pitch obtenue avec le même son que précédemment et un peigne à 10 dents d'amplitude décroissant en fonction inverse de la racine carrée de leur rang. Les deux pics parasites les plus dangereux sont à l'octave (2,1) et la sous-octave (1,2) du pic principal. Le recul de ces pics par rapport au pic principal n'est pas très grand (respectivement 6 et 3 dB environ). En soi le fait que l'émergence du pic principal soit faible n'empêche nullement le système de bien fonctionner si le signal ne contient qu'une voix et si la mise en œuvre garantit que le maximum est unique et correspond à F_0 .

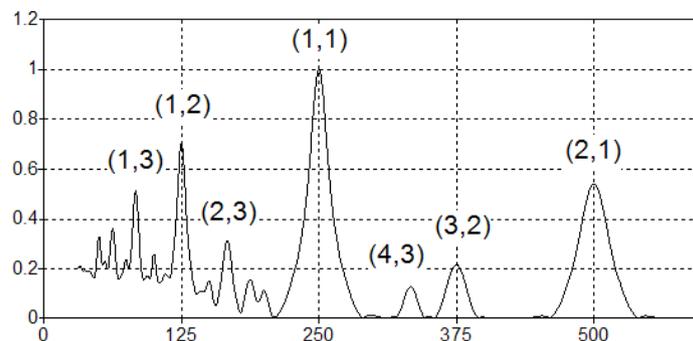


Figure 3: FP d'un son harmonique à $F_0=250$ Hz analysé par un peigne à 10 dents d'amplitude décroissant selon l'inverse de la racine carrée de leur rang.

3.5 La mise en œuvre du peigne

La manière dont le spectre est préparé avant son analyse par le peigne a pour objet de garantir l'unicité du maximum, d'obtenir une mesure précise et de limiter le temps de calcul. Ici comme partout ailleurs le diable est dans les détails. Dans ses premiers articles Philippe Martin utilise des amplitudes en décibels, un seuillage pour éliminer les composantes trop faibles, une interpolation parabolique dans l'entourage des maxima locaux. De plus, la largeur de la fenêtre temporelle gouverne la résolution fréquentielle, la complexité du calcul doit être adaptée à la puissance du processeur disponible pour atteindre le temps réel, les problèmes de précision de calcul imposent de choisir un algorithme plutôt qu'un autre pour calculer la fonction souhaitée, etc. C'est un travail d'artisan, qui ne peut être effectué que par quelqu'un qui a une profonde connaissance de tous les aspects du processus, et qui sait exactement à quel résultat il veut aboutir.

3.6 Universalité de la fonction de pitch

Comme cela a été évoqué plus haut (§3.1) une forme de fonction de pitch est présente dans pratiquement tous les systèmes contemporains de détermination de F_0 , même si cela n'apparaît pas explicitement. Comme la fréquence fondamentale d'un signal est habituellement unique, physiquement et auditivement (on peut toutefois construire artificiellement des signaux à périodicité multiple), il paraît tout naturel de choisir un extrémum unique, souvent le premier d'une série, les autres étant considérés comme parasites.

En réalité, ces autres extrémums ne tombent pas du ciel. Ils sont produits par un processus d'interférence entre deux objets mathématiques périodiques: le signal analysé d'une part, avec sa fréquence fondamentale fixe, et le signal - ou la structure de calcul récurrente mise en œuvre pour la mesure - dont on fait varier la fréquence ou la cadence pour déterminer sa ressemblance avec le signal. Cette ressemblance peut être à caractère cumulatif, et on cherche alors un maximum de la fonction de pitch, ou à caractère compensatoire, et on cherche alors un minimum. Mais quelles que soient les variations entre méthodes, on retrouve toujours les mêmes propriétés interférentielles qui se traduisent par les pics de la FP, repérables par deux indices tels que les entiers p et q définis ci-dessus. Le point important ici est que les pics parasites n'ont pas une origine aléatoire; ils sont la conséquence de la comparaison de deux processus périodiques. Au lieu de subir cette contrainte, on peut peut-être essayer d'en tirer parti pour améliorer la mesure. C'est ce que nous allons exposer maintenant.

4. Nouveaux problèmes, nouveaux peignes

4.1. De nouveaux problèmes ?

Aujourd'hui les meilleurs détecteurs de pitch donnent des taux d'erreur (erreurs dites "grossières", c'est-à-dire s'écartant de plus de 20% de la fréquence locale du signal) inférieurs à 1% sur des bases de données de référence (voir p.ex. Camacho & Harris 2008 ou Signol 2009), pour autant que les comparaisons aient un sens car chacun peut donner satisfaction dans une application et s'avérer insuffisant dans une autre. Cet ordre de grandeur paraît minime mais masque de graves

insuffisances. En particulier en ce qui concerne le voisement et la résistance au bruit.

Outre sa signification du point de vue de la phonation (vibration des cordes vocales), le terme de voisement a deux acceptions bien distinctes. Du point de vue phonologique c'est un trait distinctif, c'est-à-dire une qualité abstraite, binaire, caractérisant deux grandes classes de phonèmes. Du point de vue acoustique c'est un degré de périodicité approximative, apprécié sur un segment de quelques dizaines de millisecondes, dans un intervalle convenu, par exemple 60 à 600 Hz. Les limites temporelles d'un segment voisé ne sont pas parfaitement définies: commence-t-il à la première période détectée, à la seconde, ou après un intervalle de temps défini ? Les périodes, en matière de voix parlée ou chantée, ne sont jamais exactement identiques à elles-mêmes; quel degré de ressemblance doivent-elles avoir pour conférer à l'instant la qualité "voisé" ou "non-voisé" ? Le problème est encore plus flou en fin de segment voisé, où souvent le voisement s'éteint lentement pour devenir une simple oscillation quasi-sinusoïdale et se confondre avec la réverbération ambiante.

Comme la notion de voisement acoustique est floue, on a du mal à évaluer la performance d'un système de détection du voisement - et tout système de détection du pitch est aussi et avant tout un système de détection du voisement. C'est pour cette raison que l'on s'en tient en général à la performance en matière de F_0 , comptée sur les instants indiscutablement bien voisés, au cœur des syllabes. Mais ceci conduit à un paradoxe: un système élaboré, détectant correctement le voisement aux limites des segments voisés, a toutes les chances de faire des erreurs de pitch en ces zones où, précisément, le voisement est imparfait et la fréquence fondamentale incertaine. Il sera pénalisé sous l'aspect du taux d'erreurs sur F_0 par rapport à d'autres qui ne considèrent comme voisés que les noyaux syllabiques.

Par ailleurs, à l'écoute d'un signal de parole, il est très peu usuel de percevoir l'octave, et surtout la sous-octave du fondamental. Or ces erreurs existent, au moins potentiellement, dans tous les systèmes de détection de pitch. On peut donc penser que la perception auditive procède autrement, ou dispose de mécanismes lui permettant de les ignorer ou de les neutraliser.

Un autre problème apparaît dans les situations de bruit, dont chacun sait qu'elles altèrent le fonctionnement de tous les systèmes artificiels, qu'il s'agisse de détermination du pitch et du voisement, ou de reconnaissance automatique de la parole ou du locuteur. On s'arrange habituellement pour contourner le problème en enregistrant en local insonorisé ou en pratiquant une prise de son de proximité. Mais on ne peut oublier que l'audition naturelle, elle, est normalement immergée dans du bruit au point que, bien souvent, le locuteur n'en a même pas conscience. Comment fait-elle pour saisir toutes les nuances de la parole et de la voix malgré les conditions adverses ?

Ce qui vaut pour le bruit vaut aussi pour l'extraordinaire capacité qu'a tout sujet bien entendant à suivre plusieurs conversations à la fois. On sait que l'effet de cocktail party met en jeu un processus attentionnel de haut niveau qui permet de sélectionner à volonté l'une ou l'autre des voix présentes dans le signal. Il n'en reste pas moins que plusieurs flux auditifs sont formés simultanément au bas niveau de la perception auditive. Et qu'un signal totalement imprévu mais hautement significatif peut aussi accéder prioritairement aux niveaux conscients (effet d'émergence soudaine ou pop-up). En ce qui concerne la détection de pitch, il faut bien en déduire que le système auditif est capable de former simultanément plusieurs flux vocaux à un niveau préattentionnel, et donc de mettre en oeuvre une détection de plusieurs fréquences fondamentales simultanées.

Pour toutes ces raisons nous avons été amenés à nous intéresser à la détection multipitch. Dans cette optique nous avons repris les fondements du peigne spectral en essayant d'améliorer son comportement face à des signaux mélangés.

4.2 Des peignes et des dents

Pour bien comprendre comment fonctionne le peigne spectral il n'est pas inutile de se demander quel est précisément le rôle de chaque dent.

Et d'abord, peut-on imaginer un peigne à une seule dent ? La réponse est oui, bien sûr: appliquer à un spectre quelconque un peigne à une dent d'amplitude unitaire conduit à une fonction de pitch identique au spectre lui-même. Et si l'on applique un coefficient d'amplitude, différent de 1, à cette unique dent, alors le résultat est une fonction de pitch dont toutes les valeurs sont multipliées par ce coefficient, même s'il est négatif.

Le cas du peigne à deux dents est très instructif. La fonction de pitch obtenue représente la somme du spectre avec lui-même, décalé de l'intervalle entre les deux dents. Si les fréquences des deux dents sont dans un rapport du simple au double, soit F_c et $2F_c$, et si le spectre est composé d'harmoniques multiples de F_0 , la fonction de pitch passe par des maxima principaux pour $F_c=F_0$, $F_c=2F_0$ etc, et par des maxima secondaires pour $F_c=F_0/2$, $3F_0/2$ etc. Si la seconde dent est affectée du coefficient -1, alors la fonction de pitch n'est plus le cumul, mais la différence du spectre avec sa version décalée.

D'une manière générale chaque dent ajoutée à un peigne à n dents ajoute à la FP autant de spectres décalés, chacun affecté du coefficient propre à la dent.

Nous allons maintenant utiliser ces observations pour mettre en œuvre des peignes particuliers, destinés à renforcer ou atténuer les pics indésirables de la fonction de pitch. Deux familles de peignes sont présentées ci-dessous, les peignes à dents manquantes et les peignes à dents négatives.

4.3 Peignes à dents manquantes (pdm)

Dans le peigne uniforme infini à dents manquantes de la figure 4 on a supprimé les dents de rang pair. L'effet de ce peigne noté pdm2 (ou peigne impair) est de ne prendre en compte aucune bosse spectrale lorsque le peigne est calé sur $F_0/2$, comme le montre la figure (cas M1,2). Il en va de même pour $F_0/4$, $F_0/6$ etc.

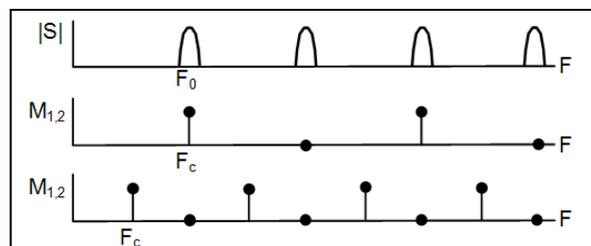


Figure 4: peigne à dents manquantes d'ordre 2, calé sur F_0 et sur $F_0/2$

Ce peigne supprime donc les sous-harmoniques indicés $(p,2k)$ de la fonction de pitch, comme le montre la figure 5 obtenue à partir du même son que précédemment (k est un entier positif).

De la même manière, le peigne à dents manquantes pdm_3 , peigne uniforme infini dont on aurait enlevé les dents 3, 6, 9... $3k$ supprimerait les pics (1,3), (1,6)... (p,3k).

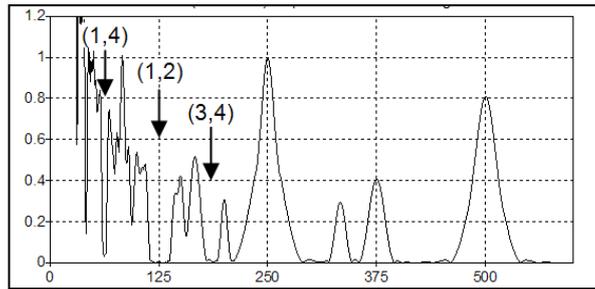


Figure 5: fonction de pitch obtenue avec le peigne impair sur un train d'impulsions de $F_0=250$ Hz. Les pics parasites (1,2), (1,4), (3,4)... (p,2k) sont supprimés.

4.4 Peignes à dents négatives (pdn)

Dans le peigne uniforme infini à dents négatives (ou alternées) on ajoute des dents négatives entre les dents positives. Par convention la fréquence F_c du peigne est celle de la première dent positive. La figure 6 montre le pdn d'ordre 2, noté $A_{2,1}$. Lorsque F_c est calé sur F_0 la valeur de la FP est la même que celle du simple peigne uniforme infini puisque les dents négatives sont alignées avec des valeurs nulles du spectre. Mais lorsqu'il est calé sur $F_0/2$ l'effet des bosses paires et impaires s'annule globalement et le pic parasite correspondant (2,1) est supprimé.

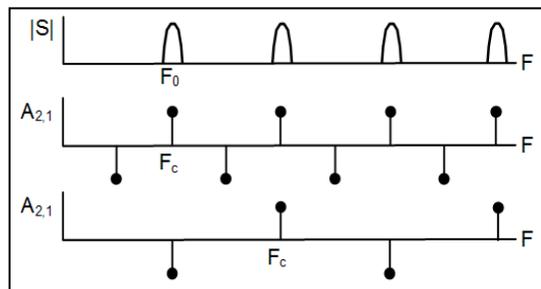


Figure 6: peigne à dents négatives d'ordre 2, calé sur F_0 et sur $2F_0$

La figure 7 montre la FP obtenue en appliquant ce peigne au même son que dans les exemples précédents. On voit que non seulement le pic d'octave (2,1) est supprimé, mais qu'il en va de même pour tous les pics (2p,q).

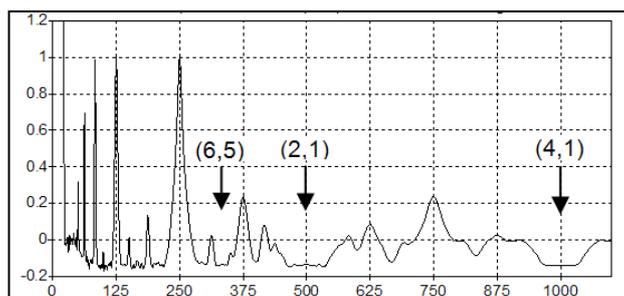


Figure 7: FP obtenue avec le son à 250 Hz analysé par un pdn d'ordre 2

Comme le précédent, ce type de peigne se généralise aisément: en intercalant deux dents négatives d'amplitude 0,5 régulièrement espacées entre deux dents positives (pdm3) les pics parasites (3p,q) sont supprimés.

4.5 Peigne à Suppression Harmonique (PSH)

Chacun des pdm et pdn présentés ci-dessus a comme caractéristique de supprimer une famille de pics parasites, tout en préservant le pic principal. Dans le Peigne à Suppression Harmonique PSH (Lienard et al. 2008 ; Signol 2009) les peignes pdm et pdn sont utilisés conjointement (produit des FP correspondantes, ordres 2, 3, 5 et 7). Le résultat sur notre son de référence est présenté ci-dessous, figure 8, à comparer avec la figure 3. Les deux pics les plus dangereux (octave et sous-octave) se trouvent maintenant à -20 dB du pic principal. Ce surcroît d'émergence permet d'assurer une plus grande fiabilité à la détection du maximum dans le cas d'un signal monovoix, et de séparer plusieurs pics principaux dans le cas d'un mélange de plusieurs voix. La détection du voisement acoustique est elle aussi grandement facilitée.

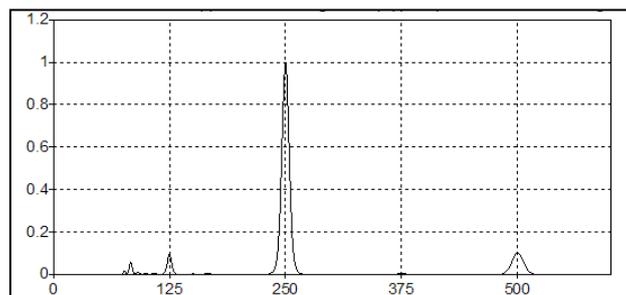


Figure 8: FP du son à 250 Hz obtenue par application conjointe des pdn et pdm d'ordres 2, 3, 5 et 7.

5. Conclusion

Depuis les années 70 Philippe Martin a mis au point et continuellement perfectionné une méthode originale d'analyse et de représentation de la fréquence fondamentale de la parole, outil essentiel pour ses recherches prosodiques. Sa méthode du peigne spectral donne un résultat immédiat et fiable, répondant à ses exigences expérimentales. Ayant à l'esprit un autre but, relevant de la perception de la parole, nous avons voulu en comprendre le fonctionnement en profondeur pour l'adapter à notre propre problème d'analyse de voix mélangées. Cette démarche nous a permis de constater l'efficacité de la méthode, et de proposer diverses extensions qui en conservent l'esprit et certaines qualités, en particulier la simplicité et la rapidité de calcul.

6. Références

- Camacho, A. and Harris, J.G. (2008), "A sawtooth waveform inspired pitch estimator for speech and music", *J. Acoust. Soc. Am.* Vol. 124, No 3, pp 1639-1652.
- Hess, W. (1983), *Pitch determination of speech signals*, Springer-Verlag.
- Liénard, J.S., Signol, F. and Barras, C. (2007), *Speech fundamental frequency estimation using the Alternate Comb*, *InterSpeech*, Anvers.

- Liénard, J.S., Barras, C. and Signol, F. (2008): "Using Sets of Combs to Control Pitch Estimation Errors", *ASA Proceedings Of Meetings on Acoustics*, Vol 4, 060005.
- Martin, Ph (1973), *Réalisation d'un analyseur de mélodie fonctionnant en temps réel en vue de la recherche phonétique et l'enseignement*, Dissertation, Université Libre de Bruxelles.
- Martin, Ph. (1981), "Extraction de la fréquence fondamentale par intercorrélation avec une fonction peigne", *12e Journées d'Etude sur la Parole*, Montréal, pp 221-232.
- Martin, Ph. (1982), "Comparison of pitch detection by cepstrum and spectral comb analysis", *Icassp*, Paris, pp 180-183.
- Martin, Ph. (1996), "WinPitch: un logiciel d'analyse temps réel de la fréquence fondamentale fonctionnant sous Windows", *21e Journées d'Etude sur la Parole*, Avignon.
- Martin, Ph. (2000), "Peigne et brosse pour F0: mesure de la fréquence fondamentale par alignement de spectres séquentiels", *23e Journées d'Etude sur la Parole*, Aussois, pp 245-248.
- Martin, Ph. (2008), "Crosscorrelation of adjacent spectra enhances fundamental frequency estimation", *InterSpeech*, Brisbane.
- Martin, Ph. (2012), "Multi methods pitch tracking", *ISCA Speech Prosody*, Shanghai.
- Schroeder, M.R. (1968), "Period histogram and product spectrum: new methods for fundamental frequency measurement", *J. Acoust. Soc. Am.* Vol. 24, pp 456-465.
- Signol, F. (2009), *Estimation de fréquences fondamentales multiples en vue de la séparation de signaux de parole mélangés dans un même canal*, thèse de doctorat, Univ. Paris XI.

The effect of rhythm and pitch contour manipulation on intelligibility in L2 Italian informative speech

Massimo Pettorino, Anna De Meo, Marilisa Vitale

mpettorino@unior.it, ademeo@unior.it, marilisavitale@hotmail.it

University of Naples L'Orientale

Abstract

This study aims at investigating the role played by syllable duration and pitch contour in the intelligibility of speech in L2 Italian. Three news, not requiring any specific tonal variation, were read and recorded by a native Italian speaker and were manipulated using WinPitch into two ways: 1) syllables were made isochronous; 2) tonal range was flattened. The corpus was administered to 35 native Italian listeners and to 132 speakers of L2 Italian, in order to verify if the manipulation of pitch contour and syllable durations may alter the understanding of textual content. The results of the test show that pitch contour and rhythm changes do not interfere with the L1 Italian listeners' intelligibility, while they play a negative role for L2 listeners, in particular the pitch contour flattening has a stronger effect than syllabic isochrony.

1. Introduction

Fundamental frequency contour, syllabic duration and intensity level are the three main prosodic features that the speaker uses in order to convey a large amount of the message in the communicative exchange. The continuous accelerations and decelerations of the speech, together with the different location of the F0 and intensity peaks, not only mark the portions of the utterance on which the speaker wants to draw the listener's attention, but also give the interlocutor the necessary tools for the correct decoding of the text. In other words, the prosodic features of the utterance, expressing intentions, attitudes, and emotional states and allowing the correct identification of the constituents of the sentence, help the listener in the semantic processing of the message. In this perspective, it seems appropriate to verify if fundamental frequency contour, syllabic duration and intensity level play an equally important role in the intelligibility of an utterance or if one of them prevails over the other two (Munro 2008). If a difference exists, is it stable or does it depend on the communicative context?

To answer these questions we decided to carry out an investigation in order to verify if the total or partial lack of one of the three prosodic features may cause a significant impediment to the understanding of the message. To this end, we focused on intralingual behaviour, that is speech perception within the same language community (L1 Italian), and interlingual dimensions of speech perception (L1 Italian speakers – L2 Italian listeners). In the latter communicative context listening is certainly more complex as the listener, who does not have the same cognitive and linguistic background of his interlocutor, must recourse to all components of the acoustic signal in order to correctly decode the text. For this reason, the artificial manipulation of any acoustic parameter could cause a significant variation in intelligibility.

In this study we will limit our attention to two of the three parameters mentioned above, namely

syllabic duration and pitch contour.

2. The experiment

For the purpose of this study, Three news (40 words each), not requiring any specific prosodic pattern expressing emotions, attitudes or communicative intents, were read and recorded in Italian by a native Italian speaker (female voice). Each utterance was manipulated using WinPitch ([WinPitch. http://www.winpitch.com/](http://www.winpitch.com/)) in two different ways: 1) syllables were made isochronous; 2) tonal range was completely flattened. The corpus thus consisting of 9 news (3 original, 3 duration-modified and 3 tone-flattened) was administered to 35 native Italian listeners and to 132 non-native speakers of Italian, having different L1s. Listeners were divided into 3 groups, each one corresponding to a different level of competence of Italian (A2, B1 and C1 of the CEFR). Each group was split into three subgroups, in order to avoid the multiple administration of the same piece of news (original or manipulated) to the same listeners, but at the same time to ensure that each group of listeners would evaluate the whole corpus. All A2 level speakers were L1 Spanish, while B1 and C1 levels were speakers of different native languages (Arabic, Chinese, English, French, German, Japanese, Polish, Portuguese, Russian, Spanish, Turkish, Ukrainian, Vietnamese).

For each utterance, listeners were asked to answer two questions in order to verify to which extent the text of the news had been understood. News and questions have all been selected by an expert in L2 Italian evaluation tests, to guarantee a good level of intelligibility even for the A2 non-native listeners.

3. Manipulation

Through the latest version of WinPitchW7, that allows to manage the duration of each syllable and to manipulate the pitch contour, all the three original news underwent both kinds of manipulation: in the first case, the resulting speech was completely isochronous (160 ms each syllable), while in the second case the utterance was tonally flattened (160 Hz) (fig. 1).

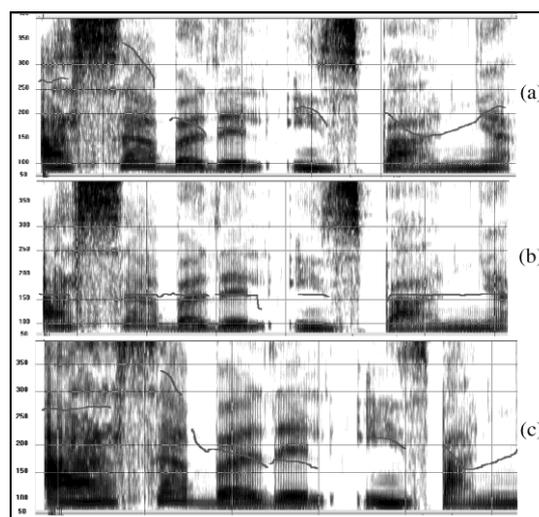


Figure 1: Spectrograms of original (a), pitch flattened (b) and modified duration (c) sample.

4. The results

The results of the perceptive tests carried out on the synthetically modified utterances clearly show a marked difference between the *intra*lingual and the *inter*lingual speech perception behaviour. In the first case, the changes made to the acoustic signal do not involve any lowering of the percentage of correct understanding of the informative read text: the native Italian listeners always answered properly after each stimulus, demonstrating that a change in the pitch contour or in the level of syllabic isochrony does not impair the intelligibility of the text. However, this consideration should be limited to the particular type of corpus on which the experiment has been conducted. It is quite evident that, for instance, in a spontaneous spoken dialogue, containing different speech acts, the test results could have been quite different.

The analysis of the data related to non-native listeners shows a different scenario: the changes made to the acoustic signal lead to a significant drop in the level of intelligibility, as regards both the tone and the syllabic duration. However, the data show that the negative effects are not the same for all non-native listeners, but vary depending on their level of competence and their L1s. Figure 2 shows the percentage of correct answers for each competence level of the CEFR (fig. 2).

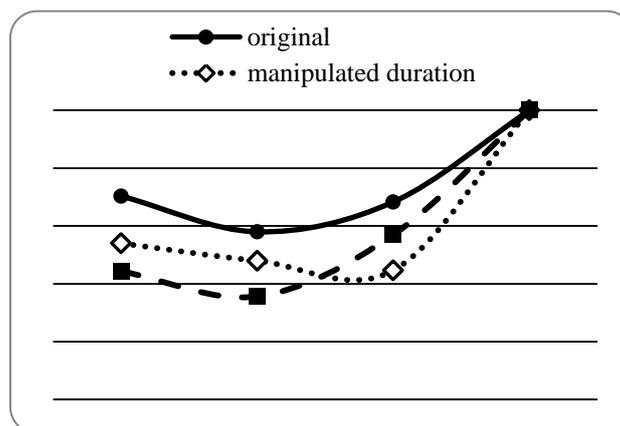


Figure 2: Percentage of correct answers for native and non-native listeners.

As it can be seen, the original signal always obtained the highest value of correct answers, even if it is remarkable that the A2 and C1 levels got comparable results. In this regard it should be noted that the A2 listeners were all Spanish and for this reason the data should be evaluated separately, since the rate of oral intercomprehension between the two involved Romance languages is very high (see Benucci 2005; Blanche-Benvensite & Valli 1997 et Doyle 2009).

As regards the effect of each synthetic modification, it can be noticed that its relevance depends on the degree of language competence of the subject: for lower levels, i.e. A2 and B1, the flattening of the pitch contour plays a more important role than the modified syllabic duration, whereas the reverse is true for the C1 level. These data seem particularly interesting, since they point out the existence of an acquisitional sequence of the L2 suprasegmental level. Indeed in the early stages of interlanguage the pitch contour helps to highlight the information units of a text. In contrast, for advanced interlanguage levels, intelligibility is mostly affected by altered rhythmic structure: at this

level the continuous accelerations and decelerations constitute an essential tool for processing the message.

Figures 3-6 show the percentage of correct answers for the largest groups of L1 listeners, that are Spanish, Chinese, Vietnamese and Slavic languages (32, 52, 14 and 7 listeners respectively).

The overall data highlight some general trends and allow us to make some considerations:

1. In no case the change of the syllabic duration implies a more evident effect than the one produced by the F0 manipulation;
2. For Spanish listeners, who have a language more similar to Italian, the pitch flattening produces a lowering of intelligibility greater than the one determined by a forced syllabic isochrony. Anyway for these listeners the level of correct answers is in all cases very high, from nearly 90% in the case of the original signal, to 63% in the case of a flattened pitch (fig. 3);
3. Chinese listeners do not seem to be sensitive at all to the carried out prosodic changes, as if the suprasegmental level was completely cut off from the communicative exchange (fig. 4);
4. Vietnamese listeners, though they have an L1 which is very different from Italian, seem to catch the rhythmic-intonational variations better than the Chinese listeners (fig. 5);
5. Slavic listeners show a high level of intelligibility for the original utterances (71%) and the effects given by the synthesized speech allow to attribute such a good result more to the perception of the pitch contour than to rhythmic characteristics of the utterance (fig. 6).

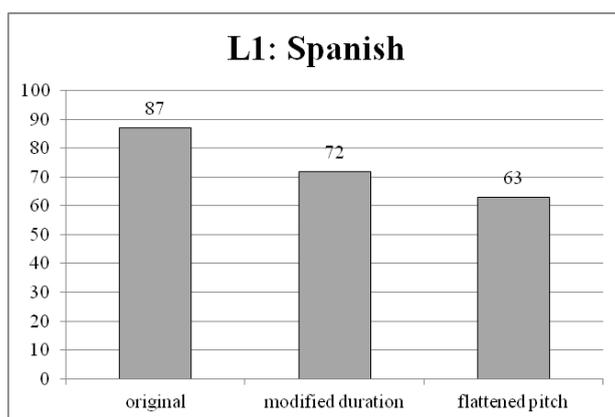


Figure 3: Percentage of correct responses for Spanish L1 listeners.

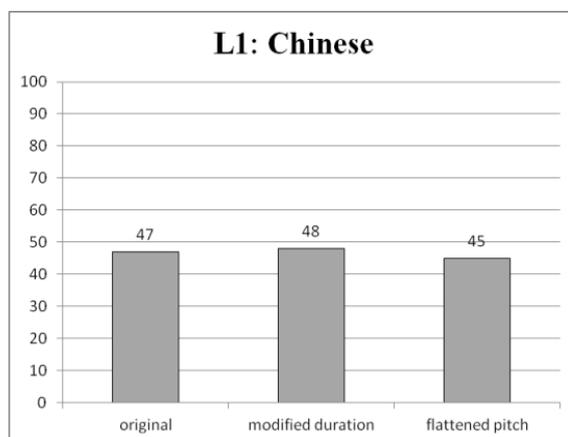


Figure 4: Percentage of correct responses for Chinese L1 listeners.

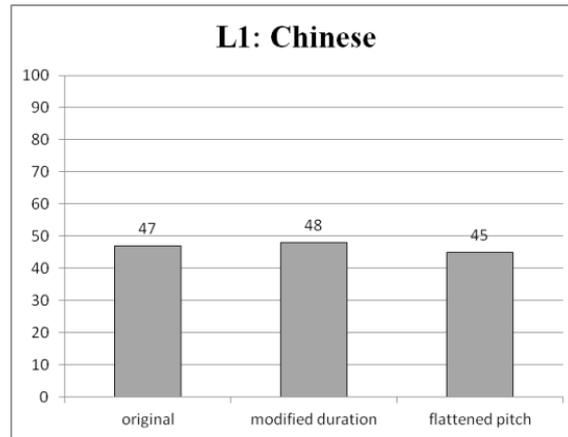


Figure 5: Percentage of correct responses for Vietnamese L1 listeners.



Figure 6: Percentage of correct responses for Slavic L1 listeners.

6. Conclusions

The latest version of WinPitch allowed us to verify experimentally the different role played by intonation and rhythm in the understanding of an informative L2 message. The analysis outcome showed that, in general, the pitch contour affects the intelligibility of a message to a greater extent than its rhythmic characteristics. However, both parameters interfere with communication processing though at different degrees, according to the L1s and to the levels of competence of the listeners. The degree of mutual intelligibility between L1 and L2 also facilitates the linguistic exchange, regardless of the competence level. Finally, the study showed the existence of an acquisitional sequence that implies for the lower levels of interlanguage (A2 and B1) the use of pitch contour movements as a clue for the identification of the information units, while for more advanced levels the message is mainly processed on the basis of the rhythmic structure.

Références

Benucci A. ed. (2005), *Le lingue romanze*. Una guida per l'intercomprensione, Torino, Utet.

Blanche-Benveniste C. et Valli A. (1997) "L'Intercompréhension: le cas des langues romanes, *Le Français dans le monde: Recherches et applications*, numéro spécial, janvier 1997.

- Doye P. (2004), "A Methodological Framework for the Teaching of Intercomprehension", *Language Learning Journal*, 30/1, pp. 59-68.
- Munro M. (2008), "Foreign accent and speech intelligibility", Hansen Edwards J.G., Zampini M. (eds.) *Phonology and Second Language Acquisition*, John Benjamins Publishing Company, Amsterdam, pp. 199-224.

La perception des traits phonologiques: des logatomes à la parole spontanée

Willy Serniclaes

willy.serniclaes@ulb.ac.be

Unité de Neurosciences Cognitives (UNESCOG), Université Libre de Bruxelles

Abstract

Most current models of speech perception are based on data collected with elementary units, often single syllables without definite meaning ("logotemes"). Here I ask whether these models are compatible with the results of intonation studies, especially the cognitive model promoted by Philippe Martin (Functional Cognitive model, FC). A throughout comparison between logoteme and sentence models, indicates that both have in common the points of being elementarist, interactive and hierarchical. Also, both models are robust enough to cope with spontaneous speech.

1. Introduction

Les modèles actuels de perception des segments de parole se fondent en première instance sur des données recueillies avec des segments de parole sans signification, ou « logatomes ». Ces modèles ont ensuite fait l'objet de généralisation à des segments de parole avec signification (aux mots), parfois également aux phrases et de manière beaucoup plus rare à de la parole spontanée. L'objectif de cet article est de fournir une brève description de l'un de ces modèles, celui que je préconise, et de le comparer avec celui proposé par Philippe Martin pour la perception des traits prosodiques. Objectif motivé par un commentaire de Philippe qui a attiré mon attention sur la similitude entre nos deux approches. Dont acte.

2. Modèles

2.1. Modèles segmentaux

Les unités les plus élémentaires que l'on puisse imaginer ne sont pas des segments mais bien des oppositions entre segments : les traits distinctifs (cf. Jakobson 1973). Les traits sont des unités différentielles irréductibles aux segments: ce qui est commun aux oppositions ba/pa, da/ta, ga/ka, ...gi/ki, ... igi/iki ... ce ne sont pas les diverses propriétés acoustiques communes aux logatomes /ba, da, ga .../ ou /pa, ta, ka, .../ mais bien les différences communes à l'ensemble de ces oppositions. Si le phonème est éminemment variable, le caractère différentiel du trait le rend susceptible de se réaliser de manière invariante dans de multiples contextes différents. Mais la question de l'invariance des traits reste néanmoins posée, et ce depuis longtemps (p.ex. Fromkin 1979).

Le problème soulevé par l'invariance des traits trouve son origine dans un hiatus entre perception et production. Les traits perceptifs universels mis en évidence chez l'enfant pre-linguistique (pour une revue, cf. Hoonhorst et al. 2009), et qui correspondent à des seuils psychoacoustiques élémentaires (p.ex. le changement de direction ascendante/descendante d'une transition de fréquence), ne sont pas produits de manière indépendante en raison de la coarticulation. Pour être perçus de manière invariante ces traits doivent être intégrés de manière spécifique au sein de chaque langue pour former des traits phonologiques (Serniclaes 2011).

L'intégration phonologique entre traits universels fait appel à deux types de processus : l'intégration de leurs corrélats acoustiques ('compensation for coarticulation', pour une revue, voir Mitterer 2006) et les couplages percept-percept (Epstein 1983) par lesquels la perception d'un trait affecte celle d'un autre trait (pour une revue, voir Serniclaes & Wajskop 1992). Ce dernier processus s'accompagne d'un traitement hiérarchique ('top-down'), la perception de traits les plus clivants affectant ceux qui le sont moins (Figure 1).

Le dispositif mis en oeuvre pour percevoir les traits est donc éminemment complexe mais il est performant. Par exemple, l'étude de la perception du trait de voisement dans des échantillons sans signification de parole spontanée (trop brefs pour pouvoir y reconnaître des mots) montre que le taux erreurs d'identification est minime (moins de 2% : Saerens et al. 1989). Dans les rares cas où le trait n'est pas identifié correctement dans le segment prélevé, l'information manquante n'est pas de nature acoustique mais bien de nature lexicale. Lorsque l'on élargit la fenêtre de prélèvement, ce n'est que si l'auditeur perçoit des mots que le trait de voisement est identifié correctement (comme illustré dans le Tableau 1). L'identification correcte du trait ne dépend donc pas d'indices distribués de manière plus large mais bien du contexte lexical.

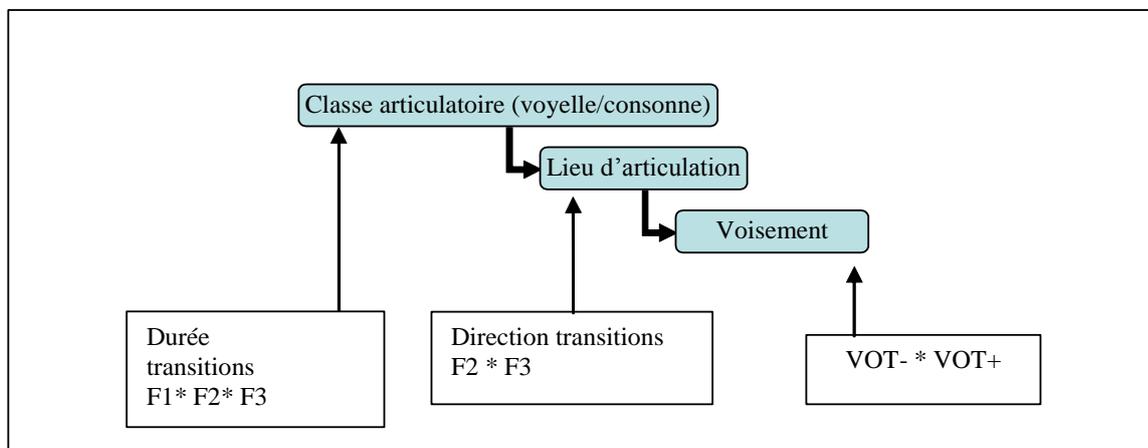


Figure 1 : Processus de perception des traits segmentaux. Couplages entre traits acoustiques (*) et traitement hiérarchique des traits phonologiques.

segment	auditeur 1	auditeur 2
V[p]V	'ibvegasses'	'euba'
+100 ms	'renzest'	'deubas'
+200 ms	'keeves'	'creuvase'
+300 ms	'reux <u>parce que</u> '	'rebasque'
+400 ms	' <u>heureux parce</u> <u>qu'en</u> '	'dereubasquant'

Tableau 1. Extrait des résultats de Saerens et al 1989. Exemple de transcription de deux participants en réponse à des extraits de la portion de phrase 'dangereux par ce qu'en fait' prononcée spontanément et dont l'extrait /oepa/ n'a pas été identifié correctement (réponse /oeba/). Les réponses lexicales sont soulignées.

Ces résultats montrent bien que les critiques que l'on adresse souvent à l'égard du modèle de perception en traits ne sont pas fondées. La variabilité contextuelle des segments phonémiques a souvent servi d'argument pour justifier une approche globaliste (p.ex. Hawkins 2011). Mais l'exemple du trait de voisement suggère que la perception des traits segmentaux dépend d'indices concentrés dans des portions relativement brèves du signal, de l'ordre de une à deux syllabes. Lorsque ces indices font défaut, ce qui est relativement rare, seul le contexte lexical voire sémantique est susceptible de rectifier la perception.

2.2. Modèles prosodiques

Elémentarisme, interactivité, traitement hiérarchique, tels semblent être les ingrédients de la formule perceptive des traits segmentaux. Formule gagnante d'ailleurs si l'on en juge par sa robustesse dans la perception de la parole spontanée. Mais qu'en est-il des traits prosodiques ?

L'approche fonctionnelle cognitive du traitement prosodique (ci-après modèle FC) proposée par Philippe Martin (Martin 2010 ; Martin 2011) comporte plusieurs points communs avec le modèle de perception en traits phonologiques que je viens d'esquisser. Contrairement au modèle autosegmental-métrique (p.ex. Ladd 2008), le modèle FC postule une subdivision de la phrase en événements prosodiques (Figure 2).

Il s'agit donc bien d'un modèle élémentariste et il n'est donc pas surprenant de le voir faire appel à des processus de recomposition. Avec ce modèle, la concaténation des événements prosodiques s'opère de manière hiérarchique, au cours duquel la perception de l'un de ces événements affecte celles des autres.

Le modèle FC est élémentariste, interactif et hiérarchique (concaténations). Enfin, le traitement de l'intonation par ce double mouvement de segmentation puis concaténation postulé par le modèle FC est robuste. Ce modèle rend mieux compte des contours mélodiques dans des échantillons de parole lue et de parole spontanée (cf. Martin 2011; pour des exemples spectrographiques d'extraits de parole spontanée : Martin 2010).

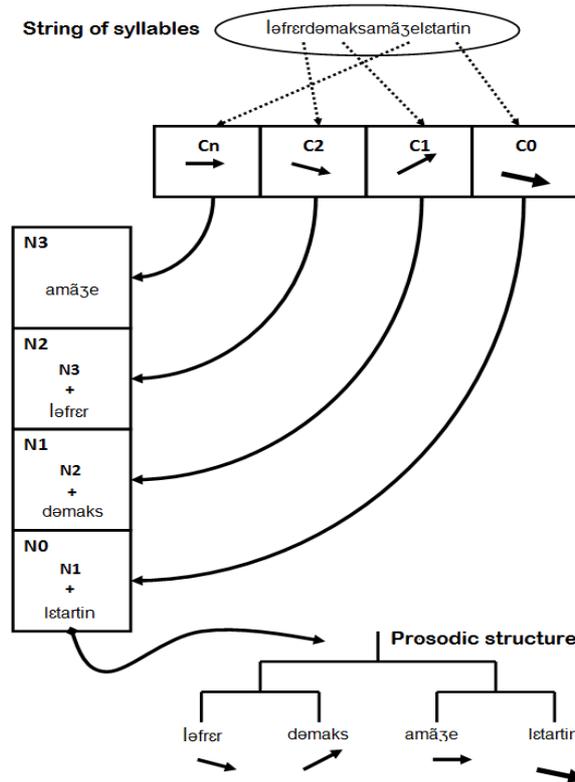


Figure 2. Processus de perception des traits prosodiques. Extrait de Martin 2010, Fig.2 ('Le frère de Max a mangé les tartines').

3. Discussion et conclusions

Le Tableau 2 résume les similitudes entre de perception des traits segmentaux et de la prosodie, dans le cadre des modèles considérés dans cet essai.

modèle	segmental	prosodique
unités	traits	événements (EP)
interactions	couplages entre traits psychoacoustiques	couplages entre F0, intensité et durée ?
hiérarchie	dépendances entre traits	concaténation d'EP

Tableau 2. Similitudes entre la perception des traits segmentaux et de la prosodie.

Les unités postulées par les deux modèles sont des traits dynamiques, de type transitoire (transitions de fréquence) ou relationnel (VOT) et non des segments.

Comme pour les traits segmentaux, la perception des traits prosodiques dépend de différents traits acoustiques, ceux relatifs à la hauteur (F0), à la force sonore (intensité), à la durée (temps) ou encore à la qualité vocale comme rappelé dans Martin 2010. Des couplages entre ces différents traits sont dès lors possibles, bien qu'il n'y ait pas de données à ce sujet (du moins à ma connaissance).

Enfin, chaque modèle fait appel à un traitement hiérarchique. Dans le cas des traits

segmentaux, il s'agit de dépendances en chaîne, la perception du trait de classe affectant celle du lieu, la perception de ce dernier affectant celle voisement (pour se limiter à quelques traits ; voir Figure 1). Le traitement hiérarchique des traits prosodiques se présente de manière assez différente. Ici ce sont des séquences de valeurs du même trait- des séquences de PE- qui sont concaténées, opération qui est pilotée par le PE qui occupe une position dominante.

Le traitement hiérarchique est la partie la plus complexe du processus de perception. Dans le cas des traits segmentaux, on sait que les processus hiérarchiques haut-bas sont doublés de processus bas-haut par lesquels des indices acoustiques secondaires ajustent la perception d'un trait en fonction des traits environnants (p.ex. l'incorporation du F1 dans la perception du voisement ajuste la frontière de VOT en fonction du contexte vocalique). L'incorporation des indices secondaires résulte d'un processus développemental très lent qui s'étend jusqu'à l'adolescence (cf. Medina et al. 2010 ; Hoonhorst et al. 2011).

Le traitement hiérarchique joue probablement un rôle dans l'acquisition de percepts phonémiques, notamment sous l'influence de l'apprentissage de la lecture (Morais et al. 1979) ; pour une discussion sur ce point cf. Springer-Charolles et al. 2006. Il pourrait être intéressant de voir si la concaténation des événements prosodiques, notamment son mode de pilotage, de se développe de manière semblable pour déboucher sur des patrons intonatifs intégrés.

En conclusion, dans le cadre d'une approche analytique, percevoir les traits prosodiques fait appel à des processus semblables à ceux mis en évidence pour les traits segmentaux. Cette approche s'est révélée robuste à l'épreuve de la parole spontanée et ne prête donc pas le flanc aux critiques globalistes.

4. Références

- Epstein, W. (1982) "Percept-percept coupling" *Perception*, Vol. 11, pp. 75-83.
- Ladd, D.R. (2008) *Intonational Phonology* Cambridge University Press.
- Fromkin, V.A. (1979) "Persistent questions concerning distinctive features". In B.Lindblom and S.Öhman (Eds.) *Frontiers of Speech Communication Research*. (pp. 323-334). Academic Press.
- Hawkins, S (2011). "Phonetic perspectives on modelling information in the speech signal" *Sadhana*, Vol.36, pp. 555-586.
- Hoonhorst, I., Colin, C., Markessis, E., Radeau, M., Deltenre, P., & Serniclaes, W. (2009) "French native speakers in the making: from language-general to language-specific voicing boundaries" *Journal of Experimental Child Psychology*, Vol. 104, pp. 353-366.
- Hoonhorst, I., Medina, V., Colin, C., Markessis, E., Radeau, M., Deltenre P., & Serniclaes, W. (2011). The development of categorical perception: comparisons between voicing, colors and facial expressions. *Speech Communication*, Vol. 53, pp.417-430.
- Jakobson, R. (1973) *Essais de Linguistique Générale*. Editions de Minuit.
- Martin Ph. (2010) "Prosodic structure revisited: a cognitive approach. The example of French ", *Proc. Speech Prosody 2010, Chicago*, pp. 100-120.
- Martin Ph. (2011). "Prosodic characteristics of read and spontaneous speech in French" *Proc. 17th International Congress of Phonetic Sciences (ICPhS XVII), Hong Kong*. pp. 1330-1333.

- Medina, V., Hoonhorst, I., Bogliotti, C., & Serniclaes, W. (2010) "Development of voicing perception in French: Comparing adults, adolescents, and children" *Journal of Phonetics*, Vol. 38, pp. 493-503.
- Morais, J., Cary, L., Alegria, J. and Bertelson, P. (1979) "Does awareness of speech as a sequence of phones arise spontaneously?" *Cognition*, Vol. 7, pp. 323-331.
- Mitterer, H. (2006) "On the causes of compensation for coarticulation: Evidence for phonological mediation" *Perception and Psychophysics*, Vol. 68, pp.1227-1240.
- Saerens, M., Serniclaes, W., & Beeckmans, R. (1989) "Acoustic versus contextual factors in stop voicing perception in spontaneous French" *Language and Speech*, Vol. 32, pp. 291-314.
- Serniclaes, W. (2011) "Features are phonological transforms of natural boundaries" In G.N. Clements and R. Ridouane (Eds.) *Cognitive, physical and developmental bases of distinctive speech categories* (pp. 237-257). John Benjamins.
- Serniclaes, W., & Wajskop, M. (1992) "Phonetic versus acoustic account of feature interaction in speech perception" in *Analytic Approaches to Human Cognition, Proc. of the Conference in Honour of Paul Bertelson*, Brussels June 1991. J. Alegria, D. Holender, J. Junça de Morais, and M. Radeau eds.(pp.77-91). North-Holland.
- Sprenger-Charolles, L., Colé, P., & Serniclaes, W. (2006) « *Reading Acquisition and Developmental Dyslexia* » Psychology Press (Developmental Essays).