

# Pour une approche phonologique de la prosodie dans l'interaction

Roxane Bertrand, Cristel Portes

DANS **LANGUE FRANÇAISE** 2012/3 n°175 , PAGES 19 À 36  
ÉDITIONS **ARMAND COLIN**

ISSN 0023-8368

ISBN 9782200927790

DOI 10.3917/lf.175.0019

Date de mise en ligne : 14/11/2012

Article disponible en ligne à l'adresse

<https://shs.cairn.info/revue-langue-francaise-2012-3-page-19?lang=fr>



Découvrir le sommaire de ce numéro, suivre la revue par email, s'abonner...  
Scannez ce QR Code pour accéder à la page de ce numéro sur Cairn.info.



**Distribution électronique Cairn.info pour Armand Colin.**

Vous avez l'autorisation de reproduire cet article dans les limites des conditions d'utilisation de Cairn.info ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Détails et conditions sur [cairn.info/copyright](http://cairn.info/copyright).

Sauf dispositions légales contraires, les usages numériques à des fins pédagogiques des présentes ressources sont soumises à l'autorisation de l'Éditeur ou, le cas échéant, de l'organisme de gestion collective habilité à cet effet. Il en est ainsi notamment en France avec le CFC qui est l'organisme agréé en la matière.

---

# Pour une approche phonologique de la prosodie dans l'interaction

## 1. INTRODUCTION

Dès les premières études en *Analyse Conversationnelle* (Sacks *et al.* 1974 ; Schegloff 1988) puis au sein des différents courants qui se sont développés autour de la problématique de l'interaction (la « contextualisation », Gumperz 1982 ; la « phonologie pour la conversation », Local *et al.* 1985, ou encore la « linguistique interactionnelle », Selting & Couper-Kuhlen 2001 ; Ford, Fox & Thompson 2002), les phénomènes prosodiques sont considérés, au même titre que les phénomènes syntaxiques, comme des ressources dont usent les participants à une interaction pour atteindre des objectifs interactionnels spécifiques. La prosodie joue ainsi un rôle crucial pour signaler et interpréter certaines activités ou certaines actions (comme l'initiation de séquences de réparation par exemple, voir Selting 2010), dont la description et la compréhension constituent le focus de ces études.

Dans cette perspective interactionnelle <sup>1</sup> (terme qui résumera désormais les courants cités *supra*), la prosodie est définie par l'ensemble de ses paramètres acoustiques (fréquence fondamentale, intensité et durée) et par certaines des fonctions qu'elle véhicule (notamment son rôle dans l'organisation des tours de parole). De manière générale, dans la mesure où les indices considérés sont *pertinents* interactionnellement, à savoir s'ils influent sur le déroulement du discours et l'alternance des locuteurs par exemple, les interactionnistes considèrent qu'il faut les prendre en compte (Szczepek Reed, 2006 : 9). Cette vision, partagée par

---

1. Nous parlerons indifféremment au cours de l'article, d'interaction ou de conversation pour qualifier la parole en situation d'interaction, sans toutefois distinguer les différents types d'interactions (débat, interview, conversation, etc.). Conversationnalistes et interactionnistes seront également utilisés indifféremment.

certain phonéticiens dont l'intérêt pour la parole en interaction s'est manifesté très tôt (Local, Wells & Sebba 1985), a conduit ces mêmes auteurs à regretter que les approches plus traditionnelles de la prosodie ne s'intéressent pas davantage aux phénomènes prosodiques en situation de parole naturelle (authentique) et se cantonnent à l'analyse de parole de laboratoire (très contrôlée et non interactionnelle). Ceci a entraîné l'élaboration de modèles extrêmement sophistiqués, mais souvent inadéquats, selon eux, pour rendre compte de la prosodie dans l'interaction.

Nous souhaitons montrer que les modèles existants, s'ils doivent être adaptés à la parole continue en situation d'interaction, peuvent cependant se révéler très utiles pour appréhender et démêler les différents phénomènes prosodiques à l'œuvre en conversation. Nous défendons aussi l'idée que des études expérimentales rigoureuses permettront de valider des résultats liés à des phénomènes propres aux interactions mis en lumière par les interactionnistes. À titre d'illustration, les signaux backchannels – signaux voco-verbaux et mimo-gestuels émis par les auditeurs pour montrer leur intérêt et leur participation à l'interaction en cours – ont été définis très tôt dans la perspective interactionnelle (voir Schegloff 1982 ; Laforest 1992). Ils font désormais l'objet d'études systématiques relatives à leur contexte d'apparition notamment (Koiso *et al.* 1998 ; Caspers 2003 ; Bertrand, Portes & Sabio 2007 ; Bertrand *et al.* 2007 ; Heldner, Edlund & Hirschberg 2010). Ces études ont permis de confirmer leurs implications interactionnelles et, plus généralement, de souligner leur rôle tant dans le processus de compréhension que dans celui de production des discours (Fox Tree 1999), leur présence/absence influant directement sur le discours produit par le locuteur principal. Ces études ont également des répercussions intéressantes dans les systèmes de dialogue homme-machine (Gravano & Hirschberg 2010).

Cet article promeut une approche fondée sur des principes relevant à la fois de la perspective interactionniste et de la perspective prosodique. Pour rendre notre proposition audible par les deux communautés, nous déclinons cet article en deux parties. Dans la première, nous présentons quelques principes d'analyse de l'approche phonologique prosodique tels que la séparation des différents niveaux de l'organisation phonologique ou la distinction entre unités discrètes et variations graduelles. Dans la seconde, nous tentons d'illustrer l'intérêt d'opérer de telles distinctions en contexte conversationnel, en montrant les rôles divers que jouent les variations prosodiques dans l'organisation des tours de parole ou dans l'organisation interne d'activités telles que la narration.

## 2. L'APPROCHE PHONOLOGIQUE DE LA PROSODIE

L'approche phonologique défend l'existence d'une organisation prosodique abstraite, indépendante de la syntaxe et soumise à des contraintes internes d'organisation. Elle opère trois types de distinction qui s'avèrent très fructueux dans l'étude des données conversationnelles : i) la distinction entre structure

prosodique et intonation d'une part, tonal et métrique d'autre part au sein de l'organisation phonologique de la prosodie ; ii) la distinction entre unités discrètes et variations prosodiques graduelles ; iii) le traitement séparé des pauses et des disfluences.

## **2.1. La distinction et l'articulation de plusieurs niveaux d'organisation phonologique**

Au sein même de la dimension phonologique de la prosodie, l'approche phonologique propose plusieurs niveaux d'organisation : par exemple, la structure prosodique est distinguée de l'intonation, elle-même organisée en deux aspects partiellement dissociables, la variation tonale d'une part et la structuration métrique d'autre part.

### **2.1.1. La séparation entre structure prosodique et intonation**

Dans l'approche phonologique, la structure des phrases en constituants prosodiques et les variations des contours intonatifs sont souvent étudiées séparément. L'étude de la structure phonologique est développée dans le cadre de la *phonologie prosodique* (Selkirk 1984 ; Nespor & Vogel 1986) qui définit une structure de constituants, hiérarchique, abstraite, généralement conçue comme non récursive et dépendante de la langue étudiée.

De leur côté, les variations pertinentes des contours intonatifs sont décrites et expliquées par l'étude de l'*intonation*. L'articulation de ces deux dimensions est bien sûr indispensable dans la mesure où les contours intonatifs s'élaborent sur les domaines définis par la structure prosodique. Toutefois, sur l'empan d'une même unité prosodique, le locuteur a le choix entre plusieurs contours intonatifs au sein d'un paradigme distinctif. Ainsi, en français, au moins au niveau du syntagme intonatif (l'unité majeure), un paradigme minimal de quatre contours distinctifs est étudié (Post 2000 ; Beyssade & Marandin 2007 ; Bertrand, Portes & Sabio 2007). En outre, on décrit dans plusieurs langues, des épisodes désaccentués, dépourvus de variation tonale, mais dont la structure de constituants est préservée par la persistance des marques d'allongement (voir Di Cristo & Jankowski 1999 pour le français).

Ce type d'épisodes désaccentués illustre également la dissociation partielle qui peut exister entre la structure tonale et la structure métrique au sein même de l'intonation puisque les proéminences métriques restent marquées malgré l'absence d'événement tonal (voir section suivante).

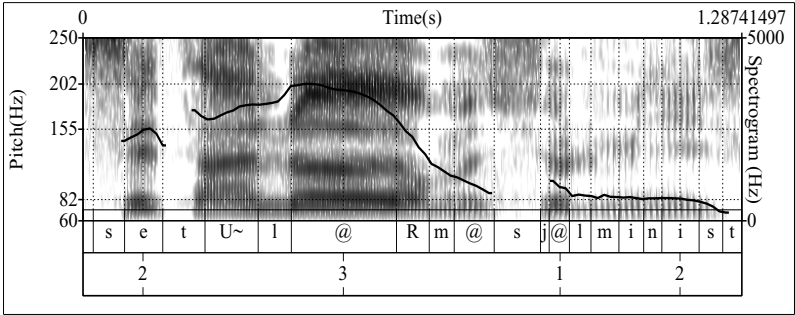
### **2.1.2. La séparation entre structure tonale et structure métrique**

La notion d'*intonation* articule en effet deux dimensions partiellement indépendantes : la structure métrique qui définit les relations entre différents degrés de proéminence des unités syllabiques et la structure tonale qui définit les unités de construction des contours intonatifs. L'intrication complexe de ces deux niveaux est reflétée dans le nom de la théorie qui sert de cadre général au

développement de cette approche : la *théorie autosegmentale-métrique*, la notion d'*autosegment* désignant ici les tons tels qu'ils furent conçus dans les théories phonologiques non linéaires (Goldsmith 1976 ; cf. également § 2.2 *infra*).

L'exemple (1) illustre les épisodes désaccentués évoqués § 2.1.1 et qui témoignent de la dissociation entre structure tonale et structure métrique :

- (1) C'est un leurre, Monsieur le Ministre.



**Figure 1 : Spectrogramme, courbe de f0, phonèmes et grille métrique de *C'est un leurre, Monsieur le Ministre* <sup>a</sup>**

a. Spectrogramme, courbe de f0, phonèmes et grille métrique (dont les proéminences de force relative sont codées 1, 2 et 3) de la phrase *C'est un leurre, Monsieur le Ministre*, dont le segment détaché à droite est désaccentué.

Dans la figure 1, le segment détaché à droite *Monsieur le Ministre* apparaît comme dépourvu de variation tonale et réalisé de façon monotone dans le bas du registre du locuteur (autour de 82 Hz). Malgré cela, les proéminences métriques associées aux syllabes /sj@/ de *Monsieur* (force 1, accent initial) et /nistR/ de *Ministre* (force 2, accent final) sont perceptivement conservées grâce au maintien de l'allongement syllabique. Le marquage du syntagme accentuel est ainsi conservé malgré la désaccentuation (Astesano, Bard & Turk 2007).

Cette distinction paraît d'autant plus importante pour l'étude des conversations qu'il a été montré que la désaccentuation joue un rôle dans le marquage de la structure informationnelle des énoncés dont on sait qu'elle reflète notamment la relation au sein des paires adjacentes (pour le français, voir Di Cristo 1999 ; Beysade *et al.* 2009).

## 2.2. La séparation entre les unités discrètes et les variations graduelles

L'approche phonologique distingue clairement la partie de la prosodie qui relève de l'organisation systématique d'unités discrètes (structure prosodique, contours intonatifs) et le domaine de la variation graduelle des paramètres prosodiques (comme le registre tonal ou la vitesse d'élocution : voir Portes & Bertrand 2011). Ces variations graduelles de la fréquence fondamentale ou f0, de l'intensité, et de la durée y sont étudiées dans leurs relations avec l'expression des émotions et

des attitudes. Par exemple, K. Scherer, R. Ladd et K. Silverman (1984) montrent le rôle des variations paramétriques graduelles et des contours intonatifs (unités discrètes) dans la perception des émotions. C. Gussenhoven (2002) propose, quant à lui, que les unités discrètes proviennent de la grammaticalisation de trois codes biologiques exploitant les variations paramétriques graduelles. Ces travaux montrent l'importance de distinguer, mais pour mieux les articuler, les dimensions phonologiques et graduelle de la prosodie. En voici une illustration en (2).

- (2) Je sais également ce que cela signifie **pour vos familles**, que je veux saluer **particulièrement**, dont j'imagine qu'elles sont souvent confrontées à l'absence et parfois à l'angoisse.

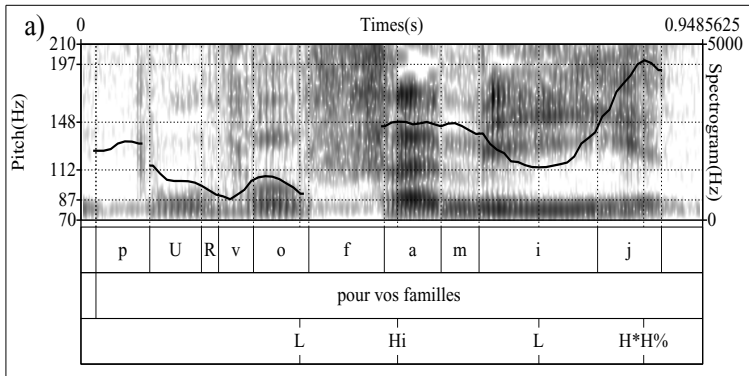


Figure 2 : Spectrogramme, courbe de f0 et annotation tonale de la séquence *pour vos familles* de l'exemple (2)

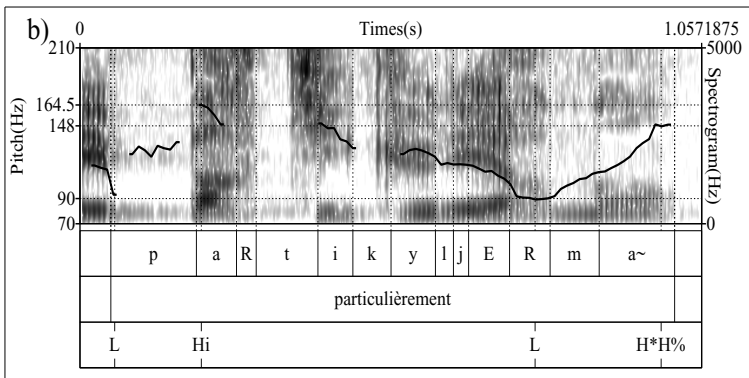


Figure 3 : Spectrogramme, courbe de f0 et annotation tonale de la séquence *particulièrement* de l'exemple (2)

L'exemple (2) est extrait d'un discours de N. Sarkozy (Avanzi *et al.* 2010, corpus C-PROM). Chacune des deux séquences soulignées dans le texte *pour vos familles* (Fig. 2) et *particulièrement* (Fig. 3) se déploient sur l'empan d'un syntagme accentuel, final d'un syntagme intonatif (les deux constituants prosodiques consensuels en français). Toutes deux utilisent le même contour intonatif codé LH<sub>i</sub>LH\*H% où L (low) et H (high) figurent les tons bas et haut qui composent la courbe mélodique : un même ton T a un statut différent selon qu'il est associé avec le début du mot T<sub>i</sub>, avec la syllabe proéminente T\* ou avec la frontière du constituant T% (Jun & Fougeron 2000, 2002). Mais ce contour est employé sur des structures morphosyntaxiques différentes : le ton haut initial Hi est associé à la troisième syllabe « fa- » (Fig. 2) tandis qu'il est associé à la première syllabe « par- » (Fig. 3). Dans les deux cas, Hi est associé à la première syllabe du premier mot de contenu du syntagme accentuel, comme différents modèles phonologiques du français le prédisent (Di Cristo 1999 ; Post 2000). La décomposition du contour intonatif en tons permet ainsi de comprendre pourquoi des courbes de f<sub>0</sub> différentes (la première montée mélodique représentée par Hi n'est pas placée au même endroit dans les deux courbes globales) sont des réalisations du même contour intonatif. Ici, la variation phonétique s'explique par la différence de localisation dans le syntagme de la première syllabe du premier mot de contenu.

Cet exemple illustre l'intérêt de postuler des unités phonologiques abstraites et de dissocier l'analyse phonologique de celle des variations paramétriques graduelles. Les Figures 2 et 3 montrent, en effet, des différences du registre tonal (niveau de hauteur de voix) : LH\*H% assume dans les deux cas la même fonction de liage entre deux propositions consécutives, mais elle se réalise à un *niveau* plus haut (ton L à 112 Hz *vs* 90 Hz) et avec une *amplitude* plus grande (différence entre L et H\*H% : 85 Hz *vs* 58 Hz) entre 2 et 3. D'autre part, on constate que la montée initiale LHi est, au contraire, plus *ample* en 3 qu'en 2 (74 Hz *vs* 61 Hz). Il est possible de tirer au moins deux conclusions : d'abord, le paramètre graduel appelé *registre tonal* peut varier : i) de façon très locale ; ii) indépendamment du contour intonatif ; iii) indépendamment pour la montée finale LH\* et pour la montée initiale LHi en français. Ensuite, on voit l'intérêt d'une modélisation fine du registre tonal comme celle que propose D. R. Ladd (1996-2008) qui distingue deux dimensions partiellement indépendantes : le *niveau* du registre (basé sur les valeurs de f<sub>0</sub> des cibles basses L) et l'*amplitude* du registre (écart entre les valeurs de f<sub>0</sub> des cibles basses L et des cibles hautes H). Une simple moyenne aurait confondu ces deux dimensions.

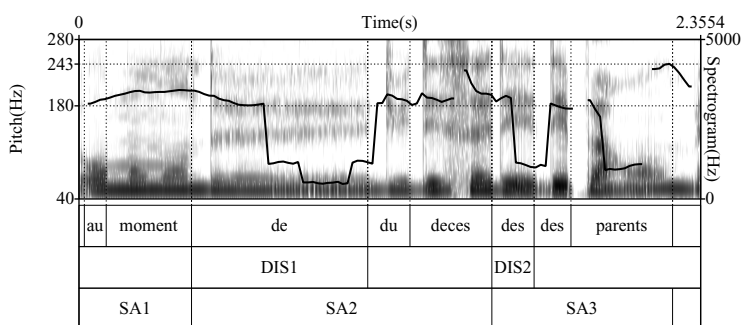
### 2.3. Le traitement séparé des pauses et des disfluences

Disposer d'une conception détaillée et articulée des unités prosodiques phonologiques permet également de rendre compte de façon plus satisfaisante de différents faits prosodiques traités dans la littérature au travers des notions de pause et de disfluence.

La notion de *pause* regroupe un ensemble de faits fonctionnellement disparates, mais similaires de par leur nature acoustique d'interruption du flux de parole (à l'exception des silences liés à la phase d'occlusion des consonnes occlusives sourdes). Certains de ces silences servent de marqueur facultatif aux unités majeures de la structure prosodique ou de l'organisation des tours de parole ; d'autres (parfois les mêmes) permettent au locuteur de reprendre son souffle (pauses dites « respiratoires ») ; d'autres, fréquentes dans les discours politiques ou pédagogiques, permettent de ménager différents effets stylistiques (Duez 1999) ; d'autres, enfin, représentent dans le signal acoustique des traces de l'activité d'encodage, de construction « en ligne » du message : cette dernière catégorie de pause entre dans la typologie des disfluences, où l'on trouve également des répétitions, reformulations et interruptions d'unités linguistiques.

Nous voudrions souligner ici l'importance de concevoir pauses et disfluences comme séparées de la structure phonologique. Cela permet d'expliquer, par exemple, pourquoi les pauses ne sont pas obligatoires pour marquer la frontière des syntagmes phonologiques quand le débit de parole est trop rapide pour le permettre, ou de comprendre pourquoi certaines pauses parfois très longues ne marquent aucune frontière prosodique. En effet, les unités de la structure prosodique phonologique peuvent être interrompues, reprises, reformulées, et même co-construites comme le sont les unités morpho-syntaxiques. C'est pourquoi les pauses et les disfluences peuvent les couper ou les interrompre sans en marquer les bornes. Cela a une conséquence radicale sur la notion même d'*unité prosodique* : la présence d'une pause dans le signal ne suffit plus à identifier la frontière d'une unité prosodique. L'exemple 3 (extrait du CID, voir § 3.2) illustre ce point :

(3) Au moment de du décès des des parents



**Figure 4 : Spectrogramme, courbe de f0, transcription orthographique, disfluences (DIS) et syntagmes accentuels (SA) de *au moment du décès des parents* <sup>a</sup>**

a. Contenu dans un unique syntagme intonatif terminé par un contour majeur de continuation (culminant à 243 Hz). Les syntagmes accentuels SA2 et SA3 sont tous deux interrompus par des disfluences DIS1 et DIS2.

La figure 4 montre comment le syntagme intonatif dans lequel se développe la séquence morpho-syntaxique *au moment du décès des parents* est interrompu deux fois par des disfluences qui ont lieu en début de production des syntagmes accentuels SA1 *du décès* (DIS1) et SA2 *des parents* (DIS2) : chacun de ces deux syntagmes accentuels est initié, interrompu, puis repris tout comme le sont les syntagmes prépositionnels qu'ils mettent en forme prosodiquement. Dans cet exemple, ces interruptions sont signalées acoustiquement par un allongement très marqué de la voyelle finale du segment interrompu (*de, des*)<sup>2</sup>. Dans d'autres cas, elles peuvent l'être par des pauses très longues, souvent remplies par le *euh* d'hésitation. Selon nous, les pauses et ces allongements de segments interrompus ne doivent pas être considérés comme des « unités » prosodiques, mais comme les marqueurs acoustiques du travail d'encodage. Seul un modèle abstrait des unités prosodiques permet d'expliquer comment il est possible de les reconstituer dans les corpus malgré les interférences occasionnées par les épisodes disfluents.

### 3. APPLICATION À L'ANALYSE DE CONVERSATION

Cette section est consacrée à l'illustration, en contexte, de l'intérêt de distinguer les dimensions évoquées précédemment pour l'analyse des conversations. L'application de modèles existants à une parole aussi différente que celle analysée communément par les prosodistes sous-tend la conception de l'existence d'une seule grammaire, quelles que soient les conditions de production, et implique donc une certaine pérennité des catégories établies (voir Portes & Bertrand 2011).

Nous portons notre attention sur l'organisation des tours de parole et, plus précisément, sur l'alternance des locuteurs. Dans la première sous-section, nous rappelons le rôle du composant prosodique dans le système d'organisation des tours de parole. Dans la seconde, nous présentons le corpus dont sont issus les exemples examinés dans la troisième sous-section.

#### 3.1. Les unités de tours

L'analyse des interactions a déplacé le curseur de la phrase ou de l'énoncé, pertinent pour les monologues et l'écrit, vers le tour de parole. Depuis H. Sacks, E. A. Schegloff et G. Jefferson (1974), le tour est défini par des unités de construction de tour (*turn-constructional-unit*, TCU) qui reposent sur des constructions de nature variée. Quand ils sont complets, ils projettent une place potentielle de changement de locuteur (*transition-relevance place*, TRP). E. A. Schegloff (1988) a souligné très tôt l'importance du critère prosodique dans la délimitation des TCUs, par la mise en évidence d'un pic mélodique qui projetterait leur fin imminente. Mais c'est le travail de C. E. Ford et S. A. Thompson (1996) qui établit

---

2. On constate des petites erreurs de détection liées à la présence de *creaky voice*.

le rôle crucial de la prosodie, particulièrement de l'intonation, dans cette projection. Lorsqu'un énoncé est considéré comme complet syntaxiquement, c'est le critère intonatif qui sélectionne et détermine sa complétude en projetant ou non une fin imminente de tour (un point potentiel de changement de locuteurs) (*op. cit.* : 157). Dans le prolongement de ce travail, M. Selting (2000) propose de clarifier davantage encore le rôle de l'intonation dans la définition des tours en distinguant les TCUs des TRPs, jusqu'alors souvent confondus : le TCU devient un tour complet si et seulement si il atteint un point de complétion et donne lieu à une TRP. Ce TCU-final (*vs* TCU-non final) peut être construit avec un (ou plusieurs) TCU(s), lui-même conçu avec une ou plusieurs unités intonatives (définies principalement par la présence d'un contour intonatif majeur).

La segmentation du discours, quel que soit le type de parole, est une étape cruciale de l'analyse (Szczepiek Reed 2011). Mais ce découpage, qu'il soit fondé sur le niveau prosodique, syntaxique ou pragmatique, pose encore de grandes difficultés. Concernant la segmentation prosodique du côté interactionniste, la définition de l'unité intonative principalement liée au contour majeur n'est pas suffisante. Du côté des prosodistes, il s'agit de prendre en compte des phénomènes inhabituels puisqu'ils ont travaillé essentiellement avec de la parole de laboratoire. Mais un aller-retour entre phonologie prosodique et linguistique interactionnelle paraît opérationnel pour améliorer cette étape de segmentation prosodique.

### 3.2. Les données

L'un des apports de la perspective interactionnelle concerne la variété des situations, des populations et des langues considérées. Les études portent autant sur des interactions conversationnelles entre amis que sur des entretiens thérapeutiques, médiatiques, en situation de travail, etc. Les populations considérées sont des enfants, mais aussi des adultes sains ou atteints d'une pathologie (autisme, schizophrénie, etc.). Enfin, les études sont menées sur de nombreuses langues (anglais, japonais, allemand, suédois, finnois, français, entre autres) et s'inscrivent dans une visée de plus en plus comparative.

Mais il existe aussi de nombreux dispositifs favorisant la production de parole non préparée, en situation d'interaction, « semi-naturelle » (Ito & Speer 2006, pour une revue). Le cas du *Map-Task* (Anderson *et al.* 1991, pour la présentation du corpus originel) est le plus connu et le plus répandu puisqu'il a été reproduit sur de nombreuses langues. Ce type de corpus offre une base de données homogène favorisant la comparaison et la systématisation des phénomènes élicités.

De la même manière, l'extrait analysé *infra* est issu du CID (*Corpus of Interactional Data*, Bertrand *et al.* 2008), constitué au Laboratoire Parole et Langage. Le CID est né de la volonté d'éliciter des phénomènes discursifs analysés préalablement dans des corpus « authentiques » (Bertrand & Espesser 1998, 2002 ;

Priego-Valverde 1999), en vue de mener des études quantitatives et systématiques. Par ailleurs, le CID a été constitué pour pallier certaines difficultés liées à l'étude prosodique (et donc acoustique) de la parole en interaction. Cette analyse est rendue possible par son dispositif expérimental : une qualité audio optimale (en chambre sourde pour éviter tout bruit parasite) et l'enregistrement des voix de chaque locuteur sur piste séparée, afin de faciliter l'analyse acoustique des phases de chevauchement de parole, phases dont on sait l'importance dans l'alternance des tours de parole (cf. Depperman 2010). Enfin, le CID a été entièrement transcrit et annoté à différents niveaux linguistiques (voir le projet OTIM<sup>3</sup>) pour permettre l'étude des liens entre ces derniers durant les activités diverses auxquelles se livrent les participants au cours de leur dialogue.

### 3.3. Illustrations

Les extraits présentés *infra* sont issus de récits d'un dialogue entre deux locutrices ayant pour consigne de se raconter des anecdotes insolites. Nous avons transcrit le premier récit dans sa quasi-totalité car il nous permet d'illustrer plusieurs points, tandis que nous ne transcrivons que partiellement les autres.

CM, la locutrice principale, raconte une anecdote qui a eu lieu lors de son premier congrès.<sup>4</sup>

- CM 49 alors on était parti c'était le premier congrès que j'ai fait  
avec euh mh  
CM 50 euh à Montréal  
⇒ CM 51 j'étais avec P. R. et euh donc on était à Montréal et après le  
congrès on est parti en Gaspésie tu vois à peu près où c'est en  
fait euh ouais mh tu connais  
AB 35 oui  
⇒ AB 36 c'est # une région perdue du Canada + oui un peu isolée  
CM 53 voilà ouais c'est ça  
CM 54 et euh et en fait euh  
⇒ CM 55 en f c'était complètement loufoque on arrive y a que des tout  
petits patelins où y a rien quoi en fait où y a un bistrot qui  
AB 37 mh  
fait tout euh etc  
CM 56 et on était p on avait loué une voiture on était avec Y et P en  
fait on était et le soir avec P on s'est dit  
CM 57 \$on va aller\$ on devait pleuter dans un une auberge de jeunesse  
je sais pas com on s'est dit \$on va aller boire un coup dans le  
café du coin\$  
CM 58 et on est tombé  
CM 59 ouais mh côté insolite ça me fait vraiment penser à ça on est  
tombé  
⇒ CM 60 on est rentré dans un bar + qui faisait boîte  
AB 38 @  
CM 61 y avait  
CM 62 que nous  
CM 63 y avait la musique + à fond les ballons  
CM 64 et y avait un espèce d'écran géant donc un matériel d'enfer  
CM 65 qui projetait un film de De Funès avec une musique de boîte @@ à  
fond les ballons @@ @ et on s'est dit \$mais  
CM 66 où on est là qu'est-ce qu'on fait\$  
CM 67 on était complètement euh complètement  
AB 39 ils nous attendent hein  
CM 68 désynchronisés en fait ouais c'est ça de euh

3. <http://www.lpl-aix.fr/~otim>

4. Dans cet extrait, les conventions utilisées sont les suivantes : les chevauchements de parole sont soulignés ; les rires sont notés @ ; les paroles dites en riant bornées par @@ ; les noms propres sont bornés par \$ ; les + renvoient aux pauses inférieures à 200 ms ; les locuteurs sont à gauche ; et chaque numéro de ligne correspond à un groupe phonatoire (borné par une pause supérieure à 200 ms).

Nous examinons *infra* les variations prosodiques liées aux chevauchements de parole et la question des pauses.

### Les chevauchements de parole

La locutrice principale (tire 1 : CM 51) s'engage dans son récit. On remarque (Fig. 5) que la partie finale de l'énoncé, à partir de *Gaspésie*, est marquée par une compression de la gamme tonale (*tu vois à peu près où c'est en fait ouais*). Ce type d'épisode compressé (ou à l'inverse élargi) est fréquent en conversation. Nous attribuons cette variation de la gamme tonale à la nature incidente de l'énoncé. Par ce dernier, la locutrice s'assure de l'état des connaissances partagées, qui peut s'avérer garant de la poursuite de son récit. La réponse de l'interlocutrice (l. 4 : *oui*), bien qu'en chevauchement avec la parole de la locutrice principale, apparaît immédiatement après la fin d'une unité majeure (codée 3). Ce niveau 3 correspond au syntagme intonatif, évoqué avec le syntagme accentuel (codé 1) dans la première partie. Nous utilisons également un niveau intermédiaire de frontière (codé 2) utile pour la conversation. Nous constatons que la distinction concernant ce qui relève du niveau du phrasé et ce qui relève de la variation continue (cf. § 2.2) permet, d'une part, d'expliquer la prise de parole de l'interlocutrice en ce lieu spécifique et, d'autre part, de caractériser la nature de l'énoncé de la locutrice principale.

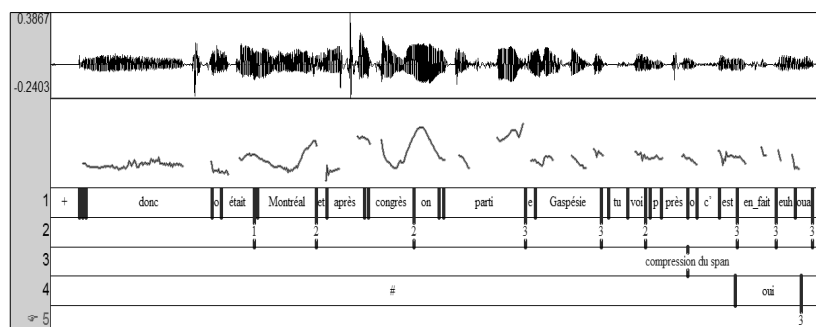


Figure 5 : Distinction entre les unités discrètes et les variations graduelles (CM51 et AB35) <sup>a</sup>

a. Chaque tire correspond à un niveau d'annotation.

La figure 6 est la suite immédiate de la conversation entre les deux interlocutrices. On constate que la prise de parole de AB apparaît après une unité codée 3. De la même manière, CM reprend la parole après une unité majeure de AB (locutrice 2, tire 4). Ce tour, produit en chevauchement, est de nature coopérative, d'autant qu'il s'agit de signaux backchannel (BC) marquant l'accord de CM. Ce point confirme les résultats de R. Bertrand *et al.* (2007) selon lesquels les BC sont produits de manière significative après des unités intonatives majeures. Dans le même temps, on peut noter qu'AB reformule la première partie de son discours,

n'ajoutant ainsi aucune information à ce qu'elle vient de déclarer. Ceci confirme la nature coopérative du chevauchement.

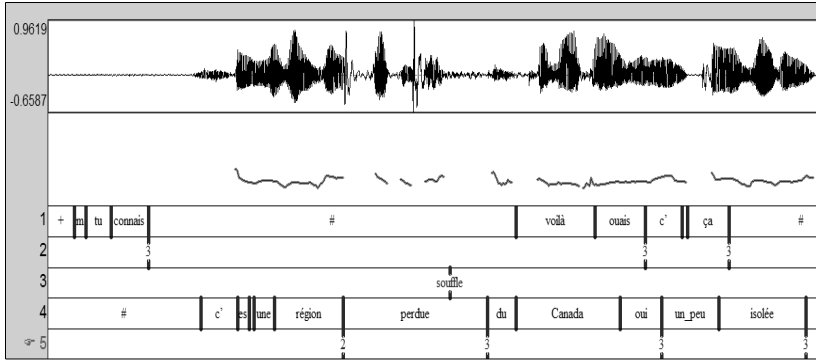


Figure 6 : Phrasé et chevauchement coopératif (AB 36, CM 53) <sup>a</sup>

a. Le signal et la courbe de fréquence fondamentale correspondent à la parole de la locutrice principale (l. 4).

Par opposition, les figures suivantes (7 et 8) illustrent des cas de chevauchement non coopératif issus de deux autres récits. La transcription est présentée avant chaque figure respectivement.

La prise de parole de AB (Fig. 7, l. 1) apparaît au milieu d'une unité prosodique de la locutrice principale (l. 4 : *y en avait un autre*, bornée à gauche par une frontière de niveau 3 et à droite de niveau 2). Ce point fait écho au travail de B. Wells et S. MacFarlane (1998) selon lequel un chevauchement est coopératif ou compétitif selon sa localisation dans le TCU, mais aussi selon qu'il apparaît avant ou après l'accent nucléaire (défini par ces auteurs comme l'accent principal, final de l'énoncé). Ainsi, un chevauchement est jugé comme compétitif s'il apparaît non seulement avant la fin d'un TCU, mais aussi avant l'accent nucléaire. S'il apparaît avant la fin du TCU, mais après l'accent nucléaire, il peut être considéré comme une simple anticipation de la fin du tour.

AB 88 on se demandait ce qu'on faisait là en fait et ils nous ont  
servi les repas typiques euh  
AB 90 style de le la dinde avec euh de la gelée à je sais pas trop  
quoi y en a un autre en face qui était  
CM 142 ah c'est aussi typique en Angleterre  
AB 91 oui oui

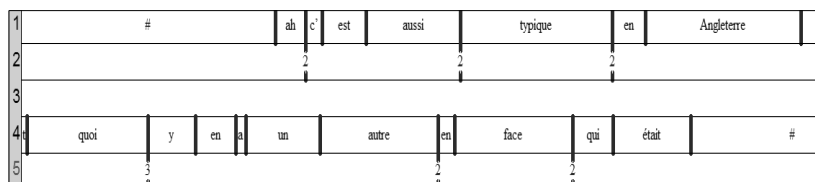


Figure 7 : Phrasé et chevauchement de parole compétitif

Dans la figure 8, la prise de parole par l'interlocutrice AB (l. 4) apparaît peu après une frontière d'unité qui ne présente pas un niveau de finalité majeur (2). La configuration prosodique qui se déploie alors chez la locutrice interrompue (l. 1) illustre une variation du registre vers le haut qui signale son intention de garder le tour (voir French & Local 1983, qui y ajoute le rôle de l'intensité). On constate encore l'intérêt de dissocier unités du phrasé et variations continues pour rendre compte de l'alternance des tours, bien qu'en l'occurrence les deux indices prosodiques concourent à signaler la nature compétitive du chevauchement.

CM 249 très bizarre c'est que je savais plus à qui je m'adressais c'est-à  
 -dire qu'y avait des gens  
 CM 250 qui étaient en blouse blanche etc mais je savais pas si c'était  
du personnel ou pas du personnel  
 AB 130 tu savais pas si c'étaient des vrais ou des faux oui @

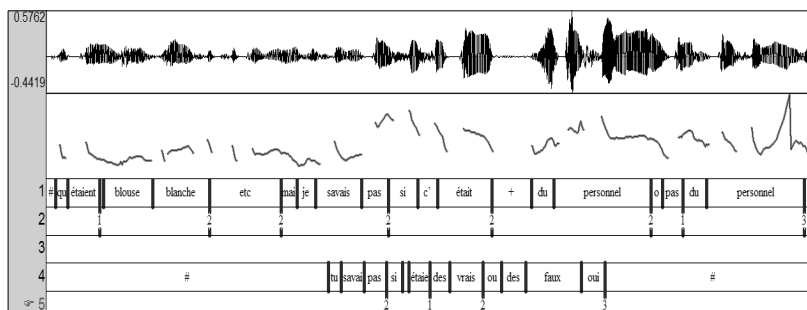


Figure 8 : Chevauchement de parole et variation du registre <sup>a</sup>

a. Le signal et la courbe de fréquence fondamentale correspondent à la parole de la locutrice principale (l. 1).

### À propos des pauses

À la différence des modèles traditionnels de la prosodie, dans lesquels la pause est souvent considérée comme un des corrélats les plus importants de la présence d'une frontière, l'analyse de la parole en interaction implique de prendre en compte le rôle variable des pauses et de certains allongements, et de ne pas les réduire à leur seul rôle de marqueur de frontière. S'ils peuvent marquer la fin

d'un tour, les pauses et les allongements peuvent aussi révéler un problème de conceptualisation (hésitation) ou contribuer à produire un effet stylistique.

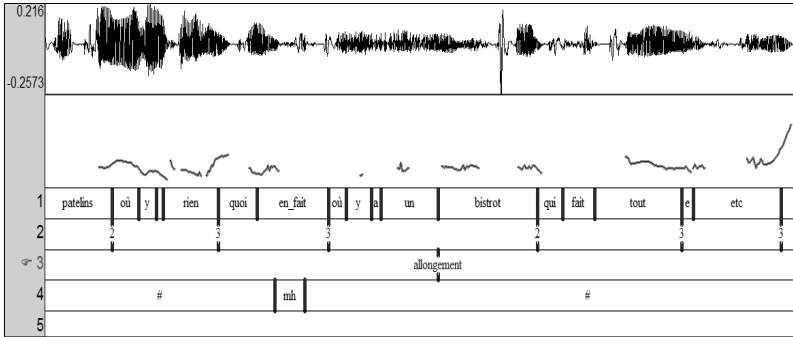


Figure 9 : Phrasé prosodique et allongement disfluent sur le premier récit (CM 55)

Outre le fait que cette figure illustre une nouvelle fois l'importance du phrasé dans la production des signaux BC (le *mh* apparaît après une unité prosodique codée 3), elle illustre surtout la présence d'un allongement sur le déterminant *un* précédant le mot *bistrot*. Cet allongement, considéré comme disfluent, n'a pas donné lieu à la pose d'une frontière (voir fig. 4).

Plus loin dans ce récit, nous observons la présence de pauses (notées #) qui jouent un autre rôle et ont un lien au phrasé quelque peu différent.

Le passage considéré a reçu le codage en unités prosodiques suivant :

- 1 on est rentré]1 dans un bar]2 # qui faisait boîte]3 #
- 2 y avait]2 # que nous]3 #
- 3 y avait]1 la musique]2 # à fond les ballons]3 #
- 4 et y avait]1 un espèce]1 d'écran]1 géant]2 donc]1 un matériel]2 d'enfer]3 #
- 5 qui proj(e)tait]1 un film]2 de De Funès]3 avec une musique]3 de boîte]3 à fond les ballons]3 #

On constate trois réitérations de la même séquence d'unités prosodiques : la première unité a une force de frontière codée 2 et la seconde 3. Entre les deux unités, on observe une pause silencieuse (extrait 1, 2, 3). Mais cette dernière ne constitue pas le paramètre principal pour décider de la présence d'une frontière en ce lieu puisque l'on observe la même constituance prosodique dans les séquences suivantes : *matériel]2 d'enfer]3*, *un film]2 de De Funès]3* sans que celles-ci ne comportent de pause.

La pause considérée est stylistique et pragmatique. Elle correspond à une nouvelle phase dans le récit de la locutrice : la phase de complication doublée de l'apex (point culminant de l'histoire) selon les phases constituantes du récit chez W. Labov (2007). Cette phase narrative est donc signalée par une prosodie spécifique marquée par la réitération d'une structuration typique, qui met en lumière

cette séquence particulière. Les pauses permettent de renforcer ce qui suit, qui est notamment marqué par un accent d'insistance à l'initiale contribuant au renforcement. La figure suivante illustre la configuration prosodique de l'unité après la pause, qui conjugue donc un accent initial (*fond*) et un accent final (*ballons*) qui crée un « arc accentuel » (Di Cristo 1999) qui contribue à la cohésion de l'unité. Notons que cette configuration est similaire à celle des figures 2 et 3. Réalisée alors sur un seul mot, elle s'applique ici sur une expression comportant plusieurs mots, lesquels sont précisément reliés par cette configuration spécifique.

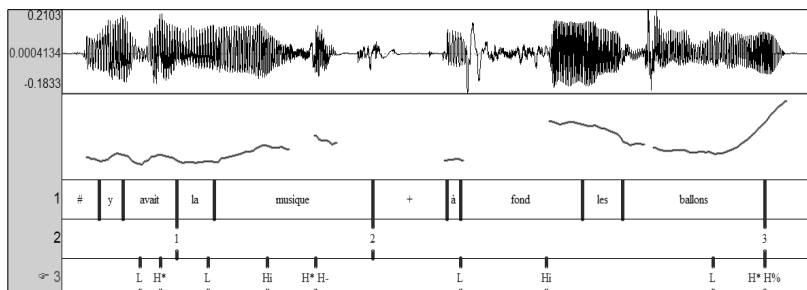


Figure 10 : Arc accentuel (LHi LH\* H%) marquant l'expression à fond les ballons

L'analyse de cet exemple a montré le rôle spécifique de certaines pauses dans la réitération de certaines structures prosodiques qui permettent d'illustrer la mise en relation de l'organisation prosodique avec d'autres niveaux d'organisation du discours comme les phases de narration. Il est à noter également que le changement de phrasé de la part de la locutrice est saillant puisqu'il produit un effet sur l'interlocutrice qui, dès la première itération de la séquence en question, rit, comme par anticipation de la suite du récit dont elle sait que la chute sera drôle. La réitération de ce phrasé sur plusieurs occurrences permet de rendre la phase en question cohérente et saillante par rapport aux autres phases du récit.

#### 4. CONCLUSION

Dans cet article, nous avons cherché à montrer les apports d'une approche phonologique fondée sur des catégories d'analyse préétablies, qui permet de distinguer les différents phénomènes prosodiques à l'œuvre dans l'interaction, afin de mieux cerner leurs implications interactionnelles à travers le rôle des uns et des autres dans l'organisation des tours de parole. Nous espérons avoir montré qu'une telle approche permettrait d'améliorer la définition des unités d'analyse de la conversation, notamment les unités de construction des tours. En distinguant clairement unités discrètes et variations graduelles, on peut notamment améliorer notre compréhension des phases de chevauchement de parole qui sont cruciales dans cette organisation des tours de parole. L'approche phonologique de la prosodie dans l'interaction permet aussi de valider certains modèles de

la langue tout en ne niant pas la variabilité à laquelle celle-ci peut être soumise. De ce point de vue, l'analyse de données de parole continue, en interaction, ne peut qu'enrichir les modèles existants. Les pauses ou les allongements syllabiques, reconnus principalement comme marqueurs de frontières, doivent être revisités d'un point de vue fonctionnel. Enfin, les principes de l'approche phonologique semblent pertinents pour favoriser des études sur les différents niveaux de l'analyse linguistique. En effet, la prosodie est envisagée alors comme une organisation autonome, soumise à des règles et à des contraintes internes. Cette conception semble également pertinente pour aborder la question de la multimodalité en favorisant d'abord l'étude des contraintes inhérentes au niveau prosodique avant de considérer les contraintes résultant des liens mutuels existant entre les différents niveaux linguistiques.

### Références bibliographiques

- ANDERSON A.H. *et al.* (1991), "The HCRC map task corpus", *Language and Speech* 34, 351-366.
- ASTESANO C., BARD E. & TURK A. (2007), "Structural influences on Initial Accent placement in French", *Language and Speech* 50 (3), 423-446.
- AVANZI M. *et al.* (2010), "C-PROM. An annotated corpus for French prominence studies", *Proceedings of Prosodic Prominence, Speech Prosody*, Chicago (Cederom).
- BERTRAND R. & ESPESSER R. (1998), « Prosodie et discours rapporté : la mise en scène des voix », in J. Verschuere (éd.), *Pragmatics in 1998. Selected papers from the 6th International Pragmatics Conference*, vol. 2, Antwerp : International Pragmatics Association, 45-56.
- BERTRAND R. & ESPESSER R. (2002), "Voice diversity in conversation: a case study", in B. Bel & I. Marlien (eds), *Proceedings of the First International Conference on Speech Prosody*, Aix-en-Provence: Laboratoire 'Parole & Langage', 171-174.
- BERTRAND R., PORTES C. & SABIO F. (2007), « Distribution syntaxique, discursive et interactionnelle des contours intonatifs du français dans un corpus de conversation », *Travaux Neuchâtelois de Linguistique – TRANEL* 47, 59-77.
- BERTRAND R. *et al.* (2007), "Backchannels revisited from a multimodal perspective", *Proceedings of Auditory-visual Speech Processing*, Hilvarenbeek: The Netherlands, 1-17 (Cederom).
- BERTRAND R. *et al.* (2008), « Le CID – Corpus of Interactional Data – Annotation et Exploitation Multimodale de Parole Conversationnelle », *Traitement Automatique des Langues* 49 (3), 105-134.
- BEYSSADE C. & MARANDIN J. M. (2007), "French Intonation and Attitude Attribution", in P. Denis *et al.* (eds), *Proceedings of the 2004 Texas Linguistics Society Conference: Issues at the Semantics-Pragmatics Interface*, Somerville (MA): Cascadilla Proceedings Project, 1-12.
- BEYSSADE C. *et al.* (2009), "Prosodic Markings of Information Focus in French", in Y. Hiyon & E. Delais-Roussarie (éds), *Actes de 'Interface Discours & Prosodie'*, 109-122.
- BLACHE P., BERTRAND R. & FERRÉ G. (2009), "Creating and exploiting multimodal annotated corpora: the ToMA project", in M. Kipp *et al.* (eds), *Multimodal Corpora: From Models of Natural Interaction to Systems and Applications*, Berlin: Springer-Verlag, 38-53.
- CASPERS J. (2003), "Local speech melody as a limiting factor in the turn-taking system in Dutch", *Journal of Phonetics* 31, 251-276.
- COUPER-KUHLEN E. & FORD C. E. (2004), *Sound patterns in Interaction*, Amsterdam: John Benjamins.

- DEPPERMAN A. (2010), "Future prospects of research on prosody: the needs for publicly available corpora: comments on Margret Selting 'Prosody in Interaction: State of the art'", in D. Barth-Weingarten, E. Reber & M. Selting (eds), *Prosody in Interaction*, Amsterdam: John Benjamins, 41-47.
- DI CRISTO A. (1999a), « Le cadre accentuel du français contemporain : essai de modélisation », *Langues* 3 (2), 184-205.
- DI CRISTO A. (1999b), « Le cadre accentuel du français contemporain », *Langues* 4 (2), 258-267.
- DI CRISTO A. & JANKOWSKI L. (1999), "Prosodic organisation and phrasing after focus in French", *Proceedings of the XIVth International Congress of Phonetic Sciences*, San Francisco: University of California, 1565-1568.
- DUEZ D. (1999), « La signification des pauses dans la production et la perception de la parole », *Parole* 3 (4), 273-299.
- FORD C. E. & THOMPSON S. A. (1996), "Interactional units in conversation: syntactic, intonational, and pragmatic resources for the management of turns", in E. Ochs, E. Schegloff & S. A. Thompson (eds), *Interaction and grammar*, Cambridge: Cambridge University Press, 134-184.
- FORD C. E., FOX B. A. & THOMPSON S. A. (eds) (2002), *The language of turn and sequence*, Oxford: Oxford University Press.
- FOX TREE J. E. (1999), "Listening in on monologues and dialogues", *Discourse processes* 27, 35-53.
- FRENCH P. & LOCAL J. (1983), "Turn competitive incomings", *Journal of Pragmatics* 7, 17-38.
- GOLDSMITH J. A. (1976), "An Overview of Autosegmental Phonology", *Linguistic Analysis* 2 (1), 23-68.
- GRAVANO A. & HIRSCHBERG J. (2010), "Turn-taking cues in task-oriented dialogue", *Computer Speech and Language* 25, 601-634.
- GUMPERZ J. J. (1982), *Discourse strategies*, Cambridge: Cambridge University Press.
- GUSSENHOVEN C. (2002), "Intonation and interpretation: Phonetics and Phonology", in B. Bel & I. Marlien (eds), *Proceedings of the First International Conference on Speech Prosody*, Aix-en-Provence: Laboratoire 'Parole & Langage', 47-57.
- HELDNER M., EDLUND J. & HIRSCHBERG J. (2010), "Pitch similarity in the vicinity of backchannels", *Proceedings Interspeech 2010*, Makuhari: Japan, 3054-3057.
- ITO K. & SPEER S. R. (2006), "Using interactive tasks to elicit natural dialogue", in P. Augursky & D. Lenertova (eds), *Methods in Empirical Prosody Research*, Berlin/New York: Mouton de Gruyter, 229-257.
- JUN S.-A. & FOUGERON C. (2000), "A phonological model of French intonation", in A. Botinis (ed.), *Intonation: Analysis, modelling and technology*, Boston: Kluwer, 209-242.
- JUN S.-A. & FOUGERON C. (2002), "Realizations of accentual phrase in French", *Probus* 14, 147-172.
- KOISO H. (1998), "An analysis of turn-taking and backchannels based on prosodic and syntactic features in Japanese Map Task dialogs", *Language and Speech: Special issue on Prosody and Conversation* 41 (3/4), 295-321.
- LABOV W. (2007), "Narrative pre-construction", in M. Bamberg (ed.), *Narrative -State of the Art*, Amsterdam: John Benjamins, 47-56.
- LADD D. R. (1996-2008), *Intonational phonology*, Cambridge: Cambridge University Press.
- LAFOREST M. (1992), « L'influence de la loquacité de l'informateur sur la production de signaux backchannel par l'intervieweur en situation d'entrevue sociolinguistique », *Language Variation and Change* 4, 163-177.

- LEHISTE I. (1970), *Suprasegmentals*, Cambridge (MA): MIT Press.
- LOCAL J., WELLS W. & SEBBA M. (1985), "Phonology for conversation: phonetic aspects of turn delimitation in London Jamaican", *Journal of Pragmatics* 9, 309-330.
- NESPOR M. & VOGEL I. (1986), *Prosodic Phonology*, Dordrecht: Foris.
- OGDEN R. (2004), "Non modal voice quality and turn-taking in Finnish", in E. Couper-Kuhlen & C. E. Ford (eds), *Sound patterns in Interaction*, Amsterdam: John Benjamins, 29-62.
- PORTES C. & BERTRAND R. (2006), "Some cues about the interactional value of the 'continuation' contour in French", in C. Auran et al. (eds), *Discours & Prosody as a Complex Interface (IDP)*, Cederom (14 pages).
- PORTES C. & BERTRAND R. (2011), « Permanence et variation des unités prosodiques dans le discours et l'interaction », *Journal of French Language Studies* 21 (1), 97-110.
- POST B. (2000), *Tonal and phrasal structures in French intonation*, The Hague: Holland Academic Graphics.
- PRIEGO-VALVERDE B. (1999), *L'humour dans les interactions conversationnelles : jeux et enjeux*, Thèse de doctorat de l'Université Aix-Marseille I.
- ROSSI M. et al. (1981), *L'intonation. De l'acoustique à la sémantique*, Paris : Klincksieck.
- SACKS H., SCHEGLOFF E. A. & JEFFERSON G (1974), "A simplest systematics for the organization of turn-taking for conversation", *Language* 50, 696-735.
- SCHEGLOFF E. A. (1982), "Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences", in D. Tannen (ed.), *Analyzing discourse: Text and talk*, Washington D.C.: Georgetown University Press, 71-93.
- SCHEGLOFF E. A. (1988), "Discourse as an interactional achievement II: an exercise in conversation analysis", in D. Tannen (ed.), *Linguistics in Context: Connecting Observation and Understanding*, Norwood (NJ): Ablex, 135-158.
- SCHERER K., LADD R. & SILVERMAN K. (1984), "Vocal cues to speaker affect: testing two models", *JASA* 76, 1346-1356.
- SELKIRK E. (1984), *Phonology and syntax: the relation between sound and structure*, Cambridge (MA): MIT Press.
- SELTING M. (2000), "The construction of units in conversational talk", *Language and Society* 29, 477-517.
- SELTING M. (2010), "Prosody in interaction: State of the art", in D. Barth-Weingarten, E. Reber & M. Selting (eds), *Prosody in Interaction*, Amsterdam: John Benjamins, 3-40.
- SELTING M. & COUPER-KUHLEN E. (2001), *Studies in Interactional Linguistics*, Amsterdam: John Benjamins.
- SZCZEPEK REED B. (2006), *Prosodic orientation in English Conversation*, Basingstoke: Palgrave Macmillan.
- SZCZEPEK REED B. (2011), *Analysing conversation. An introduction to prosody*, Basingstoke: Palgrave MacMillan.
- TANAKA H. (2004), "Prosody for marking transition-relevance places in Japanese conversation: The case of turns unmarked by utterance-final objects", in E. Couper-Kuhlen & C. E. Ford (eds), *Sound Patterns in Interaction*, Amsterdam: John Benjamins, 63-96.
- WELLS B. & MACFARLANE S. (1998), "Prosody as an interactional resource: turn projection and overlap", *Language and Speech* 41 (3/4), 265-294.
- XU Y. (2005), "Speech melody as articulatorily implemented communicative functions", *Speech Communication* 46, 220-251.