

# Similarités prosodiques entre langues romanes

Philippe MARTIN

LLF, UFR Linguistique  
Université Paris Diderot  
philippe.martin@linguist.univ-paris-diderot.fr

## RESUME

La théorie Autosegmentale-Métrique, dominante depuis quelques dizaines d'années, définit la structure prosodique d'un énoncé comme une organisation hiérarchique non récursive (ou partiellement récursive). Je me propose ici d'aborder la description prosodique des langues romanes dans une toute autre perspective théorique, permettant entre autre de montrer la similitude phonologique entre les différents systèmes y compris le français, présenté souvent comme le vilain petit canard du groupe roman du fait de l'absence d'accent lexical dans cette langue.

**Mots-clés:** prosodie, phonétique française, théories prosodiques.

[Recibido, septiembre 2013; aprobado, diciembre 2013]

## Prosodic similarities among Romance languages

## ABSTRACT

The Autosegmental-Metric theory, the dominant one in the past decades, defines the prosodic structure of a statement as a non-recursive hierarchical organization (or partly recursive). I propose here to address the prosodic description of the Romance languages in a different theoretical perspective, allowing among other things to show the phonological similarity between different systems including the French one, often presented as the ugly duckling in the Romance group because of the absence of lexical accent in that language.

**Keywords:** prosodics, French phonetics, prosodic theories.

## 1. Introduction

La théorie Autosegmentale-Métrique, dominante depuis quelques dizaines d'années, définit la structure prosodique d'un énoncé comme une organisation hiérarchique non récursive (ou partiellement récursive), regroupant en niveaux successifs des syllabes en mots lexicaux (*content words*) et mots grammaticaux (*grammatical words*), les mots lexicaux et grammaticaux en groupes accentuels AP (*accent phrases*), les groupes accentuels en syntagmes intonatifs intermédiaires ip (*intermediate phrases*), les ip en syntagmes intonatifs IP (*intonation phrases*), et finalement les IP en structure prosodique R. Dans ce cadre, les événements prosodiques relèvent de trois catégories distinctes: les tons de frontière BT (*boundary tones*), les accents mélodiques PA (*pitch accents*) et les accents secondaires ou emphatiques (*secondary accents*). Ces événements prosodiques sont transcrits au moyen du système de notation ToBI, combinant une ou plusieurs cibles tonales Haute et Basse et leurs variantes.

Une conférence récente tenue à Tarragona en 2012 a réuni plusieurs équipes de prosodistes se proposant de comparer les caractéristiques prosodiques de plusieurs langues romanes dans le cadre Autosegmental-Métrique (Romance ToBI workshop 2011). Je me propose ici d'aborder la description prosodique des langues romanes dans une toute autre perspective théorique, permettant entre autre de montrer la similitude phonologique entre les différents systèmes y compris le français, présenté souvent comme le vilain petit canard du groupe roman du fait de l'absence d'accent lexical dans cette langue.

## 2. Le modèle incrémental de stockage-concaténation

Une des différences essentielles du modèle incrémental de stockage concaténation ISC (Martin 2009) porte sur les accents mélodiques, censés dans le modèle autosegmental métrique ne pas interagir entre eux ni participer à l'indication de la hiérarchie prosodique, alors que ces interactions sont essentielles dans le modèle utilisé ici. Dans cette dernière approche, les accents mélodiques participent avec les tons de frontière à l'indication de la structure prosodique. De plus, au lieu de la notation ToBI, c'est un système de traits contextuels spécifiant les contours mélodiques qui est utilisé, de manière à rendre compte des oppositions et contrastes nécessaires et suffisants que doivent maintenir entre eux les événements prosodiques dans une configuration donnée de la hiérarchie prosodique.

Les réalisations des événements prosodiques sont décrites en terme de contours mélodiques et non de tons: plus la structure est complexe (en terme de nombre de groupes accentuels et de niveaux hiérarchiques), plus le nombre de traits des contours nécessaires et suffisants pour se différencier des autres contours qui pourraient apparaître à leur place (donc dans le même contexte) est élevé.

Enfin, une autre différence essentielle caractéristique du modèle incrémental est de considérer les contrastes entre contours mélodiques opérant non pas statiquement sur l'ensemble de la phrase mais dynamiquement sur l'axe temporel relativement au passé et au futur. Du point de vue de l'auditeur en effet, les événements prosodiques sont perçus les uns après les autres dans une fenêtre temporelle très courte, et non pas dans leur ensemble comme le donnent à penser les analyses linguistiques habituelles.

Le processus de stockage concaténation incrémental opère séquentiellement «de gauche à droite», c'est-à-dire d'un événement prosodique à l'événement prosodique suivant, jusqu'à la perception par l'auditeur d'un contour mélodique particulier, conclu-

sif, terminant l'énoncé. Dans ce processus, les événements prosodiques successifs sont perçus et classés par l'auditeur selon un certain nombre de catégories dépendant de chaque système linguistique. Cette identification est contextuelle, c'est-à-dire que chaque contour mélodique est comparé au contour suivant pour opérer le classement, ce qui permet aux locuteurs une assez grande variabilité dans leurs réalisations acoustiques. L'exemple suivant permettra d'illustrer le mécanisme de ce modèle.

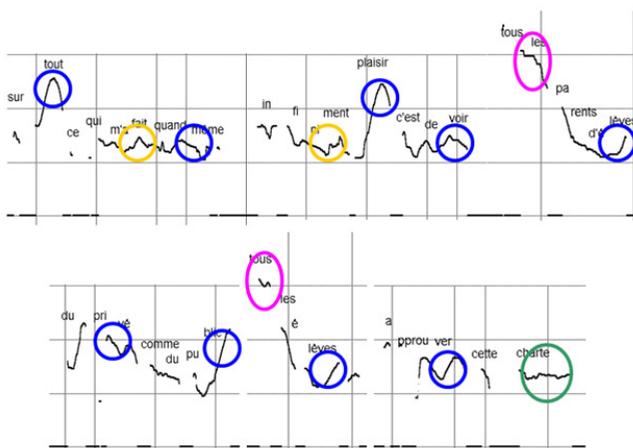
### 3. Un exemple

Claude Allègre, ministre de l'éducation nationale en France dans les années 1997-2000, est interviewé dans une émission de télévision à propos d'une réforme qu'il a proposée récemment. À une question d'un journaliste, il répond: « Surtout ce qui m'a fait quand même infiniment plaisir c'est de voir tous les parents d'élèves du privé comme du public tous les élèves approuver cette charte. » Cet énoncé, spontané mais prononcé par un professionnel de la parole publique et donc selon les normes de l'écrit (sans répétitions, abandons, euh d'hésitation, reformulations, etc.), peut être segmenté en un certain nombre de groupes accentuels (c'est-à-dire des séquences ne comportant qu'une seule syllabe accentuée, réalisée en français en fin de groupe):

| Surtout | ce qui m'a fait | quand même | infiniment | plaisir | c'est de voir | tous les parents d'élèves | du privé | comme du public | tous les élèves | approuver | cette charte |

Dans cette segmentation, les syllabes accentuées sont soulignées en caractère gras. En plus des accents de groupe, en position finale, on remarque également deux accents d'insistance placés sur *tous*.

À l'écoute de cet extrait, on se rend compte que les réalisations acoustiques des syllabes accentuées successives ne sont pas identiques, et se différencient par des caractéristiques prosodiques qui peuvent être faciles à déterminer pour une oreille musicale, mais que les phonéticiens aiment à visualiser sous la forme d'une courbe mélodique (Fig. 1).



**Figure 1.** Courbe mélodique de l'exemple *Surtout ce qui m'a fait quand même infiniment plaisir c'est de voir tous les parents d'élèves du privé comme du public tous les élèves approuver cette charte*. Les syllabes accentuées sont entourées d'un cercle.

Dans la Fig. 1, les segments de courbe mélodique correspondant aux syllabes accentuées sont entourés d'un cercle, ce qui permet de constater que les variations mélodiques sur ces syllabes ne sont pas identiques: certaines sont montantes, d'autres descendantes, d'autres encore plates. Ces variations ne sont pas dues au hasard, mais résultent d'un système qui permet à l'auditeur, en évaluant leurs différenciations, de classer les groupes accentuels successifs et de les assembler sélectivement. Cette opération opère la reconstitution de la structure prosodique voulue par le locuteur, structure facilitant à l'auditeur l'accès à la structure syntaxique de l'énoncé dans une étape ultérieure.

Le lexique prosodique dont dispose l'auditeur est relativement simple: un contour terminal conclusif, présentant (presque) toujours les mêmes caractéristiques: mélodie descendante ou plate atteignant un niveau de fréquence bas. Les autres contours de ce lexique sont appelés C1, montant, C2, descendant, Cn contour neutralisé, court et généralement de peu d'empan mélodique. Le contour d'emphase (appelé aussi d'insistance), caractérisé en général par une montée mélodique sur la première syllabe des mots non grammaticaux, est noté Cemph. Les réalisations de ces contours sont contextuelles, c'est-à-dire que les traits acoustiques qui les caractérisent dans les détails varient selon les contrastes nécessaires et suffisants à maintenir dans un énoncé donné pour se différencier des autres contours qui pourraient apparaître à leur place, déterminant ainsi une autre hiérarchie prosodique. Ainsi dans une structure simple ne comportant que deux groupes accentuels, le contour C1 dans une séquence ne devra se différencier que du contour terminal conclusif C0 (et de ses variantes de commandement, d'évidence, interrogatives, de doute et de surprise). C1 pourra être réalisé par le locuteur de manière à se différencier minimalement avec C0, par exemple par une montée mélodique de faible empan, ou par une durée vocalique plus faible que celle relative à la syllabe accentuée finale. Par contre, dans un énoncé complexe, pourvu d'une structure prosodique à plusieurs niveaux, le système de contrastes nécessaires et suffisants utilisera un plus grand nombre de traits, comme par exemple une pause à la suite d'une montée mélodique importante (comme dans l'exemple de la Fig. 1). La sélection de traits nécessaires et suffisants pour différencier les contours mélodiques appartenant à des classes phonologiques différentes rend compte de la diversité des réalisations, diversité qui a rendu perplexes beaucoup de chercheurs dans ce domaine.

La courbe mélodique de l'exemple annotée selon ces catégories prosodiques est présentée Fig. 2.

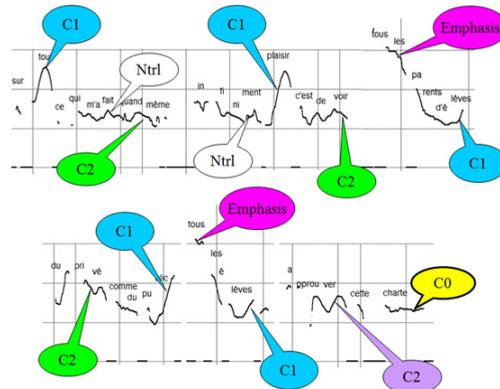


Figure 2. Classification des événements prosodiques de l'exemple en classes C0, C1, C2, Cn et Cemph.

La table 1 donne les détails des opérations de stockage-concaténation incrémental synchronisés par les contours mélodiques successifs. Ces contours déclenchent les opérations de mémorisation dans trois mémoires tampon C2, C1 et C0:

- a) *surtout* terminé par C1 stocké dans la mémoire C1;
- b) *ce qui m'a fait quand même* terminé par C2 stocké dans la mémoire C2. Cn neutralisé ne participe pas à ce mécanisme;
- c) *infiniment plaisir* terminé par C1, est concaténé avec la séquence mémorisée dans C2 (*ce qui m'a fait quand même*) et stockée dans la mémoire C1, qui contient à cette étape *qui m'a fait quand même infiniment plaisir*. La mémoire C2 est remise à zéro. Cn neutralisé ne participe pas à ce mécanisme;
- d) *c'est de voir* terminé par C2 stocké dans la mémoire C2;
- e) *tous les parents d'élèves* terminé par C1, est concaténé avec la séquence mémorisée dans C2 (*c'est de voir*) et stockée dans la mémoire C1, qui contient à cette étape *c'est de voir tous les parents d'élèves*. La mémoire C2 est remise à zéro;
- f) *du privé* terminé par C2 stocké dans la mémoire C2;
- g) *comme du public* terminé par C1, est concaténé avec la séquence mémorisée dans C2 (*du privé*) et stockée dans la mémoire C1, qui contient à cette étape *du privé comme du public*. La mémoire C2 est remise à zéro;
- h) *tous les élèves* terminé par C1, est stocké dans la mémoire C1;
- i) *approuver* terminé par C2 stocké dans la mémoire C2;
- j) *cette charte* terminé par le contour conclusif Co, est concaténé avec toutes les séquences syllabiques restantes stockées dans la mémoire C1 pour reconstituer la structure prosodique complète de l'énoncé. Toutes les mémoires sont remises à zéro pour faire place au traitement de l'énoncé suivant.

Lors des différentes étapes de la concaténation des séquences syllabiques mises en mémoire, la mémoire C0 ne peut contenir qu'un nombre limité de groupes accentuels. On peut imaginer que le locuteur ne retienne que des mots clés, comme par exemple *surtout, plaisir, voir, élèves, privé, public, charte*, c'est-à-dire les mots terminés par un accent de groupe ou porteurs d'un accent d'insistance, ou en tout cas procède à une interprétation sémantique du groupe en question.

Séquence syllabique	Évènement prosodique	Mémoire C2	Mémoire C1	Mémoire C0
<i>Surtout</i>	C1		<i>Surtout</i>	
<i>ce qui m'a fait quand même</i>	Cn C2	<i>ce qui m'a fait quand même</i>		
<i>infiniment plaisir</i>	Cn C1		<i>ce qui m'a fait quand même +infiniment plaisir</i>	
<i>c'est de voir</i>	C2	<i>c'est de voir</i>		
<i>tous les parents d'élèves</i>	Cemph C1		<i>c'est de voir + tous les parents d'élèves</i>	
<i>du privé</i>	C2	<i>du privé</i>		
<i>comme du public</i>	C1		<i>du privé + comme du public</i>	
<i>tous les élèves</i>	Cemph C1		<i>tous les élèves</i>	
<i>approuver</i>	C2	<i>approuver</i>		
<i>cette charte</i>	C0			<i>Surtout + ce qui m'a fait quand même infiniment plaisir + c'est de voir tous les parents d'élèves + du privé comme du public + tous les élèves + approuver cette charte</i>

**Table 1.** Le processus de stockage-concaténation incrémental synchronisé par la séquence des contours mélodiques de l'exemple

#### 4. Différences entre le français et les autres langues romanes

À la différence des autres langues romanes, bien que dérivant du latin, le français a perdu progressivement les syllabes postérieures aux syllabes accentuées (sauf celles contenant un *e* muet), et seul l'accent de groupe fait aujourd'hui partie du système. Comme on l'a vu dans l'exemple ci-dessus, ce sont les tons de frontière instanciés par des contours mélodiques spécifiques qui encodent la structure prosodique et permettent à l'auditeur de reconstituer la hiérarchie des groupes accentuels. On notera que l'occurrence de ces contours, et donc des syllabes accentuées (à l'exception des accents d'insistance) est contrainte dans le temps. Deux syllabes accentuées successives ne peuvent être réalisées et perçues que séparées d'au moins 250 ms environ. D'autre part, deux syllabes accentuées successives ne peuvent être distantes de plus de 1250 ms environ (Martin 2013). Cette dernière contrainte rend compte de l'impossibilité de prononcer quelque sept ou huit syllabes d'affilée sans en accentuer au moins deux, comme dans des mots très longs comme *anticonstitutionnellement oula paraskevidékatriaphobie* (la peur du vendredi 13). Cette limite se manifeste surtout pour le français du fait de l'absence d'accent lexical.

Les autres langues romanes, l'italien, l'espagnol, le catalan, le portugais, le roumain, etc. possèdent à la fois un accent lexical et un des tons de frontière. On a montré, à l'inverse de ce que disent les tenants du modèle autosegmental-métrique, que ces deux

types d'événements prosodiques participent à l'indication de la structure prosodique pour l'auditeur. Dès lors, en se reportant aux mêmes mécanismes liés au processus de stockage-concaténation incrémental, la seule différence caractéristique des autres langues romanes que le français porte sur l'existence d'un contour mélodique complexe Cc, inexistant en français, réalisé par une variation mélodique légèrement descendante ou plate sur la syllabe accentuée, et une montée relativement forte sur la syllabe finale du groupe accentuel. Si les syllabes accentuée et finale se confondent, le contour complexe est réalisé sur une seule syllabe (en fait sur la seule voyelle accentuée).

Les observations expérimentales permettent d'établir que les langues romanes disposent d'un lexique de contours prosodiques composé de Cn, contour neutralisé, C2 contour descendant, C1 contour montant, Cc contour complexe et C0 contour conclusif descendant et bas. Ce sont ces contours qui permettent à l'auditeur de reconstituer la structure prosodique encodée par le locuteur selon le processus de stockage-concaténation incrémental décrit plus haut.

### 5. Les contours mélodiques dans les phrases lues

Les figures qui suivent donnent quelques exemples de réalisations des contours mélodiques pour des phrases similaires des points de vue syntaxique et lexical dans six langues romanes. Ces phrases, extraites du corpus EuRom4 (1991-1997), ont été lues par des locuteurs natifs. On a entouré sur chacun de ces graphiques par une ellipse les segments de courbe mélodique correspondant à une syllabe accentuée, et par un rectangle ceux correspondant à une syllabe finale d'un contour complexe Cc. La structure prosodique telle qu'indiquée par les séquences de contours est indiquée dans chaque exemple par un parenthésage [ ].

Une première série d'exemples montre que pour chacune de ces phrases, la séquence de contours est la même C1 C0, à l'exception du roumain Cc C0. Cette dernière différence est due à un facteur phonostylistique donnant plus d'importance à la frontière prosodique terminant le premier groupe accentué. Les phrases analysées sont:

Français: *L'information est essentielle*

Italien: *L'informazione è essenziale*

Espagnol: *La información es esencial*

Catalan: *La informació és esencial*

Portugais: *A informação é essencial*

Roumain: *Informația este esențială*

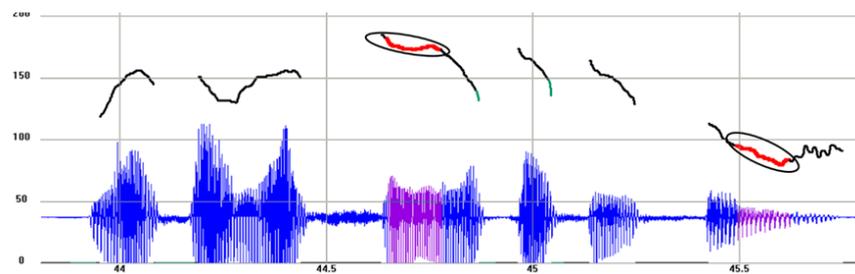


Figure 3. Français frfn5F [L'information C1] [est essentielle C0]

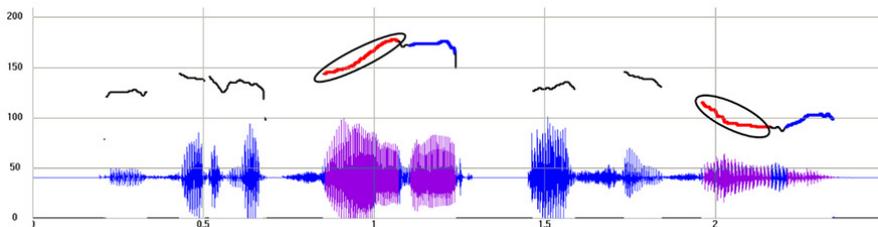


Figure 4. Italien itfn5I [L'informazione C1] [è essenziale C0]

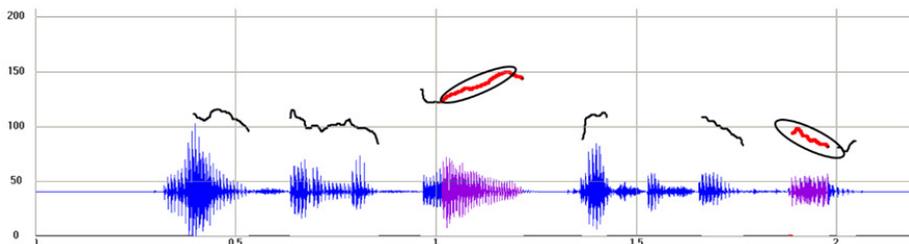


Figure 5. Espagnol esfn5E [La información C1] [es essencial C0]

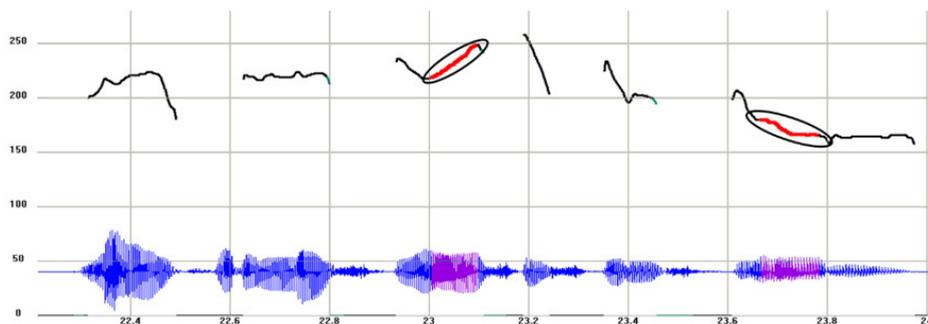


Figure 6. Catalan cafn5C [La informació C1] [és essencial C0]

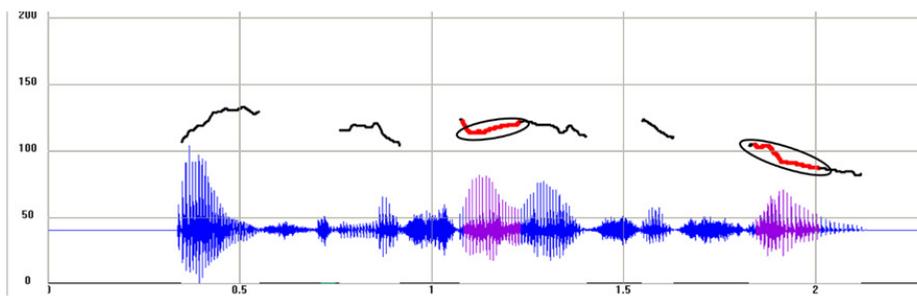


Figure 7. Portugais ptfn5P [A informação C1] [é essencial C0]

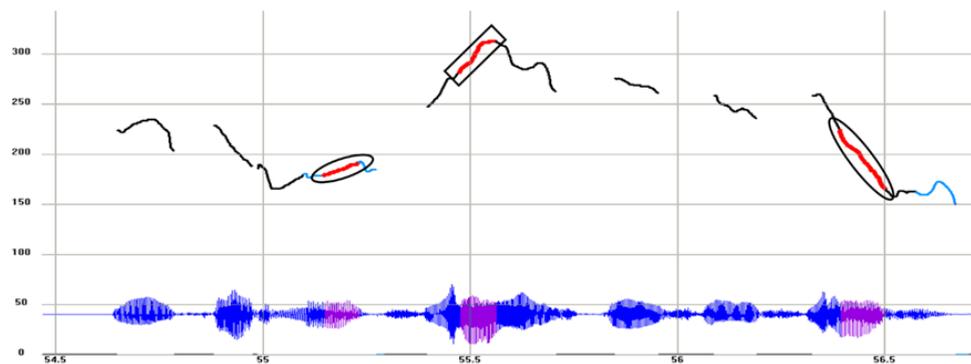


Figure 8. Roumain rofn5R [Informația Cc] [este esențială C0]

On constate également que la réalisation en français du contour C1 est presque totalement neutralisée quant à la variation mélodique, très restreinte.

Les séquences de contour entre le français et les autres langues romanes se différencient lorsque le nombre de groupes accentuels devient plus important et que la structure qui les organise hiérarchiquement devient plus complexe, comme dans les réalisations suivantes qui mettent en jeu le contour complexe Cc, absent en français.

Français: *La violence raciste a dépassé en Allemagne la limite*  
 Italien: *La violenza razzista ha superato in Germania il limite*  
 Espagnol: *La violencia racista ha superado en Alemania el límite*  
 Catalan: *La violència racista ha superat a Alemanya el límit*  
 Portugais: *A violência racista na Alemanha ultrapassou o limite*  
 Roumain: *In Germania violența rasistă a depășit limita*

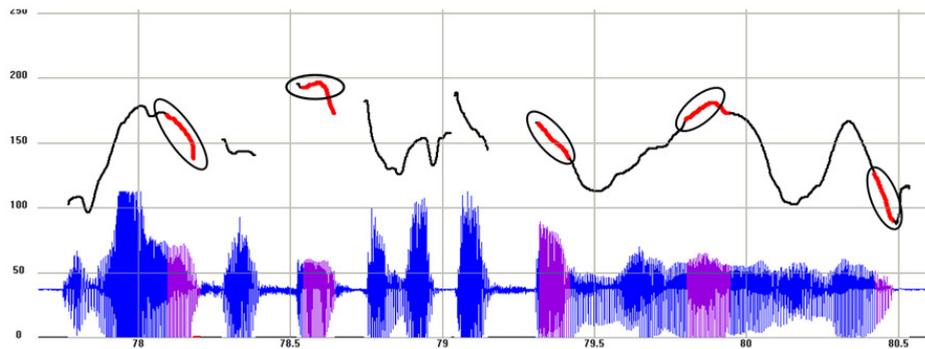
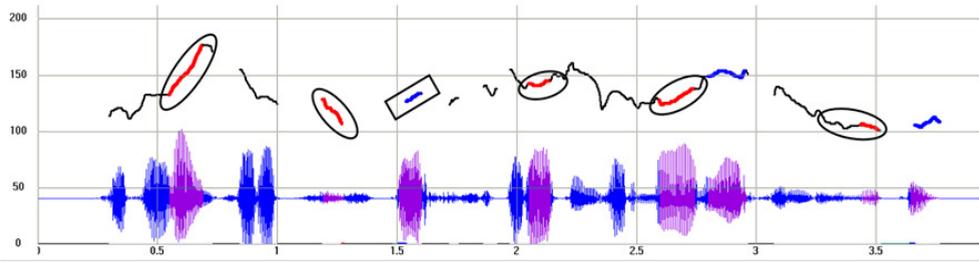
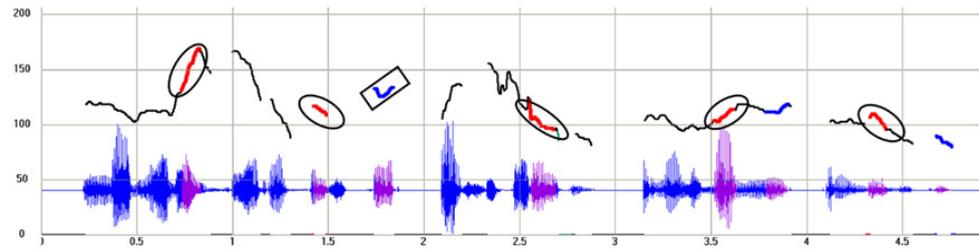


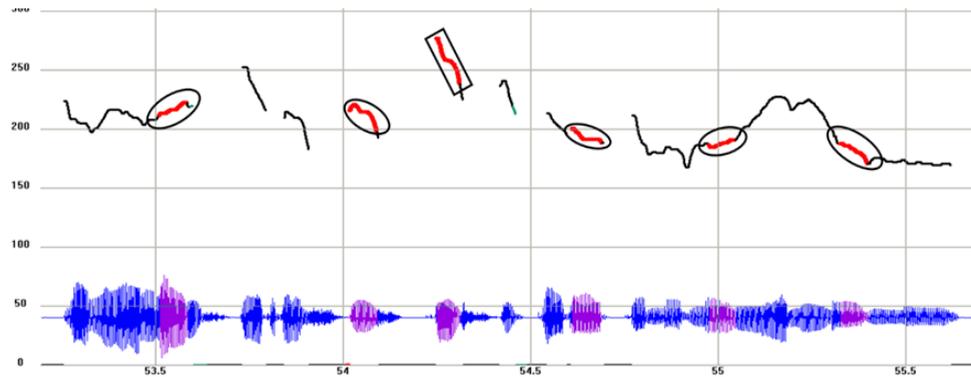
Figure 9. Français frfn15F. [La violence C2 raciste C1] [a dépassé C2 en Allemagne C1] [la limite C0] (Le deuxième contour mélodique est effectivement montant mais ne se manifeste que partiellement sur la voyelle de la syllabe, précédée d'une consonne fricative non voisée). Le système de contraste de pente mélodique C2 C1 caractéristique du français se manifeste également en deux endroits dans cet exemple.



**Figure 10.** Italien itfn15I. [La violenza C1 razzista Cc] [ha superato C1] [in Germania C1] [il limite C0]



**Figure 11.** Espagnol esfn15E. [La violencia C1 racista Cc] [ha superado C2 en Alemania C1] [el límite C0]



**Figure 12.** Catalan cafn15C. [La violència C1 racista Cc] [ha superat Cn a Alemanya Cc] [el límit C0]

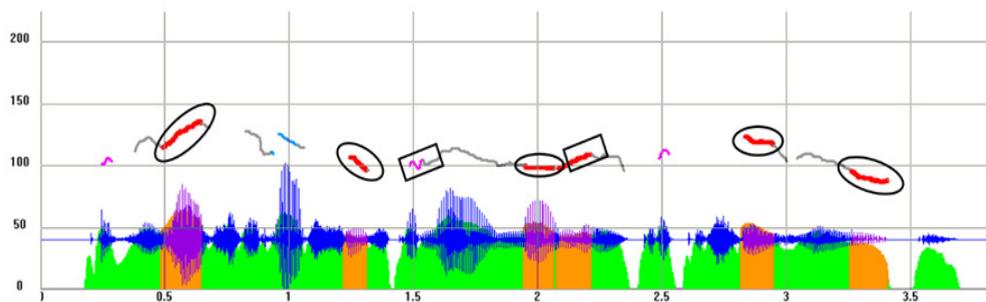


Figure 13. Portugais ptfn15P. [A viol<sup>ên</sup>cia C1 rac<sup>is</sup>ta Cc] [na Ale<sup>ma</sup>nha Cc] [ultrapassou Cn] [o lim<sup>ite</sup> C0]

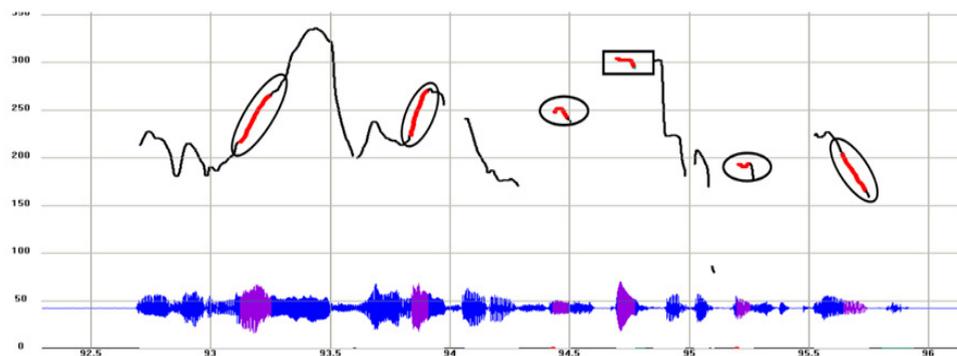


Figure 14. Roumain rofn15R. [In Ger<sup>ma</sup>nia C1] [viol<sup>en</sup>ta C1 ras<sup>is</sup>tă Cc] [a dep<sup>ă</sup>șit Cn] [lim<sup>ita</sup> C0]

Les séquences de contours mélodiques dans le groupe prosodique correspondant à [violence raciste] est donc le même pour toutes les langues romanes C1 Cc, à l'exception du français, C2 C1. Pour les autres groupes, les réalisations varient, mais en utilisant le même système de contrastes entre contours mélodiques.

Un dernier exemple illustre la différence dans la séquence de contours mélodiques lorsque le locuteur encode une hiérarchie plus simple. Dans ces réalisations (à l'exception du roumain, qui adopte, du moins dans cet exemple, une configuration semblable aux exemples précédents), le premier contour mélodique est descendant et contraste avec un contour montant (comme en français).

Français: L'attentat de Molln a relancé la mobilisation contre le racisme  
 Italien: L'attentato di Molln ha rilanciato la mobilizatione contro il razzismo  
 Espagnol: El atentado de Molln ha relanzado la movilización contra el razismo  
 Catalan: L'atemptate de Molln ha rellançat la movilització contra el racisme  
 Portugais: O atentado de Molln relançou a mobilização contra o racismo  
 Roumain: Atentatul de la Molln a relansat mobilizarea împotriva rasimului

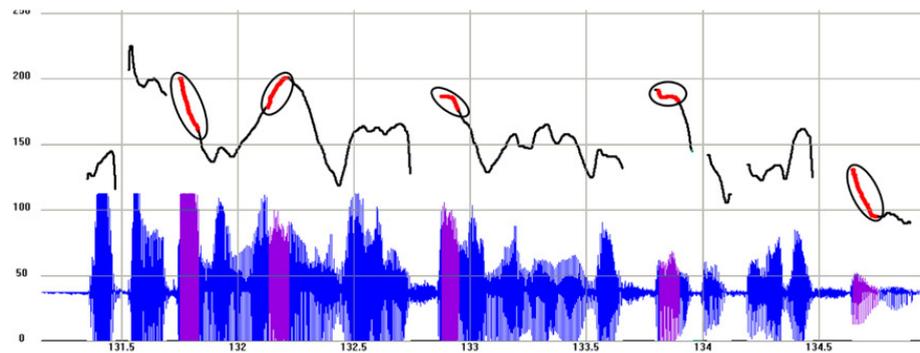


Figure 15. Français frfn26F. [L'attentat C2 de **Molln** C1] [a relancé Cn la mobilisation Cn contre le racisme C0]

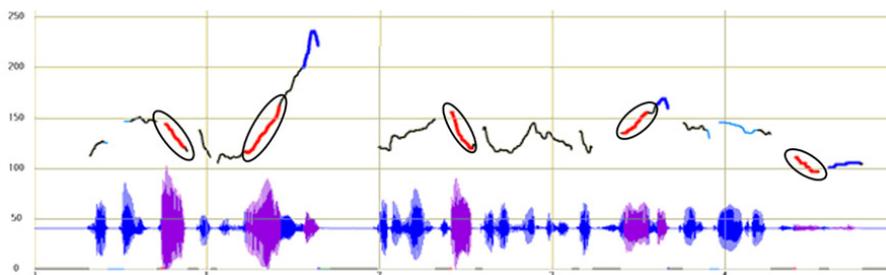


Figure 16. Italien itfn26I. [L'attentato C2 di **Molln** C1] [ha rilanciato C2 la mobilitazione C1] [contro il razzismo C0] (**Molln** est prononcé [molne] avec deux syllabes).

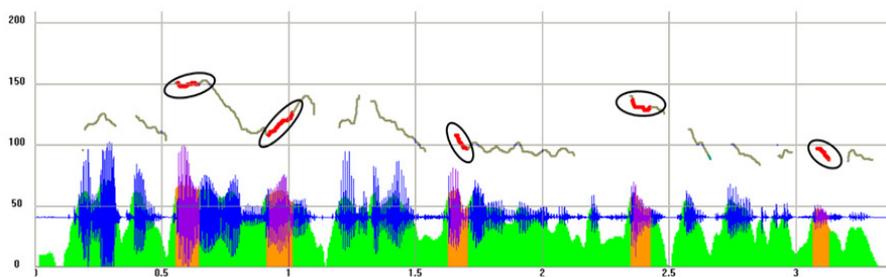


Figure 17. Espagnol esfn26E. [El atentado C1 de **Molln** Cc] [ha relanzado Cn la movilización C1] [contra el racismo C0]

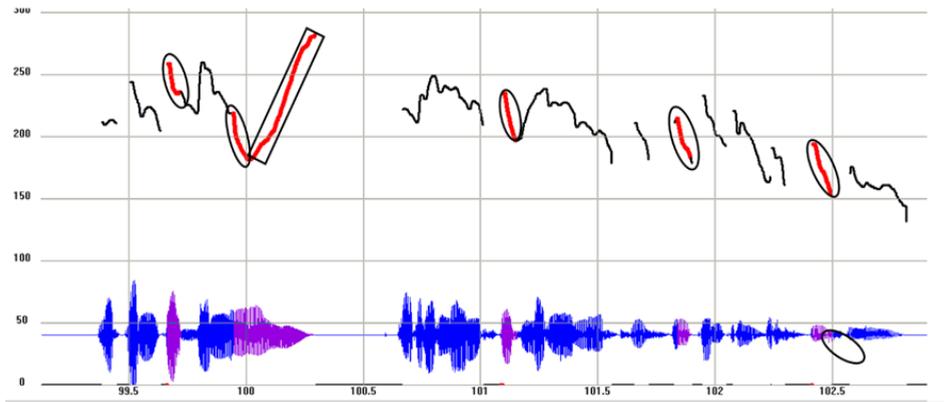


Figure 18. Catalan cafn26C. [L'atemptat C1 de **Molln** Cc] [ha rellançat C2 la movilització C2 contra el racisme C0]

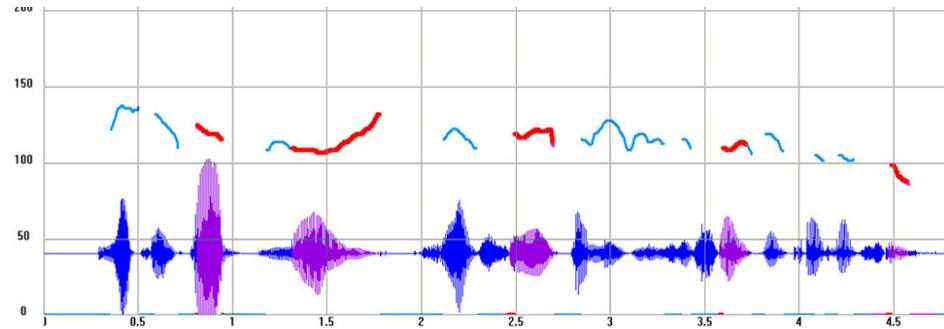


Figure 19. Portugais pfn26P. [O atentado Cn de **Molln** Cc] [relançou Cn a mobilização Cn contra o racismo C0]

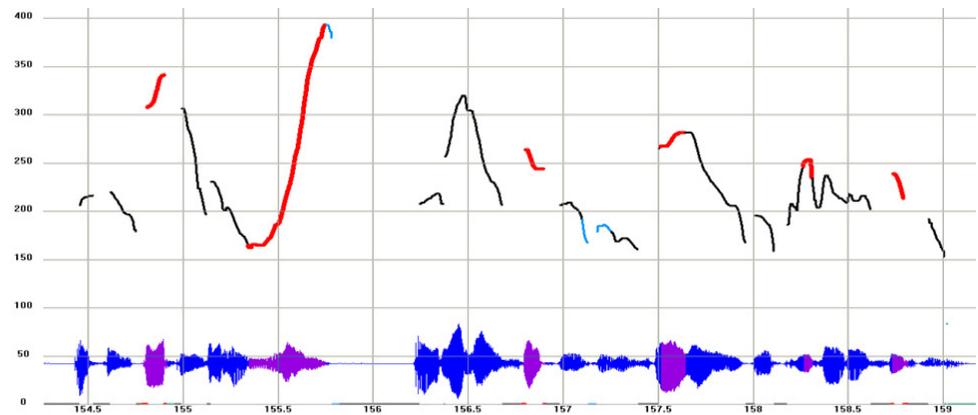


Figure 20. Roumain rofn26R. [Atentatul C1 de la **Molln** Cc] [a relansat Cn mobilizarea Cn împotriva Cn rasismului C0].

Le choix du premier contour mélodique varie dans ces différentes réalisations, mais implique dans chaque exemple un contour de rang supérieur C1 ou Cc pour le deuxième contour: C1 Cc, C2 Cc ou C2 C1.

## 6. Conclusion

On constate que les langues romanes font appel au même mécanisme de contraste de pente mélodique que le français, en contrastant C2 descendant avec C1 montant (cf. Fig. 16). Lorsque le groupe prosodique se compose de deux groupes accentuels terminés par un contour complexe Cc, le locuteur a le choix entre les séquences C1 Cc, C2 Cc ou C2 C1. Par contre, si le groupe comprend trois groupes accentuels, les configurations [[A B] C] et [A [B C]] sont différenciés par les séquences C1 C2 Cc pour [[A B] C] et C1 Cn Cc pour [A [B C]]. Dans les deux cas, le choix d'un contour montant C1 sur la première syllabe accentuée implique la réalisation d'un contour complexe Cc en fin de groupe.

## 7. Bibliographie

- EUROM4 (1991-1997). *Projet européen Lingua (CEE)*. Institut National de la Langue Française (INALF), Responsable du programme: Claire Blanche-Benveniste. Partenaires: Università Terza di Roma, Universidad de Salamanca, Universidade de Lisboa, Université de Provence.
- MARTIN, Philippe (2009). *Intonation du français*. Paris, Armand Colin.
- \_\_\_\_\_ (2013). «Contraintes phonologiques de l'intonation de la phrase réinterprétées à la lumière des recherches récentes en neurophysiologie». *La Linguistique* 49(1): 97-113.
- ROMANCE TOBI WORKSHOP (2011). *Satellite workshop PaPI 2011*. Universitat Rovira i Virgili (Tarragona), June 23, 2011.