

Structures prosodiques des langues romanes

Philippe Martin

LLF, UFRL, Université Paris Diderot Sorbonne Paris Cité

Place Paul Ricœur, 75013 Paris, France

philippe.martin@linguist.univ-paris-diderot.fr

RESUME

La description phonologique de la structure prosodique des langues romanes apparaît similaire lorsque les interactions entre les accents mélodiques est prise en compte (ce qui n'est pas le cas dans la théorie autosegmentale-métrique). L'analyse acoustique de plus de 2600 énoncés lus et spontanés suggère que la réalisation des accents mélodiques, décrits en termes de contours mélodiques plutôt que de cibles tonales, indiquent avec les contours de frontière, des relations de dépendance « vers la droite » entre groupes accentuels. Ces relations permettent par incrémentation successive dans l'axe du temps la reconstitution par l'auditeur de la structure prosodique voulue par le locuteur. Dans ce cadre théorique, les langues romanes (italien, espagnol, catalan, portugais, roumain) utilisent les mêmes contours phonologiques pour indiquer les relations de dépendance menant au codage de la structure prosodique. Le français, dépourvu d'accent lexical, utilise un système de contours différent.

ABSTRACT

Prosodic Structures of Romance Languages

Prosodic phonologic description of Romance Languages (except French) appears surprisingly similar, once the interaction between pitch accents as melodic contours rather than tonal targets is considered. Acoustic analysis performed on more than 2,600 sentences of both read and spontaneous speech suggests that the realizations of pitch accent, together with boundary tones, do indicate a relation of dependency from one accent phrase towards another accent phrase “on its right” (i.e. occurring in the future of the sentence). This process eventually leads to a reconstitution by the listener of the prosodic structure intended by the speaker through an incremental process along the time scale. French, deprived from lexical stress, uses another set of melodic contours.

MOTS-CLES : intonation, langues romanes, français, structure prosodique.

KEYWORDS: sentence intonation, Romance language, French, incremental prosodic structure.

1. Introduction

La comparaison des caractéristiques prosodiques des langues romanes a fait l'objet de plusieurs études depuis quelque dix ou quinze ans. Le projet AMPER par exemple (Contini et al., 2002) décrit les différences prosodiques de phrases comparables dans de nombreuses variétés de langues romanes.

Plus récemment, dans un ouvrage édité par S. Fróta et P. Prieto (2015), des différences prosodiques détaillées ainsi que les variations liées à la modalité sont analysées et comparées non seulement pour les langues romanes nationales (espagnol et portugais européen, italien, roumain, français), mais aussi pour des réalisations régionales comme le catalan, le friulien, l'occitan et le sarde. Le cadre théorique adopté est autosegmental-métrique, et la transcription des données utilise le système ToBI.

On relate ici un travail similaire, mais qui utilise un cadre théorique totalement différent dans le but de mieux différencier les traits phonologiques des détails phonétiques dans les réalisations d'énoncés lus et spontanés pour six langues romanes, le français, l'italien, le catalan, le roumain ainsi que l'espagnol et le portugais européens. De cette approche résulte une grande similitude phonologique dans l'indication de la structure prosodique des langues considérées, à l'exception du français. En effet, le français, dépourvu d'accent lexical, utilise un système différent pour indiquer la structure prosodique des énoncés.

2. Cadre théorique

Alors que l'approche autosegmentale-métrique rend compte de l'intonation de l'énoncé en collectant des séquences tonales bien formées telles que transcrites avec le système de notation ToBI, le cadre théorique utilisé ici est focalisé sur les caractéristiques des événements prosodiques considérés dans leur séquence temporelle. Au lieu d'être envisagée comme un objet statique, la structure prosodique est décrite du point de vue phonologique comme un processus dynamique dans lequel les contours mélodiques (i.e. les variations de hauteur) à l'endroit des syllabes accentuées et en fin de groupe prosodique assurent par des relations de dépendance entre contours la construction incrémentale de la structure prosodique voulue par le locuteur et reconstruite par l'auditeur.

Une relation de dépendance entre deux objets phonologiques s'établit lorsque l'occurrence d'un des objets ne peut survenir que si l'autre objet est ou sera présent dans l'énoncé. Ainsi par exemple, dans le cas d'un processus dynamique comme l'élaboration par le locuteur et la reconstitution par l'auditeur de la structure prosodique, un contour dit de continuation majeure présuppose la réalisation future d'un contour terminal terminant l'énoncé. C'est donc le locuteur, dans la planification de la structure prosodique, qui tient compte dans la réalisation des contours successifs marquant la structure prosodique de l'existence dans le futur immédiat de contours dont ils dépendent, c'est à dire qu'un contour ne peut apparaître que si la réalisation de celui dont il dépend est planifiée par le locuteur dans le déroulement de l'énoncé.

Il en résulte que les contours mélodiques indiquent nécessairement une relation de dépendance envers un autre contour situé « à droite », c'est-à-dire dans le futur du signal de parole, à l'exception de contour terminal conclusif, qui constitue la racine de la structure prosodique.

On peut montrer (Martin, 2009, 2015) que ces relations de dépendance définissent de manière incrémentale des regroupements successifs d'unités prosodiques minimales constituées par les groupes accentuels. Ces regroupements constituent la structure prosodique de l'énoncé. Un groupe accentuel, équivalent du *Accent Phrase* de la théorie Autosegmentale-Métrique, est constitué d'une séquence de syllabes dont une seule est accentuée, hors effet d'insistance.

Pour les langues romanes à l'exception du français, la syllabe accentuée correspond à l'accent lexical. Pour le français, elle correspond à la syllabe finale des groupes accentuels, dont la durée d'énonciation est limitée à quelque 1250 ms (Martin, 2015). Un exemple de regroupements successifs de groupes accentuels tels qu'indiqués par les contours mélodiques qui les terminent est donné Fig. 1. Il n'est peut-être pas inutile de rappeler que les contours prosodiques apparaissent séquentiellement dans le temps, et ne sont pas perçus simultanément par l'auditeur comme pourrait le suggérer la représentation planaire de la Fig. 1.

Pour que ce processus fonctionne, il faut (et il suffit) non seulement que les contours prosodiques soient différenciés, c'est-à-dire qu'il existe dans leurs réalisations une différenciation acoustique utilisant des traits de montée, de descente et/ou d'empan mélodique, mais aussi qu'ils soient ordonnés, c'est-à-dire qu'on puisse établir des relations de dépendance entre eux. L'observation des données mène à définir différentes classes de contours prosodiques nécessaires et suffisantes pour assurer l'indication des relations de dépendance entre groupes accentuels, classes

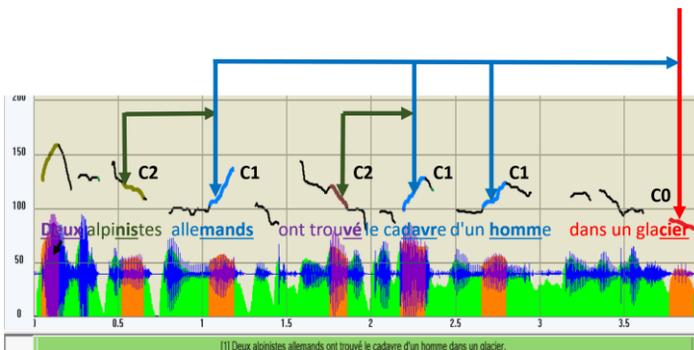


FIGURE 1 : Exemple d'un réseau de relations de dépendance (indiquées par les flèches horizontales) liées aux contours prosodiques définissant la structure prosodique de l'énoncé. On a donc C2 dépendant de C1, et C1 dépendant de C0. Le parenthésage dérivé [deux alpinistes C2 allemands C1][ont trouvé C2 le cadavre C1][d'un homme C1][dans un glacier C0] indique les regroupements successifs des groupes accentuels. On peut noter dans cet exemple lu que la structure prosodique n'est pas totalement congruente avec la structure syntaxique (deux alpinistes allemands) (ont trouvé (le cadavre d'un homme)) (dans un glacier).

instanciées par des variations (contours) mélodiques placés sur les syllabes accentuées (et essentiellement sur leur noyau vocalique) :

C0 : contour terminal conclusif (cas déclaratif), descendant et bas.

Cc : contour mélodique complexe, plat ou descendant sur la syllabe accentuée du groupe accentuel, et montant sur la syllabe finale. Si cette dernière est accentuée, les deux mouvements successifs descendant et montant sont réalisés sur cette même syllabe (n'existe pas en français).

C1 : contour montant perçu comme tel et non comme un ton statique, c'est-à-dire phonétiquement supérieur au seuil de glissando (Rossi, 1971).

C2 : contour descendant perçu comme tel et non comme un ton statique, c'est-à-dire phonétiquement supérieur au seuil de glissando.

Cn : contour neutralisé, descendant, plat ou montant perçu comme un ton statique (donc dont la variation mélodique est phonétiquement inférieure au seuil de glissando).

Le seuil de glissando est défini à partir des valeurs de fréquence fondamentale de début et de fin de voyelle des syllabes accentuées. Bien que ce seuil ait été établi à l'origine sur la perception de voyelles synthétiques, et qu'il implique, mais dans une moindre mesure, d'autres paramètres que la vitesse de variation mélodique, il constitue un indice parmi d'autres pour différencier à partir de leurs réalisations phonétiques les contours C1 et C2 du contour neutralisé Cn.

En établissant par l'analyse des données expérimentales les relations de dépendance transitives entre ces contours prosodiques, on obtient, pour le français, dépourvu d'accent lexical et donc de contour complexe Cc :

Cn -> C2 -> C1 -> C0 (dépendance de Cn par rapport à C2, de C2 par rapport à C1, de C1 par rapport à C0 le contour terminal de fin d'énoncé).

et pour les autres langues romanes :

Cn -> C1 -> C2 -> Cc -> C0 (dépendance de Cn par rapport à C1, de C2 par rapport à C2, de C2 par rapport à Cc, de Cc par rapport à C0 le contour terminal de fin d'énoncé).

La hiérarchie des classes de contours est alors, pour les langues romanes autres que le français Cn < C1 < C2 < Cc < C0, et pour le français, dépourvu d'accent lexical et donc de contour complexe Cc,

$C_n < C_2 < C_1 < C_0$. La hiérarchie des contours montant C_1 et descendant C_2 est donc inversée pour le français par rapport aux autres langues romanes.

En comparant deux contours mélodiques successifs (donc placés sur deux syllabes accentuées successives ne constituant pas d'accents d'insistance), et tenant compte de leur ordonnancement, les regroupements de groupes accentuels s'opèrent à parti des trois règles suivantes :

Soient C_x et C_y des contours appartenant à l'ensemble C_0, C_c, C_1, C_2, C_n ,

Si $C_x < C_y$ structure partielle [$C_x C_y$] les groupes prosodiques dont le dernier groupe accentuel contient les contours C_x et C_y forment un groupe plus grand. Par exemple la séquence $C_2 C_c$ indique une dépendance de C_2 relativement à C_c et déterminent le regroupement des unités prosodiques (groupes accentuels ou regroupement de groupes accentuels) porteurs de C_2 et C_c .

Si $C_x = C_y$ structure partielle [$C_x C_y \dots$] les groupes prosodiques dont le dernier groupe accentuel contient les contours C_x et C_y font partie d'une liste, terminée plus tard par l'apparition d'un contour de rang plus élevé. Ainsi, la séquence $C_1 C_1$ forme une liste partielle terminée par un contour C_2 ou C_c (pour les langues romanes) et par C_0 pour le français.

Si $C_x > C_y$ structure partielle [$C_x [C_y \dots$] les groupes prosodiques dont le dernier groupe accentuel contient les contours C_x et C_y ne forment pas un groupe plus grand. Une séquence de contours telle que $C_c C_2$ ne détermine donc pas un groupe prosodique plus grand dont les syllabes accentuées de l'unité accentuelle finale seraient C_c et C_2 .

Ce processus est donc local, et implique également la transitivité. Ainsi le contour neutralisé C_n peut indiquer une relation de dépendance aussi bien avec C_1, C_2, C_c ou C_0 , puisque ces contours sont situés à un rang supérieur dans la hiérarchie des contours. Ce n'est que quand la complexité relative de la structure prosodique le requiert que des contours de rang immédiatement inférieur sont obligatoirement réalisés par le locuteur, sinon ils sont optionnels.

Comme signalé plus haut, les relations de dépendance opèrent « à droite » relativement au futur de la séquence de contours mélodiques. Ainsi un contour C_1 de continuation majeure, en indiquant non seulement que l'énoncé n'est pas terminé, présuppose l'apparition d'un contour terminal C_0 signalant la fin de l'énoncé et permettant à l'auditeur de rassembler les différents groupements de syntagmes pour accéder au sens de l'énoncé. L'existence de C_1 dépend donc de l'apparition future d'un contour C_0 . Il en va de même pour le contour descendant C_2 , qui dépend de l'apparition future d'un contour C_1 , etc.

3. Analyse expérimentale

Le corpus d'analyse comporte :

1. Un ensemble de 60 phrases de complexité syntaxique croissante lues par deux locuteurs, extrait du corpus EuRom4 (1991-1997). Les phrases à l'origine en français, italien, espagnol et portugais européen et ont été adaptées en catalan et en roumain.
2. Le corpus EuRom4 (français, italien, espagnol, portugais européen), contenant 24 histoires courtes de longueur variable lues par 4 locuteurs dans chaque langue. Le nombre moyen de phrases pour chaque langue est d'environ 200.
3. Le corpus EuRom5 (Bonvino 2011) français, italien, espagnol, catalan, portugais européen, composé de 20 histoires courtes lues par deux locuteurs professionnels, avec environ 300 phrases pour chaque langue.

L'analyse acoustique a été réalisée avec le logiciel WinPitch (1996-2016). Les transcriptions de texte étant préalablement disponibles pour les trois corpus, l'alignement texte-parole a été réalisée la fonction d'alignement de volée, permettant un alignement rapide et efficace des transcriptions de texte. Cet alignement permet l'analyse des informations prosodiques retrouvées en un clic à partir de

n'importe quelle partie du texte. Une toute petite partie des résultats est présentée ci-dessous, afin d'illustrer quelques aspects de la procédure de découverte des relations de dépendance entre contours. Des données plus complètes sont disponibles dans Martin (2015).

Les figures qui suivent présentent des courbes mélodiques dont les segments continus plus épais correspondent aux voyelles accentuées et donc aux contours C0, C1 et C2. Le contour Cn est représenté en trait plus fin, et les traits interrompus correspondent à la variation mélodique montante de frontière relative aux contours Cc.

Comment alors déterminer la structure prosodique d'un exemple donné, celle-ci étant par hypothèse indépendante des autres structures de l'énoncé, et en particulier de la structure syntaxique ? Ayant posé que les groupes accentuels en constituent les unités minimales, on procède par expansions successives du nombre de ces groupes accentuels dans l'analyse des données. Ainsi avec un seul groupe accentuel on ne peut former qu'une seule structure prosodique, terminée par un contour terminal conclusif C0 indiquant également la modalité déclarative ou interrogative (et ses variantes) de l'énoncé (on se limite ici à la modalité déclarative).

Deux groupes accentuels peuvent constituer deux structures prosodiques distinctes, indiquées par la séquence de contours C0 Cx et Cy C0, Cx et Cy devant être déterminés. On a donc soit un contour terminal C0 identifié sur le premier groupe accentuel, soit sur le second, soit encore sur les deux (cas d'une épexégèse). L'identification d'un contour terminal se fait par test de perception de la finalité en segmentant éventuellement l'énoncé avec un éditeur de signal. On trouve alors que dans le premier cas C0 Cx, le contour Cx est plat et sera noté Cn (configuration classique dite *propos-thème* ou *noyau-postnoyau*).

Dans la seconde configuration Cy C0, le contour Cy peut être réalisé de plusieurs manières que l'on peut recenser dans les données : Cc (contour complexe), C1 (continuation majeure) ou Cn (contour neutralisé). Puisqu'un seul trait acoustique suffit à différencier le contour Cy de C0, le locuteur peut choisir parmi les contours disponibles Cn, ou utiliser plus de traits acoustiques de différenciation avec C1 ou Cc pour des raisons stylistiques par exemple.

Trois structures prosodiques distinctes peuvent hiérarchiser trois groupes accentuels A, B et C : [A] [B] [C], [A B] [C] et [A] [B C]. Pour s'assurer de la réalisation effective d'une de ces structures dans un exemple donné, et ainsi déterminer les contours qui l'indiquent, on peut utiliser des énoncés dépourvus de structure syntaxique et informationnelle, telle que les énumérations, les numéros de téléphone, etc.

C'est le cas de la Fig. 2, montrant la dépendance en italien de C2 par rapport à Cc, et de C1 par rapport à C0. Tous les groupes sont structurés avec le contours C2 Cc, sauf le dernier par C1 C0, illustrant les relations de dépendance C2 -> Cc et C1 -> C0. (de Italo Svevo, *La Coscienza Di Zeno*, lu par Moro Silo, Il narratore, 2006).

Italien [B C2 *antiparasitari* Cc] [C C2 *anticoncezionali* Cc]...[M C1 *antilopi* C0].

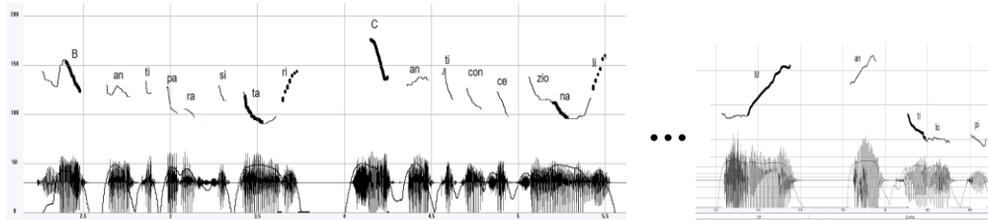


FIGURE 2 : **B** *antiparasitari* **C** *anticoncezionali*,..., **M** *antilopi* "B, antiparasitaire, C, anticonceptionnel,..., M, antilope.", exemple extrait d'une longue énumération faite à partir d'une séquence de groupes de deux mots prosodiques, terminé par Cc, sauf le dernier se terminant par un contour C0 conclusif placé sur la syllabe accentuée *ti* de *antilopi*.

La configuration ci-dessus se retrouve non seulement en italien, mais aussi dans les autres langues romanes comme le montrent les figures 3 à 5.

Espagnol (Eurom5 E12-1) C2 Cc : [cuando C2] [se constate Cc]

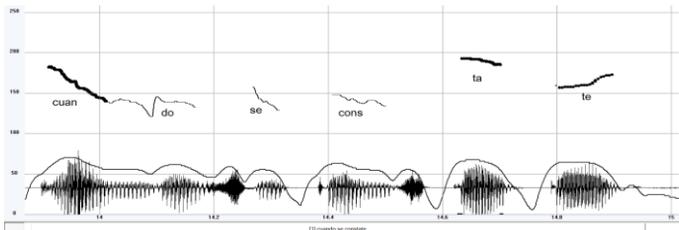


FIGURE 3. *una pide asus 47 Estados miembros que establezcan algún tipo de sanción cuando se constate que hay una "provocación a la discriminación" en los mensajes publicitarios* "Une demande à ses 47 États membres de mettre en place tout type de sanction lorsqu'il est constaté qu'il existe une "incitation à la discrimination" dans la publicité".

Portuguais (Eurom5 P04-1) C2 Cc : [apelidado C2] [de Óscar Cc] "appelé Oscar"

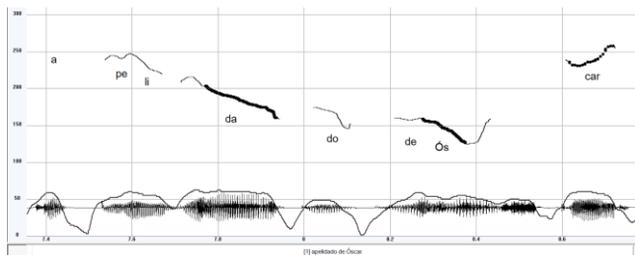


FIGURE 4. *um cão de raça terra nova apelidado de Óscar, cujo dono é um socialite, vai ser submetido a um lifting aos olhos* "un chien de race terre-neuve surnommé Oscar, dont le propriétaire est un mondain, va subir un lifting des paupières"

Catalan (Eurom5 C19-1) C2 Cc : [Com C2] [les formigues Cc]

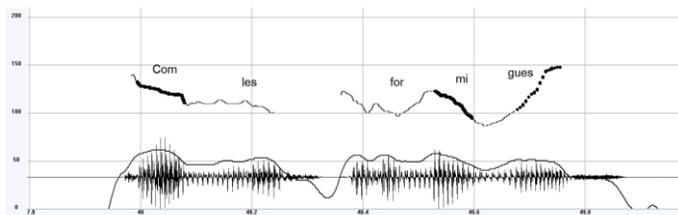


FIGURE 5. *Com les formigues, els cucs o els escarabats, les meduses són en algunes cultures un element bàsic de l'alimentació.* "Comme les fourmis, les coléoptères et les vers, les méduses sont dans certaines cultures un aliment de base".

Tous ces exemples illustrent la relation de dépendance C2 → Cc et Cc → C0, établie en partant de l'hypothèse d'une congruence au moins partielle portant sur les segments analysés des structures prosodiques et syntaxiques de surface. Cette congruence est généralement plus fréquente dans les phrases lues, le locuteur s'efforçant de faire correspondre les frontières prosodiques avec les frontières syntaxiques importantes dans la lecture. On note aussi que si ce sont les mêmes contours phonologiques, et donc les mêmes contrastes, qui sont utilisés dans ces langues romanes, leur réalisations phonétiques peut être différente ainsi que le montrent les tracés d'analyse acoustiques.

Portugais (Eurom5 P05-1)

[é permitido C2] [matar C2] [um escocês Cc] "Il est permis du tuer un Écossais".

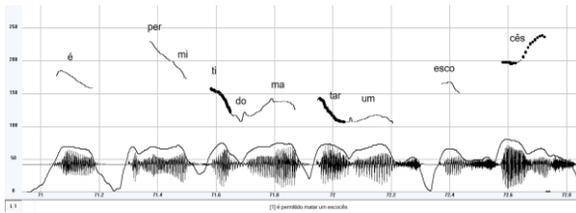


FIGURE 8. Courbe mélodique et contours de l'exemple *é permitido matar um escocês*.

Français

Le français ne possède pas de contour complexe Cc, et les relations de dépendance sont illustrés par les exemples des figures 6 et 7. L'exemple suivant illustre le contraste prototypique d'inversion de pente mélodique en français. [*les garçons C2 de piste C1*] (Eurom5 F03-1)

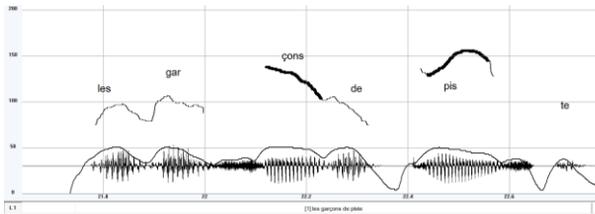


FIGURE 6. *Les garçons de piste se transforment en porteur, voltigeur, funambule ou cascadeur...* Cet exemple illustre la relation de dépendance C2 -> C1, prototypique du français avec C1 -> C0. Les mêmes relations se trouvent dans des exemples plus complexes, comme illustré Fig. 7. [*ainsi C1*] [*sa nouvelle gamme C2 de combinés C2 présentés Cn lundi C1*] (Eurom5 F04-1)

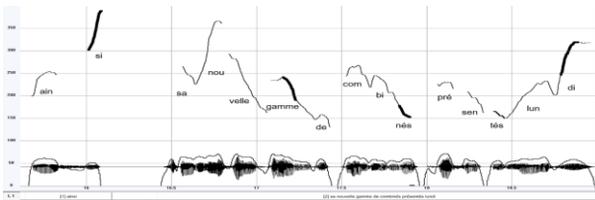


FIGURE 7. *ainsi sa nouvelle gamme de combinés présentés lundi*.

Les deux derniers exemples (Fig. 9 et 10) illustrent le fonctionnement des contours de relations de dépendance dans l'indication de deux regroupements différents II et III dans des lenagues romanex autres que le français.

Roumain (Eurom4 rofn40)

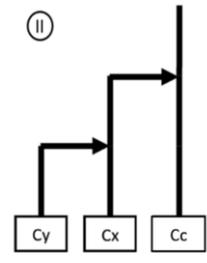
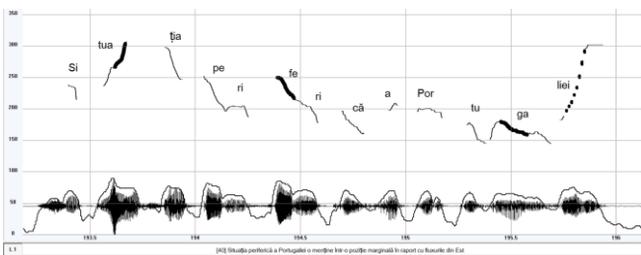


FIGURE 9. Un exemple du contraste C1 -> C2 -> Cc : *Situația periferică a Portugaliei o menține într-o poziție marginală în raport cu fluxurile din Est*, réalisant la structure II.

Italien (Eurom4 I_23_06) [[[che C2] [[trasferirsi C1] [in USA Cc]]] “lequel, transféré aux USA...”

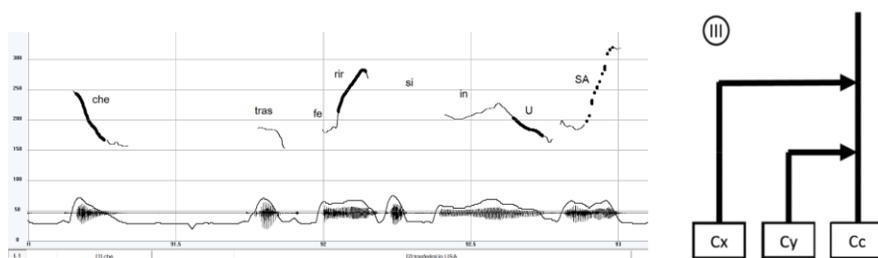


FIGURE 10. Un exemple du contraste C2 -> Cc et C1 -> Cc : che trasferirsi in USA

Ce dernier exemple est prototypique d'un contour mélodique C2 sur le mot monosyllabique *che* juste avant le début (la limite gauche) de la parenthèse *trasferirsi in USA*. Elle montre clairement que C2 fait contraste avec Cc et non avec le contour suivant C1, formant un groupe avec toute la parenthèse.

4. Conclusions

Le modèle de structure prosodique incrémentale est basé sur un principe clairement défini, l'indication d'une relation de dépendance "à droite" entre les événements prosodiques instanciés par des contours mélodiques des syllabes accentuées des groupes accentuels et en position finale de syntagmes intonatifs (cas du contour complexe Cc). À partir d'un inventaire des classes de contours mélodiques en termes de relations de dépendance, une description unifiée de la structure prosodique émerge pour toutes les langues romanes à accent lexical. Grâce à cette approche, la seule différence avec le français dépourvu d'accent lexical se rapporte à l'absence de contour complexe Cc, ce qui modifie le classement des contours mélodiques en préservant le principe de l'inversion de la pente mélodique.

En effet, dans les langues romanes le contraste de pente mélodique, à savoir montant -> descendant et descendant -> montant est appliqué dans la relation de dépendance C1 -> C2, C2 -> CC et CC -> C0, le contour complexe apparaissant phonologiquement comme une montée mélodique. En français, les relations existantes sont C2 -> C1 et C1 -> C0. Les langues romanes ont donc la possibilité de différencier un niveau supplémentaire dans la structure prosodique.

Cette conclusion va clairement à l'encontre du cadre théorique dominant basé sur la phonologie métrique-autosegmentale. En niant ex abrupto le rôle des accents mélodiques (pitch accents), cette approche ne peut évidemment rendre compte d'une quelconque interaction entre les contours mélodiques. De même, la définition des groupes accentuels n'a pas de rapport direct avec les prééminences métriques.

Une dernière remarque. Le cadre autosegmental-métrique ne propose en fait aucune explication quant à la fonction de la structure prosodique autre qu'actualiser la syntaxe, en se basant sur le concept du « bien-formé » emprunté à l'approche générative transformationnelle. Or on peut montrer (Martin, 2015) que la structure prosodique n'est pas dérivée de la syntaxe (si ce n'est dans une certaine mesure dans l'oralisation de l'écrit), elle la précède aussi bien pour le locuteur que pour l'auditeur. Du reste, il ne saurait y avoir de production langagière sans structure prosodique, même en lecture silencieuse, alors que l'inverse est possible.

Références

BONVINO E. et al. (2011). *EuRom5*, Ulrico Hoepli, Milano.

CONTINI M. et al. (2002). Un projet d'atlas multimédia prosodique de l'espace roman. *Proceedings of the 1st International Conference of Speech Prosody*, Aubenas d'Ardèche, Lienhart, 227-231.

EUROM4 (1991-1997). Projet européen Lingua (CEE) Institut National de la Langue Française (INALF), Responsable : Claire Blanche-Benveniste.

FROTA S., Prieto P. eds. (2015). *Intonation in Romance*. OUP Oxford, 400p.

MARTIN Ph. (2009) *Intonation du français*. Armand Colin, Paris, 256p.

MARTIN Ph. (2015) *The Structure of Spoken Language. Intonation in Romance*. Cambridge University Press, 340p.

ROSSI M. (1971). Le seuil de glissando ou seuil de perception des variations tonales pour la parole, *Phonetica* (23) 1-33.

WINPITCH. (2016). www.winpitch.com