



HAL
open science

La prosodie au cœur du verbal

Anne Lacheret

► **To cite this version:**

Anne Lacheret. La prosodie au cœur du verbal. *Rééducation orthophonique*, 2011, 246, pp.87-104.
halshs-00636556

HAL Id: halshs-00636556

<https://shs.hal.science/halshs-00636556>

Submitted on 18 Nov 2011

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La prosodie au cœur du verbal

Anne Lacheret

Laboratoire MODYCO, Paris Ouest Université de Nanterre la Défense

anne@lacheret.com

Prosody: At the heart of language

Abstract

From the very beginning of life, before the acquisition of articulated language, prosody is essential in the development of human intersubjectivity. It is the principal initiator of language, making the prosody of one's mother tongue difficult to escape in the acquisition of second languages. Finally, prosody is often the only remaining linguistic component in language disorders (speech loss). This irreducibility of prosody is basically due to its semiotic dimension and its linguistic functions: far from being merely an optional musical component of speech, linked only to the paraverbal domain, prosody is at the heart of verbal communication and absolutely necessary for the construction of meaning in discourse.

Résumé

Avant l'accès au langage articulé, dès les premières semaines de la vie, la prosodie est essentielle dans le développement des relations intersubjectives. Initiatrice principale du langage, c'est aussi elle qui perdure lors de l'apprentissage de langues secondes, celle, enfin, qui nous ramène vers ce langage primitif, d'où, dit-on, la double articulation est absente. Cette irréductibilité de la prosodie est fondamentalement associée à sa dimension sémiotique : contrairement au seul rôle qu'on a voulu lui attribuer pendant longtemps, la prosodie n'est pas une simple musique vocale qui accompagne librement les modulations de la pensée et qui, de ce fait, doit rester cantonnée dans la sphère du paralinguistique, elle s'inscrit d'emblée au cœur de la communication langagière.

1. La prosodie du message parlé

La prosodie constitue un objet d'étude complexe par rapport à la composante segmentale de la parole, ceci pour des raisons liées d'une part à sa spécificité même, d'autre part au champ disciplinaire dont elle relève, *i.e.* l'histoire des idées en linguistique et en phonétique. Concernant le premier point, soulignons les deux caractéristiques essentielles de la prosodie : sa dimension continue qui rend difficile le repérage et l'extraction d'unités discrètes, ses fonctions linguistiques plurielles : elle interagit avec les différents niveaux de représentation du langage : phonologique syntaxique et sémantico-pragmatique¹. De cette singularité découle le second point : c'est sans doute cette complexité formelle et fonctionnelle, d'où la double articulation est absente, qui a conduit pendant longtemps à écarter la prosodie du champ de la linguistique structurale (pour une revue, voir Lacheret 2010). Il faudra donc attendre les travaux en traitement automatique du langage, en linguistique appliquée (enseignement d'une langue étrangère) et plus récemment dans le champ des sciences cognitives pour que la prosodie trouve ses lettres de noblesse. Néanmoins, pour une première approche du domaine, il reste souvent naturel d'assimiler la prosodie au domaine du para-verbal et de l'extralinguistique (Boutard & Guillon, 2010), le verbal ne pouvant désigner que le matériel segmental, (phonèmes, morphèmes et syntagmes). C'est contre cette vision réductrice qu'il est nécessaire d'argumenter scientifiquement en expliquant pourquoi et comment la prosodie appartient à part entière à la sphère linguistique :

¹ Par *sémantico-pragmatique*, nous désignons la construction du sens en contexte.

composante incontournable de la construction du sens en contexte, elle est indispensable pour une communication réussie. Pour les orthophonistes et d'une manière plus générale, les différents acteurs impliqués dans la rééducation langagière, il semble important de se saisir des différentes composantes linguistiques de la prosodie, dans sa matérialisation sonore, sa dimension formelle et son rôle fonctionnel et son implication majeure dans la communication verbale. Ce point constitue un préalable nécessaire pour aborder sur des bases solides, avec les bons outils méthodologiques et conceptuels les dysprosodies auxquelles ils peuvent être confrontés dans leurs pratiques professionnelles, quelle que soit l'origine du trouble et son expression sonore.

2. Un objet d'étude complexe

Sous l'angle cognitif, la prosodie, qui correspond aux faits intonatifs et accentuels, assume deux fonctions : une fonction de **groupement** (intonation) et une fonction de **mise en relief** (accentuation)². D'une part, les constructions prosodiques fournissent des indices à l'interlocuteur pour segmenter le continuum sonore et identifier des unités linguistiques de rang et de portée variable ; il s'agit d'autre part de rendre perceptivement saillants les éléments importants du message. A ces fonctions font écho deux principes élémentaires que nous proposons de nommer respectivement **principe de contraste** et **principe gabaritique**. Selon le premier, la sémantique du message n'émerge que par la mise en contraste perceptive des formes (faible vs. forte)³. Selon le second principe, les contrastes doivent être distribués sur la chaîne parlée relativement à l'ensemble des contraintes fonctionnelles sous-jacentes (lexicales, syntaxiques, sémantico-pragmatiques) et à leur combinaison éventuelle ; le gabarit prosodique mobilisé (enveloppe temporelle et mélodique) étant censé rendre compte de la nature et de la force des contraintes en jeu, et de la façon dont elles se distribuent sur la chaîne parlée.

Du point de vue de l'analyse, l'objectif est d'expliquer comment s'organisent et interagissent ces contraintes et comment elles se réalisent acoustiquement⁴. Trois angles d'attaque sont convoqués de façon complémentaire : (i) le traitement phonétique concerne l'analyse acoustique et perceptive des données, (ii) l'analyse phonologique propose des représentations abstraites qui permettent d'unifier la description et de rendre compte des variations accentuelles et intonatives observées en parole, (iii) l'interprétation fonctionnelle met au jour les contraintes linguistiques qui sous-tendent la structuration prosodique du message parlé. Si les processus de groupement et de mise en relief prosodique sont sans doute des universaux du langage, leurs marquages varient selon les types de langue, les méthodes et les outils de description et d'analyse doivent bien sûr en tenir compte ; dans le cas présent nous nous focaliserons sur la prosodie du français parlé (pour une revue, voir Lacheret & Beaugendre 1999).

3. Perspective sonore

Trois paramètres sont mobilisés de façon première dans l'organisation prosodique du signal de parole (figure 1) : la **fréquence fondamentale** (estimation du son laryngien à un instant donné sur le signal acoustique), la **durée** (mesure d'un intervalle de temps nécessaire pour émettre un segment sonore), enfin, **l'intensité** (relative à l'énergie contenue dans le signal). A ces éléments, s'ajoute la qualité vocale, *i.e.* les caractéristiques spectrales de la parole qui peuvent, entre autres choses, constituer des indices de frontières de segments

² Elle relève donc des composantes syntagmatique et paradigmaticque du message parlé.

³ Ce principe n'étant sans doute pas spécifique à la prosodie ni même au langage mais correspondant à une fonction cognitive élémentaire d'organisation de formes (Laks 1997).

⁴ La détection d'une variation mélodique ample accompagnée d'un allongement temporel est ainsi considérée comme un indice de frontière de groupe en français ; la forme du contour (convexe ou concave), sa direction (montante ou descendante) donne des indications sur sa modalité (figure 4).

(Gendrot 2009). Parmi ces paramètres, correspondant respectivement, sous l'angle perceptif, à la **mélodie**, la **longueur**, la **sonie** et le **timbre**, c'est sans nul doute la mélodie (variations de la hauteur de la voix) et la durée (variations temporelles) qui ont été les plus étudiées et qui, en conséquence, sont les mieux référencées dans la littérature. L'intensité quant à elle demande des conditions d'enregistrement très contrôlées pour produire des mesures significatives, quant à la qualité vocale, ce n'est que récemment qu'elle a commencé à faire l'objet d'investigations fines en parole (D'Alessandro 2006). Nous nous centrerons donc ici sur les variations mélodiques et temporelles, et partirons des questions suivantes : (i) dans quelle mesure ces variations constituent des indices acoustiques fiables pour segmenter le continuum sonore et identifier des unités de traitement linguistique pertinentes ? (ii) à quelles fonctions linguistiques sont associées ces variations (morphologique, lexicale, syntaxique, etc.) ? Les réponses seront évidemment variables en fonction du type de langue (Hyman 2006), celles pour lesquelles la prosodie joue un rôle lexical (langues à tons, langues à accent morphologique)⁵ et celles dans lesquelles elle intervient essentiellement au niveau post-lexical⁶ dans l'organisation globale du discours. Indépendamment des différences typologiques, on peut se fixer comme première tâche de localiser les proéminences syllabiques distribuées sur la chaîne parlée. Cette étape constitue en effet la pierre angulaire du traitement prosodique : c'est en fonction des proéminences accentuelles perçues que l'auditeur va générer et interpréter la structure prosodique. Pour revenir au français, selon la littérature, les accents dans cette langue ont un rôle démarcatif puisqu'ils permettent d'identifier des fins de groupes syntactico-sémantiques. Le français serait donc plutôt une langue à frontière qu'à accent. En pratique, l'observation des données montre qu'en français contemporain, la distribution de l'accent est très variable, le marquage d'un groupe pouvant également être réalisé à sa frontière gauche⁷, en outre, comme on le verra plus loin, on observe de nombreuses proéminences non démarcatives (ni finales, ni initiales de groupes), motivées par des contraintes rythmiques et/ou pragmatiques.

Si la distribution des accents est très variable et reste finalement difficilement prédictible en français, on constate en revanche que les proéminences syllabiques perçues sont généralement marquées par des variations significatives de la mélodie et/ou de la durée, donc facilement repérables dans la chaîne parlée et modélisables (figure 3). Il est classique de représenter ces variations par des contours intonatifs, *i.e.* des patrons mélodiques prototypiques qui se réalisent dans un empan temporel donné et que l'on peut caractériser en termes de traits (Martin 1981) ou, par une suite de tons (Mertens 2004) qui spécifient la durée du contour (long ou court), sa direction (plat, montant, descendant ou en cloche)⁸, son amplitude (ample ou restreint), les niveaux de hauteur traversés, etc.

⁵ Dans les langues à tons, les variations de hauteur mélodique (ou variations tonales) sont encodées et réalisées au niveau lexical (niveau du mot) pour opposer des paires minimales, *i.e.* des mots sémantiquement différents (ex. [bønér] en norvégien, prononcé respectivement avec les tons HLH vs LLH signifie dans le premier cas *paysan* dans le second *haricots*). Dans les langues à accent morphologique, si la valeur sémantique reste la même, en revanche la catégorie syntaxique du mot contraint sa distribution accentuelle ; ainsi en anglais, Record (verbe : *enregistrer* ou nom : *disque*), ne portera pas le même schéma accentuel : position finale vs. initiale de l'accent. A noter que (i) la complexité des systèmes tonaux et le nombre de tons distinctifs varient selon le type de langue, (ii) certaines langues peuvent combiner un système tonal et un système accentuel, c'est le cas du norvégien, pour une raison d'économie du système, les oppositions tonales restent alors relativement rudimentaires.

⁶ Opposition technique pour distinguer les processus qui opèrent au niveau du mot (lexical) et ceux qui affectent le niveau du discours (postlexical). L'unité d'analyse n'est donc pas la même (*mot vs groupe de mots*).

⁷ En conséquence, certains parlent d'arc accentuel, d'accent bipolaire, voire d'arc accentuel complexe (Fonagy 1979, Di Cristo 1999, Astesano 1999).

⁸ Les contours montants et descendants sont les mieux décrits dans la littérature prosodique, ils sont appelés respectivement continuatifs et conclusifs (Rossi & al. 1981). En fonction de l'amplitude du geste, ces contours sont dits mineurs (amplitude faible) ou majeurs (amplitude forte).

Pour le chercheur, il s'agit donc de définir des modes de représentation de la prosodie produite et perçue, de repérer et d'annoter ses corrélats acoustiques sur le signal de parole en montrant par exemple dans quelle mesure ces corrélats varient en fonction de la prééminence syllabique (terminale ou non de groupe prosodique)⁹. Le problème auquel on se heurte ici est, rappelons-le, la dimension continue de la prosodie, continuité qui s'exprime sous l'angle syntagmatique (axe linéaire de la chaîne parlée), mais également par la nature scalaire des prééminences. Du point de vue syntagmatique d'abord, il faut choisir un domaine segmental de description, ou point d'ancrage temporel, pour l'analyse des contours intonatifs. S'il est classique de manipuler des contours syllabiques, on peut également travailler avec des contours de mot, de syntagme, voire d'énoncé¹⁰. Tout dépend de ce dont on souhaite rendre compte fonctionnellement mais également du type de parole produite. Ainsi, dans la communication préverbale précoce, jusqu'à au moins 8 mois, il semble illusoire de penser que la syllabe constitue un point d'ancrage robuste et les chercheurs travaillent sur des segments plus larges (Gratier 2010, Martel & Leroy-Collombel 2010). Pour résumer ce premier point, en fonction de l'ancrage segmental de la mélodie, le degré de discrétisation du continuum sonore n'est pas le même. Concernant le second point, une syllabe est perçue prééminente, non pas dans l'absolu mais relativement à son contexte et de façon parfois très variable d'un auditeur à un autre. Par ailleurs, la perception n'est pas catégorielle (contrairement aux phénomènes segmentaux) mais graduelle et continue, *i.e.* une syllabe n'est pas perçue accentuée ou non mais plus ou moins accentuée, c'est pour cela que l'on parle de degré accentuel (Dell 1984, Avanzi & al. 2010).

Dans ce contexte, les analyses empiriques (oreille, papier, crayon) ne sont plus tenables. Il est plus que jamais nécessaire de se doter d'outils de mesure adéquats pour déterminer objectivement les variations temporelles et mélodiques significatives : seuils d'allongement et/ou de hauteur, registre¹¹, amplitude d'un mouvement mélodique nécessaires pour considérer qu'un segment est prosodiquement saillant et peut être porteur d'un effet de prééminence, débit de parole¹², etc. De tels outils présupposent des méthodes d'annotation prosodique de corpus dont la nature et le degré de granularité varient en fonction de ce à quoi va servir le matériel ainsi annoté. Quel que soit l'objectif, le point de départ est toujours le même : il s'agit, à partir d'un enregistrement sonore et d'une transcription associée, de produire un alignement texte-son et sur ces bases de réaliser une annotation prosodique des données. Parmi les divers outils disponibles, le logiciel PRAAT, gratuitement téléchargeable et très bien documenté (Boersma P. & D. Weenink 2011) constitue une bonne ressource pour amorcer un tel type de travail. Par ailleurs, des protocoles d'annotation (annotation des prééminences accentuelles, repérage des frontières prosodiques de différents rangs, schèmes tonals associés aux contours observés, étiquetage des disfluences¹³, etc) voient progressivement le jour et peuvent constituer une base solide pour dériver sa propre méthode d'annotation en fonction de son objet d'étude.

⁹ Dans la figure 2, on observe que seules la syllabe terminale est porteuse d'un contour dynamique (HF).

¹⁰ De ce point de vue, la notion de contour est ambiguë (Avanzi 2011) puisqu'elle renvoie à deux types d'objets : elle désigne soit les segments de courbe mélodique portés par les syllabes (figure 2), soit des patrons mélodiques globaux (de mots, groupe ou énoncé (figure 4)).

¹¹ Défini entre une ligne mélodique de base et une ligne haute. On peut utiliser 3 (bas, moyen, haut) ou 4: (infragave, grave, aigu, suraigu) niveaux de hauteur.

¹² Différentes mesures peuvent donner des indications précieuses quant au débit de parole : nombre total de phonèmes articulés dans une séquence, temps de locution (ou temps passé à prononcer un énoncé), temps de pause, nombre de pauses (pour des indications de mesure, voir Lacheret & Beaugendre 1999, 255-256).

¹³ On définit une disfluence comme un élément qui brise le déroulement syntagmatique dans la chaîne parlée, une sorte de trébuchement vocal. Il peut prendre différentes natures et correspond souvent à un allongement syllabique excessif associé au travail de formulation en cours.

4. Une structure hiérarchique

Les phonologues posent comme hypothèse centrale que les constructions prosodiques rencontrées dans la parole, *i.e.* la structure de surface, sont associées à des représentations formelles abstraites (structure profonde) qui encodent deux types d'informations, la première de nature tonale est associée aux variations d'intonation, la seconde, de nature métrique, aux variations accentuelles. Autrement dit, il existerait des patrons tonaux et métriques sous-jacents encodés dans la grammaire mentale des locuteurs (Jun & Fougeron 1995). Ces patrons appartiennent à des domaines, *i.e.* des constituants prosodiques qui s'organisent hiérarchiquement en vertu de contraintes morphosyntaxiques de base (Selkirk 1984, Nespor et Vogel 1986, Delais 2005). On distingue minimalement deux domaines à l'intérieur d'un énoncé respectivement dénommés : constituants *majeurs*, ou *syntagme intonatifs* (au plus haut dans la hiérarchie) et constituants *mineurs*, ou *syntagmes accentuels* (au plus bas dans la hiérarchie). Selon une seconde hypothèse, les variations prosodiques observées dans la structure de surface rendent compte de cette représentation abstraite de la prosodie. Ainsi, il a été montré très vite que plus le niveau du constituant prosodique actualisé dans la structure de surface est élevé dans la hiérarchie, plus les frontières intonatives et les proéminences syllabiques qui les réalisent sont importantes (Rossi & al 1981), et l'occurrence des pauses fréquentes (Grosjean & Dommergues 1983). Selon des travaux récents, il existerait une corrélation entre les variations temporelles des segments aux frontières des constituants et leur place dans la hiérarchie prosodique (allongement moindre pour un constituant mineur notamment, Fougeron 2001). En conséquence, la perception des variations temporelles et mélodiques dans le signal de parole permet à l'auditeur de reconstruire la structure prosodique sous-jacente et le cas échéant de repérer des dysfonctionnements dans les cas où par exemple les degrés de frontières ne s'alignent pas sur les domaines des constituants (une frontière forte qui ponctue un constituant syntaxique mineur par exemple, cf. *supra* principe gabaritique, *infra* figure 5). Reste à savoir quand-même si ces dysprosodies révèlent un dysfonctionnement dans les représentations sous-jacentes stockées par le locuteur où sont exclusivement corrélées à des problèmes moteurs.

5. Des fonctions plurielles

Partant du principe que les structures du langage sont déterminées fondamentalement par les fonctions qu'il sert (Robert 2002), nous nous pencherons sur les deux fonctions centrales de la prosodie en français : (i) une fonction de structuration du message (segmentation, intégration, hiérarchisation, groupement, équilibre), (ii) une fonction pragmatique, qui s'organise autour de contraintes communicatives plurielles : traitement de l'information, modalités énonciatives (marquage de la force illocutoire d'un énoncé), expression des émotions.

- **Fonction syntaxique**

La prosodie assume une fonction syntaxique, on parle **d'intonosyntaxe**¹⁴, dans la mesure où en premier lieu, le profil prosodique global d'un énoncé (une phrase) donne des informations sur sa modalité syntaxique (affirmation, interrogation, question, ordre, etc., cf. *infra*, figure 4). Ensuite, on l'a vu, à l'intérieur de l'énoncé, l'intonation permet d'indiquer la hiérarchie des constituants qui se construisent sur les bases de contraintes morphosyntaxiques minimales. Ainsi, (i) un syntagme accentuel est formé d'une tête lexicale et de ses dépendants, (ii) le degré et la distribution des frontières prosodiques sont corrélés à la

¹⁴ Au sein de ce champ intonosyntaxique, les chercheurs se donnent comme objectif principal de cerner les corrélations régulières et d'en expliquer les fondements discursifs et cognitifs.

hiérarchie syntaxique¹⁵. Autrement dit, la structure intonative a une fonction d'actualisation de la hiérarchie syntaxique. Ainsi, dans les deux énoncés suivants (Cauvet & al. 2010) :

(le petit chien mort) (sera enterré)

(le petit chien) (mord la laisse)

La segmentation intonative guide le calcul syntaxique on line des énoncés : « mort » adjectif vs. « mord » verbe.

- **Fonction rythmique**

Au rôle syntaxique de la prosodie viennent se greffer des contraintes rythmiques propres à la dimension sonore du message, articulées autour de deux principes dénommés respectivement **principe d'eurythmie** et **principe de progression**, qui conduisent à une réorganisation partielle de la structure intonosyntaxique de base.

Selon le principe d'équilibre rythmique, ou *eurythmie*, a priori universel : (i) deux syllabes contiguës dans la chaîne parlée ne peuvent pas être accentuées, de tels contextes, dénommés contextes de *clash accentuel*, seront jugés aprosodiques ; (ii) les groupes prosodiques qui se succèdent dans la chaîne doivent être de taille équivalente (couramment 3 à 4 syllabes), peuvent difficilement se constituer autour d'une seule syllabe et ne peuvent a priori pas excéder 7 syllabes à débit moyen. Ainsi, à l'énoncé *l'autobus de Serre Chevalier va arriver* devrait correspondre la segmentation intonosyntaxique (*l'autobus de Serre Chevalier*) (*va arriver*) mais la suite (*l'autobus*) (*de Serre Chevalier*) (*va arriver*) sera rythmiquement plus optimale (3+5+4 vs. 8+4 syllabes) et donc préférée. Toutefois, le marquage prosodique de la structure syntaxique reste ici respecté si la proéminence accentuelle terminale qui frappe le premier groupe est, toute chose égale par ailleurs,¹⁶ moindre que les proéminences subséquentes (voir note 15). Cela n'est pas toujours le cas : la segmentation intonosyntaxique (*papa*) (*dîne*) en vertu des principes 1 et 2 n'est pas acceptable dans le cadre d'un modèle prosodique standard et soit (i)° les deux groupes seront fusionnés en un seul (*papa dîne*), soit (ii)° il y a report du premier accent sur la syllabe initiale du mot (*papa dîne*).

En vertu du principe de progression (Lacheret & Beaugendre 1999), la taille des groupes n'est pas équivalente mais, au contraire, augmente progressivement au fur et à mesure de l'énoncé (*ex. (pas d'train) (pas d'bus) (prenez un aribus)*).¹⁷

En pratique donc, les réorganisations rythmiques sont nombreuses et expliquent partiellement pourquoi la place de l'accent est si difficilement prédictible en français (cf. *supra*, §.3.). Cela dit, le rythme n'explique pas tout : il existe dans le message parlé des contraintes beaucoup plus fondamentales encore auxquelles viennent également se modeler les constructions rythmiques, il s'agit des contraintes sémantico-pragmatiques. Elles sont de trois ordres : (i) le traitement de l'information et l'encodage de la structure communicative (voir note 17), (ii) le champ de la co-énonciation, (iii) : la dimension expressive du message (expression des émotions, sentiments et attitudes).

- **Fonction sémantico-pragmatique**

En termes de traitement de l'information, un énoncé type est articulé autour de deux éléments centraux : un élément donné dont on dit quelque chose, accessible par le cotexte

¹⁵ En français, la force d'un accent terminal serait proportionnelle à la force de la coupe syntaxique qui le suit (Dell 1984).

¹⁶ Indépendamment de contraintes pragmatiques qui pourraient venir modifier encore ce patron accentuel de base.

¹⁷ On peut supposer que cette progression rythmique est, dans certains genres de discours en tous cas, liées à la nature même dont est encodée l'information (de la moins essentielle à la plus importante) ; selon Givon (1994), plus la séquence informationnelle sera dense, plus le constituant qui la code sera long.

discursif ou le contexte situationnel¹⁸, cet élément est appelé *thème*, ou *topic*, et une information nouvelle dénommée *rhème* ou *focus*. A ces catégories conceptuelles répondent des positions dédiées dans la chaîne parlée. Ainsi en français ordinaire, le thème, qui actualise minimalement une entité référentielle, est extrait en position initiale (on parle aussi de position détachée), le rhème est encodé par une proposition subséquente (ex. *le cheval, il est parti*). Face à cette organisation prototypique, on trouve des variantes de construction dont on pose l'hypothèse qu'elles ont toutes une motivation fonctionnelle précise : thème extrait à droite (*il est parti, le cheval*), clivage comme dispositif syntaxique idéal pour focaliser un élément (*c'est hier que le cheval est parti*), etc. Ainsi, une construction donnée en entrée s'articule autour de différents champs que l'on peut voir comme des *zones topologiques*. Chaque champ convoque un profil prosodique spécifique : une frontière majeure actualisée par un contour montant, ou *continuatif*, marque la fin du champ gauche, *i.e.* le thème détaché en position initiale d'énoncé ; dans un dispositif clivé, le champ focal est réalisé par une saillance prosodique double : proéminence initiale sur le focus et contour terminal descendant, ou *conclusif*, à la fin de la zone focale ; le champ droit, qui instancie un thème repris, est indiqué par une intonation parenthétique, *i.e.* basse et peu modulée, etc. Par conséquent, la distribution d'un élément et son profil prosodique donnent des indications très précises à l'interlocuteur sur son statut pragmatique (accessibilité référentielle du topic, degré informationnel du focus, etc). D'une manière plus générale, une telle organisation vaut pour toutes les unités, *i.e.* pas seulement celles qui encodent des entités référentielles ou des actes prédicatifs. Ainsi : dans *en avril, ne te découvre pas d'un fil, en mai, fait ce qu'il te plaît*, les circonstants en tête d'énoncé (on parle de *cadrage thématique*), sont soumis à la même contrainte de marquage de champ gauche et donc ponctués par une frontière continuative majeure. Il en va de même pour les adverbes modaux (*franchement, tu exagères*). Au fond, l'essentiel pour le locuteur, qui opère par touches successives, n'est pas tant d'insister sur ce qu'il dit de nouveau que de bien montrer le point d'où il part (Morel & Danon-Boileau, 1998), la saillance prosodique de cette zone d'ancrage énonciatif est donc essentielle. A l'inverse, comme pour les thèmes repris dans le champ droit, les parenthèses discursives sont prototypiquement dotées d'une intonation basse et peu modulées (*Martin arrive par le train de 8h, il me semble*).

Comme l'illustrent les derniers exemples ci-dessus, le locuteur n'est pas une simple machine à traiter de l'information, il le fait en fonction d'un but communicationnel précis et, dans cette mise en forme de l'information, il exprime un certain point de vue. Aussi, pour expliquer le rôle de la prosodie dans l'encodage de la structure communicative, nous proposons de voir l'activité de langage comme la construction devant un auditoire d'une représentation d'une certaine réalité et selon un certain éclairage, qu'on nommera à l'instar de Grize (1990) *schématisation discursive*. D'où l'emprunt à Victorri & Fuchs (1996) de la notion de *scène verbale*. Dans le cadre théorique que les auteurs développent, l'acte d'énonciation de base serait constitué d'au moins deux éléments : la *description d'une scène* et la *donnée d'un point de vue sur cette scène*, ce deuxième élément étant essentiel pour le partage de la "vision". En suivant cette conception de l'interlocution, la structuration prosodique contribue largement à la construction de la scène verbale et de ses différents éléments, puisqu'elle sert, par le jeu des saillances syllabiques et des frontières intonatives, soit à évoquer des entités ou des événements sur cette scène, soit à en modifier le point de vue. L'expression du point de vue est étroitement associée au concept de *co-énonciation*. Ce dernier met en jeu l'ensemble des représentations que l'énonciateur se fait de la pensée qu'il prête à son co-énonciateur, des connaissances partagées et de l'univers de référence commun. Selon Morel & Danon-Boileau (1998), l'organisation intonative des constituants indique

¹⁸ La notion d'accessibilité référentielle a fait l'objet de travaux abondants dans les grammaires cognitives et fonctionnelles. Pour une revue et son application à la prosodie, voir Lacheret & François 2003.

clairement la relation que l'énonciateur tente d'établir avec son co-énonciateur - consensuelle ou non. Plus généralement et au-delà de ce cadre théorique particulier, le marquage du point de vue dans le discours explique l'occurrence des accents d'insistance internes de groupes (*c'est incroyable*) mais également certains contours terminaux encodant des valeurs illocutoires précises (assertion avec implication, évidence, doute etc.), voire des contenus expressifs et émotionnels. On se trouve là au cœur de la prosodie expressive et affective dont il faut également pouvoir rendre compte.

Les émotions peuvent être définies simplement comme des réactions affectives intenses, des réponses de l'organisme à des situations données (Darwin 1872, Damasio 2010). Des travaux déjà anciens ont montré comment l'intonation et l'accentuation fonctionnent à la fois comme des signaux expressifs dans une situation de communication et des indices relatifs à l'émotion d'un locuteur (Faure 1979, Fonagy 1983, Scherer & al. 1986)¹⁹. Pour une synthèse, on pourra consulter Léon (1993) qui distingue les marqueurs acoustiques suivants pour opposer prosodiquement les émotions : la variation de registre, l'amplitude et la nature d'un contour mélodique, son intensité, sa durée, et le tempo associé au changement de débit. En suivant Léon (1993), on peut distinguer au moins 3 types d'émotions de base : (i) les émotions brutes peu ou pas contrôlées, (ii) les émotions mieux canalisées, (iii) les attitudes contrôlées et codées linguistiquement au plan de l'expression. Certains types émotionnels peuvent selon les cas relever de telle ou telle catégorie. Par exemple, si on s'intéresse à l'expression de la colère, sa manifestation prosodique ne sera pas la même selon qu'il s'agit d'une réaction immédiate à une situation (émotion brute) ou de l'expression d'un sentiment bien maîtrisé, *i.e.* attitude (Grichkovtsova & al 2009). Indice dans le premier cas : elle relèvera plus de la gesticulation glottique immédiate, signal dans le second : on pourra y reconnaître des schémas mélodiques et temporels caractéristiques de l'émotion en question. Ainsi, s'il existe bien un système prosodique conventionnel, *i.e.* codé en langue, il possède également une part de motivation irréductible à la composante médiale, *i.e.* phonétique²⁰.

Aujourd'hui, l'essor des travaux sur la parole affective, qui intéresse au plus haut point la recherche clinique, témoigne de la dynamique du domaine et des enjeux pour les sciences cognitives, au carrefour de la linguistique, de la psychologie, des neurosciences et de la modélisation informatique. Trois axes majeurs rendent compte de cette dynamique et illustrent chacun à leur manière l'ancrage linguistique de la prosodie affective. Dans les travaux en neurosciences, où la question fondamentale est l'étude de la spécialisation hémisphérique des fonctions cognitives, l'étude des prosodies affectives permet d'enrichir les connaissances sur le réseau neural de la prosodie affective en appréhendant notamment la coopération hémisphérique mise en jeu au cours du traitement du discours (Beaucousin 2006). De ce point de vue, elle nous aide à mieux comprendre l'interaction entre sphère grammaticale et sphère affective dans les processus langagiers (Di Cristo 2004). La thématique constitue également un enjeu important pour la linguistique fonctionnelle et cognitive, qui interroge sous l'angle sémiotique la notion d'expressivité vue comme la manifestation d'un rapport affectif et émotionnel du sujet percevant à un contenu, via les modalités prosodiques et sémantiques. Dans quelle mesure les constructions prosodiques sont-elles dépendantes du domaine segmental (lexique et constructions syntaxiques) ? Telle sont les questions qui sous-tendent l'ensemble des travaux consacrés au sujet. (Hancil 2009). Enfin, les recherches interlangues sur la dimension expressive de la communication parlée constituent un terrain récent et fécond pour mieux saisir les traits caractéristiques des types de langues, des types de discours et, d'une façon plus générale, la dimension culturelle. De

¹⁹ Déjà chez Diderot, on trouve l'idée que les émotions se reflètent dans les articulations mélodiques de la parole et trahissent même les mouvements secrets de l'âme.

²⁰ Voir chez Fonagy (1970, 1971), la notion de *bases pulsionnelles* de l'intonation.

nouveaux projets ont donc vu le jour dans l'optique de mieux comprendre et de mieux simuler les mécanismes de perception et de production sous-jacents à l'expression prosodique des affects sociaux (Abelin & Allwood 2002).

Référence bibliographiques

- Abelin Å., & Allwood J. (2002). *Cross linguistic interpretation of emotional prosody*, Gothenburg, Papers in Theoretical Linguistics, Göteborg, Sweden.
- Astésano C. (1999). *Rythme et discours : invariance et sources de variabilité des phénomènes accentuels en français*, Thèse de doctorat, Université de Provence.
- Avanzi M., Lacheret-Dujour A. & Victorri B. (2008). « ANALOR, A Tool for Semi-Automatic Annotation of French Prosodic Structure », in *Proceedings of Speech Prosody'08*, 119-122.
- Avanzi M., Lacheret-Dujour A. & A. C. Simon (2010). *Proceedings of Prosodic Prominence: Perceptual and Automatic Identification, Speech Prosody 2010 Satellite Workshop, Chicago, May 10*. <http://www2.unine.ch/speechprosody-prominence>.
- Avanzi M., Lacheret-Dujour A. & B. Victorri (2010). « A corpus-based learning method for prominence detection in spontaneous speech », in *Prosodic Prominence: Perceptual and Automatic Identification, Proc. Speech Prosody 2010 Workshop, Chicago*.
- Beaucousin V. (2006). *Bases neurales de la compréhension affective*, Thèse de doctorat, Université de Caen.
- Boersma P. & D. Weenink (2011). *Praat: doing phonetics by computer (Version 5.2)*. www.praat.org.
- Boutard C. Guillon A. (2010). Le langage élaboré chez l'enfant dysphasique, évaluation et perspective de rééducation », in *Le langage oral : données actuelles et perspectives en orthophonie*, Th. Rousseau & F. Valette-Fruhinsholz (eds.), 255-286.
- Cauvet E., Brusini P., Fiévet A.C., Millotte S., Christophe A. (2010). « Démarrer l'acquisition de la syntaxe », in *Rééducation orthophonique, 48^e année, n°244 : l'émergence de la communication et du langage*, F. Coquet (ed.), 95-109.
- Chafe W. (1994). *Discourse, Consciousness and Time. The Flow and Placement of Conscious Experience in Speaking and Writing*, Chicago, University of Chicago Press.
- D'Alessandro, Ch. (2006). « Voice source parameters and prosodic analysis », in Sudhoff, S. et al. (eds) *Methods in Empirical Prosody Research*. Berlin, New York: De Gruyter, 63-87.
- Damasio, A. (2010). *L'Autre moi-même. Les nouvelles cartes du cerveau, de la conscience et des émotions*, Paris, O. Jacob.
- Darwin Ch. (1872). *The Expression of the Emotions in Man and Animals*, Londres, Murray, Chicago University Press, édition de 1989.
- Delais-Roussarie E. (2005). *Phonologie et Grammaire : Études et modélisation des interfaces prosodiques*, Mémoire d'Habilitation à Diriger des Recherches (HDR), Université de Toulouse-le Mirail.
- Delattre P. (1966). « Les dix intonations de base du français », in *French Review*, 40, American Association of Teachers of French, Illinois, 1-14.
- Dell F. (1984). « L'accentuation dans les phrases françaises », in F. Dell & al. (1984 eds.), 65-122.
- Dell F., Hirst D. & Vergnaud J.R. (1984 eds.). *Forme sonore du langage, structure des représentations en phonologie*, Hermann, Paris.
- Di Cristo A. (1999). « Le cadre accentuel du français : essai de modélisation », in *Langues*, vol. 2, no. 3 & 4, 184-205 & 258-269.
- Di Cristo A. (2004). « La prosodie au carrefour de la phonétique, de la phonologie, et de l'articulation formes-fonctions. », in *Travaux interdisciplinaires du Laboratoire parole et Langage d'Aix-en-provence*, 67-211.
- Faure G. (1970). « Contribution à l'étude du statut phonologique des structures prosodématiques », in *Studia Phonetica*, 3, P. Léon, G. Faure & A. Rigault (eds.), Montréal, Didier, 93-108.
- Fonagy I. (1979). « L'accent en français : accent probabilitaire », in *L'accent en français contemporain*, *Studia Phonetica*, 15, Ivan Fonagy et Pierre Léon (eds.), Paris: Didier, 123.233.
- Fonagy Y. (1970, 1971). « Les bases pulsionnelles de la phonation, I. Les sons, II. La prosodie », in *Revue française de psychanalyse*, 34, Société psychanalytique de Paris, 101-136 & 543-591.
- Fonagy Y. (1983). *La vive voix*, Paris, Payot.
- Fougeron C. (2001). « Articulatory properties of initial segments in several prosodic constituents in French », in *Journal of Phonetics* 29:2, 109-135.
- Gendrot C. & Gerdes K. (2009). « Prosodic hierarchy and spectral realization of vowels in French », in *Interface Discours & Prosodie*, Paris, 9-11 septembre 2009, 191-205.
- Givon T. (1994). « Isomorphism in the Grammatical Code: Cognitive and Biological Considerations », in *Iconicity in Language*, R. Simone (ed.), Benjamins, Amsterdam, 47-76.
- Goldman J.-P. (2008). *EasyAlign: a semi-automatic phonetic alignment tool under Praat*, <http://latlcui.unige.ch/phonetique>.

- Gratier M. (2010). *La dimension tacite de l'interaction sociale : communication mère-bébé et improvisation musicale*, mémoire d'habilitation à diriger des recherches, université Paris-Ouest Nanterre.
- Grichkovtsova I. (2006). *A Cross Linguistic Study of Affective Prosody Production by Monolingual and Bilingual Children*, PhD, Université d'Edinburgh.
- Grichkovtsova I., Morel M. Lacheret A. (2009). « Perception of Affective Prosody in Natural and Synthesized Speech : which Methodological Approach », in Hancil (2009éd.), 371-390.
- Grize J.B. (1990). *Logique et langage*, Paris, Ophrys.
- Grosjean F. & Dommergues Y. (1983). « Les structures de performance en psycholinguistique », in *L'Année Psychologique*, 183, Paris, PUF, 513-536.
- Hancil S. (2009éd.). *The role of Prosody in Affective Speech*, Linguistics Insights, 97, Peeter Lang, Bern.
- Hyman L.M. (2006). « Word-prosodic typology », in *Phonology* 23, Cambridge University press, 225-257.
- Jun S.A. & Fougeron C. (1995). « The accentual phrase and the prosodic structure of French », in *Proceedings of the 13th ICPHS (Stockholm)*, vol. 2, 722-725.
- Lacheret A. & Beaugendre F. (1999). *La prosodie du français*, Paris, éditions du CNRS.
- Lacheret A. & François J. (2003). « De la notion de détachement topical à celle de constituant thématique extrapositionnel », in *Cahiers de Praxématique*, 40, Linguistique du détachement, F. Neveu (ed.), 167-198.
- Lacheret A. (2010). « Petit parcours prosodique de Grenoble à Aix-en-Provence : pionniers, dissidents et fédérateurs », in *Un siècle de phonétique expérimentale, fondation et éléments de développement*, L.J. Boë & C.E. Vilain (eds.), ENS Editions, 277-297.
- Laks B. (1997). *Phonologie accentuelle – Métrique, autosegmentalité et constituance*. Paris, Editions du CNRS.
- Lambrecht K. (1994). *Information Structure and Sentence Form*, Cambridge University Press.
- Lenormand M.T. & Lacheret A. (2010). « Prosodie chez des enfants implantés cochléaires », in *Le langage oral : données actuelles et perspectives en orthophonie*, Th. Rousseau & F. Valette-Fruhinsholz (eds.), 63-88.
- Léon P. (1993). *Précis de phonostylistique, – parole et expressivité –* Paris, Nathan.
- Martel K. & Leroy-Collombel M. (2010). « Du gazouillis au premier mot : rôle des compétences préverbales dans l'accès au langage », in *Rééducation orthophonique, 48^e année, n°244 : l'émergence de la communication et du langage*, F. Coquet (ed.), 75-94.
- Martin Ph. (1981). « Pour une théorie de l'intonation : l'intonation est-elle une structure congruente à la syntaxe ? », in M. Rossi & al. (1981 eds.), 234-271.
- Mertens P. (2004). « Un outil pour la transcription de la prosodie dans les corpus oraux », in *Traitement Automatique des Langues* 45 (2), 109-130.
- Morel M.A. & Danon-Boileau L. (1998). *Grammaire de l'intonation – l'exemple du français*, Paris, Ophrys.
- Nespor M. & Vogel I. (1986). *Prosodic phonology*, Dordrecht, Foris.
- Robert S. (2002). « Modèles linguistiques de production », in M. Fayol (éd.), *Traité des Sciences Cognitives, vol. Production du langage*, Paris, Hermès, 66-86, 2002.
- Rossi M. & al. (1981 eds.). *L'intonation : de l'acoustique à la sémantique*, Paris, Klincksieck.
- Scherer K., Wallbott H.G. & Summerfield A.B. (1986). *Experiencing emotion, A cross-cultural study*, Cambridge, Cambridge University Press.
- Selkirk E.O. (1984). *Phonology and Syntax: The Relation between Sound and Structure*. Cambridge, MIT Press.
- Victorri B. & Fuchs Ch. (1996). *La polysémie. Construction dynamique du sens*, Paris, Hermès, 1996.

Annexe : 5 figures²¹

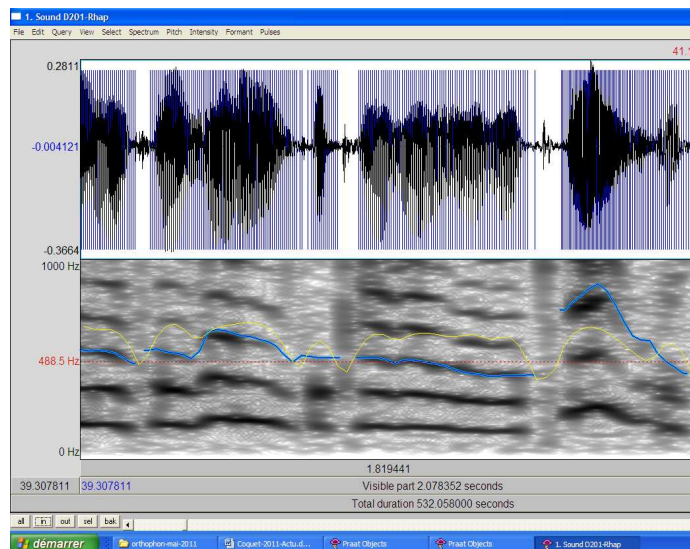


Figure 1. Représentation tridimensionnelle des paramètres prosodiques sous PRAAT de l'énoncé *il n'y a que les hommes qui travaillent avec moi*

Avec, dans la fenêtre du haut, le signal acoustique, dans la fenêtre du bas : le spectre, la ligne d'intensité en décibels, et les modulations de la fréquence fondamentale en hertz, en repérant par des valeurs nulles les zones non voisées ; sur l'axe des abscisses : le temps en secondes.

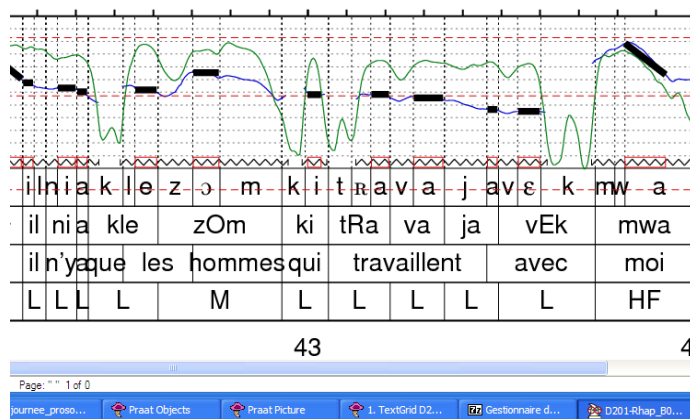


Figure 2. Contours syllabiques de l'énoncé *il n'y a que les hommes qui travaillent avec moi* représentés par une chaîne de tons dans le modèle Prosogram (Mertens 2004)

Avec dans la fenêtre du haut : (i) la ligne d'intensité et la ligne mélodique brutes, (ii) les segments de droite en gras qui représentent les contours mélodiques résultant du modèle de stylisation perceptive ; dans la fenêtre du bas : les tires de transcription : (i) les phonèmes, (ii) les syllabes, (iii) les mots orthographiques, (iv) les segments tonaux associés à chaque syllabe : L = ton bas statique (low) ; M = ton moyen statique séparé du niveau bas par un petit intervalle mélodique ; HF : ton haut (high) séparé du niveau L par un grand intervalle mélodique et suivi d'une descente ample (fall)

²¹ Les transcriptions phonétiques sont en alphabet SAMPA, http://fr.wikipedia.org/wiki/Symboles_SAMPA.

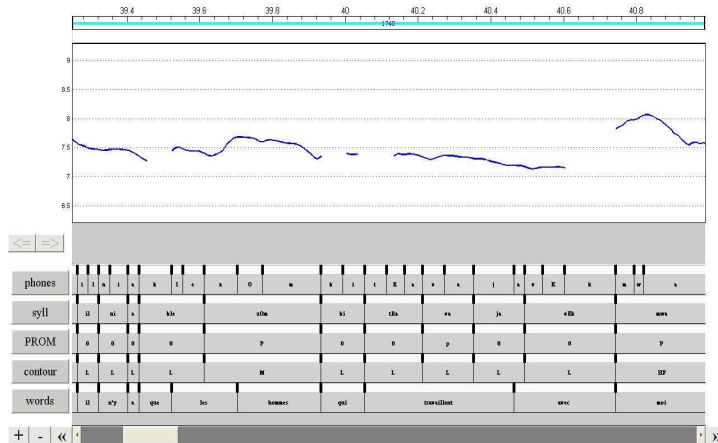


Figure 3. Etiquetage des proéminences syllabiques avec le modèle ANALOR (Avanzi & al. 2008, Avanzi & al 2010) de l'énoncé *il n'y a que les hommes qui travaillent avec moi*

Avec dans la fenêtre du haut : sur l'axe des abscisses, les durées syllabiques indiquées en millisecondes. Sur l'axe des ordonnées, les variations de la fréquence fondamentale représentées par une échelle logarithmique ; dans la fenêtre du bas : les tires de transcription avec de haut en bas : les étiquettes phonétiques, puis syllabiques, **l'annotation des proéminences (P pour les proéminences fortes sur les mots « hommes » et « moi », p pour les proéminences faibles : sur la syllabe terminale du mot « travaillent »)** et des contours (issus du program, figure 2), les mots orthographiques. La segmentation semi-automatique de ce passage a été réalisée avec le logiciel easy-align (Goldman 2008)²².



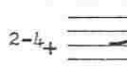
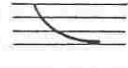
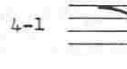
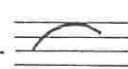
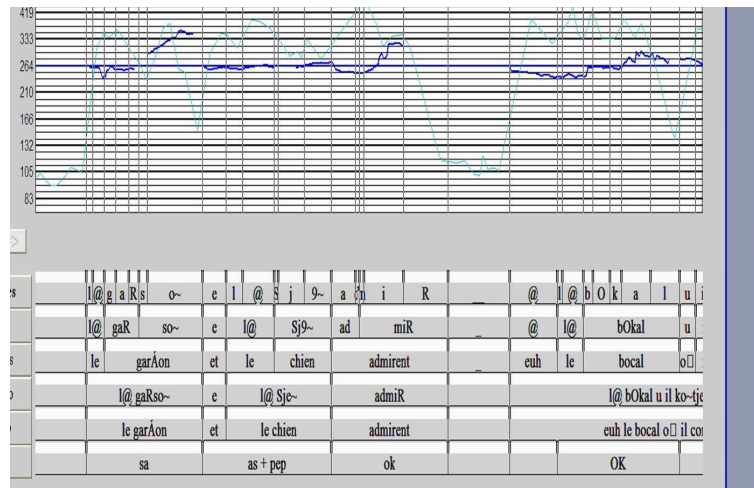
2-1 	4-1 
affirmation	Ordre
2-4+ 	4-1 
Question (réponse oui/non)	interrogation
4-1 	2-4- 
Exclamation	Evidence

Figure 4 Représentation schématique des intonations modales en français à l'aide de 4 niveaux de hauteur : niveau infra-grave, grave, aigu, suraigu (d'après Delattre 1966)

²² Les logiciels Analor et Easy-align sont téléchargeables gratuitement sur les sites suivants XX.



**Figure 5. Exemple de dysprosodie : mauvais appariement intonosyntaxique
Lenormand et Lacheret (2010) :**

Où les degrés de frontières ne s'alignent pas sur les domaines syntaxiques : frontière majeure ponctuant un constituant syntaxique mineur (*le garçon*), contour plat ponctuant le syntagme nominal sujet, pause décalée après le prédicat verbal.