

Existent-ils des indices intonatifs de segmentation en unités macro-syntaxiques ?

Elisabeth Delais-Roussarie^o et Injoo Choi-Jonin*

UMR 5610 – ERSS (Equipe de Recherche en Syntaxe et Sémantique)

Université de Toulouse – le Mirail, 5 allées Antonio Machado – 31058 TOULOUSE Cedex, France

Tél.: ++33 (0)5 61 50 36 02 - Fax: ++33 (0)5 61 50 50 96

Mél: ^o Elisabeth.Roussarie@wanadoo.fr et * choijoni@univ-tlse2.fr ou injoochoijonin@aol.com

^o <http://www.univ-tlse2.fr/erss/textes/pagespersos/delais/index.html> et * <http://www.univ-tlse2.fr/erss/textes/pagespersos/choijoni/index.html>

ABSTRACT

This paper reports on a pilot study which aims at evaluating the role of intonation in the segmentation of spontaneous speech in macro-syntactic units. A corpus of 3500 words, including two distinct styles of speech (unformal conversation and formal presentation), has been analyzed by both a syntactician and a phonologist in order to see if some tonal phenomena are associated with the edges of the various macro-syntactic units, e.g. the *periodes*, the *enlarged clauses* and the *clauses*. Preliminary results show that a reset usually takes place at the left edges of *periodes*, while it does not at the edges of *enlarged clauses* and *clauses*. Concerning right edges, they are marked by a falling tone (L or B) in the case of *periodes* and by a rising tone (H or T) in the case of *enlarged clauses* and *clauses*.

1. INTRODUCTION

En lecture de texte, la mélodie fournit fréquemment des indices permettant de déterminer certains aspects de la structure syntaxique, à savoir : la fin des phrases, les frontières de syntagmes syntaxiques majeurs ou indépendants (éléments disloqués, etc.). En parole spontanée, et cela quel que soit le style (présentation formelle, entretien, conversation familière), le rôle des indices mélodiques dans la segmentation syntaxique a fait l'objet d'un nombre plus restreint de travaux (cf., entre autres, Di Cristo [6] et Fonagy et al. [7]). Cela s'explique pour plusieurs raisons parmi lesquelles :

- l'impossibilité d'utiliser pour traiter l'oral spontané les unités syntaxiques communément admises (en particulier la phrase) ;
- l'absence de cadre formel d'analyse fournissant des constituants intonatifs et prosodiques valides pour l'oral spontané : dans la théorie prosodique par exemple, les constituants prosodiques (mot prosodique, groupe accentuel, etc.) sont surtout définis à partir d'une analyse syntaxique ayant pour extension maximale la phrase.

Ceci étant, une meilleure compréhension des liens entre syntaxe et prosodie en parole spontanée offre de nombreux intérêts, et cela dans plusieurs domaines. D'une part, cela pourrait permettre de mieux appréhender comment l'intonation (et plus largement la prosodie) joue son rôle de guidage syntaxique en parole spontanée (cf.

Di Cristo [6]). D'autre part, en traitement automatique et en linguistique de corpus, la mélodie pourrait faciliter la segmentation du continuum sonore (et de la transcription orthographique associée) en sous-unités qui peuvent ensuite être utilisées pour mener à bien des analyses syntaxiques automatiques (segmentation en constituants, calcul des dépendances, etc.). A l'heure actuelle, les analyseurs syntaxiques ne fournissent généralement pas d'analyse satisfaisante de l'oral, du fait notamment de l'absence de ponctuation (et donc d'unités syntaxiques « cadres » pré-définies) et de certains phénomènes tels que les hésitations (1), les reprises (2) ou les faux départs (3).

(1) *il euh il a fait des stages euh des stages en Allemagne*

(2) *alors mon grand-père était représentant de de commerce...*

(3) *justement il y a qu'est-ce qui quand vous êtes arrivée à Toulouse, est-ce que vous avez trouvé que c'était un bol d'air...*

Dans le cadre d'un projet de recherche ayant pour but le développement d'un analyseur syntaxique permettant l'annotation de corpus oraux, nous avons mené une expérience pilote dont le but est de voir s'il existe des indices intonatifs robustes qui permettent une pré-segmentation du corpus en unités macro-syntaxiques (cf., entre autres, Béguelin [1], Berrendonner [2] et Blanche-Benveniste et al. [5]). Supposer que de tels indices existent se justifie par le fait que, dans les situations de communication, les auditeurs arrivent très facilement à traiter les données qui leur sont proposées, et donc à en fournir, si besoin est, une analyse syntaxique pour accéder au sens du message.

Dans ce papier, les résultats de ce travail vont être présentés. Dans un premier temps, nous exposerons la méthodologie utilisée, en insistant en particulier sur les cadres retenus pour les descriptions syntaxiques et mélodiques. Dans un second temps, les résultats obtenus seront présentés.

2. MÉTHODOLOGIE

Pour tenter d'appréhender si des indices prosodiques, et plus particulièrement intonatifs, jouent un rôle important dans la détermination des unités macro-syntaxiques, nous avons travaillé sur un corpus de parole spontanée extrait du corpus ACSYNT et nous avons procédé comme suit :

- un expert syntacticien a travaillé sur les transcriptions orthographiques alignées des données afin d'en fournir une analyse macro-syntaxique ;
- parallèlement, une étude intonative des données a été faite à partir d'une stylisation de la courbe mélodique par points cibles et d'un codage INTSINT (cf. Hirst et al. [8] et [9]).
- les points cibles (et le codage symbolique INTSINT associé) qui sont réalisés à l'initiale et à la fin des unités macro-syntaxiques ont été relevés et classés sous forme de tableau de façon à voir si des régularités émergent.

2.1. Corpus utilisé

Pour mener cette étude pilote nous avons travaillé sur une partie du corpus ACSYNT, un corpus oral de 114 000 mots qui a les caractéristiques suivantes :

- l'ensemble du corpus est transcrit orthographiquement, et la transcription est entièrement alignée sur le signal. Cette tâche a été faite avec le logiciel PRAAT ;
- les données se composent d'enregistrements effectués auprès de 17 locuteurs (15 femmes, 2 hommes), âgés de 20 à 59 ans (12 personnes de moins de 25 ans) et étant étudiants (13 personnes) ou enseignants (une enseignante, deux enseignants-chercheurs et un chercheur) ;
- les données comprennent trois types de documents : des textes lus (13 437 mots, soit 11,83 % du total), des présentations formelles sous forme de monologue (19 945 mots, soit 17,6 %) et des entretiens sociolinguistiques (80 064 mots, soit 70,57 %)

Pour l'analyse des indices mélodiques et des découpages macro-syntaxiques, nous avons travaillé sur un sous-corpus de 3500 mots qui comprend :

- une partie de l'entretien enregistré par une locutrice (21 ans, étudiante) ;
- une partie de la présentation formelle faite lors d'une conférence par une des locutrices (40 ans, chercheur).

2.2 Pré-analyse syntaxique : cadre théorique et étude des données

Les découpages en unités syntaxiques fournis par l'expert syntacticien reposent sur une analyse macro-syntaxique. Ils ont pour but de déterminer les unités macro-syntaxiques pour les données du corpus, et d'utiliser ensuite ces unités pour automatiser en leur sein l'analyse syntaxique plus « fine ».

2.2.1 Cadre théorique : la macro-syntaxe

La macro-syntaxe est fondée sur la cohésion discursive. La combinatoire des unités macro-syntaxiques repose, selon Berrendonner et Reichler-Béguelin [4], sur des conditions d'appropriété, et non sur des rapports

rectionnels. Ces auteurs (cf. aussi Béguelin [1], Berrendonner [2] et [3]) proposent comme unité minimale de la macro-syntaxe la *clause*, et comme unité maximale, la *période*. Une *clause*, d'après ces linguistes, a pour fonction de servir à l'accomplissement d'un acte énonciatif. Il s'agit d'une unité minimale discursive, qui implique un certain état de l'information : « son emploi requiert que soient valides, donc présents en mémoire discursive, certains éléments de connaissance, faute desquels l'acte énonciatif qu'elle accomplit paraîtra impropre ou déplacé » (Berrendonner & Reichler-Béguelin [4] : 114). Par ailleurs, toujours selon les mêmes linguistes, toute clause incorpore des traits prosodiques, et plus particulièrement un contour intonatif. Une *période* est « une suite d'énonciations formant un programme discursif complet, qui est marqué par la présence sur son dernier terme d'un intonème conclusif » (Berrendonner [3] : 22). Dans une période, les clauses, sauf la clause finale, seraient alors marquées, selon ces linguistes, par un contour intonatif ascendant ou plat.

2.2.2 Segmentation des données

Si l'on applique les principes d'analyse proposés par Berrendonner et Béguelin, le passage suivant d'un entretien extrait de notre corpus forme une période, constituée de neuf clauses¹ :

(4) [*Oui / ben par exemple la semaine dernière / j'étais à l'opéra / à Castres / pour avoir un opéra / il faut vraiment [rire] chercher loin / donc là / non non / c'est c'est euh ce genre de chose / on le trouve pas dans une ville euh une petite ville.*]

Or, nous constatons que parmi les clauses, certaines peuvent former un énoncé autonome (*oui ; non non ; j'étais à l'opéra ; il faut vraiment chercher loin ; on le trouve pas dans une ville euh une petite ville*), alors que d'autres n'ont pas cette autonomie à elles seules, mais s'adjoignent à celles qui peuvent constituer un énoncé autonome (*ben par exemple la semaine dernière / j'étais à l'opéra ; à Castres / pour avoir un opéra / il faut vraiment chercher loin ; donc là / non non ; c'est c'est euh ce genre de chose / on le trouve pas dans une ville euh une petite ville*). L'ensemble formé par une clause autonome et par une (ou deux) clause(s) non autonome(s) correspond à la macro-syntaxe, telle qu'elle est définie par Blanche-Benveniste et al. [5]. Ces auteurs la considèrent en effet comme des relations de dépendance ou d'interdépendance, qui ne sont pourtant pas descriptibles en termes de rection ; « C'est le lieu des *compléments ambiants* de Damourette et Pichon, ou de ce qu'on a souvent appelé « les cohésions discursives », où les marques intonatives jouent un rôle discriminant » (Blanche-Benveniste et al. [5] : 114). L'unité minimale de la macro-syntaxe est appelée par ces auteurs *noyau*, unité permettant de former un énoncé autonome ; les séquences qui n'ont pas d'autonomie d'emploi, et qui se placent avant ou après le noyau sont appelées *préfixe*, *suffixe* ou *postfixe*. Le regroupement entre noyau et affixe(s) donne

¹ Une clause est délimitée par /, et une période, par [].

lieu à un énoncé *fini*. D'après ces auteurs, le passage (4) peut être segmenté comme suit² :

(5) [*Oui // ben par exemple la semaine dernière / j'étais à l'opéra // à Castres / pour avoir un opéra / il faut vraiment [rire] chercher loin // donc là / non non // c'est c'est euh ce genre de chose / on le trouve pas dans une ville euh une petite ville.//*]

Dans notre analyse macro-syntaxique, nous avons adopté à la fois les unités proposées par Berrendonner et Béguelin, et celles proposées par Blanche-Benveniste et al. Suivant Berrendonner et Béguelin, l'unité minimale est appelée *clause* et l'unité maximale, *période*; nous introduisons l'unité intermédiaire, *clause élargie*, afin de rendre compte, à l'instar de Blanche-Benveniste et al., de la relation de dépendance ou d'interdépendance observable entre certaines clauses. Dans la notation, une clause élargie, ainsi qu'une clause qui peut former seule un énoncé fini seront délimitées par //, et une clause qui constitue un composant d'une clause élargie sera délimitée par /; enfin, une période sera délimitée par []³.

2.3 Pré-analyse intonative

L'analyse intonative des données a consisté en une étude de la mélodie (les pauses et les allongements sont actuellement en cours d'analyse). Elle a été menée à partir d'une stylisation de la courbe de F0 par points cibles effectuée sous PRAAT en utilisant MOMEL (cf. Hirst et al. [8])⁴. Les points cibles ont ensuite été codés symboliquement avec INTSINT, un système de transcription formelle de la prosodie qui repose sur un inventaire de huit symboles. Trois d'entre eux indiquent des propriétés absolues liées au registre des locuteurs (**Bottom**, **Mid** et **Top**); cinq autres codent des informations de hauteur relative (**Up** ou ascendant, **Down** ou descendant, **Same**, **Higher** et **Lower**). Le codage INTSINT a été enregistré sous forme de fichier 'textgrid' lisible sous PRAAT.

L'étiquetage mélodique effectué avec MOMEL sous PRAAT a été utilisé pour étudier la valeur des points cibles associés aux frontières des différents constituants macro-syntaxiques. Pour chaque type de document du corpus (présentation formelle et entretien) et pour chaque type de constituants macro-syntaxiques, nous avons construit des tableaux ayant la forme suivante :

Table 1 : Grille d'analyse syntactico-mélodique.

| CLAUSES ÉLARGIES | | | | | |
|---------------------------------------|---------------|-------------|------------------------|-----------------------|---------------------|
| Séq. ortho. | Symbole Début | Symbole Fin | Val F0 _{deb.} | Val F0 _{fin} | Comment. |
| Ben par exemple...étais à l'opéra | D | U/H | 196 | 216 | |
| À Castres pour avoir....chercher loin | D | H | 185 | 247 | |
| Donc là non non | L | U | 173 | 206 | éléments discursifs |

De telles grilles ont été élaborées pour l'ensemble des données du corpus. Elles ont permis de voir si des régularités sont observables et peuvent constituer des indices prosodiques en vue d'une segmentation en unités macro-syntaxiques.

3. RÉSULTATS PRÉLIMINAIRES

L'étude des différentes tables de synthèse a fait émerger des régularités qu'il est important de noter, afin ensuite d'en évaluer la robustesse, voire éventuellement de les croiser avec d'autres indices reposant sur d'autres paramètres prosodiques (localisation et durée des pauses, allongements syllabiques, etc.). A ce stade, nous allons présenter les régularités observées en faisant une distinction entre les deux styles de parole spontanée, la présentation formelle et l'entretien. En effet, bien que les mêmes types d'indices se retrouvent dans les deux cas, leur portée et leur robustesse est plus nette dans le cas de la présentation.

3.1 Indices prosodiques dans les présentations formelles

Dans la présentation formelle, la mélodie fournit des indices importants d'aide à la segmentation. En ce qui concerne les frontières gauches des unités macro-syntaxiques, il est important de noter qu'un *reset* est réalisé au début de chaque période. Un 'ton' absolu **M** (associé à une valeur moyenne de 230 Hz) est réalisé dans 80 % des cas à l'initiale de chaque période, tandis que ce sont principalement les tons **L** et **B** (valeur moyenne de 190 Hz) qui sont réalisés au début des clauses et des clauses élargies (respectivement 180 Hz et 190 Hz de hauteur moyenne). En outre, la période a un contour global essentiellement descendant : des tons **L** ou **B** (avec une valeur moyenne de 153 Hz situé dans le bas du registre de la locutrice) sont toujours réalisés sur la dernière syllabe des périodes. En revanche, les clauses élargies et les clauses portent un contour montant.

La clause élargie et la clause se distinguent principalement en fonction de la hauteur sur laquelle elles s'achèvent. En effet, la frontière droite des clauses élargies est marquée par un ton haut majoritairement réalisé dans le haut du registre de la locutrice (ton **Top**),

² Un affixe est délimité par /, et un noyau, éventuellement affixé, par //

³ Les clauses parenthétiques, qui « servent soit à mettre à jour la mémoire discursive par une manœuvre de rattrapage informationnel, soit à commenter métadiscursivement l'énoncé » (Béguelin [1] : 253), sont provisoirement analysées comme des clauses faisant partie d'une clause élargie, mais il va sans dire qu'elles méritent un traitement spécifique. Ce sujet fera l'objet d'une autre étude.

⁴ Momel sous PRAAT a été développé par Cyril Auran et peut être téléchargé du site suivant : <http://www.lpl.univ-aix.fr/~auran/francais/index.html>

alors que seul un ton haut relatif est réalisé sur la dernière syllabe des clauses (ton **H**). Ces différences sont cependant moins robustes que celles observées pour indiquer le début des périodes et des autres constituants.

Les tons associés aux frontières gauche et droite des différentes unités syntaxiques sont synthétisés dans le tableau ci-après :

Table 2 : Contours et macro-syntaxe.

| Unité macrosyntaxique | Contour mélodique associé |
|-----------------------|---------------------------|
| Période | [M.....L/B] |
| Clause élargie | [L.....H/T] |
| Clause | [L..... H] |

3.2 Indices intonatifs dans l'entretien

Bien que les régularités et les contours globaux observés pour chaque unité dans l'entretien sont similaires à ceux définis pour la présentation formelle, les résultats sont moins nets.

Alors que dans la présentation un *reset* est réalisé au début des périodes dans plus de 80 % (ton **M** vs. ton **L** ou **B**), cela n'est fait que dans 55 % des cas lors de l'entretien. La hauteur moyenne des initiales de période est de 190 Hz, contre 180 Hz pour les clauses et les clauses élargies !

En ce qui concerne le ton final, là aussi les résultats ne sont pas aussi nets que pour la présentation. Le ton final associé à la dernière syllabe des clauses est un ton **H** dans 41 % des cas et un ton **U** (upstep) dans 37 % des cas (le registre haut **T** n'étant présent que dans 7 %). Pour les clauses élargies, le registre haut de la locutrice est plus souvent atteint (**T** dans 22 % des cas, et **H** dans 42 %).

Pour synthétiser, alors que la mélodie fournit des indices de segmentation macro-syntaxique assez robustes pour la présentation, elle est loin de le faire pour les entretiens. Il semble donc nécessaire de croiser les informations mélodiques avec d'autres informations prosodiques et linguistiques, notamment :

- la distribution et la durée de pauses et des allongements syllabiques ;
- la réalisation lexicale des unités discursives conclusives (*enfin, voilà*). La mélodie associée aux clauses se terminant par ces éléments lexicaux est en effet généralement réalisée à un registre de niveau globalement inférieur.

Ce travail est actuellement en cours et permettra, espérons-le, d'affiner la portée et la robustesse des indices prosodiques.

4. CONCLUSION

Bien que cela soit moins net qu'en lecture, les phénomènes mélodiques fournissent en parole spontanée

des indices de segmentation macro-syntaxique. Ces indices portent principalement sur la frontière gauche des unités macro-syntaxiques : un *reset* est généralement réalisé à l'initiale des périodes. Croisés avec d'autres types d'informations prosodiques et linguistiques, ces indices peuvent sans doute permettre de pré-segmenter en unités macro-syntaxiques des données de parole spontanée.

BIBLIOGRAPHIE

- [1] M-J. Béguelin. *De la phrase aux énoncés : grammaire scolaire et descriptions linguistiques*. De Boeck Duculot, Bruxelles, Belgique, 2000.
- [2] A. Berrendonner. Pour une macro-syntaxe. In *Travaux de linguistique*, 21 : 25-36, 1990.
- [3] A. Berrendonner. La phrase et les articulations du discours. In *Le français dans le monde : Recherches et applications*, février-mars 1993, pages 20-26, 1993.
- [4] A. Berrendonner et M-J. Reichler-Béguelin. Décalages: les niveaux de l'analyse linguistique. *Langue française*, 81 : 99-125, 1989.
- [5] Cl. Blanche-Benveniste, M. Bilger, Ch. Rouget et K. Van den Eynde. *Le français parlé. Études grammaticales*. Ed. du CNRS, Paris, 1990.
- [6] A. Di Cristo. La problématique de la prosodie dans l'étude de la parole dite spontanée. *Revue Parole*, 15&16 : 189-250, 2000.
- [7] I. Fonagy et J. Fonagy. L'intonation et l'organisation du discours. *Bulletin de la Société de Linguistique (BSL)*, LXXVIII/1 : 161-209, 1983.
- [8] D. Hirst et A. Di Cristo. *Intonation Systems : a survey of Twenty languages*. Cambridge University Press, Cambridge, UK, 1998.
- [9] D. Hirst, A. Di Cristo et R. Espesser. Levels of Representation and Levels of Analysis for the description of Intonation Systems. In M. Horne (ed), *Prosody : Theory and Experiment*. Kluwer Academic Publisher, Dordrecht, Pays-Bas, pages 51-88, 2000.
- [10] B. Post. *Tonal and Phrasal Structures in French Intonation*. Thesus, La Haye, Pays-Bas, 2000.
- [11] M. Rossi. A model for predicting the prosody of spontaneous speech. *Speech Communication*, 13 : 87-107, 1993.