La structure prosodique opère-t-elle avant ou après la syntaxe ?

Does the prosodic structure operates before or after syntax?

Philippe Martin

Édition électronique

URL : <https://journals.openedition.org/tipa/943>
DOI : 10.4000/tipa.943
ISSN : 2264-7082

Éditeur

Laboratoire Parole et Langage

Référence électronique

Philippe Martin, « La structure prosodique opère-t-elle avant ou après la syntaxe ? », *TIPA. Travaux interdisciplinaires sur la parole et le langage* [En ligne], 29 | 2013, mis en ligne le 17 décembre 2013, consulté le 10 mars 2023. URL : <http://journals.openedition.org/tipa/943> ; DOI : <https://doi.org/10.4000/tipa.943>

Ce document a été généré automatiquement le 10 mars 2023.

Creative Commons - Attribution - Pas d'Utilisation Commerciale - Pas de Modification 4.0 International
- CC BY-NC-ND 4.0
<https://creativecommons.org/licenses/by-nc-nd/4.0/>

La structure prosodique opère-t-elle avant ou après la syntaxe ?

Does the prosodic structure operates before or after syntax?

Philippe Martin

Introduction

- 1 Que ce soit dans la production orale spontanée, lue, ou même en lecture silencieuse, la structure prosodique est toujours présente. On ne peut produire un énoncé sans accompagner les séquences de syllabes successives encodant les morphèmes et les groupes syntaxiques de marques prosodiques localisées, impliquant des paramètres acoustiques définis. Ces marques, réalisées par des allongements syllabiques, des mouvements mélodiques, des variations d'intensité vocalique, apparaissent nécessairement dans des intervalles temporels variant relativement peu (de l'ordre de 250 ms à 1000 ms), intervalles nécessaires, comme on le verra, pour le traitement de l'information stockée dans la mémoire à court terme de l'auditeur. Même en lecture silencieuse, ou en mode chuchoté, des variantes de réalisation apparaissent, mais elles établissent toujours un système de contrastes nécessaires et suffisants pour permettre à l'auditeur ou au lecteur silencieux d'un texte écrit de reconstituer une organisation hiérarchique d'unités prosodiques minimales, les groupes accentuels, appelés aussi mots prosodiques ou groupes temporels.
- 2 Dans la conception traditionnelle, la structure prosodique, regroupement hiérarchique des unités prosodiques, participe à l'organisation syntaxique de l'énoncé, et agirait comme une béquille pour aider une syntaxe qui serait défaillante dans des soi-disant cas d'ambiguïté syntaxique. Ainsi dans l'exemple souvent cité *le petit garde les montre(s)*, la position de la première syllabe accentuée sur *petit* ou sur *garde* détermine l'identification de *garde* comme verbe ou comme nom, entraînant les interprétations respectives (*le petit*) ((*garde* (les *montres*)) ou (*le petit garde*) (les *montre*). Selon ce type d'exemples, la structure prosodique serait alors toujours congruente à la structure

syntactique et fonctionnerait en parallèle avec les autres structures qui organisent l'énoncé.

- 3 Cette conception fait bon marché, entre autres, d'une caractéristique de la structuration prosodique évidente mais souvent négligée : leur dimension temporelle, c'est-à-dire le fait que les événements prosodiques, portés par les syllabes successives, sont générés par le locuteur et perçus par l'auditeur séquentiellement, l'un après l'autre lors du déroulement du paramètre temps. Partant, la représentation graphique des événements prosodiques dans l'espace bidimensionnel d'une feuille de papier obscurcit considérablement cet aspect, pourtant essentiel, qui a trait à la linéarité du langage. Cette représentation suggère aussi que la structure prosodique serait aussi connue et analysée par l'auditeur en une seule fois, immédiatement, tout comme d'ailleurs les objets morphologiques, syntaxiques ou informationnels et leurs structures. C'est l'observation de données fournies par l'analyse de la parole spontanée (donc non lue et non préparée) qui permet de reconsidérer ces faits dans le cadre de leur dimension temporelle, reflétant ainsi au mieux les processus de génération de l'énoncé par le locuteur et sa compréhension par l'auditeur.
- 4 Se pose alors une question : les différentes structururations de l'énoncé identifiées par les linguistes, prosodiques, morphologiques, syntaxiques, informationnelles, etc. opèrent-elles toutes en même temps, ou existe-t-il un ordre présidant aux différentes structururations élaborées dynamiquement au cours du déroulement de l'énoncé ?
- 5 Un argument pour établir par exemple la priorité de la syntaxe sur la morphologie pourrait porter sur les corrections d'accord, de genre ou de nombre souvent observées dans la parole spontanée. Ces reprises éventuelles du locuteur pour modifier la flexion d'un verbe, d'un nom ou d'un adjectif semblent démontrer la préséance de la syntaxe puisque l'ajustement de la flexion en est une conséquence (ex. : *quand je pense à cette/ ce problème*). Mais qu'en est-il de la structure prosodique par rapport à la structure syntaxique ?

La structure prosodique est toujours présente

- 6 On ne peut ni parler ni lire même silencieusement sans produire en même temps une structure prosodique alors qu'on peut réaliser une structure prosodique sans parole donc sans syntaxe, en fredonnant par exemple. On connaît l'exemple du petit Philibert, cancre avéré parmi les personnages évoqués dans les années 60 par le chansonnier Jacques Bodoïn, qui connaît la mélodie associée à la table de multiplication, sans s'en rappeler les paroles. Ceci implique que la structure prosodique peut exister sans structure syntaxique actualisée, alors que l'inverse n'est pas vrai. La structure syntaxique dépend donc au sens de la glossématique de la structure prosodique, et cette relation n'est pas symétrique (il n'y a pas de relation de solidarité entre ces deux structures au sens de la glossématique). Ceci constitue un premier argument en faveur de la préexistence de la structuration prosodique par rapport à la syntaxe.

L'énonciation est structurée par des groupes de souffle

- 7 Un deuxième point porte sur les paramètres liés au mécanisme même de la phonation, dépendant directement du processus de production des sons de la parole. Que ce soient des voyelles, des consonnes voisées ou des consonnes sourdes (et mêmes des clics), leur production ne peut se faire que durant la phase d'expiration du cycle respiratoire inspiration-expiration. Il en résulte que le locuteur doit estimer avant chaque groupe de souffle (i.e. intervalle de temps entre deux phases inspiratoires successives) le volume d'air à inspirer dans ses poumons de manière à pouvoir disposer d'une quantité suffisante pour pouvoir produire la séquence de syllabes planifiées. Ceci implique que le locuteur doit effectuer une estimation raisonnable de l'effort phonatoire à fournir, et qui comprend comme paramètres l'intensité acoustique, les durées syllabiques, les mouvements mélodiques, le débit syllabique, accompagnant les syllabes (accentuées ou non). Des paramètres non linguistiques sont également impliqués dans cette estimation, tels que l'effort physique indépendant de la phonation à fournir (marche, nage, transport d'objets lourds, station assise ou non, etc.), et aussi l'état émotionnel pouvant intervenir dans la dépense d'air pulmonaire (peur, joie, tristesse, etc.).
- 8 Si cette estimation est en général approximative, divers procédés d'ajustement et de correction sont disponibles. Ainsi, le locuteur peut accélérer le débit syllabique pour réaliser la planification prévue, et d'économiser ainsi l'air pulmonaire emmagasiné. Il peut aussi diminuer sa dépense d'air pulmonaire en réduisant l'intensité de la parole produite et la fréquence de vibration laryngée nécessaire à la production des voyelles et des consonnes voisées. Il peut également modifier la structure prosodique prévue en simplifiant les contrastes des marques prosodiques à réaliser, par exemple en produisant une énumération de séquences de syntagmes prosodiques organisées en un seul niveau. L'inverse est aussi possible lorsque le locuteur dispose de plus d'air pulmonaire que nécessaire, et décide de ralentir le débit syllabique par exemple. Les professionnels de la parole, acteurs, enseignants, politiciens, sont eux entraînés à gérer au mieux l'adéquation air pulmonaire groupes de souffle, d'autant qu'ils disposent d'un large lexique de séquences syllabiques toutes préparées.
- 9 Tous ces mécanismes gérables par le locuteur déterminent le volume d'air à inspirer mais permettent aussi un ajustement dynamique en cours d'élocution de chaque groupe de souffle. Or les pauses respiratoires provoquent une interruption de la phonation, et leur position dans le flot syllabique doit être signalée par un événement prosodique spécifique, qui peut être un contour conclusif, ou un contour de continuité mineure ou majeure.
- 10 On conclura donc à ce stade, du moins pour la parole oralisée et non pour la lecture silencieuse, que la ou les structures prosodiques résultant de ces processus sont contraintes tout autant que les structures syntaxiques par le cycle de respiration qui met en place une ébauche de cadre dans lequel viendront s'inscrire les différentes unités morphologiques et syntaxiques au cours de la phonation.

L'énoncé est structuré en groupes temporels

- 11 Des travaux récents en neurocognition (Gilbert et Boucher, 2009, Gilbert, 2012) ont montré que pour être perçues et mémorisées, les syllabes doivent être organisées en segments (chunks) de 3 ou 4 syllabes si ces segments ne correspondent pas à des entrées lexicales (en réalité à des entrées de groupe accentuels connus de l'auditeur et stockés dans sa mémoire à long terme). Il s'agit donc de groupes sans signification de type *bisraktoubzachdujpermas*. À l'écoute de groupes plus longs, l'auditeur va organiser ces séquences en sous-groupes de 3 ou 4 syllabes (par exemple 7 syllabes décomposées en deux séquences de 3 et 4, ou 4 et 3).
- 12 L'allongement, donc la proéminence de syllabes à l'intérieur des groupes de plus de 4 syllabes va déterminer le découpage résultant, par exemple *bisraktoub* et *zachdujpermas* si la troisième syllabe est accentuée, *bisraktoubzachdujpermas*, et *bisraktoubzach* et *dujpermas* si la quatrième syllabe est accentuée *bisraktoubzachdujpermas*. La présence d'une syllabe accentuée détermine donc le découpage des syllabes. Dans l'examen des tracés électroencéphalographiques (EEG), on s'aperçoit du reste que les ondes Thêta, dont la gamme de variation va de 4 Hz à 10 Hz, soit de 10 ms à 250 ms, synchronisent la perception des syllabes, ce qui conduit à faire une hypothèse concernant les ondes Delta, qui elles varient de 1 Hz à 4 Hz, soit de 250 ms à 1000 ms : les ondes Delta synchroniseraient le transfert de la mémoire à court terme stockant les séquences de syllabes (Martin, 2013).
- 13 Cette hypothèse validée par des tracés EEG récents (Boucher, Gilbert, Martin, 2013), permet d'expliquer plusieurs contraintes observées auxquelles est soumise la structure prosodique (Martin, 2009) :
 - a) la « collision d'accent », qui impose un certain intervalle de temps entre deux syllabes accentuées successives (de type *il boit son café # chaud* vs. *Il boit son café chaud* avec un recul du premier accent) ;
 - b) la limite du nombre de syllabes non accentuées successives (le mot *paraskevidekatriaphobie* (la peur du vendredi 13) demande – au moins – deux syllabes accentuées, par exemple *paraskevidekatriaphobie*) ;
 - c) l'eurythmie, tendance à équilibrer le nombre de syllabes de groupes prosodiques de même niveau dans la structure ([*Marie adore*] [*les îles Fidji*] plutôt que [*Marie*] [*adore les îles Fidji*]), formant ainsi des groupes de 4 + 4 syllabes, plutôt que 2 + 6 ;
 - d) l'« alignement syntaxique », empêchant des découpages tels que [*Marie adore les*] [*îles Fidji*], le groupe accentuel *Marie adore les* étant inattendu. Cette dernière contrainte pourrait apparaître comme un contre argument, mais il n'en est rien, si l'on considère que c'est dans le « moule » prosodique préexistant que doivent s'insérer les groupes syntaxiques « bien formés », c'est-à-dire faisant partie d'un lexique de formes connues par l'auditeur. Ce lexique n'est pas limité aux mots (au sens traditionnel des dictionnaires), ni non plus, du moins en français, à une seule unité de classe ouverte (nom, verbe, adjectif, adverbe), autour duquel gravitent des unités de classe fermée (conjonction, pronom...) qui en dépendent. Il peut s'agir aussi bien d'une syllabe (*je ne vais pas te le ré-pé-ter deux fois*) que d'un groupe plus grand (*la ville de Meaux*). La sélection d'un groupe syntaxique se fait donc après l'installation mentale du groupe temporel, le locuteur devant adapter le débit syllabique en le ralentissant ou en accélérant selon le nombre de syllabes impliquées dans sa sélection.

- 14 La fonction de déclenchement attribuée aux ondes Delta des séquences de syllabes rend compte de ces 4 contraintes. En effet, les périodes successives de ces ondes ne peuvent d'une part être inférieures à 250 ms (ordre de grandeur), ce qui empêche l'apparition de syllabes accentuées distantes de moins de cette valeur temporelle. Elles ne peuvent pas non plus excéder 1000 ms, ce qui explique la nécessité d'accentuer une ou plusieurs syllabes supplémentaires dans des mots très longs dont l'élocution demanderait plus de 1000 ms, sans quoi les séquences de syllabes correspondantes ne pourraient pas être transférées dans une autre zone mémorielle et l'information serait perdue. Enfin, les périodes successives des ondes Delta ne peuvent varier énormément d'une période à la suivante, ce qui explique la contrainte d'eurythmie : l'énonciation des groupes accentuels successifs va alors se réguler soit par un choix du locuteur de nombre de syllabes sensiblement égal, soit par une variation du débit syllabique d'un groupe à l'autre. Le résultat sera le même, un équilibre dans les durées des groupes accentuels (ou mots prosodiques), qu'on devrait plutôt appeler groupes temporels (Gilbert, 2012).
- 15 Ces différentes contraintes d'ordre neurocognitif préexistent évidemment à la structure syntaxique et aux autres structures de l'énoncé et imposent une segmentation des séquences de syllabes en groupes temporels. Les groupes de plus de 4 syllabes qui ne correspondent pas à une entrée « lexicale » sont, on l'a vu, réorganisés en groupes plus petits. Par contre, ceux qui correspondent à une entrée présente dans la mémoire à long terme de l'auditeur sont limités seulement par les périodes les plus longues des ondes Delta qui assurent leur synchronisation pour leur mise en mémoire, et par bootstrapping si la syllabe accentuée est en position non finale.
- 16 Le fait que les événements prosodiques soient toujours présents pour permettre à l'auditeur de décoder la structure prosodique même en lecture silencieuse procède d'un mécanisme cognitif lié à la mémorisation à court terme. De même que les séquences syllabiques, synchronisées par les ondes Delta, ne peuvent excéder un certain nombre de syllabes, de même les séquences de groupes temporels (les mots prosodiques) doivent être en nombre limité de manière à ne pas dépasser les capacités de mémorisation de l'auditeur. Un regroupement hiérarchique en niveaux successifs est donc indispensable pour assurer un décodage sans perte d'information pour cause d'informations trop nombreuses, et ce regroupement est d'abord indiqué par des événements prosodiques, et ensuite par des propriétés syntaxiques et des traits morphologiques.

Les reprises, reformulations, répétitions et abandons de l'oral spontané sont structurées en groupes temporels

- 17 Outre les procédés de régulation évoqués plus haut, le locuteur peut en cours de production de parole modifier la structure prosodique prévue par divers moyens, dont certains relèvent des « scories » relevées dans la parole spontanée (et plus rarement dans la parole lue, le texte étant imposé au lecteur). Parmi ces « scories », on trouve la répétition, l'hésitation (par le *euh* d'hésitation, l'allongement de voyelles, etc.), la reformulation, l'abandon. Ces événements modifient aussi bien la structure syntaxique que la structure prosodique.

- 18 Si cet événement n'est pas réalisé, n'a pas eu lieu du fait d'un abandon du groupe accentuel en cours de formation, abandon suivi ou non d'une répétition ou d'une reformulation, le groupe n'est pas mémorisé par l'auditeur et se trouve donc effacé de la séquence constituant l'énoncé final. Il y a donc raturage d'un segment de l'oral. De plus, la reprise en groupe accentuel, séquence de syllabes ne comprenant qu'un accent (lexical) s'accompagne éventuellement d'un ajustement phonétique comme l'élision d'une voyelle de l'élément de classe fermée devant l'unité lexicale commençant par une voyelle.

Alors **la** l'infirmière de temps en temps me m'humectait euh les lèvres (I.92 Vallier 6,8)

- 19 L'article **la** ne comportant pas d'évènement prosodique indiquant la fin d'un groupe accentuel n'est pas mémorisé par l'auditeur et est remplacé par le groupe *l'infirmière*.

Je je j'ai eu les jambes qui ont tremblé (I.92 Selin 11,12)

- 20 De même, les répétitions de *je je* ne porte pas d'évènement prosodique relatif à la structure prosodique, ne seront pas mémorisées par l'auditeur, même si celles-ci comportent un allongement vocalique ou sont suivies d'un *euh* d'hésitation.
- 21 Dans l'interprétation de ces données par C. Blanche-Benveniste (2003) ces hésitations-répétitions indiqueraient une séparation entre le traitement de la syntaxe et celui du lexique : le cadre syntaxique est posé, prévu, mis en place mais le choix de l'unité de lexique approprié reste en suspens. Ce qui est révélateur du point de vue prosodique c'est que même s'il n'y a pas d'ajustement phonétique par élision dans la reprise, c'est l'entière du groupe accentuel qui sera reprise et non le seul élément lexical enfin retrouvé par le locuteur. On a alors une répétition, comme dans

Je je **je vais le faire** bientôt (CL 96,4,8)
Et on les **on les cultive** comme ça (Choix 54, 98)

- 22 Les événements prosodiques dont les instanciations constituent les marques indicatrices de la structure prosodique, apparaissent dans la séquence syllabique constituant l'énoncé comme autant de déclencheurs de conversion en unités linguistiques de rang supérieur qui seront mises en mémoire par l'auditeur. Cette opération de conversion porte sur des unités qui ne sont pas des mots mais des groupes temporels constitués comme on l'a vu par une (ou plusieurs) unités de classe ouverte dont dépendent éventuellement une ou plusieurs unités de classe fermée. Selon la classe de chaque événement prosodique identifié par l'auditeur et donc telle que l'a codé le locuteur, les groupes temporels déjà stockés dans la mémoire de l'auditeur sont concaténés avec le groupe temporel terminé (en français) par le dernier événement prosodique identifié.

La structure prosodique est générée dynamiquement par le locuteur

- 23 Mais la structuration prosodique ne se limite pas à un découpage en groupes temporels qui seraient au même niveau d'une structure « plate » d'énumération. On peut montrer que ce sont les modulations des paramètres d'accentuation des syllabes par des variations de durée, de mélodie ou d'intensité qui permettent d'encoder des groupements hiérarchisés des groupes temporels, constituant ainsi une structure prosodique. La linéarité du langage impose à ce processus d'opérer selon un mécanisme dit de « stockage-concaténation » (Martin, 2009), par lequel les séquences de groupes temporels sont regroupées hiérarchiquement de manière à reconstituer dynamiquement, c'est-à-dire en se construisant au cours du temps, la structure prosodique voulue par le locuteur. Ces regroupements sont déclenchés par des événements prosodiques spécifiques instanciés essentiellement par des contours mélodiques, des allongements vocaliques, etc.
- 24 La génération de la structure prosodique est donc un processus dynamique, dont l'empan ne saurait dépasser une certaine limite fixée par les capacités de mémorisation à court terme, limite sans doute de l'ordre de quelques groupes temporels. Dans ce processus, le locuteur dispose de choix limités dans l'encodage des fragments de structure prosodique successifs (correspondant aux ip – syntagme intonatif intermédiaire – et IP – syntagme intonatif – selon les définitions autosegmentales-métriques de la structure prosodique SP).
- 25 Il y a en français 4 types de tons de frontières (et aucun accent lexical). Appelons les C0 (frontière de SP), C1 (frontière de IP), C2 (frontière de ip), et Cn (frontière des AP). Des définitions commodes et plus intuitives que celles utilisant la notation ToBI sont données par les traits - Haut, - Montant, - Ample pour C0, + Haut, + Montant, + Ample pour C1, - Haut, - Montant, + Ample pour C2 (le trait *Ample* se référant à l'amplitude de variation mélodique). Cn est un contour neutralisé qui ne partage aucun de ces traits.
- 26 À partir d'un contour Cx (x = 0, 1, 2, n), le locuteur doit nécessairement planifier une séquence de groupes temporels qui se terminera nécessairement par C0, c'est-à-dire un contour conclusif terminant la séquence de groupes temporels à énoncer (on écartera les abandons ainsi que la possibilité d'avoir un ajout macrosyntaxique (i.e. un complément différé) pour ne pas compliquer la suite de la discussion). En fait le choix doit se porter sur une relation de parataxe (juxtaposition des unités) ou de rection (dépendance des unités) envers le contour mélodique qui suit. Il y a donc nécessairement planification, au moins locale jusqu'à l'apparition du contour C1 suivant, puisque par définition les classes de contour définissent des relations de dépendance envers un contour situé dans la suite de la séquence (i.e. « à droite » dans une représentation traditionnelle).
- 27 Le processus de *stockage-concaténation* détermine l'utilisation des séquences de contours par l'auditeur à partir des relations de dépendance indiquées par les contours. Ainsi, Cn indique une relation de dépendance par rapport à un contour C2, C1 ou C0 qui le suit, et une relation de juxtaposition si le contour suivant est de même classe Cn. Dans le premier cas, le groupe temporel terminé par Cn formera un groupe plus grand avec celui terminé par C2, C1 ou C0 (rection). Dans le second, les groupes terminés par Cn seront simplement réunis dans une énumération (parataxe). Le même principe

s'applique aux contours C2 et C1 (en même C0 ...).

Il résulte aussi de ce mécanisme que les contours fonctionnent par des relations locales (i.e. par rapport aux contours qui les suivent) et ne nécessitent pas de caractéristiques de réalisations acoustique invariantes, pourvu que les contrastes locaux puissent être identifiés par l'auditeur.

Parataxe

- 28 Pour réaliser une relation de parataxe à partir d'un contour Cx, le locuteur devra donc choisir un contour de même classe Cx pour un groupe prosodique de même niveau.
- 29 C0 -> C0 Deux contours conclusifs successifs dans le même énoncé (cas du complément différé – Bailly, 1944)
- 30 Dans l'exemple *...bloqués à Londres en Angleterre dans le métro* (C0) *qu'on devait prendre pour rejoindre notre ville d'accueil* (C0) (nar-fr, C-PROM) la locutrice termine une première séquence *...en Angleterre dans le métro* par un contour conclusif sur la dernière syllabe de *métro*, puis ajoute un complément différé introduit par la conjonction *que*, ajout terminé par un deuxième contour conclusif terminal C0 placé sur la dernière syllabe de *accueil*. Il s'agit bien d'un ajout, l'élimination de la séquence de complément différé par un éditeur de signal permettant de constater que la première partie de l'énoncé est bien formée à la fois du point de vue syntaxique et du point de vue prosodique.
- 31 C1 -> C1 Parataxe de premier niveau dans la structure prosodique
- 32 L'exemple suivant illustre un cas similaire d'ajout mais cette fois avant le contour conclusif C0 :

je confirme que le premier ministre Elio Di Rupo m'a parlé de cette situation et euh de l'inquiétude des brasseurs (C1) belges (C1) par rapport à ce qu'était la consommation française qui va rester je vous le confirme à un bon niveau (C0) (FH).

- 33 Cette réalisation contient un ajout, où le locuteur a séparé par deux contours montants de même classe les unités *des brasseurs* [belges au lieu de les regrouper syntaxiquement en *l'inquiétude (des brasseurs belges)* (cf. Godement-Berline et Martin, 2010). Cet exemple démontre assez bien que la structure prosodique préexiste dans la planification du locuteur, qui ne peut qu'ajouter un groupe temporal supplémentaire pour accommoder l'adjectif *belges* non prévu au départ de sa formulation (on peut aussi penser que le locuteur FH, fréquent utilisateur de cette construction, a réalisé cet ajout pour des raisons stylistiques...).
- 34 C2 -> C2 Un exemple est donné par
- oui si vous voulez ce qui me frappe (C2) c'est que depuis un demi-siècle (C2) a commencé ce qu'on appelle la construction européenne (C1)...(PM).
- 35 Ici le locuteur réalise un contour C2 sur *frappe* et a choisi une séquence terminée par C1 (*européenne*) mais composé de deux groupes prosodiques dont le premier est terminé

par C2. Il y a donc planification d'une subordination prosodique, cadre dans lequel vient s'inscrire la syntaxe.

bien que j'ai eu trois enfants (C1) déjà à l'époque (C1) que j'avais déjà construit ma petite maison à Montpellier (C1) ma petite villa (C0) (PRI-NAR-2).

- 36 Cet énoncé montre trois IP terminés par C1 donc en relation de parataxe.

Rection

- 37 Selon la grammaire prosodique vue plus haut, pour une rection, à partir de C1, le choix est entre Cn, C2 : C1 Cn ou C1 C2. En voici quelques exemples à partir de C1, C2 C1.

- 38 C2 -> C1

nous avons vu une euh euh un crépuscule (C2) euh **alors que** nous étions d- au au sommet de la mosquée (C1) (PRI-AIX-1).

- 39 Il y a ici subordination prosodique, congruente à la subordination syntaxique introduite par *alors que*. La séquence de contours C2 C1 ne permet pas d'interpréter la séquence comme parenthèse.

bon à Compiègne (C2) où j'avais une très très très grosse clientèle (C2) **puisque** sur cinq paroisses (C2) j'en avais quatre (C1) eh /bien, ben/ à Saint-Antoine de Compiègne \ (PRI-AMI-3).

- 40 Autre exemple de subordination prosodique correspondant à la subordination syntaxique.

alors que là (C1) c'est des gens de passage (C2) qui passent dans la rue (C1) (PRI-BES-2).

- 41 Autre type d'exemple de subordination prosodique [*c'est des gens de passage (C2)*] et [*qui passent dans la rue (C1)*].

- 42 Le cas du français est particulièrement révélateur de la planification nécessaire de la part du locuteur. En effet en français, le mécanisme de contraste de pente, dans des conditions de non neutralisation, impose d'inverser la pente mélodique de l'avant dernier contour de l'énoncé, ou alors de le neutraliser (par un contour Cn), selon les séquences C1, C1,..., C2, C0 ou C1, C1,..., Cn, C0.

- 43 La génération de la structure prosodique se fait syntagme prosodique par syntagme prosodique (ip par ip et IP par IP). Le simple fait que le codage des relations de parataxe et de rection envisagées par le locuteur se fait par rapport aux segments prosodiques futurs, puisque les relations de dépendance prosodiques se font « à droite », implique que la structuration prosodique précède nécessairement la structuration syntaxique et a fortiori la sélection lexicale et morphologique.

Conclusion : la structure prosodique (dynamique) préexiste à la structure syntaxique

- 44 Si on revient à l'aspect temporel, essentiel pour une meilleure compréhension de la génération des structures syntaxiques et prosodiques (du moins dans la parole spontanée), on s'aperçoit que, en plus des « scories » (abandons, reformulations, hésitations, répétitions, ponctuants,...) susceptibles d'interrompre le flux de l'énonciation « bien formée », le locuteur procède par génération de segments à différents niveaux la structure prosodique (ip et IP dans la théorie Autosegmentale-Métrique). Des observations précédentes, on peut au moins conclure que la structure prosodique partielle précède la structure syntaxique dans le processus d'énonciation.
- 45 Dans la parole lue, l'existence préalable du texte implique évidemment celle de la syntaxe, imposée au lecteur, mais qui peut y associer une structure prosodique relativement indépendante (ou pas, selon qu'il s'agisse d'un bon ou mauvais lecteur...).
- 46 On peut aussi remarquer que l'estimation du volume d'air pulmonaire par le locuteur pour la production d'un groupe de souffle ne détermine pas nécessairement la configuration de la structure prosodique, mais seulement le nombre approximatif de syllabes composantes, et éventuellement de groupes accentuels. Avec un même volume d'air inspiratoire, le locuteur pourra tout aussi bien produire une structure prosodique plate (i.e. une énumération, c'est à dire une concaténation des groupes accentuels à un seul niveau) qu'une structure à deux ou même trois niveaux, la différence d'effort articulatoire étant négligeable relativement à la réalisation de contours plus ou moins montants ou descendants.
- 47 De même, le tronçonnage en chunks et les reprises observées en parole spontanée en groupes temporels de l'ordre de 7 syllabes au maximum ne définit pas la structure prosodique mais seulement ses unités composantes. Pour ce qui est de la structure prosodique elle-même, son processus de génération, ainsi que le mécanisme de décodage par stockage-concaténation par l'auditeur, demandent une planification de la part du locuteur qui ne préjuge pas directement de l'organisation syntaxique ni de la sélection lexicale, et qui donc leur préexiste.
- 48 Le plus surprenant dans cette conclusion est que l'objet langagier le plus négligé des études linguistiques apparait de plus en plus comme le plus essentiel.

BIBLIOGRAPHIE

Bally, Ch. (1944) *Linguistique générale et linguistique française*, Francke, Berne, 410 p.

Blanche-Benveniste, C. (2003) La naissance des syntagmes dans les hésitations et répétitions du parler, in Araoui J.L. ed. *Le sens et la mesure. Hommages à Benoît de Cornulier*, Honoré Champion, Paris, 40-55.

- Blanche-Benveniste, C. et Ph. Martin (2011) Structuration prosodique, dernière réorganisation avant énonciation, *Langue française*, 170 (2/2011), 127-142.
- Boucher, V., A. Gilbert et Ph. Martin (2014) Les ondes cérébrales Thêta et delta et la perception de la structure prosodique dans la parole. (à paraître).
- Godement-Berline, R. et Ph. Martin (2010) Suffixes complexes : quand c'est fini, ça recommence... JEP 2010 (*XVIIIèmes Journées d'Étude sur la Parole*), Mons, Belgique. 25-28 mai 2010.
- Gilbert, A. & V. Boucher (2007) What do listeners attend to in hearing prosodic structures? Investigating the human speech-parser using short-term recall, *Proc. Interspeech 2007* : 430-433.
- Martin, Ph. (2009) *Intonation du français*, Armand Colin, Paris.
- Martin, Ph. (2012) Souveraineté-Association en linguistique : L'exemple de l'intonation et de la (macro)syntaxe, 2012, R. Druetta (éd.), *Claire Blanche-Benveniste. La linguistique à l'école de l'oral*, Sylvain-les-Moulins : Éditions GERFLINT, 2012 (Coll. Essais Francophones, n° 1).
- Martin, Ph. (2012) La structure prosodique dynamique : Rature et insertion de texte dans l'oral spontané, 2012, *Penser les langues avec Claire Blanche-Benveniste*, S. Caddéo, M-N. Roubaud, M. Rouquier et F. Sabio, éd., Presses Universitaires de Provence, coll. Langues et langages, 2012, 256 p., 117-125.
- Martin, Ph. (2013) Neurophysiological research explains prosodic structures constraints, 2013, *Revista da linguagem*, (20) 2, UFMG, Belo Horizonte, Minas Gerais, Brésil, 13-22.

RÉSUMÉS

Que ce soit dans la production orale spontanée, lue, ou même en lecture silencieuse, la structure prosodique est toujours présente. On ne peut produire un énoncé sans accompagner les séquences de syllabes successives encodant les morphèmes et les groupes syntaxiques de marques prosodiques qui déterminent une structure prosodique, c'est-à-dire un regroupement hiérarchique de groupes accentuels (séquence de syllabes ne comprenant qu'une seule syllabe accentuée qui n'ait pas de fonction iconique – tel que l'accent d'insistance).

Pour établir que la structuration prosodique intervient dans le processus d'énonciation avant des autres structurations de l'énoncé, syntaxiques, morphologiques, informationnelles, on présente différents arguments d'ordre neurocognitif, syntaxique et prosodique. On en conclut que la structuration prosodique est la première à intervenir dans le processus de génération de l'énoncé par le locuteur. Elle est aussi la première à être décodée par l'auditeur, lui permettant ensuite l'accès aux autres structures organisant les unités de l'énoncé, et en particulier la structure syntaxique.

Whether in read, silent reading or spontaneous speech, the prosodic structure is always present as an obligatory linguistic object in order to allow the listener to process the information brought by the flow of syllables and access the syntactic information contained in the sentence. The aim of this paper is to demonstrate that the elaboration of the prosodic structuration necessarily present in the sentence actually precedes the elaboration of the other structures and particularly of the syntactic structure, whether in the generation process by the speaker or the perception process by the listener.

Arguments favoring this conclusion are of various order and are based on the following facts:

- The prosodic structure can exist without any words or any syntax whereas the opposite is not true. Syntax depends on the presence of prosody, but prosody does not depend on the presence of syntax;

- The phonation process requires a flow of expiratory air flow which is segmented by phases of inspiration in the speaker respiration cycle;
- The flow of syllables must be segmented in chunks in order to be processed by Delta brain waves. Delta waves synchronize the transfer of sequences of syllables stored in short-term memory;
- In spontaneous speech, reformulations are always realized by retaking a complete stress group and not a selected word (except perhaps in stylistic applications);
- The dynamic process of the prosodic structure generation shows that the speaker has to choose between a relation of dependence (rection) or independence (paratax) between the actual prosodic group (ip or IP in autosegmental-metrical terminology). This is done by specifying prosodic contours indicating a dependency relation towards another contour to occur in the immediate future (i.e. to “the right”).

All these observations lead to a conclusion suggesting that the prosodic structure operates before the syntactic and the other structures of the sentence. The usual graphic representation and analysis of the prosodic structure obscures considerably this aspects, leading to believe that intonation acts as a supplement to syntax, to be processed by the listener (in reality only the reader) as another set of syntactic features.

INDEX

Mots-clés : structure prosodique, syntaxe, parole spontanée, reformulations, répétitions

Keywords : prosodic structure, syntax, spontaneous speech, reformulations, repetitions

AUTEUR

PHILIPPE MARTIN

LLF UMR 7110, UFRL, Université Paris Diderot Sorbonne Paris Cité

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