



HAL
open science

Oralisation de structures visuelles : de la lexico-syntaxe à la prosodie

Fabrice Maurel, Julie Lemarié, Nadine Vigouroux

► **To cite this version:**

Fabrice Maurel, Julie Lemarié, Nadine Vigouroux. Oralisation de structures visuelles : de la lexico-syntaxe à la prosodie. Interfaces Prosodiques (IP 2003), Mar 2003, Nantes, France. pp.137-142. hal-03627623

HAL Id: hal-03627623

<https://hal.science/hal-03627623v1>

Submitted on 4 Apr 2022

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.

Oralisation de structures visuelles : de la lexico-syntaxe à la prosodie

Fabrice Maurel^{*}, Julie Lemarié^{**} & Nadine Vigouroux^{***}

Abstract

In the general framework of a research about the aural transposition of text visual structures, we have an interest in the relationship between syntax and prosody. Indeed, after the study of a first solution consisting in the recomposition under a discursive form, of the information conveyed by the text visual structure, a second solution considered is to make appear a functional equivalence between the lexico-syntactical marks of this discursive form and specific prosodic patterns playing an "architectural" role too. This approach was tested for aural transposition of enumerative structures. The discursive form construction rests on a textual linguistic-based model (the "Modèle d'Architecture Textuelle"), whereas the elaboration of a form tapping prosody is based both on the M. Rossi's theoretical framework of pragmatic intonation and empirical data. Thus, the establishment of a relation between lexico-syntactical and prosodic marks permits us to propose a prosodic pattern for enumerative structures *a priori* equivalent to their discursive version.

1 Introduction

La complexité du domaine de recherche qui se consacre à la prosodie est en partie due à la difficulté de modéliser les relations entre intonation et syntaxe (congruence, indépendance, inclusion...) (Martin, 1979), (Lacheret-Dujour, Beaugendre, 1999). Cette problématique a été jusqu'alors abordée essentiellement au niveau de la phrase dans le parler spontané. Nous nous proposons d'étudier cette relation dans un cadre différent, celui de la transmodalité écrit/oral appliquée aux structures visuelles des textes.

Un texte ne peut être réduit à une suite linéaire de mots qui se constituent en phrases. En particulier, la nature même du support physique de présentation du texte implique à la fois des contraintes (ne pas sortir du cadre défini par ce support, choisir une police de caractère, ...) et des possibilités (en particulier par l'utilisation de procédés spécifiques tels que la mise en saillance typographique, la ponctuation, l'espacement horizontal ou vertical, ...). Ces procédés peuvent ainsi se constituer en configurations particulières, intentionnelles et signifiantes (les *Objets Textuels* : titres, exemples, insistances, ...) auxquelles les technologies actuelles de synthèses de parole à partir de textes ne savent pas (ou savent mal) réagir (Maurel, 2002a). L'objet de cet article est d'apporter des éléments de réponse à la problématique des relations entre syntaxe et intonation dans le cadre de la transposition automatique à l'oral des structures visuelles des textes.

2 Représentation de la structure visuelle des textes

Dans l'objectif de modéliser la structure visuelle des textes et afin d'envisager une représentation des documents écrits qui tiennent compte des « intentions architecturantes » du rédacteur, nous nous sommes appuyés sur le Modèle d'Architecture Textuelle (MAT), (Virbel, 1989), (Pascual, 1996). Celui-ci fait l'hypothèse d'une équivalence fonctionnelle entre différents types de marques « architecturantes » (dites de Mise en Forme Matérielle ou MFM), qu'elles soient de nature typographique, dispositionnelle ou lexico-syntaxique. Dans ce modèle, les propriétés morfo-dispositionnelles des textes peuvent être décrites à l'aide de phrases méta-textuelles, en ce sens que leurs référents sont des éléments du texte, et non le segment du monde dont le texte parle par ailleurs. Ainsi, cette approche a pour conséquence l'existence d'un continuum entre des formes entièrement discursives et des formes visuelles : on passe d'un pôle à l'autre de ce continuum en réduisant (grâce à des marques visuelles) ou en reconstruisant (par interprétation) le métalangage architectural.

L'objet de la section suivante est d'appliquer les principes du Modèle d'Architecture Textuelle à la problématique de la transmodalité écrit/oral des structures visuelles des textes, pour obtenir dans un premier temps une forme dite « discursive interprétative » (FDI). Nous construirons ensuite de nouvelles formes intégrant des marques prosodiques.

* IRIT (UMR CNRS 5505), Université Paul Sabatier (Toulouse III), France – fmaurel@irit.fr

** IRIT (UMR CNRS 5505), Université Paul Sabatier (Toulouse III), France – lemarie@irit.fr

*** IRIT (UMR CNRS 5505), Université Paul Sabatier (Toulouse III), France – vigourou@irit.fr

3 Principes de construction de formes "oralisables"

L'interprétation des marques typographiques et dispositionnelles des textes et leur reformulation systématique par des marques lexico-syntaxiques, permet de construire une forme discursivement développée du texte et de son « inscription spatiale ». Cette forme discursive interprétative, est par la suite considérée comme une forme commune à l'écrit et à l'oral et pourra, de ce fait, être directement oralisée ou bien transformée à nouveau, grâce à des procédés spécifiques à l'oral (Maurel, 2001). Dans le cadre de cet article nous nous intéresserons à l'exploitation de marques prosodiques. A l'instar de l'hypothèse d'équivalence fonctionnelle proposée par MAT (Cf 2.), nous stipulons une équivalence fonctionnelle entre certaines marques lexico-syntaxiques et des configurations prosodiques. Ainsi, la FDI peut être considérée comme le pivot central pour l'élaboration d'un continuum de nouvelles formes. Celles-ci peuvent être générées par réduction de tout ou partie des marques de MFM de nature lexico-syntaxique, au profit de marques prosodiques, équivalentes du point de vue de leur fonction. A titre d'illustration de l'élaboration de nos différentes formes, nous avons opté pour une étude en profondeur d'un objet particulier du texte : les structures énumératives.

4 Application aux structures énumératives

Nous définissons les différents Objets Textuels suivants (Luc, 2001) :

- (1) un item : c'est une entité co-énumérée ;
- (2) une amorce : c'est une phrase introductrice précédant et introduisant l'énumération. Une amorce peut être syntaxiquement complète ou incomplète ; dans ce dernier cas, le ou les constituants manquants sont fournis par un ou des items ;
- (3) une énumération : c'est l'ensemble des items ;
- (4) une structure énumérative : c'est l'ensemble amorce + énumération.

4.1 Forme discursive interprétative des structures énumératives

Le métalangage architectural relatif aux structures énumératives peut être instancié de manière à respecter les contraintes du *pattern* décrit dans le cadre ci-dessous (les "[...]" indiquent l'optionnalité, les "()" fournissent des exemples ou des précisions et la graisse met en saillance la catégorie grammaticale des différents éléments du métalangage utilisés) :

Amorce :

Article (défini, pluriel) + **numéral** (>1) + **nom classifieur** + [complément(s) de nom] + [proposition(s)] + **Groupe Verbe Spécifique** (e.g. : "sont énumérés" ou "sont énumérées") + **Adverbe Spécifique** (e.g. : "ci-après").

Énumération :

1^{er} item

Article (défini) + **organisateur textuel** ("premier" ou "première") + **nom classifieur** + **verbe introducteur** ("est", "comprend", "regroupe",...) + groupe nominal + [proposition(s)].

i^{ème} item (1 < i < n)

Article (défini) + **organisateur textuel** ("i^{ème}") + **nom classifieur** + **verbe introducteur** ("est", "comprend", "regroupe",...) + groupe nominal + [proposition(s)].

n^{ème} item

Article (défini) + **organisateur textuel** ("n^{ème}") + **marque fin item** ("et dernier" ou "et dernière") + **nom classifieur** + **verbe introducteur** ("est", "comprend", "regroupe",...) + groupe nominal + [proposition(s)].

Nous décrivons dans le tableau 1, à partir d'un exemple, comment MAT, par l'exploitation de la morpho-disposition des structures énumératives, offre un principe de reconstruction de la structure énumérative en unités phrastiques. Nous précisons que l'automatisation de cette étape, facilitée par une phase manuelle de normalisation typo-dispositionnelle des structures énumératives, est en cours de développement (Maurel, 2002b).

Version normalisée	Version discursive interprétative ¹
Les documents sont : <ul style="list-style-type: none"> - les 5 derniers avis d'imposition ; - le formulaire de demande dûment rempli ; - une fiche individuelle d'état civil ; - un relevé d'identité bancaire; - un certificat médical détaillé, descriptif de l'état de santé de la personne dépendante. 	Les cinq documents sont énumérés ci-après . Le premier document regroupe les 5 derniers avis d'imposition. Le second document est le formulaire de demande dûment rempli. Le troisième document est une fiche individuelle d'état civil. Le quatrième document est un relevé d'identité bancaire. Le cinquième et dernier document est un certificat médical détaillé, descriptif de l'état de santé de la personne dépendante.

Tableau 1 : versions "architecturée" et discursive d'une structure énumérative

L'étape suivante, décrite dans la sous-section subséquente, consiste à substituer certaines marques lexico-syntaxiques de cette forme discursive par des marques prosodiques *a priori* fonctionnellement équivalentes.

4.2 Forme prosodique des structures énumératives

Notre démarche pour constituer qualitativement ces marques est basée sur de l'introspection et des écoutes de discours de conférenciers (cf 4.2.6 et 4.2.7) ainsi que sur un cadre théorique de l'intonation pragmatique (Cf. 4.2.2 à 4.2.5). Quant à leur réalisation en terme de paramètres acoustico-prosodiques, nous nous sommes appuyés sur ce même cadre théorique et des écoutes critiques par des juges non naïfs.

La fonction représentative de la grammaire de l'intonation pragmatique du français telle qu'elle est définie dans (Rossi, 1999) permet, à l'instar des marques de MFM à l'écrit, de hiérarchiser l'information du message, c'est-à-dire de "déterminer la relation du locuteur avec ce qu'il dit en organisant le contenu du message en information d'arrière-plan et de premier plan". L'exploitation des propriétés de cette grammaire intonative permet de formaliser des marques prosodiques dont la signification participerait à la structure hiérarchique du texte oralisé. Nous décrivons dans les sous-sections suivantes, à partir de l'exemple du tableau 1, la manière dont l'exploitation des marques de MFM de nature lexico-syntaxique, permet de construire une nouvelle forme énumérative, intégrant des marques prosodiques.

4.2.1 Substitution du numéral et du classifieur

Le numéral ne sera pas substitué en raison de l'information sémantique qu'il porte en sus de sa fonction architecturante. Le fait de conserver le numéral implique la conservation du classifieur pour des raisons syntaxiques (de plus, le classifieur, à l'instar du numéral, est porteur d'une information sémantique lorsqu'il n'est pas générique).

4.2.2 Substitution de « énuméré(e)s »

Le groupe verbal « sont énumérées » annonce de manière anticipée dans le *pattern* choisi :

- par sa construction syntaxique, la partie de l'amorce qui est concernée directement par l'énumération ;
- par sa morphologie et sa classe sémantique, la pluralité de ces informations (le participe passé du verbe « énumérer » est très rare au singulier) et la co-énumérabilité des hyponymes du classifieur.

La conservation de « sont » permet de conserver le rôle indiqué par le premier point énuméré ci-dessus. L'accent de focalisation (AF) sera utilisé sur le numéral de l'amorce pour insister sur la pluralité et la co-énumérabilité.

Nous relevons dans (Rossi, 1999:122) 4 possibilités d'utilisation de AF :

1. Si AF est amalgamé avec l'accent lexical sous-jacent (**as**) sa portée est circonscrite au mot domaine de l'accent lexical et a valeur de proéminence sémantique ; l'intonème continuatif qui clôt le domaine est inversé ;
2. Si AF est initial :

¹ La partie en gras représente le texte ajouté à la version originale.

- a. et si l'intonème continuatif qui clôt le domaine est inversé alors **AF** a une valeur contrastive,
- b. et si l'intonème continuatif qui clôt le domaine n'est pas inversé, alors **AF** a une valeur focalisatrice pleine ; dans ce cas la portée focalisatrice de **AF** est le domaine circonscrit par **AF** et le continuatif. La portée de **AF** se confond alors avec son domaine et détermine un focus étendu. Ce domaine est inclus dans un arc prosodique. Le contenu de l'arc prosodique est intégralement focalisé,
- c. et si l'intonème qui clôt le domaine de **AF** est celui du morphème conclusif majeur (**CC**), alors la portée de **AF** est circonscrite au mot porteur de **AF**.

Le cas 2.c. ne correspond pas à notre *pattern* car il semble peu naturel d'envisager un conclusif pour clore le domaine de **AF**.

Le cas 1. n'est pas approprié pour un numéral dans notre cadre car il sous-entendrait qu'il y a des numéraux qui peuvent être très importants et d'autres non.

Le cas 2.a. mettrait en contraste le numéral en reléguant à l'arrière plan le nom classifieur ce qui engendrerait potentiellement un conflit avec l'objectif d'anticipation de la co-énumérabilité des hyponymes du classifieur.

Le cas 2.b. est intéressant si l'arc prosodique est réalisé sur un domaine de **AF** comportant numéral + classifieur. Ce cas paraît être le plus adapté puisqu'il permettrait, par sa valeur focalisatrice pleine, d'insister sur la pluralité des informations liées au classifieur et sur le classifieur lui-même (et donc globalement sur la co-énumérabilité de ses hyponymes).

4.2.3 Substitution de « ci-après »

(Rossi, 1999:56) propose une nouvelle définition du topique : "constituant de tête qui acquiert une valeur contextuelle référentielle ou inférentielle à laquelle s'ajoute, comme une conséquence nécessaire de sa valeur première et de sa place dans l'énoncé, une valeur thématique (ce dont on parle)".

Dans ce cadre, le SN (Syntagme Nominal) de l'amorce peut être considéré comme un topique et "ci-après" montre que le locuteur va parler de SN de manière imminente. L'effacement de "ci-après" peut ainsi donner le statut d'opération de topicalisation (extraction en tête du topique) à l'opération de mise en amorce incomplète ; ainsi, l'effacement de "ci-après" dans la structure énumérative peut être intonativement contrebalancé par l'utilisation d'un morphème intonatif de topique contextuel référentiel (noté **CTr**) à la fin du SN de l'amorce : dans ce cas, le SN est chargé d'une valeur contextuelle référentielle, il "montre du doigt" ce dont on va parler et l'allocutaire sait qu'il prend connaissance du référent *juste* avant de recevoir l'information.

4.2.4 Substitution de l'information de début d'item portée par les marques lexico-syntaxiques d'organisateur textuel et de reprise du classifieur

Nous considérerons qu'il existe une relation entre la hiérarchie des objets du texte et la stratégie pausale utilisée pour en rendre compte. Autrement dit, en déclinant l'objet du texte *Structure Enumérative* comme composé de deux objets (*Amorce* et *Énumération*), l'énumération étant elle-même composée d'objets appelés *Items*, nous utiliserons une durée de **pause** d'autant plus courte que l'objet considéré est bas dans cette "architecture" (1 UP² par niveau). Par exemple à une durée de 800 ms entre deux paragraphes (de même niveau que la structure énumérative elle-même considérée comme un paragraphe) correspondront une durée de 670 ms entre l'amorce et l'énumération, une durée de 560 ms entre chaque item et, le cas échéant, 460 ms entre les phrases qui composent l'item.

4.2.5 Substitution de la reprise du classifieur

La reprise du classifieur, dans chaque item, a également pour rôle de souligner l'importance informative de l'hyponyme auquel il réfère. Ainsi, les groupes hyponymes (syntagmes contenant les hyponymes du classifieur) des items de l'énumération peuvent être considérés comme les noyaux rhématiques des items, c'est à dire véhiculant l'information nouvelle. Or (Rossi, 1999:106-107) indique que "dans une langue comme le français, l'identification de l'information nouvelle est fondamentalement opérée par le morphème intonatif conclusif **CC** : le rhème (à distinguer du focus) est constitué des éléments syntaxiques qui demeurent devant le morphème intonatif conclusif majeur **CC**, lorsqu'on a extrait toutes les dislocations représentées par le topique et les thèmes".

² UP : Unité de Perception – une unité de perception est égale à une fois le seuil de perceptibilité (ou seuil différentiel) du paramètre considéré (fréquence, intensité, temps).

Parmi les 3 modes de présentation de l'information nouvelle identifiés par (Rossi, 1999:112-113), nous pouvons retenir celui de l'*identification* : "lorsque le prédicat verbal est extraposé dans les structures RT [Rhème/Thème] les arguments ou les adjoints devant CC sont présentés non plus par simple indication, ils sont identifiés hors de l'acte de prédication. L'identification est un procédé de reconnaissance à la fois par indication et par contraste, mais où domine l'indication. L'identification est un mode d'indication marqué". Ainsi, le morphème CC à la fin du noyau rhématique de l'item aurait le double rôle de créer une dislocation intonative en thématissant le cas échéant le reste de l'item, et d'identifier les arguments du prédicat extraposé par sa mise en facteur dans l'amorce ("sont"). Lorsque la fin de l'item n'est pas congruente à la fin du noyau rhématique, l'item sera terminé par morphème conclusif mineur (cc) donnant ainsi à la suite de l'item le statut d'une sorte de parenthétique (de thème externe).

4.2.6 Substitution des organisateurs textuels

Les organisateurs textuels semblent indiquer de manière absolue la position de l'item dans l'énumération et à ce titre, ils paraissent *a priori* non effaçables au profit d'une marque prosodique fonctionnellement équivalente. Malgré tout, le nombre exact d'items étant explicité par la conservation du numéral dans l'amorce, une combinaison de marques prosodiques indiquant le positionnement relatif des items serait peut-être suffisante pour se repérer dans la structure (est-on plutôt en début, au milieu ou proche de la fin de l'énumération ?). La marque prosodique pourrait être une déclinaison intonative régulière lors de chaque transition vers un nouvel item ; le dernier item étant prononcé sur la ligne mélodique de base du locuteur ; nous exploitons ici le changement de registre du locuteur à chaque nouveau départ de *downdrift*³, au début des items (figure 1 : 3 premiers items). Cette stratégie pourrait être accompagnée d'une "prise d'inspiration" en début d'énumération, dont la force et la durée seraient fonction de la longueur de l'énumération.

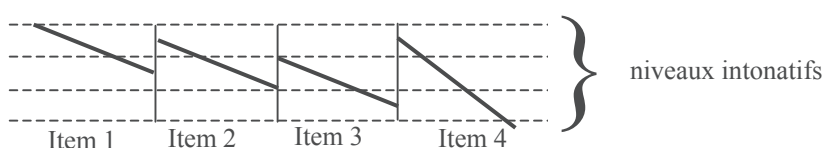


Figure 1. Changement de registre à chaque déclinaison intonative (3 premiers items pour une structure énumérative de 4 items) et augmentation de la pente de la déclinaison intonative du dernier item

4.2.7 Substitution de la marque de dernier item

Les marques prosodiques pourraient être un morphème conclusif majeur CC à la fin de l'item et un "étirement" de la pente de fréquence fondamentale de manière à ce que la fréquence initiale soit réalisée plus haut dans le registre du locuteur mais que la fréquence finale reste celle du conclusif "normal" du locuteur (figure 1 : 4^{ème} item).

5 Conclusion et perspectives

Dans le cadre que nous nous sommes fixés, celui de la transmodalité écrit/oral appliquée aux structures visuelles des textes, notre étude montre qu'il est *a priori* possible d'envisager une relation entre des aspects lexico-syntaxiques de la phrase et des marques prosodiques sous l'angle d'une équivalence fonctionnelle. Ainsi, certains des paramètres prosodiques du discours proviendraient non pas de la construction syntaxique qui le sous-tend ou d'influences sémantico-pragmatiques, mais au contraire de l'effacement de marques lexico-syntaxiques spécifiques dont ils seraient la trace. Ce travail est logiquement poursuivi actuellement par une tentative de validation psycholinguistique. Ainsi, dans l'objectif de statuer sur le compromis oral pertinent (niveau d'explicitation, longueur minimale du message, anticipation des intentions de l'auditeur), afin de favoriser la compréhension et la mémorisation de l'information (à l'instar du rôle joué par l'inscription spatiale des textes à l'écrit), notre prochaine étape consistera à évaluer l'impact de différentes formes orales des structures énumératives sur le traitement cognitif (compréhension, mémorisation) pour deux types de population d'utilisateurs (jeunes versus âgés).

³ downdrift : phénomène physiologique d'abaissement mélodique progressif observé sur l'énoncé

6 Remerciements

Ce travail est partiellement soutenu par le Conseil Régional de Midi-Pyrénées (dans le cadre de la Convention DAER-99008531) et le programme interdisciplinaire, Société de l'Information du CNRS et de l'ACI Cognitive.

Les auteurs remercient Jean-Luc Nespoulous, Jacques Virbel et Mustapha Mojahid pour la relecture de ce travail et les améliorations suggérées.

7 Références

- Lacheret-Dujour A., Beaugendre F., 1999, "La prosodie du français", Paris, CNRS Editions.
- Luc C., 2001 "Une typologie des structures énumératives basée sur les structures rhétoriques et architecturales du texte", in Actes de TALN'2001, p. 263-272.
- Martin Ph., 1979, "Sur les principes d'une théorie syntaxique de l'intonation", in P. Léon et M. Rossi (1979 éd.), p. 91-101.
- Maurel F., Vigouroux N. & Nespoulous J.-L., 2001, "Problématique, enjeux et perspectives de la présentation orale de documents électroniques", in Actes de CIDE 01, Europia., p. 247-263.
- Maurel F., Luc C., Mojahid M., Nespoulous J.-L., Vigouroux N. & Virbel J., 2002a, "De l'influence de la variabilité des textes sur leur transposition automatique à l'oral", in Actes de CIDE'2002, Le Chesnay , INRIA, p. 211-225.
- Maurel F., Luc C., Mojahid M., Nespoulous J.-L., Vigouroux N., Virbel J., 2002b, "Etude des structures visuelles pour leur présentation oralisée", in Actes de DVP 2002, Brest, ENST Bretagne, p. 11-24.
- Pascual E., 1996, "Integrating Text Formatting and Text Generation", in Trends in Natural Language Generation : an Artificial Intelligence Perspective. Réds : Adorni M. et Zock M., Eds: Springer-Verlag , p. 205-221.
- Rossi M., 1999, "L'intonation, le système du français : description et modélisation", Paris, Ophrys.
- Virbel J., 1989, "The Contribution of linguistic knowledge to the interpretation of text Structure", in J. André, V. Quint et R. Futura eds., Structured Documents, Cambridge University Press, p. 161-181.