

La topologie comme interface entre syntaxe et prosodie : un système de génération appliqué au grec moderne

Kim Gerdes (1), Hi-Yon Yoo (2)

(1) Lattice – Université Paris 7
2, place Jussieu 75251 Paris Cedex 05
kim@linguist.jussieu.fr

(2) ARP – Université Paris 7
2, place Jussieu 75251 Paris Cedex 05
hi-yon.yoo@linguist.jussieu.fr

Résumé – Abstract

Dans cet article, nous développons les modules syntaxique et topologique du modèle Sens-Texte et nous montrons l'utilité de la topologie comme représentation intermédiaire entre les représentations syntaxique et phonologique. Le modèle est implémenté dans un générateur et nous présentons la grammaire du grec moderne dans cette approche.

In this paper, we develop the syntactic and topological modules of the Meaning-Text model. We show the use of topology as the interface of syntax and phonology. The model is implemented in a speech generator and we present its Greek grammar.

Keywords – Mots Clés

grammaires de dépendance, ordre des mots, prosodie, topologie, générateur de parole, grec
dependency grammar, word order, prosody, topology, language generation, Greek

1 Introduction

Ce papier propose une modélisation du lien entre syntaxe et prosodie. Nous considérons la représentation syntaxique comme l'un des niveaux intermédiaires entre la forme logique (le sens) et la forme phonétique (le son) et nous nous plaçons dans un modèle linéaire de la langue où la synthèse (et l'analyse) se fait par une série de correspondances entre des niveaux de représentations adjacents. De telles approches linéaires se sont avérées avantageuses pour le traitement automatique des langues, notamment la traduction automatique, depuis ses débuts jusqu'aux projets les plus récents (par exemple VerbMobil, Wahlster 2000). En effet, des modèles peu articulés comme les modèles chomskyens, qui donnent une place prépondérante à une seule structure qui encode à la fois des informations de types sémantique,

syntactique et l'ordre des mots, sont peu aptes à un traitement informatique. La considération d'une chaîne de niveaux intermédiaires permet d'élaborer des règles de correspondance plus simples et plus structurées.

Notre travail se place dans le cadre de la Théorie Sens-Texte (TST), qui propose un modèle très élaboré d'une analyse linguistique modulaire et linéaire. Néanmoins, dans ce cadre, la place de la prosodie n'a pas été clairement définie, la plupart des travaux portant sur les correspondances entre sémantique et syntaxe ou sur l'élaboration de la morphologie (cf. Mel'čuk 1988).

Dans la suite, nous présentons très brièvement les niveaux classiques de la TST ainsi que les aménagements que nous proposons. La Section 2 introduit la correspondance entre syntaxe et topologie de manière générale et spécifique au grec moderne. De même, la Section 3 présente la correspondance entre topologie et phonologie et l'instanciation pour le grec. Avant de conclure, nous indiquons très brièvement comment nous produisons le son à partir de la structure phonologique dans notre générateur DepLin.

1.1 Aménagement de la TST entre syntaxe et surface

La TST classique considère trois niveaux intermédiaires entre la structure sémantique et la structure phonétique : Syntaxe, Morphologie et Phonologie, chaque niveau étant sous-divisé en un niveau profond (en direction du sens) et un niveau de surface (en direction de la sortie phonétique). Dans notre approche, nous gardons les niveaux sémantique et syntaxique. L'arbre syntaxique de surface représente notre point de départ (dans le sens de la synthèse, la direction préférée de la TST, bien que les règles soient bidirectionnelles). Les représentations syntaxiques de la TST sont les arbres de dépendance syntaxiques. A partir de cet arbre de dépendance, nous calculons une représentation de la phrase linéairement ordonnée. Nous suivons Gerdes & Kahane 2001 en considérant une représentation morphologique profonde augmentée d'un arbre de constituants, appelé structure topologique. C'est à partir de cette structure que se fait le calcul de l'étape suivante : le niveau phonologique, constitué d'une chaîne linéaire de phonèmes dotée d'une structure prosodique arborescente. Le niveau phonologique permettra par la suite le calcul final de la sortie phonétique.

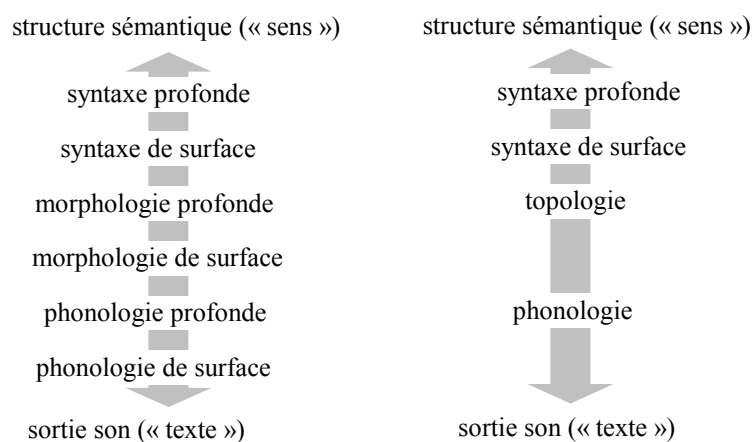


Figure 1 : Niveaux de la TST classique (gauche) et aménagement pour la prosodie (droite)

Nous ne pouvons exclure l'utilité d'une sous-division plus fine des étapes entre syntaxe et sortie son, comme le propose la TST classique, pour le traitement d'autres langues que celles testées avec notre générateur (allemand, coréen, français et grec moderne) ou pour des analyses plus fines que celles faites par nos grammaires tests. L'implémentation modulaire de DepLin (en Java) et la possibilité de sortie ou entrée en fichier XML à chaque niveau rendent facile d'affiner nos étapes quand l'analyse linguistique l'exigera.

2 Interface Syntaxe-Topologie

Notons d'abord que notre structure de départ, l'arbre de dépendance syntaxique de surface, contient un nœud pour chaque mot de la phrase, et comme il est de coutume dans les grammaires de dépendance, le verbe fini est le nœud racine de l'arbre de dépendance d'une phrase, qu'il s'agisse d'un auxiliaire, d'un verbe modal ou d'un verbe plein. Dans la TST, les nœuds de cet arbre sont augmentés d'un marquage indiquant le groupement communicatif (Mel'čuk 2001, on trouve également le terme « structure informationnelle », cf. Lambrecht 1994). Mel'čuk 2001, dans une tradition pragoise, propose un découpage principal en thème et en rhème, la notion de thème renvoyant à « ce dont on parle », et celle de rhème à « ce qu'on en dit ». Nous adoptons une répartition tripartite (Valduvi 1992, voir aussi Choi 1999) en topique, focus et tail, où le focus correspondrait au rhème, le topique au thème proéminent et le tail en thème non proéminent. Un tel arbre de dépendance diffère d'un arbre X-barre, qui contient outre l'information de dépendance entre les mots (la tête d'un syntagme X-barre gouverne syntaxiquement les autres membres du syntagme) l'information sur le regroupement des mots et sur leur ordre linéaire. Dans les grammaires de dépendance, on considère que ces différents types d'information doivent être exprimés dans des structures certes liées mais séparées. Pourtant, le calcul de l'ordre des mots à partir de l'arbre de dépendance peut s'avérer difficile pour des langues comme l'allemand où l'ordonnancement des mots par rapport à l'arbre de dépendance n'est pas projectif ; autrement dit, les mots se linéarisent pas forcément par rapport à leur gouverneur. Et pourtant, « *la position des mots dans la phrase allemande n'est pas **relative**, comme on l'enseigne trop souvent avec la théorie de l'**inversion**, ce qui brouille tout, mais bien **absolue**, puisque le verbe ne se place pas **après** tel ou tel autre mot, mais bien, selon la nature de la phrase, à une place fixe et déterminée en elle-même* » (Tesnière 1959:130).

Les analyses basées sur la notion de « place » repose sur des travaux traditionnels en grammaire germanique (commençant par le système Herling-Erdmann du 19^e siècle). Pour le processus de linéarisation d'un arbre de dépendance allemand, il est également utile de donner un statut formel à cette idée de place fixe. Des premiers pas dans cette direction ont été faits par Bröker 1998. Duchier & Debusmann 2001 proposent un modèle plus complet de linéarisation dans un contexte des grammaires à contraintes, et Gerdes & Kahane 2001 ont développé un formalisme dans le cadre de la TST. C'est cette dernière approche que nous adoptons pour notre analyse. Dans cette approche, on construit une structure hiérarchique de constituants topologiques, qui s'intègrent dans le processus de linéarisation de la TST. La structure topologique résultante représente un niveau intermédiaire entre la syntaxe et le niveau phonologique : la considération de la structure topologique facilite d'une part les règles de linéarisation, et d'autre part, le calcul de la structure prosodique du niveau de représentation suivant.

Les constituants de l'arbre topologique sont appelés *domaines*. Un domaine est une série linéaire de places fixes, qu'on appelle *champs*. Chaque champ indique ses propres restrictions de contenance, i.e. le nombre de constituants (domaines) qu'il peut ou doit accueillir pour que la structure topologique soit bien-formée. Il est important de noter que les champs ne sont pas des constituants : deux constituants qui partagent le même champ ne forment pas obligatoirement une unité topologique. Etant donné notre point de départ qui est l'arbre de dépendance de surface, la linéarisation commence par placer la racine de ce dernier dans le champ initial. Les *règles de construction de domaine* permettent à un élément d'ouvrir un domaine autour de lui et de se placer dans un champ spécifié de ce domaine. Une telle création du domaine peut être contrainte par la catégorie et la valeur communicative de l'élément ainsi que le champ dans lequel il se trouve. Les *règles de correspondance* indiquent comment un dépendant se place par rapport à la position topologique de son gouverneur : le positionnement du dépendant est contraint par le champ du gouverneur, la relation syntaxique qu'il entretient avec son gouverneur et encore par la catégorie et la valeur communicative de ce dernier. Normalement, le dépendant se place à l'intérieur du domaine de son gouverneur, mais il n'est pas exclu qu'il soit placé dans un domaine supérieur sous restriction du type de domaine dont il « s'émancipe ».

Un tel formalisme peut être instancié pour des langues de types très variés, allant du français ou de l'anglais où le placement des mots est presque exclusivement déterminé par les relations syntaxiques qu'ils entretiennent entre eux, jusqu'aux langues dites à ordre de mots libre comme le russe ou le grec moderne, où les valeurs communicatives prennent le devant sur les relations syntaxiques. L'ordre des mots en allemand est contraint par les deux facteurs, puisque certains éléments comme le verbe fini se voient attribuer une place fixe tandis que d'autres peuvent apparaître dans des positions très variées en fonction de leurs valeurs communicatives.¹ Pour la grammaire topologique de l'allemand nous renvoyons à Gerdes & Kahane 2001. Dans la section suivante, nous développons la grammaire topologique du grec moderne qui, différemment du français, se base principalement sur la valeur communicative des éléments.

2.1 La topologie du grec moderne

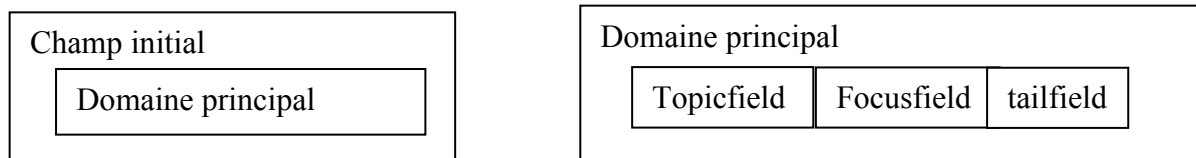
Nous illustrons dans cette partie la correspondance syntaxe-topologie par une analyse de l'ordre des mots en grec moderne, présentée dans les grammaires comme une langue « à ordre libre ». La justification que l'on trouve pour ce fait est que, par exemple, pour une phrase à deux arguments nominaux (SVO), les six permutations sont attestées (SVO, SOV, OSV, OVS, VSO, VOS). Or si l'on prend en considération les possibilités de variation prosodique pour un ordre donné, on s'aperçoit que le nombre de phrases attestées pour une structure argumentale augmente considérablement. Pour une phrase à deux arguments, on a à faire non plus à six phrases possibles, mais à 25 réalisations possibles, combinant les variations d'ordre

¹ Si la notion de topologie a été développée pour les langues germaniques, c'est que l'analyse de leur ordre des mots requiert la prise en compte des contraintes à la fois communicatives et syntaxiques ; le français écrit par exemple, s'adapte plutôt bien à l'analyse à une analyse X-barre, car les contraintes communicatives sont secondaires. Une structure topologique du français ressemblerait donc beaucoup à une structure X-barre avec la différence notable que la structure topologique ne représente pas explicitement les dépendances entre les éléments. Un calcul est nécessaire pour les obtenir.

des mots et les différentes réalisations prosodiques. Dire donc qu'une langue est « à ordre libre des mots », revient finalement à considérer que pour une structure argumentale donnée, beaucoup d'ordres des mots, mais pas tous, sont possibles, et ce avec différentes réalisations prosodiques. Néanmoins, une langue comme le grec moderne a un comportement bien différent de l'allemand car le placement des éléments de phrase est encore davantage régi par des contraintes d'ordre sémantico-communicatives que syntaxiques².

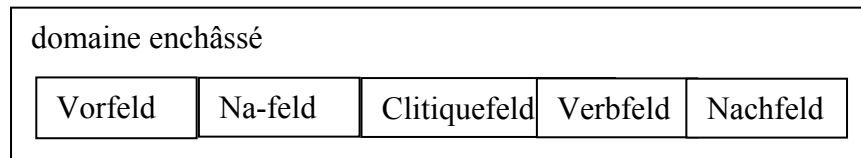
2.2 Les règles topologique du grec moderne

La structure topologique du grec moderne consiste en un champ initial qui ouvre un domaine principal. Le domaine principal consiste en trois champs, nommés à partir du marquage communicatif, compte tenu du fait que les éléments ayant le même marquage, seront susceptibles d'être placés dans ces champs. Dans l'ordre, nous avons donc le TopF (champ de topique), le FocF (champ du focus) et le tailF (champ du tail). Nous indiquons dans les règles de description de champs sur le fait que le FocF contient exactement un élément (un domaine enchâssé), tandis que le TopF et le tailF n'ont pas de restrictions quant au nombre d'éléments qu'ils peuvent accueillir.



La tête verbale se place dans l'un des champs majeurs en fonction de sa valeur communicative : s'i elle a un marquage communicatif de focus, elle sera placé dans le champ focus et ainsi de suite. S'appliquent ensuite les règles de création de domaines enchâssés. La tête verbale de la dépendance syntaxique ouvre un domaine enchâssé dans le champ majeur où il a été placé.

Le domaine enchâssé comporte 5 champs : celui de Vorfeld (qui peut contenir un nombre quelconque d'éléments), le Na-feld (la place pour le marquage de certaines formes verbales comme le futur et les verbes en *na* ; il contient zéro ou un élément), le Clitiquefeld (la place des clitiques³), le Verbfeld (qui ne contient que le verbe) et le Nachfeld (pour d'autres dépendants du verbe, sans restriction du nombre d'éléments qu'il peut contenir).



La tête verbale ouvre donc ce domaine enchâssé et se place dans le champ du verbe. Ensuite, les dépendants verbaux peuvent se placer. Si les dépendants partagent le même marquage

² Les contraintes syntaxiques peuvent également intervenir. Contrairement à l'allemand où nous avons une place syntaxiquement déterminée pour le verbe principal de la phrase (V2), le grec moderne semble plus « libre » dans le sens où c'est essentiellement le marquage communicatif porté par les éléments et hérité du niveau sémantique qui détermine leur place linéaire.

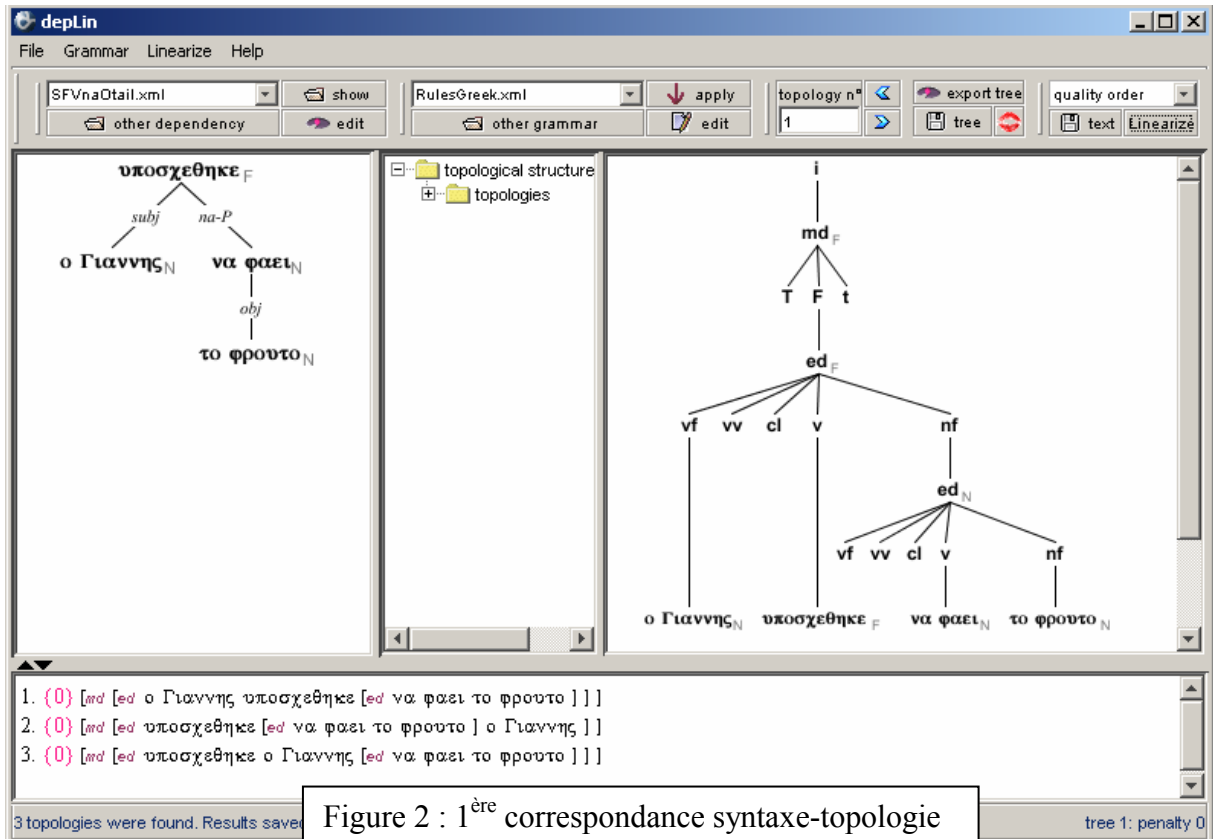
³ Comme en français, l'ordre de placement des clitiques n'est pas libre, et pour un traitement complet, il faudra envisager une séparation plus fine de ce champ.

communicatif, ils restent à l'intérieur du même domaine et se placent selon le rôle qu'ils jouent dans la dépendance syntaxique ou leur nature grammaticale : ainsi, un clitique se place dans le Clitiquefeld, le complément sujet dans le Vorfeld ou le Nachfeld, le complément objet direct dans le Nachfeld etc. Nous supposons que la structure communicative est une partition, c'est-à-dire que chaque élément appartient à un et un seul groupement communicatif. Afin de représenter à l'aide de traits simples la superposition d'une telle structure communicative sur l'arbre de dépendance syntaxique, nous introduisons la convention suivante : seul le gouverneur syntaxique d'un groupe porte la valeur communicative du groupe, ses dépendants étant marqués N comme « neutre ». Autrement dit, chaque élément qui porte une valeur autre que N ouvre un autre groupement communicatif. Par contre si un dépendant a une valeur communicative différente de son gouverneur, les règles de correspondance l'obligent généralement à s'émanciper dans le champ correspondant à sa valeur communicative. Seuls l'objet direct et les dépendants verbaux connaissent des restrictions interdisant l'émancipation dans un champ correspondant à leur valeur communicative.

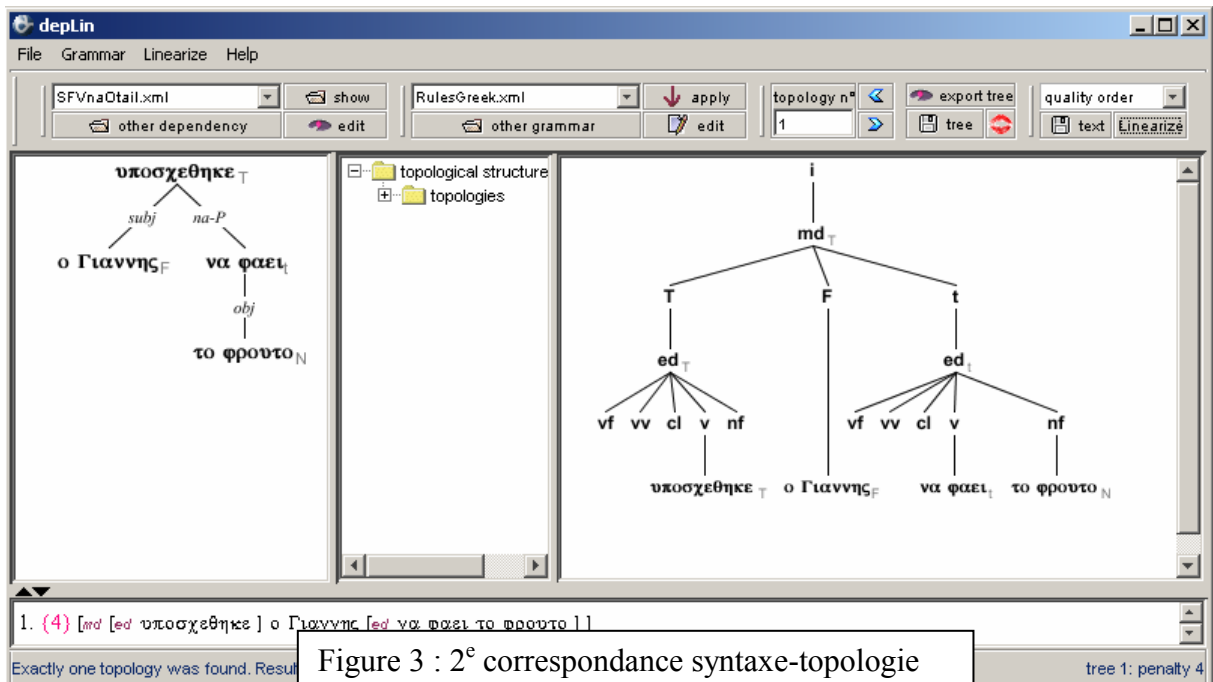
Un dépendant verbal ouvre toujours un domaine enchâssé indépendamment du champ où il a été placé. Si le dépendant verbal est un complémenteur, il ouvre un domaine pour les complétives, qui contient un champ CF pour lui-même et un deuxième champ pour le verbe régi. Ce deuxième champ diffère en fonction de l'endroit où le complémenteur a ouvert son domaine : dans le TopF, le domaine complétif consiste en CF et TopF, dans les tailF de CF et tailF et dans le Nachfeld, de CF et Nachfeld. Seul quand le complémenteur se trouve dans le focF, le domaine qu'il crée est plus complexe : il contient un CF suivi des trois champs majeurs : TopF, FocF et tailF. Les dépendants d'un gouverneur qui se trouve dans le domaine complétif doivent obligatoirement se placer à l'intérieur de ce domaine, et ne peuvent donc pas s'émanciper.

Suivant les règles décrites ci-dessus, nous obtenons à partir de l'arbre de dépendance syntaxique de surface augmenté d'un marquage communicatif à gauche dans la Figure 2, les linéarisations (en bas, un des arbres topologique s'affiche à droite). L'arbre de dépendance syntaxique représente un verbe ayant un dépendant nominal sujet et un dépendant verbal (verbe en *-na*). La tête verbale est marquée comme focus, et les dépendants, qui forment un groupe communicatif avec leur gouverneur, sont marqués neutre. Nous obtenons, après applications des règles de linéarisation pour le grec, les trois linéarisations suivantes que nous présentons dans la capture d'écran ci-dessus.

Dans la première linéarisation, la tête verbal *υποσχεθηκε* 'a promis' se place dans le FocF et ouvre un domaine offrant des places pour ses dépendants. A cause de leur marquage communicatif N, tous les dépendants restent à l'intérieur du domaine et il n'y a pas d'émancipation. Ne pouvant se placer dans le Vorfeld, le dépendant verbal *να φαι* 'de manger' ouvre un nouveau domaine dans le Nachfeld : *να* se place dans le Nachfeld et le verbe dans le VerbF. Le dépendant nominal de *να φαι* 'de manger' *το φρουτο* 'le fruit' n'aura pas d'autres choix que de se placer dans le Nachfeld de son gouverneur car il a le même marquage communicatif (donc pas d'émancipation) et qu'il ne peut pas aller dans le Vorfeld (car il est l'objet direct du verbe). Quant au dépendant sujet du verbe principal, il peut se placer dans le Vorfeld (linéarisation a.), ou dans le Nachfeld, avant le dépendant verbal (linéarisation c.) ou après le dépendant verbal (linéarisation a.).



Si l'arbre de dépendance a un autre marquage communicatif, par exemple, avec un sujet marqué comme focus, le verbe fini comme topique, et le na-verbe comme tail, alors les linéarisations obtenues ne sont pas les mêmes. Dans le cas de notre exemple, une seule linéarisation est possible :



La tête verbale qui a la valeur communicative de topique se place dans le TopF et ouvre un domaine. Cependant, ce domaine ne sera rempli que par le seul verbe car ses deux dépendants

vont s'émanciper : le dépendant nominal qui est marqué comme focus s'émancipe et se place dans le champs FocF tandis que le dépendant verbal ouvre un domaine dans le tailF (car il est marqué communicativement comme tail) et son dépendant nominal qui est neutre se place dans le Nachfeld. Figure 3 illustre la structure topologique obtenue :

3 Interface Topologie-Prosodie

A l'étape suivante de la production langagière, se construit la structure phonologique. Nous considérons que cette structure phonologique est un arbre à trois niveaux, ces derniers étant nécessaires pour déterminer la majorité des énoncés (nous écartons dans ce travail les phrases dites « spontanées » ou impliquant la prise en compte d'une émotion) :

- le niveau du mot prosodique : il s'agit de l'unité prosodique minimale qui comporte obligatoirement un et un seul mot accentué⁴. Il contient l'information lexicale, notamment la place de l'accent de mot. Par ailleurs, ce niveau est important car c'est autour de cet accent de mot que va se dessiner la courbe intonative. A ce niveau, nous retrouvons une transcription en phonèmes.
- le niveau du *groupe prosodique* : le *groupe prosodique* contient plusieurs *mots prosodiques*. Pour beaucoup de chercheurs, ce niveau est déterminé à partir de règles syntaxiques empiriques. Dans notre modèle, les groupes prosodiques coïncident avec un certain découpage communicatif. Les patrons intonatifs sont associés directement à ces groupes prosodiques.
- le niveau de la *phrase prosodique* : ce niveau indique la courbe mélodique phrastique, dépendant de la modalité de la phrase (déclarative, interrogative etc.). Ainsi, pour les phrases déclaratives que nous traitons ici, la courbe phrastique aura une allure descendante. La *phrase prosodique* regroupe une chaîne de *groupes prosodiques*.

Les règles de correspondance entre les structures topologique et phonologique sont relativement simples, car il s'agit d'une simplification de l'arbre de topologiques. La construction de l'arbre phonologique à partir de l'arbre topologique procède en deux étapes :

- Dans un premier temps, on applique les *règles des mots prosodiques* aux feuilles de l'arbre topologique. En fonction du domaine et du champ dans lequel se trouve la feuille ainsi que de l'information lexicale sur l'accentuabilité du mot, le mot va former ou non un nouveau *mot prosodique*. Dans le cas où le mot échoue dans cette tâche, la règle indique si le mot est adjacent à gauche ou à droite (i.e. il rejoint le mot prosodique qui précède ou qui suit à sa droite. Dans le cas grec, il s'agit d'une adjonction du mot à sa droite, ce qui ne serait pas le cas pour une langue comme le coréen). L'application de ces règles permet de combiner tous les mots non accentués avec un mot accentué et chaque mot, suivant la règle qu'un mot prosodique comporte une et une seule syllabe accentuée.

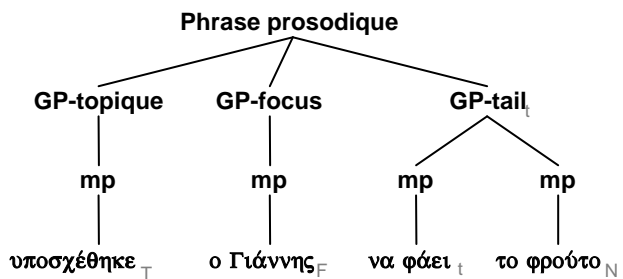
⁴ On considère qu'un mot prosodique peut contenir plusieurs mots du moment qu'il n'y a qu'un seul accent, mais qu'un mot ne peut s'étaler sur plusieurs mots prosodiques.

- Dans un deuxième temps, on applique les *règles de groupes prosodiques* aux domaines de la structure topologique. Une règle de type (D₁, C₁, F₁, D₂, C₂, GP) exprime qu'un groupe prosodique GP est construit par la contraction d'un domaine D₁ (portant la valeur communicative C₁) et du domaine D₂ (portant la valeur communicative C₂ et dominé par D₁ via le champ F). Tous les domaines qui ne sont pas contractés deviennent un groupe prosodique à part entière.

Notons que les règles de groupes prosodiques ne permettent pas de créer de nouveaux groupes prosodiques là où il n'y a pas de domaines ; on suppose donc que deux feuilles qui sont sœurs dans le même domaine font toujours partie d'un même groupe prosodique.

3.1 La correspondance topologie-phonologie en grec

Les règles des mots prosodiques du grec sont guidées par le lexique. Sachant qu'en grec moderne pratiquement tous les mots sont accentués, en principe sur l'une des trois dernières syllabes du mot (sauf les articles, les conjonctions etc.), le mot prosodique comprend le mot accentué accompagné, s'il y en a, des mots non accentués. En fait, chaque mot qui se trouve dans un *champ* de type *déterminant* ou *complémenteur* est adjacent à gauche, et se combine avec le mot ayant un accent lexical pour former un mot prosodique. Les règles des groupes prosodiques vont forcer la contraction des domaines D₁ et D₂ qui portent la même valeur communicative et où D₁ domine directement D₂. Ainsi, pour la structure topologique dans la Figure 3, nous obtenons après application de ces règles, la structure phonologique ci-contre.



4 La sortie sonore à partir de la structure phonologique

La structure phonologique contient toutes les informations nécessaires pour la synthétisation de la phrase. Il nous reste à appliquer l'information sur le type de la phrase prosodique aux contours prosodiques de chaque groupe prosodique (cf. Le Gac & Yoo 2000) ainsi que l'organisation des contours prosodiques des mots prosodiques à l'intérieur de chaque groupe prosodique. Dans le présent état de l'implémentation, ce que nous proposons est un modèle de synthèse par mots (Vaissière 1971) où les courbes mélodiques sont associées aux mots prosodiques et des règles permettant de dériver la structure intonative de la phrase. Pour l'instant, chaque mot prosodique est pré-enregistré et relissés avec le contour correspondant (Yoo 2001), ce qui pour l'instant donne trois fichiers sons par mot prosodique de nos phrases testées. DepLin combine les fichiers pour une sortie sonore. Les résultats des fichiers sonores obtenus ont été testés auprès de locuteurs natifs du grec. Les phrases synthétisées ont été jugées comme naturelles, et les auditeurs ont toujours réussi à identifier le contexte communicatif correspondant.

5 Conclusion

Les résultats prometteurs obtenus avec nos grammaires et l'implémentation test montrent que l'approche linéaire et modulaire comme le prône la TST n'est pas seulement utile pour la correspondance du sens au texte écrit, au centre de la recherche TST en particulier et TAL en général, mais aussi pour l'analyse des phénomènes phonétiques, en particulier prosodiques. La notion classique de *topologie* permet une plus grande expressivité dans la linéarisation d'une structure de dépendance et l'étude des règles topologiques pour différentes langues ouvrirait peut-être d'intéressants champs de recherche comparative.

Des travaux en cours portent sur la combinaison de courbes mélodiques abstraites avec des fichiers sons neutres, mais pour l'instant nous sommes loin de la naturalité des fichiers préenregistrés par mots prosodiques. Des travaux futures porteront sur l'extension de la couverture de nos grammaire, mais aussi sur l'intégration du rythme dans le modèle : la structure topologique semble aussi liée à la structuration temporelle de la phrase.

Références

- Bröker, N. 1998. «Separating Surface Order and Syntactic Relations in a Dependency Grammar ». Actes de *COLING-ACL 98*.
- Choi, H-W. 1999. *Optimizing Structure in Context – Scrambling and Information Structure*, CSLI, Stanford.
- Duchier, D., R. Debusmann 2001. « Topological Dependency Trees: A Constraint-based Account of Linear Precedence », dans *Proceedings ACL 2001*, Toulouse.
- Gerdes, K., S. Kahane 2001a. « Word Order in German: A Formal Dependency Grammar Using a Topological Hierarchy », *Actes d'ACL 2001*, Toulouse.
- Lambrecht, K. 1994. *Information Structure and Sentence Form*, Cambridge Studies in Linguistics 71, Cambridge University Press.
- Le Gac D. & Yoo H. 2002. "Intonative structure of focalization in French and Greek", in Beyssade & al. (eds.), *Romance languages and Linguistic Theory 2000*, John Benjamins, Amsterdam
- Mel'čuk, I. 1988. *Dependency Syntax*. New-York, State University of New York Press.
- Mel'čuk, I. 2001. *Communicative Organization in Natural Language (The Semantic-Communicative Structure of Sentences)*. Amsterdam, Benjamins.
- Tesnière, L. 1959. *Eléments de syntaxe structurale*, Klincksieck, Paris.
- Vallduví, E. 1992. *The Informational Component*. New York, Garland.
- Vaissière, J. 1971. *Contribution à la synthèse par règles du français*. Thèse de 3ème cycle, Université des langues et lettres de Grenoble
- Wahlster, W. (ed.) 2000. *VERBMOBIL – Foundations of Speech-to-Speech Translation*, Springer.
- Yoo H. 2001. « Prosodie, organisation informationnelle et ordre des mots en grec moderne », *Actes de la journée d'étude Prosodie*, Grenoble