L'amas verbal au cœur d'une modélisation topologique de l'ordre des mots

Kim Gerdes, Sylvain Kahane

▶ **To cite this version:**

Kim Gerdes, Sylvain Kahane. L'amas verbal au cœur d'une modélisation topologique de l'ordre des mots. Lingvisticae investigationes : International Journal of Linguistics and Language, 2006, 29:1, pp.75-89. halshs-00120418

HAL Id: halshs-00120418
https://shs.hal.science/halshs-00120418

Submitted on 15 Dec 2006

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.

L'amas verbal
au cœur d'une modélisation topologique du français[1]

Kim Gerdes & Sylvain Kahane

1. Introduction

Cette communication poursuit un double objectif. Premièrement, nous souhaitons montrer que les principaux phénomènes d'ordre des mots (montée des clitiques, inversion du sujet, extraction, dislocation, …) peuvent être modélisés dans une approche topologique où la linéarisation est traitée par la mise en correspondance d'une structure syntaxique de dépendance (sans ordonnancement) avec une structure ordonnée de type structure de constituants topologiques.

Deuxièmement, nous voulons mettre en évidence un constituant particulier, très rarement considéré dans les modélisations de l'ordre des mots en français et qui s'impose naturellement dans une modélisation topologique du français : ce constituant, que nous appelons *amas verbal*, est constitué d'un amas d'un ou plusieurs verbes, des éléments grammaticaux nécessaires à la formation de cet amas (prépositions et complémenteurs) et d'éléments lexicaux très contraints, c'est-à-dire en dehors des verbes eux-mêmes (ou d'éléments prédicatifs tels que adjectifs ou noms prédicatifs appartenant à des tournures verbales) des clitiques et certains adverbes.

2. Une première approche de l'amas verbal

On peut donner deux justifications à l'existence des amas verbaux. D'un point de vue onomasiologique, il y a une nécessité dans la langue à pouvoir exprimer des prédications complexes exprimant modalités, aspects, temps, causalité, mise en évidence de tel ou tel participant, etc. comme dans *devoir partir, faire mourir, chercher à réussir, avoir l'intention de réussir, avoir beau être capable de réussir*, etc. La langue doit donc offrir dans son dispositif la possibilité de gérer de telles « unités complexes ». D'un point de vue sémasiologique, on peut voir plusieurs indices dans la syntaxe du français de la présence des amas verbaux, notamment dans deux phénomènes que nous allons étudier ici : la montée des clitiques et l'inversion du sujet. Ainsi, les compléments clitiques d'une forme verbale composée ne se placent-ils pas sur le participe (qui est l'élément prédicatif qui contrôle la sous-catégorisation), mais sur le verbe auxiliaire. Or l'auxiliaire et le participe peuvent être séparés par d'autres mots (*Pierre ne lui **a** pas encore vraiment **répondu***) et à ce titre doivent être considérés comme deux mots à part entière (même si leur sort est lié du point de vue de l'interface sémantique-syntaxe). Néanmoins, les éléments qui peuvent s'interposer entre l'auxiliaire et le participe sont très contraints[2] : *Pierre a déjà répondu, Pierre à tout de suite/rapidement/*hier/ * ici / *en 3 secondes / *à toute vitesse répondu.*

---

[1] Version courte présentée aux Journées de la syntaxe, Bordeaux 2004. Version longue in K. Gerdes & C. Müller (éds), *Ordre des mots et topologie de la phrase française, Linguisticae Investigationes*, 29 : 1, 14 p.

[2] Nous laissons de côté les incises qui obéissent clairement à un autre mécanisme de placement et qui peuvent venir « couper » l'amas verbal : *Pierre a déjà, **comme tu le sais**, répondu à cette question*. Nous pensons que les incises sont des éléments qui n'obéissent pas stricto sensu au modèle topologique : contrairement aux autres éléments, elles ne viennent pas se placer dans des champs prédéfinis, mais viennent au contraire s'insérer entre deux champs.

Il nous semble que les deux propriétés sont corrélées, à savoir la possibilité pour un clitique de « monter » sur l'auxiliaire et l'impossibilité pour des éléments trop « complexes » de s'intercaler entre l'auxiliaire et le participe. En effet, l'une et l'autre sont pour nous la trace d'un même phénomène, la formation d'un « nucléus verbal » auxiliaire-participe, c'est-à-dire l'existence d'une prédication verbale complexe qui d'un certain point de vue est bien complexe (c'est-à-dire formée de plusieurs mots), mais d'un autre continue à se comporter comme un tout et garde certaines caractéristiques d'une forme simple. La modélisation que nous proposons de la « montée » des clitiques rend justement compte de cette corrélation. En effet, nous rendons possible la « montée » des clitiques par la formation d'un constituant privilégié autour de l'auxiliaire et du participe et c'est cette même formation qui rend naturelle la possibilité de contraindre les éléments qui vont apparaître dans cette formation.

L'inversion du sujet met également en évidence un amas verbal :

1. L'« inversion » du sujet ne se fait pas par rapport au seul verbe principal, mais par rapport à quelque chose que délimite justement notre amas verbal : *le livre que* ***Pierre*** *veut lire* ; ***le livre que veut*** ***Pierre*** *lire* ; *le livre que veut lire* ***Pierre***.
2. L'ordre respectif du sujet inversé et des compléments des autres verbes de la chaîne verbale est libre : *le livre que veut donner* ***Pierre*** *à sa sœur* ; *le livre que veut donner à Marie* ***son frère***.

Bien qu'il soit clair d'un point de vue syntaxique que le sujet dépende du verbe fini (accord) et l'objet indirect dépende de l'infinitif (rection), les deux verbes mettent en commun dans un même pot leurs dépendants pour la linéarisation.

3. Le modèle topologique du verbe en français

Dans une approche topologique, le calcul de l'ordre des mots est géré au niveau de la correspondance entre une structure de dépendance syntaxique (non ordonnée) et une structure de constituants ordonnée. Les nœuds de la structure syntaxique sont les mots. Lors de la linéarisation de l'arbre, chaque mot va ouvrir un *constituant topologique* (ou éventuellement plusieurs constituants enchâssés les uns dans les autres, ce qui formellement revient à peu près au même). Un tel constituant est comme une boîte dont les cases ordonnées sont appelées les *champs*. L'élément qui a ouvert un constituant en occupe un certain champ et les éléments qui dépendent de lui vont venir occuper les autres champs. On doit spécifier pour chaque champ s'il peut rester vide ou non et s'il peut accueillir ou non plus d'un élément. Un élément peut par ailleurs *s'émanciper* du constituant ouvert par son gouverneur syntaxique et venir se placer dans un constituant ouvert par un élément plus haut dans l'arbre syntaxique. Dans un tel formalisme, une règle d'ordre est donc une règle qui spécifie pour un dépendant donné quel champ il peut occuper et s'il peut ou non s'émanciper et jusqu'où en fonction de sa catégorie et/ou de sa fonction syntaxique. A ces règles s'ajoutent les règles qui disent quel type de constituant pourra ouvrir un mot placé dans un champ donné et pour chaque constituant, quelle est la liste des champs qui le composent et pour chaque champ, combien de constituants peuvent ou doivent l'occuper (cf. Gerdes & Kahane 2001 pour une présentation formelle d'une grammaire topologique).

Venons-en à la modélisation de l'ordre des dépendants du verbe et donc à la description du ou des constituants topologiques ouverts par le verbe. Nous considérons trois niveaux d'enchâssement ici. Nous commençons par présenter les constituants ouverts par le verbe principal d'une phrase. Nous discuterons ensuite des constituants ouverts par des verbes enchâssés pour lesquels des modifications peuvent être nécessaires.

- Le premier niveau est constitué par l'*amas verbal* : celui-ci comprend les champs pour les clitiques avant le verbe et deux champs après le verbe. Les champs clitiques sont au nombre de 7 : *il* < *ne* < *me* < *le* < *lui* < *y* < *en*. Les deux autres champs accueillent

successivement des adverbes et un dépendant verbal du verbe. La présence de ce dernier champ dans l'amas verbal constitue très certainement la grande originalité de notre approche[3]. A noter que l'amas verbal offre également une position pour l'enclitique en cas de construction interrogative (*Pierre lui donnera**-t-il** la réponse ?*). Le verbe à l'impératif nécessite un amas verbal spécifique où tous les pronoms sont enclitiques.

- Le deuxième niveau est constitué par le *domaine micro-syntaxique*. Ce domaine inclut tous les éléments sous-catégorisés et les modifieurs de verbe (par opposition aux « modifieurs de phrase »). Nous reprenons ici la distinction entre micro- et macro-syntaxe faite par Blanche-Benveniste 1990 et Berrendoner 1991 à qui nous renvoyons pour une définition plus précise, ainsi qu'à Deulofeu 2003. Le domaine micro-syntaxique comprend au moins un champ pour le sujet, un champ pour l'amas verbal et un champ pour les compléments. Tous les compléments sont placés dans le même champ car leur ordre respectif est libre, ou du moins, il n'est pas guidé par les mêmes paramètres que ceux que nous modélisons par la structure topologique.
- Le troisième niveau est constitué par le *domaine macro-syntaxique*. Celui-ci comprend les éléments détachés à gauche ou à droite : ***Moi, ma sœur**, il ne veut plus lui parler, **Pierre***. Nous postulons deux champs à gauche et un champ à droite.

Passons maintenant aux verbes enchâssés. Le participe passé d'un passé composé dépend du verbe auxiliaire et doit nécessairement aller dans l'amas verbal ouvert par l'auxiliaire, où il ouvrira lui-même un amas verbal. Le participe passé possède la propriété particulière de ne pas pouvoir accueillir de clitiques, ce qui peut être modélisé en bloquant les positions pour les clitiques dans l'amas verbal ouvert par un participe passé.

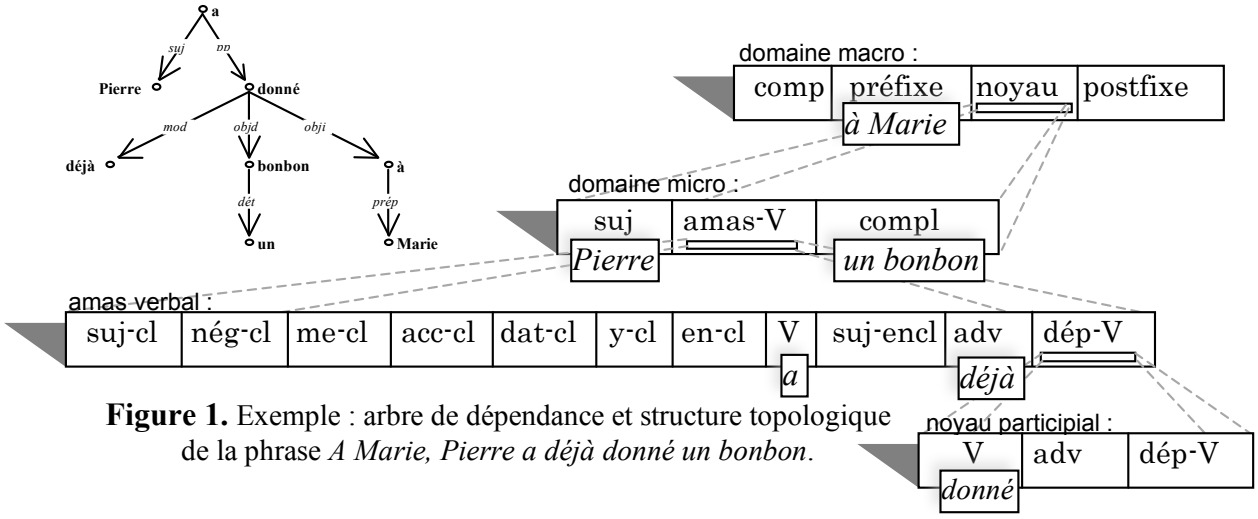

Figure 1. Exemple : arbre de dépendance et structure topologique de la phrase *A Marie, Pierre a déjà donné un bonbon*.

Montrons comment est établie la correspondance entre l'arbre de dépendance syntaxique de cette phrase et sa structure topologique[4]. Plaçons nous dans le sens de la synthèse (le locuteur parle et transforme un sens en une phrase en passant par une représentation syntaxique). Le

[3] De nombreuses approches actuelles en analyse syntaxique automatique effectuent un « tronçonnage » préalable de la phrase en tronçons appelés *chunks* (Abney 1991, Vergne 2000). Le chunk verbal s'apparente à notre amas verbal, mais s'en distingue justement par le fait qu'il ne contient toujours qu'un seul verbe.

[4] Par soucis de simplification, nous masquons totalement le traitement de la morphologie. Nous manipulons donc directement des formes fléchies sans montrer comment parallèlement à la construction de la phrase la formation des mots peut être assurée.

verbe auxiliaire *a*, qui est la racine de l'arbre de dépendance, ouvre d'abord trois constituants enchâssés : un domaine macrosyntaxique, un domaine microsyntaxique et enfin un amas verbal à l'intérieur de ce dernier. Ses dépendants vont pouvoir se placer dans les différents champs ouverts : le sujet *Pierre* ira dans le champ sujet du domaine micro, tandis que le participe passé *donné* devra occuper le champ du dépendant verbal dans l'amas verbal, où il ouvrira un amas verbal enchâssé (pour un éventuel dépendant verbal comme dans *Pierre a voulu venir*). L'adverbe *déjà* restera préférentiellement dans l'amas, où il précèdera le participe. Le complément d'objet direct *un bonbon*, qui dépend de *donné* ne pourra pas aller dans le constituant topologique ouvert par son gouverneur (un amas dans l'amas) et devra s'émanciper pour atteindre le champ complément du domaine micro ouvert par l'auxiliaire. Le complément d'objet direct *à Marie* pourrait suivre le même chemin, mais il peut aussi comme ici, en raison d'un marquage communicatif particulier, rejoindre le champ préfixe du domaine macro (c'est la topicalisation).

Contrairement à un participe passé, un infinitif a deux possibilités : il peut

- soit, comme le participe passé, aller dans le champ dépendant de l'amas verbal de son gouverneur, où il ouvrira lui-même un amas verbal (enchâssé) ;
- soit, comme les autres compléments, rejoindre le champ complément (ou même un champ du domaine macro) et ouvrir un domaine micro qui pourra accueillir tous ses dépendants.

Dans les deux cas, le verbe infinitif ouvre un amas verbal : soit directement s'il forme un amas avec son gouverneur, soit à l'intérieur de son domaine micro s'il prend un peu d'indépendance. L'amas verbal ouvert par un infinitif (le même dans les deux cas) est un peu différent de celui d'un verbe fini en raison du placement particulier du champ adverbe entre le « clitique » *ne* et les autres clitiques : *Pierre pense* [*ne* ***pas*** *le lui dire*].

Un verbe fini d'une complétive ou d'une subordonnée circonstancielle (*Pierre sait* [*qu'hier Jean n'****est*** *pas venu*] ; [*Quand tu l'****auras*** *fini, le livre*], *rapporte-le-moi*) ouvre un domaine macro comme le verbe principal d'une phrase. Ce domaine macro comporte en plus des champs pour les éléments détachés, un champ pour le complémenteur (*que, si, quand,* etc.).

Le cas des relatives est tout à fait intéressant. Elles se distinguent des complétives, du point de vue topologique, par deux propriétés :

1. L'impossibilité d'avoir un détachement à gauche :
   J'ai retrouvé le livre que je voulais prêter à Pierre
   **J'ai retrouvé le livre que, Pierre, je voulais lui prêter*
2. La possibilité d'une inversion du sujet : *le livre que lit Pierre*.

Ces deux propriétés nous semblent résulter d'une même caractéristique, qui distingue fondamentalement la relative de la complétive : le fait que le complémenteur d'une relative est aussi un pronom et qu'il joue ainsi un rôle valentiel dans la proposition qu'il introduit. Pourquoi cela aurait-il une influence sur l'ordre des mots dans cette proposition ? Comme nous l'avons vu pour la complétive, le complémenteur occupe la position la plus à gauche du domaine macro. Par contre, les éléments qui jouent réellement un rôle dans la proposition reste dans le domaine micro. Le pronom relatif, qui est les deux à la fois, va donc occuper le champ du complémenteur, mais en même temps, il va attirer ce champ dans le domaine micro, avec deux conséquences :

1. le champ préfixe est écrasé, bloquant ainsi dislocation gauche et topicalisation ;
2. un des champs du domaine micro précédant le verbe est maintenant occupé, validant ainsi l'inversion du sujet, c'est-à-dire permettant au sujet d'aller dans le champ complément.

Voyons maintenant comment nous modélisons l'inversion du sujet.

4. L'inversion du sujet

Nous allons nous intéresser à l'inversion dite stylistique[5] qui obéit à des contraintes syntaxiques : il ne doit pas y avoir de complément d'objet direct dans le champ complément (**la personne à qui donne **le livre** Pierre*)[6] et la proposition doit comprendre un extraction : interrogation en *qu*, relative, coda de clivée ou topicalisation de groupe prépositionnel (Marandin 2003). La deuxième condition se ramène, selon nous, à la présence d'un élément (non clitique) devant le verbe dans le domaine micro, soit que l'élément topicalisé occupe le champ sujet (voir section 6 de la version longue), soit que le champ complémenteur est aspiré dans le domaine micro lorsque le complémenteur est en même temps un élément valentiel.
Ici, nous souhaitons juste montrer rapidement comment notre modélisation topologique assure le bon placement du sujet inversé. En voici le principe. Le sujet inversé va dans le champ complément avec les autres compléments des verbes de l'amas verbal. Nous assurons ainsi que le sujet inversé puisse se placer de façon quelconque vis-à-vis des autres compléments des verbes de l'amas verbal : *le livre que veut donner **Pierre** à sa sœur ; … à Marie **son frère***.
Ces dernières données, facilement traitées ici par la formation d'un amas verbal, sont fort problématiques pour les grammaires syntagmatiques, qui y voit un cas de « scrambling », le sujet inversé (ici *Pierre*) venant s'insérer dans l'infinitive (ici *donner à sa sœur*) et se « mélanger » avec les compléments du verbe infinitif.
Le fait que, dans *le livre que Marie veut lire*, l'inversion soit seulement possible par rapport à *veut lire (le livre que veut lire Marie ; *le livre que veut Marie lire*) est traité en obligeant *lire* à rejoindre l'amas verbal ouvert par *veut*. Généralement, les infinitifs nus (sauf sujet ou disloqués, voir version longue) doivent aller dans l'amas ouvert par leur gouverneur. D'autres éléments peuvent y aller : les infinitifs précédés d'un marqueur *(le livre que rêve de **lire** Marie* ; *??le livre que rêve Marie de lire*), les adjectifs attributs (*le livre qu'a été **heureuse** de lire Marie* ; **le livre qu'a été Marie heureuse de lire*) ou les noms prédicatifs (*le livre qu'avait **l'intention** de lire Marie* ; **le livre qu'avait Marie l'intention de lire*). Des verbes subordonnés finis peuvent également venir dans l'amas verbal de leur gouverneur lorsque leur sujet est clitique : *le livre que veut qu'**on lise** Marie* ; **le livre que veut que Pierre lise Marie*.[7]

Une autre construction peut être rapproché de l'inversion du sujet — l'inversion du sujet logique dans les phrases telles que : *J'ai vu **ma sœur** sortir des toilettes* ; *J'ai vu sortir **ma sœur** des toilettes.* Nous proposons d'analyser le dernier cas à l'aide d'un amas verbal *ai vu sortir*, et par conséquence de rapprocher la construction en *voir* + infinitif avec « inversion » de la construction causative en *faire* + infinitif, où les clitiques doivent monter sur *faire*. Une

[5] Nous laissons de côté l'inversion dite élaborative (Marandin 2003), qui n'est possible qu'avec une classe particulière de verbes à sujets peu ou non agentifs et pour laquelle le sujet doit occuper la dernière position du domaine micro : *Sont possibles de rares inversions avec verbe initial.* ; *Rendront un devoir supplémentaire les élèves qui ont raté l'examen de chimie.*

[6] Le complément d'objet direct peut par contre être extrait (*le livre **que** lit Pierre*) ou cliticisé (*la personne à qui **le** donne Pierre*).

[7] Il semble qu'un verbe avec un sujet non clitique puisse quand même aller dans l'amas verbal. Auquel cas, son sujet ne peut aller dans l'amas et devra rejoindre le champ complément : *?le livre que Marie veut que lise Pierre*. Notons également la construction suivante : *Ce qu'a dit que ferait Luc à sa petite amie est inquiétant* (Claude Muller, communication orale). Ici le syntagme *Luc* sature en même temps la position sujet des deux verbes finis de l'amas – *a* et *ferait*. Ce phénomène ne met donc pas simplement en cause l'ordre des mots, mais met en évidence un lien, plus étroit que ne le suggère notre traitement, entre l'ordre des mots et la structure syntaxique, puisque seule l'inversion du sujet valide le partage de sujet : **Ce que Luc a dit que ferait à sa petite amie est inquiétant*.

telle analyse est corroborée par le contraste entre les deux constructions pour *voir* en ce qui concerne les clitiques : *J'ai vu ma sœur **en** sortir* [des toilettes] ; ***J'en** ai vu ma sœur sortir.* ; ??*J'ai vu **en** sortir ma sœur* ; *J'**en** ai vu sortir ma sœur.*
Notre traitement de l'inversion du sujet logique repose donc sur la formation d'un amas verbal. La formation d'un amas verbal est bloquée quand le verbe gouverneur possède un autre complément que l'infinitif car dans ce cas le « scrambling » est inacceptable :

J'ai dit à Pierre d'aller au même endroit que l'autre fois
??*J'ai dit d'aller à Pierre au même endroit que l'autre fois*.

Les verbes du type *voir* ou *entendre*, de même que le causatif *faire*, sont donc considérés comme des verbes particuliers pouvant à la fois gouverner un complément nominal et former un amas verbal avec leur complément verbal.

5. Extraction et amas verbal

Nous avons vu que notre modélisation de l'inversion du sujet nécessitait la création d'un amas verbal et que cette même inversion était déclenchée par une extraction. Nous allons maintenant voir quels sont les liens entre l'extraction et la formation d'un amas verbal.
La première chose que nous pouvons remarquer, c'est que l'extraction ne nécessite pas la formation d'un amas verbal et que l'extraction peut être possible dans des situations où l'inversion ne l'est pas, comme par exemple avec une complétive à sujet non clitique ou avec un verbe recteur avec un complément nominal :

*une réunion que **mon collègue Pierre** veut absolument que le directeur organise*
?**une réunion que veut absolument que le directeur organise mon collègue Pierre*
*une réunion à laquelle **mon collègue Pierre** menace le directeur de ne pas participer*
?**une réunion à laquelle menace le directeur de ne pas participer mon collègue Pierre*

L'existence d'un groupe extrait qui contient un complémenteur et joue en même temps un rôle valentiel dans la proposition qu'il introduit concerne l'interface sémantique-syntaxe. Il existe de très nombreux travaux sur la modélisation de l'extraction, qui reste la pierre d'achoppement de la quasi-totalité des modèles proposés depuis les années 1970. Comme on le sait, l'extraction obéit à diverses contraintes, appelées traditionnellement *contraintes d'îlots* (Ross 1967). Kahane 2000 propose une version de ces contraintes en termes de chaînes de verbes et de tournures verbales, appelées *nucléus verbaux* : l'élément extrait doit dépendre d'un nucléus verbal qui contient le verbe principal de la proposition qu'il introduit (cette chaîne est soulignée dans les exemples précédents). Les nucléus verbaux, qui s'apparentent aux amas verbaux, sont des objets de l'interface sémantique-syntaxe. Le nucléus verbal ne correspond pas nécessairement à une portion continue de la chaîne parlée (un syntagme), mais il aura néanmoins tendance à être le plus compact possible et tant que faire se peut à former un amas verbal (même si cela n'est pas toujours possible).
Reprenons le problème sous l'angle de l'ordre des mots (l'interface syntaxe-topologie) pour mieux expliquer pourquoi le nucléus verbal tend à former un amas verbal. Le placement du groupe extrait nécessite généralement une émancipation : en effet, le groupe extrait se place dans le champ complémenteur de la proposition qui est ouvert par le verbe principal de la proposition et ce verbe n'est pas nécessairement son gouverneur. Pour rejoindre le champ complémenteur, le groupe extrait devra dont « traverser » les frontières de tous les constituants ouverts par tous les éléments du nucléus verbal, c'est-à-dire par tous les verbes et éléments prédicatifs qui appartiennent à la chaîne de dépendance entre le verbe principal et le gouverneur du groupe extrait. Si ces éléments ont pu se regrouper au sein d'un amas verbal, les frontières à traverser seront donc uniquement des frontières d'amas verbaux enchâssés, c'est-à-dire des frontières relativement perméables (puisqu'elle peuvent même être traversées par des compléments ordinaires) et l'émancipation/extraction ne sera donc pas trop « coûteuse ». Par contre, si l'un des verbes du nucléus verbal n'a pas rejoint l'amas verbal de

son gouverneur et a ouvert son propre domaine, l'extraction, bien que toujours possible, devient beaucoup plus délicate. Ce deuxième type d'extraction se rencontre d'ailleurs très peu et essentiellement à l'écrit. On voit donc que l'extraction va favoriser (voire même forcer) la formation d'un amas verbal contenant l'ensemble du nucléus verbal et c'est cet amas verbal qui servira ensuite de pivot à une éventuelle inversion du sujet.

6. Conclusion

Ce travail est une première ébauche d'une modélisation de l'ordre des mots en français dans le cadre du modèle topologique, couvrant, de manière parfois un peu grossière, les principaux problèmes connus : le placement des clitiques, l'inversion du sujet, l'extraction, la dislocation, etc. (Nous renvoyons à la version longue de ce travail pour une discussion plus détaillée des phénomènes de macro-syntaxe tels que la dislocation et la topicalisation.)

La présente description est restée quelque peu informelle, mais elle s'appuie sur un formalisme mathématique complet (Gerdes & Kahane 2001) implanté par K. Gerdes (http://talana.linguist.jussieu.fr/~kim/deplin). Nous n'avons pas non plus cherché à justifier l'approche topologique, ni à mener des comparaisons avec d'autres approches. A ce sujet, on pourra consulter Gerdes & Kahane, sous presse.

Notre étude s'est centrée sur l'ordre des dépendants directs des verbes, c'est-à-dire sur la topologie du verbe, mettant en évidence la nécessité de considérer un constituant verbal particulier que nous avons appelé l'amas verbal. Une étude similaire sur la topologie du nom mettraient en évidence des parallèles intéressants, notamment sur la constitution d'amas nominaux du type « N *de* N ».

Notre étude s'est également limitée à l'étude de l'interface syntaxe-topologie (la correspondance entre arbres de dépendance syntaxiques et structures de constituants topologiques). Ce module s'inscrit dans un modèle complet de la langue comprenant également une interface sémantique-syntaxe et une interface topologie-phonologie. Cela pour dire que les structures syntaxiques et topologiques que nous considérons ne peuvent se justifier du seul point de vue de l'interface syntaxe-topologie, mais doivent être envisagées comme des structures intermédiaires entre une représentation sémantique de la phrase (incluant une structuration informationnelle/communicative) et une représentation phonologique (incluant une représentation de la prosodie et une structuration en constituants phonologiques) (cf. Gerdes & Yoo 2003, Yoo 2003 sur ce dernier point). Une validation nécessaire de notre proposition d'amas verbal doit être justement de montrer que celui-ci tend à être un constituant phonologique[8].

Enfin, notre étude s'inscrit dans une modélisation de l'ordre des mots dans la diversité des langues. Le modèle topologique vient de l'étude de l'ordre des mots dans les langues germaniques, qui pose des problèmes importants en raison d'une tension entre une grande liberté des mots (scrambling, extraposition, …) et des contraintes syntaxiques fortes (verbe final ou en deuxième position). Le modèle a également été appliqué à la description de langues à ordre plus libre comme le grec (Yoo 2003) ou moins libre comme l'arabe (El Kassas & Kahane 2004). L'application à des langues à ordre assez rigide comme le français restait à faire.

---

[8] Les phénomènes de liaison (*Pierre a bien_aimé ça* ; *Pierre ne fera pas_écouter le début* ; *Pierre ne cherche plus_à comprendre*) pourraient également servir de justification à la formation d'un amas verbal, même si clairement la liaison ne met pas en jeu la seule interface topologie-phonologie (Bonami *et al.* 2004).

Références

Abney Steven, 1991, Parsing by chunks, in R. Berwick, S. Abney and C. Tenny (eds.), *Principle-Based Parsing*, Kluwer.

Berrendoner Alain, 1991, Pour une macro-syntaxe, in Dominique Willens (éd), *Données orales et théories linguistiques*, Duculot, Paris/Louvain, pp. 25-31.

Blanche-Benveniste Claire, 1990, *Le français parlé : études grammaticales*, avec Mireille Bilger, Christine Rouget et Karel van den Eynde, Editions du CNRS, Paris.

Bonami Olivier, Gilles Boyé & Jesse Tseng, 2004, An Integrated Analysis of French Liaison, in G. Jaeger (éd.), *Formal Grammar 2004 Preproceedings*, Nancy, France.

Deulofeu José, 2003, L'approche macrosyntaxique en syntaxe : un nouveau rasoir d'Occam contre les notions inutiles, in Gagean N. (éd*.), Rencontres linguistiques en Pays Rhénan -12*, *Scolia n° 16*, Publications de l'Université de Strasbourg 2, Strasbourg.

El Kassas Dina & Sylvain Kahane, 2004, Modélisation de l'ordre des mots en arabe standard, Atelier sur le Traitement Automatique de la Langue Arabe Ecrite et Parlée, *JEP-TALN*, 259-264.

Gerdes Kim, 2002, *Topologie et grammaires formelles de l'allemand*, Thèse de doctorat, Université Paris 7.

Gerdes Kim & Sylvain Kahane, 2001, Word Order in German: A Formal Dependency Grammar Using a Topological Hierarchy, *ACL*, Toulouse, 8 p.

Gerdes Kim & Sylvain Kahane, sous presse, "Phrasing It Differently", in L. Wanner (ed.), *Papers in Meanning-Text Theory in honour of Igor Mel'cuk*, Benjamins, 34 p.

Gerdes Kim & Hi-Yon Yoo, 2003, *La topologie comme interface entre syntaxe et prosodie : un système de génération appliqué au grec moderne Actes de* TALN 2003, Batz-sur-Mer.

Kahane Sylvain, 2000, "Extractions dans une grammaire de dépendance lexicalisée à bulles", *T.A.L.*, *41:1*, 211-243.

Marandin Jean-Marie. 2003. «Inversion du sujet et structure de l'information dans les langues romanes», in Danièle Godard (éd), *Langues romanes. Problèmes de la phrase simple*, Paris: Editions du CNRS. Version longue à http://www.llf.cnrs.fr/fr/Marandin.

Ross John, 1967, *Constraints on Variables in Syntax*, PhD Thesis, MIT (publié sous le titre *Infinite Syntax*, Reidel, Dordrecht, 1985).

Vergne Jacques, 2000, *Étude et modélisation de la syntaxe des langues à l'aide de l'ordinateur - Analyse syntaxique automatique non combinatoire*, Thèse d'HDR, Université de Caen.

Yoo 2003, *Ordre des mots et prosodie : Essai de description et de formalisation pour le grec moderne*, thèse de doctorat, Université Paris 7.