

Etude du traitement de certains compléments de phrase dans le cadre d'une méta-grammaire

Éric Villemonte de la Clergerie

► To cite this version:

Éric Villemonte de la Clergerie. Etude du traitement de certains compléments de phrase dans le cadre d'une méta-grammaire. LGC'12 - 31ème 30ème Colloque international sur le Lexique et la Grammaire, Intitut de langues Romanes, University of South Bohemia, Sep 2012, Nové Hradý, République tchèque. hal-00765384

HAL Id: hal-00765384

<https://hal.inria.fr/hal-00765384>

Submitted on 14 Dec 2012

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.

ÉTUDE DU TRAITEMENT DE CERTAINS COMPLÉMENTS DE PHRASE DANS LE CADRE D'UNE META-GRAMMAIRE

Éric de LA CLERGERIE

ALPAGE, INRIA Paris–Rocquencourt & Université Paris 7, France

Eric.De_La_Clergerie@inria.fr

Résumé

Dans le cadre de FRMG, une méta-grammaire (MG) à large couverture du français, nous examinons comment les propriétés d'héritage et de modularité des MG peuvent être utilisées pour représenter les compléments de phrase, à la fois en terme de diversité et en terme de position dans la phrase.

Mots clés : méta grammaire, syntaxe, adverbiaux

Introduction

Introduites sous différentes formes (CANDITO 1999; GAIFFE & al 2002; BARRIER & al 2004; PARMENTIER & al 2005; THOMASSET & al 2005), les méta-grammaires (MG) permettent des descriptions grammaticales modulaires exploitant des propriétés d'héritage. Toutes s'appuient sur des hiérarchies de classes, qui permettent de se focaliser sur la description d'un phénomène syntaxique, comme la notion de sujet, en la raffinant progressivement au travers de l'héritage pour en capturer ses différentes formes. Un processus de compilation permet ensuite de combiner les classes pour produire les structures d'une grammaire dans un formalisme syntaxique cible, comme, par exemple, des arbres élémentaires dans le cas de grammaires d'arbres adjoints [TAG – *Tree Adjoining Grammars*] (JOSHI & al 1975).

Historiquement, les travaux autour des MG ont surtout été illustrés, avec succès, par le traitement des aspects valenciels des verbes (CANDITO 1999; CRABBE 2005) et éventuellement d'autres catégories comme les noms et les adjectifs (BARRIER 2002). Pour résumer, il s'agit de prendre en compte les diverses dimensions descriptives que forment les fonctions syntaxiques (sujet, objet, etc), les positions (canoniques, extraites, etc), et les réalisations (nominales, prépositionnelles, phrastiques, etc).

Nous nous proposons de montrer que les MG facilitent également la description des compléments de phrase, vus comme généralisant la notion d'adverbiaux. Du point de vue de la description dans le cadre de grammaires symboliques opérationnelles, ces compléments regroupent un ensemble de situations bien moins étudiées et finalement très protéiformes (et très floues). Au final, ce travail, effectué pour le français dans le cadre de la méta-grammaire à large couverture FRMG (DE LA CLERGERIE 2005), permet de capturer un large éventail de possibilités avec une forte économie en terme descriptif. Il est également l'occasion de présenter, en les illustrant, les caractéristiques de notre variante de méta-grammaires.

Nous commençons par lister quelques types de compléments que nous avons identifiés et que souhaitons décrire, en mettant en avant les problèmes potentiels. Ensuite, nous fournissons quelques éléments d'information sur les méta-grammaires utilisées dans cet article. Enfin, nous étudions comment on peut capturer de manière économique une partie de la richesse des compléments de phrase avec une méta-grammaire. Pour finir, nous indiquons quelques difficultés résiduelles.

Comment caractériser les compléments de phrase

Les compléments de phrases regroupent un ensemble très hétérogène de compléments. Dans la plupart des cas, ils sont assimilables à des adverbiaux, dans le sens où ils peuvent être remplacés, de manière plus ou moins équivalente, par des adverbes. En tant que tels, ces compléments peuvent porter sur divers aspects du temps (moment, durée, périodicité, etc), de lieu, de manière, de moyen, de

concession, de cause, de conséquence, etc. Par nature, ils sont optionnels dans la phrase et coïncident bien avec la notion d'adjoints dans le cadre des TAG. Par contre, d'un point de vue descriptif, ils souffrent des problèmes suivants:

- Les compléments de phrase peuvent prendre de très nombreuses réalisations. Certaines, comme les adverbes, les groupes prépositionnels, les participiales, ou les subordinées, dénotent des modificateurs et ne prêtent pas trop à confusion. Mais nous avons aussi les cas de groupes nominaux, de groupes adjectivaux pouvant prêter à confusion avec d'autres fonctions dans la phrase (objet et attribut par exemple). Enfin, nous avons le cas de locutions adverbiales et de constructions idiomatiques plus ou moins figées, souvent difficiles à recenser et demandant des efforts spécifiques en terme de description.
- Les compléments présentent une forte mobilité dans la phrase. Ils peuvent apparaître en tête ou fin de phrase, mais aussi autours des arguments verbaux (par exemple entre un sujet non cliticisé et un verbe) ainsi qu'entre un auxiliaire et un verbe.
- En fait, les points d'accroches sont même plus variés: à la liste précédente de points directement liés à la phrase, on peut aussi citer le cas de compléments apparaissant à droite de mots coordonnants (*il mange une pomme et, **parfois**, une poire*), de prépositions (*il part, avec, **toutefois**, une pointe de regret*) et des noms déverbaux (*L'annonce, **ce matin**, d'un remaniment a surpris tous les commentateurs*).
- Enfin, les compléments de phrases, comme d'autres types de compléments, peuvent être isolés du reste de la phrase par l'utilisation de diverses marques d'incises, en général la virgule, mais aussi les parenthèses '(X)' et les tirets '- X -'. Ces marques sont obligatoires dans certains cas, mais en général optionnelles et, de fait, souvent manquantes dans la pratique.

Ces divers points sont illustrés ci-dessous par quelques types de compléments que nous traitons dans le cadre de notre méta-grammaire et qui montrent bien cette grande diversité de situations.

Le cas le plus typique est fourni par les adverbes (*il a **parfois** envie de partir; **Désormais**, il veut partir*). Nous avons aussi les groupes prépositionnels (*Avec son ami, il a décidé de partir **sans tarder***), ainsi que les subordinées (*il est arrivé **pendant que tu parlais***). Dans les cas relativement spécifiques, nous pouvons ajouter les participiales (*déçu par ses résultats, il se retire; **Sa société n'allant pas bien**, il doit la vendre*) et adjectifs flottants (*il prend le train, **soucieux des deniers publics***).

À côté de ces cas *faciles*, nous avons des réalisations par des groupes nominaux ou des pronoms, avec

- des expressions temporelles dans « **ce matin**, il est parti, sans rien dire ».
- des indications de rang dans « il a, **le premier**, fini l'exercice ».
- des expressions nominales liés à des parties du corps humain, comme dans « **mains sur la tête**, il recule contre le mur »
- des expressions dénotant des positions spatiales (au sens large) ou des adresses
 - **couloir de droite**, vous avez la classe de Mr Louis.
 - vous trouverez, **chapitre 22**, les explications nécessaires à ce devoir.
 - il attend, **rue des Bourdonnais**, que ses amis arrivent.
- des pronoms liés au sujet de la phrase, comme dans « il a, **lui aussi**, décidé de partir »

Enfin, nous avons les cas de réalisations par des locutions, pouvant prendre des formes très variées. Ainsi, *il y a* peut être vu comme une préposition mais son origine verbale n'a cependant pas disparu comme illustré dans « il est parti, **il n'y a pas deux jours**, avec des amis ». De même, nous pouvons continuer à considérer *voici* et *voilà* comme des formes verbales (impératives) utilisables pour dénoter des expressions temporelles dans « Il est parti, **voici deux semaines**, avec des amis ». Dans cette série, nous pouvons ajouter des expressions plus exotiques comme celles construites sur le verbe *obliger*, dans « **service oblige**, je dois vous quitter ».

Nous pouvons aussi mentionner le cas des constructions concessives, comme dans « il a, **quoi que tu en penses**, toutes ses chances » et également des cas de comparatives « Paul a, **plus que son frère**, le

sens de la famille ». Pour clore cette liste, qui n'a rien d'exhaustive, nous avons les cas de compléments de phrase qui ne sont pas des adverbiaux comme les apostrophes (de l'audience) dans « *je vous annonce, chers collègues, ma démission ferme et définitive* ».

Cette liste suggère qu'il existe un cadre général commun à ses divers exemples, en terme de positions, de points d'accroche et d'incise. Par contre, un contrôle relativement fin est nécessaire pour autoriser et bloquer certaines possibilités (par exemple incise obligatoire dans certaines positions pour certains types de compléments). Et nous avons à faire face à la diversité des réalisations. Ainsi, pour les réalisations nominales, on peut constater que certaines requièrent des groupes saturés avec un déterminant (*ce soir*), d'autres non (*couloir de droite*); certaines doivent être nécessairement modifiées (*mains sur la tête*), et souvent avec des contraintes de nature sémantique. Comme nous allons le voir, les méta-grammaires fournissent un cadre possible pour cette tâche.

Rappels sur les Méta-Grammaires

La section suivante sera l'occasion de présenter concrètement certaines des caractéristiques de notre variante de Méta-Grammaire (THOMASSET & al 2005). Néanmoins, nous la présentons ici de manière plus synthétique et abstraite.

Une MG est constituée d'une hiérarchie de classes reliées par héritage (<: **SuperClass**). Une classe est un ensemble de contraintes pouvant porter sur des noeuds d'un arbre d'analyse (**node(Subject)** ou **Subject** selon les contextes) ou sur la structure (arbre) en cours de construction dans son ensemble (via le mot clé **desc**). Les contraintes sur les noeuds permettent d'exprimer la dominance (directe avec $S \gg \text{Subject}$ ou indirecte avec $S \gg+ V$) ainsi que la précédence (**Subject < V** : le sujet précède le verbe). Il est possible d'avoir des contraintes sur le contenu d'un noeud, en terme de valeur de traits dans des structures de traits (**node(V): [cat: v, bot:[mode: indicative|subjunctive]]**). On peut aussi indiquer que deux valeurs doivent être identiques (ou plus précisément unifiables) via une équation entre chemins dans des structures de traits (**node(Nc).top.number=node(Det).top.number** : accord en nombre entre un déterminant et un nom). Sans entrer dans les détails, il est également possible d'indiquer qu'un noeud est présent ou absent selon des conditions exprimées par des formules logiques sur des équations entre chemins ($\sim \text{Subject} \Rightarrow \text{node(V).top.mode} = \text{value(imperative)}$: pas de sujet pour les verbes à l'impératif).

Enfin, une classe peut fournir une fonctionnalité, vue comme une ressource (+ **agrement**) tandis qu'une autre classe peut requérir cette fonctionnalité (- **agrement**), éventuellement dans le cadre d'un espace de nom (- **det:agrement**). Ce mécanisme est une alternative plus flexible à l'héritage, dans le sens où la même fonctionnalité peut être requise plusieurs fois dans des espaces de noms différents, alors qu'on ne peut hériter qu'une seule fois des contraintes d'une classe parente.

L'héritage, les ressources et les contraintes sont utilisées pour combiner les classes de manière à obtenir des classes *neutres* (toute fonctionnalité/ressource demandée est fournie et toute ressource fournie est consommée) et *satisfiables*, du point de vue des contraintes (ainsi, deux contraintes de précédence $A < B$ et $B < A$ sont incompatibles dans une même classe). Les classes survivantes accumulent les contraintes de toutes les classes combinées, ces contraintes étant ensuite utilisées pour produire des arbres élémentaires TAG minimaux.

Rappelons qu'une grammaire TAG se compose d'un ensemble d'arbres élémentaires initiaux et auxiliaires. Les arbres initiaux sont *substitués* au niveau de noeuds feuilles étiquetés par des non-terminaux. Les arbres auxiliaires sont *adjoints* au niveau de noeuds (internes) et permettent d'adjoindre du matériel à droite et/ou à gauche du noeud d'adjonction. De manière naturelle, les arbres auxiliaires et l'adjonction sont généralement utilisés pour traiter les modificateurs (optionels). Enfin, les arbres élémentaires possèdent généralement un noeud (feuille) ancre lexicalisé, correspondant à la tête de l'arbre, par exemple un noeud ancre de catégorie *v* (verbe) pour un arbre de racine *S* (phrase).

Décrire le comportement des compléments de phrase

En premier lieu, les compléments de phrase sont des modifieurs et sont représentés comme des arbres auxiliaires dans le cadre des TAG. Ils héritent donc des propriétés des arbres auxiliaires dans les grammaires TAG.¹

La classe **auxiliaire** est très générique et indique seulement qu'un arbre auxiliaire comprend un noeud racine (de type **std** et nommé **Root**) dominant (>>+) un noeud pied (de type **foot** et nommé **Foot**), et que les deux noeuds portent la même catégorie syntaxique.

```
class auxiliary {                                % Class for TAG auxiliary trees
  Root >>+ Foot;                                % root dominates foot
  node(Root).cat = node(Foot).cat;              % same syntactic category on foot and root
  node Root : [id:Root];
  node Foot : [id:Foot];
  node(Foot).type = value(foot);
  node(Root).type = value(std);
  node(Foot).top = node(Foot).bot;
}
```

En fait, en tant qu'adjoints, les compléments se comportent comme des modifieurs assez simples, pour lesquels le noeud racine est le père du noeud fils (aucun noeud ne venant s'insérer entre la racine et le pied). Ceci est capturé par la classe **shallow_auxiliary**, qui hérite de la classe **auxiliary** et qui est utilisée par de très nombreux modifieurs. Elle se contente de préciser que **Root** est le père (>>) de **Foot**, et indique qu'elle fournit la fonctionnalité **shallow_auxiliary**.

```
class shallow_auxiliary {                       %% Generic class for shallow auxiliary trees
  <: auxiliary;
  + shallow_auxiliary;                          %% provide functionality 'shallow_auxiliary'
  Root >> Foot;                                  %% root is parent of foot
}
```

En continuant à raffiner, on constate que les compléments partagent souvent la propriété de pouvoir être encadrés par des marques d'incises, donnant lieu à la classe **modifieur_on_x** qui requiert la fonctionnalité **shallow_auxiliary** et qui introduit un nœud **Incise** de catégorie **incise** dominant le modifieur. Ce nœud **Incise** sert d'accroche potentielle pour une adjonction des marques d'incises. La propriété **incise_kind** permet de contrôler les marques d'incises autorisées et un ensemble de quelques d'arbres adjoints permettent d'adjoindre des virgules (**coma** et **comastriect**), des parenthèses (**par**) et des tirets (**dash**). Les arbres pour l'ajout des virgules sont en peu particuliers, pour tenir compte des cas de début et fin de phrase, ainsi que des cas où une virgule physiquement présente sert en fait de marque partagée pour plusieurs incises.

```
class modifieur_on_x {
  + x_modifieur;
  - shallow_auxiliary;                          % require functionality shallow_auxiliary
  Root >> Incise;                                % root parent of node incise
  node Incise : [cat:incise, id:incise, type: std];
  Incise >>+ Modifieur;                          % incise dominâtes the modifieur
  node(Incise).bot.incise_kind = value(coma|comastriect|par|dash); % incise_kind controls the marks
}
```

Il est à noter que la classe **modifieur_on_x** est en fait utilisée par de nombreux types de modifieurs comme les adjectifs sur les noms, les adverbes sur les adjectifs, les adverbes sur les adverbes, etc.

À ce stade, il devient pertinent de différencier les cas où le modifieur se situe avant ou après l'élément modifié (représenté par le nœud pied), en raffinant (par héritage) la classe **modifieur_on_x** et en

¹ À ne pas confondre avec la notion de verbe auxiliaire !

utilisant les valeurs **ante** et **post** de la propriété **position** pour permettre un contrôle.

```
class modifier_before_x {
  <: modifier_on_x;
  Incise < Foot;                                     %% Modifier on the left side of the modifiee
  node(Incise).adj = node(Incise).ante.adj;
  desc.position = value(ante);
}
class modifier_after_x {
  <: modifier_on_x;
  Foot < Incise;                                     %% Modifier on the right side of the modifiee
  node(Incise).adj = node(Incise).post.adj;
  desc.position = value(post);
}
```

Nous pouvons maintenant nous focaliser sur les compléments de phrase, qui, en première approximation, ne sont que des **modifier_on_x** avec x=S comme catégorie syntaxique de l'élément modifié. En pratique, nous considérons aussi la catégorie syntaxique artificielle **VMod** qui chapeaute tous les arguments verbaux (sujets non clitiques inclus) et qui permet d'accrocher des modifieurs de phrase à leur droite (*Paul, ce matin, est parti à la chasse*).

```
class modifier_at_S_level {
  + s_modifier;
  - x_modifier;
  - s_adj_pos;
  node(Root).bot = node(Foot).top;                   %% S-modifiers do not alter S properties
  node(Root).cat = value(S|VMod);                   %% may adjoin on S or VMod
}
class modifier_on_S {                               %% specialization for S
  + s_adj_pos;
  node(Root).cat = value(S);
}
```

Il est maintenant facile de partager ces propriétés communes pour divers types de compléments, en commençant par les adverbes, qui héritent d'autres propriétés générales des adverbes (par exemple un peu de sous-catégorisation), avec de plus la contrainte que les adverbes modifieurs de phrase ne peuvent être l'adverbe *très*, des adverbes d'intensification (comme *plus*) ou des adverbes de comparaison (comme *autant* ou *davantage*).

```
class adv_s {                                       %% Adverbs on sentences (may adjoin on both sides)
  <: adv;                                           %% inherit properties for adverbs
  - s_modifier; Adv = Modifier; node(Foot).dummy.cat = value(adv);
  node(Adv).bot.adv_kind = value(~très|intensive|equalizer); % restrictions on adverbs
  ...
}
```

Mais nous avons aussi la classe **cnoun_as_adv** qui permet d'utiliser les groupes nominaux (cat=N2) portant le trait **time** (avec une valeur non négative) comme modifieur de temps.

```
class cnoun_as_adv {
  node N2: [cat: N2, id:time_mod, type: subst];
  node(N2).top.time = value(~ -);                  % N2 should carry a time property (other than -)
  - s_modifier; N2 = Modifier; node(Foot).dummy.cat = value(adv); % N2 acts as an adv
}
```

La classe **bodypart_cnoun** (qui hérite de la classe **cnoun** des groupes nominaux) permet d'utiliser des parties du corps humain comme modifieurs phrastiques, à la condition qu'elles soient elles-mêmes modifiées (à droite), par exemple par un adjectif (*tête courbée, il avance dans le vent*).

```
class bodypart_cnoun_as_modifier {
  <: cnoun;                                         %% inherit from nominal phrase
  desc.@kind0 = value(-);
}
```

```

- s_modifier; Modifier = N2Root; node(Foot).dummy.cat = value(adv); % s_modifier acting as adv
node(Root).cat = value(~ N2); %% modifier can be attached on event nouns
node(Root).bot = node(Foot).top;
node(Anchor).bot.semtype = value(bodypart); %% semantic property bodypart
node(N2).adj = value(strict); %% right adjoining is mandatory on N2
node(N2).adjleft = value(no);
node(N2).adjwrap = value(no); }

```

De nombreuses classes requérant - **s_modifier**² existent, par exemple pour les groupes prépositionnels (**_prep_s_modifier**), les subordonnées (**_csu_s_modifier**), les participiales (**participiale**), les concessives (**concessive_relatives**), les comparatives (**comparative_as_vmod_mod**), les interjections (**_pres_s**), le rang (**cnoun_as_position_mod**), les apostrophes (**audience_on_s**), les positions spatiales (**position_on_s**, **address_on_s**), les pronoms modifieurs (**pronoun_as_mod**), certaines locutions (**verb_ilya_as_time_mod**, **verb_voici_as_time_mod**, **verb_oblige_as_mod**), etc.

Nous illustrons (partiellement) le cas de *il y a*, où nous combinons la notion de s_modifier avec les contraintes d'une structure verbale canonique (<: **verb_canonical**), qui elle-même, en cascade, va requérir les divers composants formant une phrase complète autour d'un verbe.

```

class verb_ilya_as_time_mod {
  <: _verb_canonical; % inherits canonical verb construction
  - s_modifier; S = Modifier;
  node v:[cat:v, lex: avoir];
  desc.ht.imp = value(+); % impersonal subject expected
  desc.ht.extraction = value(-); % argument extraction not allowed
  ... }

```

Par ailleurs, il est progressivement apparu que les modifieurs de phrases se retrouvent autour d'autres catégories syntaxiques que 'S' et 'VMod', comme par exemple après un mot coordonnant (catégorie 'coo'), après une préposition ('prep') ou un nom déverbal dénotant un événement. On voit que l'on force le modifieur à suivre le coordonnant (post position).

```

class modifier_on_coo {
  + s_adj_pos;
  node(Root).cat = value(coo);
  desc.position = value(post); % modifier on the right side of coo
  node(Foot).top.modifier = value(-); % restriction: no more than one modifieur per coo node
  node(Root).bot.modifier = value(+); }

```

Dans les cas des compléments venant se positionner à droite d'une préposition, nous contraignons (de manière très arbitraire) la préposition à être *avec*, *pour*, ou *contre*, ce qui semble recouvrir les cas les plus fréquents. Bien sûr, cette liste est facilement extensible, ou peut être remplacée à terme par une propriété. Il est aussi à noter que les marques d'incises sont considérées comme obligatoires, ce qui n'est pas toujours vérifié dans la réalité.

```

class modifier_on_prep {
  + s_adj_pos;
  node(Root).cat = value(vmodprep);
  desc.position = value(post); % modifier on the right side of prep
  node(Root).bot = node(Foot).top;
  node(Incise).adj = value(strict); % mandatory parenthetical marks
  node(Incise).bot.incise_kind = value(comastrict|par|dash);
  node(Foot).top.pcas = value(avec|pour|contre); %% only some prep may be modified

```

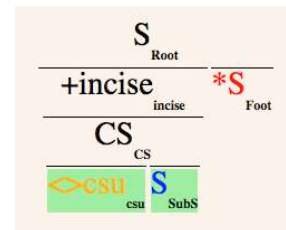
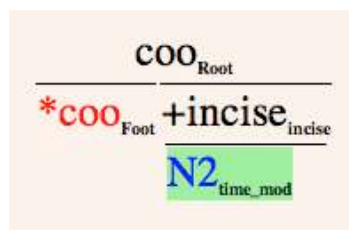
Enfin, l'application de certains adverbiaux sur des groupes nominaux est possible mais contrôlée par la présence de la valeur **event** pour le trait **semtype**, portée par les groupes nominaux dont la tête est

² ou parfois des sous-composants de **s_modifier** (plus hauts dans la hiérarchie) dans le cas de contraintes plus spécifiques.

un nom événementiel (souvent un déverbal) comme *départ* ou *annonce*. Il est à noter que ces modifieurs se distinguent des autres modifieurs du nom (adjectifs, groupes prépositionnels, relatives, etc) ainsi que des adverbes modifieurs de noms (*Deux cents francs environ ont été volés*).

```
class mod_on_N2 {          %% some event nouns may be modified like verbs
+ s_modifier;           %% example: la tenue, ce matin, d'une assemblée a évité le désastre
- x_modifier;
- s_adj_pos;
node(Root).cat = value(N2);
node(Root).bot.semtype = value(event);          %% semantic property event
}
```

Au final, l'utilisation de ces classes permet d'engendrer un ensemble conséquent d'arbres, à savoir 79 arbres, ce qui est beaucoup si on compare avec les 38 arbres de FRMG actuellement ancrés par des verbes pour couvrir les constructions canoniques et celles avec extractions (interrogatives, relatives, clivées, topicalisation), ce pour les voies actives et passives. Cependant, à la différence des arbres verbaux qui sont très complexes, les arbres pour les compléments sont en général très simples, comme illustré par les deux arbres ci-dessous, le premier pour les groupes nominaux temporels accrochés sur une coordination et le second pour les subordinées en tête de phrase.



Un décompte effectué sur les résultats d'analyse syntaxique de 10096 phrases du French TreeBank (sur du style journalistique) montre que ces 79 arbres représentent presque 8% de tous les arbres utilisés. Le plus fréquent (4,3%) correspond aux groupes prépositionnels accrochés sur les **Vmod**, le deuxième (0,8%) étant pour les groupes prépositionnels en tête de phrase. Les deux suivants sont pour les adverbes. Seulement ensuite arrivent les nominaux temporels (0,35%) et subordinées (0,13%). Il semble donc nécessaire de multiplier les arbres pour capturer ces divers types de compléments, en notant cependant que beaucoup de ces compléments ont un taux d'occurrence très faible.

Limitations

Malgré l'aide apportée par les classes présentées ci-dessus (particulièrement celles fournissant **+s_modifier**), il reste difficile de capturer la diversité des compléments de phrase, en particulier ceux liés à des locutions. Il faut effectuer un travail de recensement long et laborieux, pouvant conduire à un ensemble important d'arbres très rarement utilisés. Pour éviter de trop multiplier le nombre d'arbres, il faut continuer à rechercher des généralisations derrière ces locutions.

Par ailleurs, certains compléments sont faciles à décrire mais sont en fait contrôlés par des propriétés sémantiques sur le modifieur (temps, spatial, partie du corps humain, événements, etc). La non-prise en compte de ces contraintes peut avoir un impact important en terme de sur-génération et en terme d'efficacité pour un analyseur syntaxique. Or, la mention de ces propriétés sémantiques au sein des entrées lexicales est une tâche difficile et de longue haleine. Notons cependant que l'utilisation de versions non contraintes sur un treebank ou sur un large corpus peut permettre d'apprendre les classes sémantiques en question, pour les injecter dans le lexique ou, de manière moins stricte, pour les utiliser en tant que préférences pendant la phase de désambiguïsation.

De manière plus générale, il y a nécessité d'un dialogue entre les développeurs de FRMG et du lexique Lefff (SAGOT 2010) pour recenser, implanter et exploiter les informations lexicales utiles. Ainsi, Lefff fournit des informations assez précises sur divers types d'adverbes (SAGOT B & al 2009), informations non encore complètement prises en compte pour le contrôle des adverbes dans FRMG. Encore plus

précisément, on peut espérer que certaines de ces informations puissent indiquer si les marques d'incise sont obligatoires ou non pour un adverbe donné dans certains contextes (par exemple entre l'auxiliaire et le participe).

Régulièrement, de nouvelles positions d'accroche des compléments de phrase émergent, comme par exemple sur les noms événementiels. Comme nous l'avons vu, l'ajout dans la méta-grammaire d'un nouveau lieu n'est pas très complexe, en s'appuyant sur les classes existantes. Néanmoins, l'impact sur les temps d'analyse peut être important.

Enfin, il est à noter que l'accroche des compléments sur certains lieux comme les coordinations et les prépositions relève plutôt d'une approche de syntaxe superficielle, résultant des limites des TAG. Il serait plus judicieux d'avoir un traitement plus sémantique avec l'indication d'une portée (*scope*) pour ces compléments. Ceci sera peut-être rendu possible par une transformation des sorties syntaxiques (sous forme de dépendances) vers des structures plus profondes. Une autre alternative envisagée est l'utilisation de TAG multi-composants (MC-TAG) comme formalisme, en place des TAG actuelles.

Conclusion

Nous avons montré comment la description modulaire de diverses propriétés syntaxiques, comme la notion de modifieur et de marques d'incises, permet une gestion plus facile et finalement assez économique des compléments de phrase, qui recouvrent un ensemble large et hétéroclite de constructions. La définition de fonctionnalités de base (**shallow_auxiliary**, **x_modifyer**, **s_modifyer**, **incise**, etc) associée à des mécanismes de contrôle (via des valeurs dans des structures de traits comme **desc.position = value(post)**) fournit rapidement une bibliothèque dans laquelle on peut piocher pour décrire de nouvelles constructions syntaxiques. En jouant le jeu, parfois un peu contraignant, de la recherche du bon niveau de généralité, on facilite également à terme l'application à des cas connexes non initialement prévus. Ainsi, l'extension de l'accroche des compléments de phrase sur les coordonnants, les prépositions et noms événementiels s'est faite très naturellement. Ce travail nous semble au final une confirmation de l'utilité des méta-grammaires pour des descriptions syntaxiques relativement fines, même si nous restons conscient des difficultés des langues et du besoin d'informations complémentaires (sémantiques en particulier).

Références

- BARRIER S.; BARRIER N. (2004), *Metagrammars: a new implementation for FTAG*. *Actes du colloque TAG+7, Vancouver*.
- BARRIER N. (2002), *Une Métagrammaire pour les adjectifs du français*. *Actes du colloque TALN 2002, Nancy*.
- CRABBE, B. (2005), *Représentation informatique de grammaires fortement lexicalisées - Application à la grammaire d'arbres adjoints*. *Ph. D. thesis, Université Nancy 2*.
- CANDITO M.-H. (1999), *Organisation modulaire et paramétrable de grammaires électroniques lexicalisées*. *Thèse de doctorat de l'Université Paris 7*.
- DE LA CLERGERIE E. (2005). *From metagrammars to factorized TAG/TIG parsers*. In **Proceedings of IWPT'05** (poster), p. 190–191, Vancouver, Canada.
- GAIFFE B.; CRABBE B; ROUSSANALY A. (2002), *A New Metagrammar Compiler*. In **Proceedings of the 6th International Workshop on Tree Adjoining Grammars and Related Frameworks (TAG+6)**, Venice.
- JOSHI A. K.; LEVY L.; TAKAHASHI M. (1975), *Tree Adjunct Grammars*. *Journal of Computer and System Science* 10, 10(1), 136–163.
- PARMENTIER Y; LE ROUX J. (2005), *XMG: a Multi-formalism Metagrammatical Framework*. - **ESSLLI'2005**.
- SAGOT B. (2010), *The Lefff, a freely available and large-coverage morphological and syntactic lexicon for French*. In **Proceedings of the 7th Language Resources and Evaluation Conference (LREC'10)**, La Valette, Malte.
- SAGOT B.; FORT K.; VENANT F. (2009), *Extension et couplage de ressources syntaxiques et sémantiques sur les adverbes*. In **Linguisticae Investigationes** 32(2) (pages 305-315)
- THOMASSET F ; DE LA CLERGERIE E. (2005), *Comment obtenir plus des méta-grammaires*. In **Proceedings of TALN'05**, Dourdan, France.