

Traitement des verbes dans une grammaire catégorielle du français : le coup de GraCE

Treating verbs in a French categorial grammar with GraCE

Frédérique Segond



Édition électronique

URL : <http://journals.openedition.org/praxematique/1892>

DOI : [10.4000/praxematique.1892](https://doi.org/10.4000/praxematique.1892)

ISSN : 2111-5044

Éditeur

Presses universitaires de la Méditerranée

Édition imprimée

Date de publication : 1 janvier 1994

Pagination : 13-33

ISSN : 0765-4944

Référence électronique

Frédérique Segond, « Traitement des verbes dans une grammaire catégorielle du français : le coup de GraCE », *Cahiers de praxématique* [En ligne], 22 | 1994, document 2, mis en ligne le 01 janvier 2013, consulté le 08 septembre 2020. URL : <http://journals.openedition.org/praxematique/1892> ; DOI : <https://doi.org/10.4000/praxematique.1892>

Ce document a été généré automatiquement le 8 septembre 2020.

Tous droits réservés

Traitement des verbes dans une grammaire catégorielle du français : le coup de GraCE¹

Treating verbs in a French categorial grammar with GraCE

Frédérique Segond

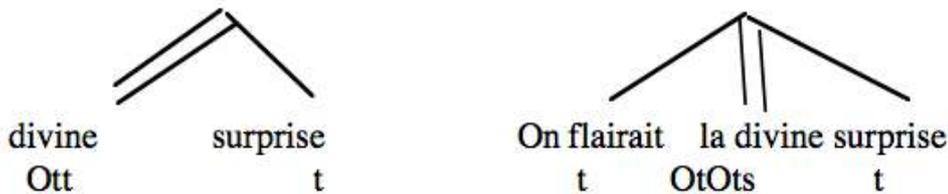
1. Introduction

- ¹ L'origine des grammaires catégorielles remonte au début de ce siècle avec les travaux de logiciens tels que Lesniewski (1929) et Ajdukiewicz (1935). Dans les années cinquante Bar-Hillel (1953) et Lambek (1958) reprennent et développent ces idées. De son côté Shaumyan (1965) développe un modèle cousin : la Grammaire Applicative Universelle.
- ² Après avoir été un modèle peu utilisé, le modèle de la grammaire catégorielle connaît un renouveau avec les travaux de Montague (1970). Enfin, depuis les années 80, c'est un modèle qui est très en vogue dans le milieu de la linguistique informatique (cf. les travaux de Bach, Steedman, Moortgat, Bès, Desclés et de leurs équipes). Il existe, en outre, des liens de parenté entre les grammaires catégorielles et les grammaires d'unification². Nous avons pour notre part développé une grammaire catégorielle informatique du français³. C'est le traitement des verbes dans le cadre de ce travail que nous nous proposons de décrire ici.

2. Les grammaires catégorielles

- ³ Il existe différents modèles de grammaires catégorielles selon qu'ils utilisent ou non l'opération de montée du type, les combinateurs de la logique combinatoire, l'unification, le calcul de Lambek, la superposition, etc. Cependant, si ces différents modèles se regroupent sous le nom de grammaires catégorielles, c'est qu'il est bien évidemment possible de dégager des points communs.

- 4 Toute grammaire catégorielle est fortement lexicalisée au sens où l'essentiel de l'information se trouve dans le lexique ; les règles de la grammaire à proprement parler sont en très petit nombre, ce qui rend les grammaires catégorielles attrayantes pour le traitement automatique de la langue.
- 5 Tous les modèles de grammaires catégorielles partent d'un ensemble d'atomes, les noms (notés t) et les phrases (notées s), à partir desquels sont récursivement engendrées toutes les autres catégories grammaticales selon le principe de base suivant :
- 6 Tous les atomes sont des types :
Soient x et y deux types ; alors Oxy est un type.
- 7 L'idée de base est donc de « repenser » la notion de catégorie grammaticale en termes d'une relation opérateur-argument. On considérera, par exemple, la catégorie des adjectifs comme l'ensemble des opérateurs qui s'appliquent à un syntagme nominal pour former un syntagme nominal plus complexe. Les adjectifs reçoivent donc le type Ott⁴ (Figure 1). La catégorie des verbes transitifs est l'ensemble des opérateurs à deux arguments qui s'appliquent à deux syntagmes nominaux (le complément d'objet direct et le sujet) pour former une phrase. Les verbes transitifs reçoivent donc le type OtOts⁵ (Figure 1).

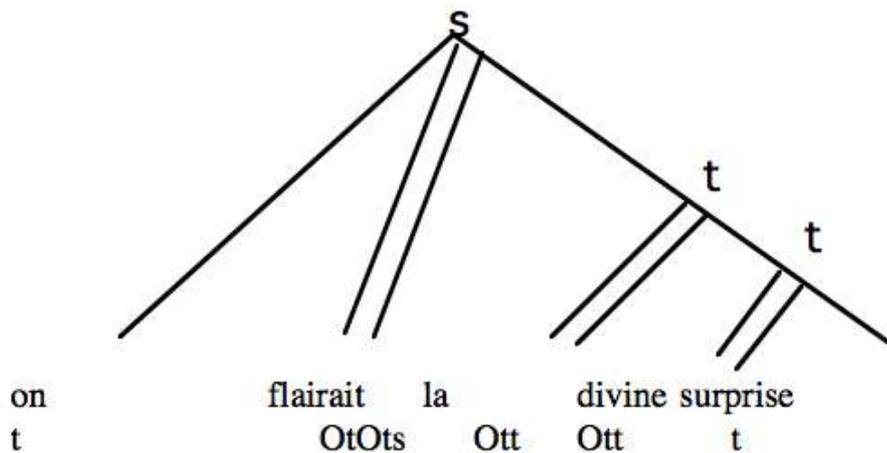


- Figure 1 -

- 8 Une fois que chaque mot a été associé, dans le dictionnaire, avec un ou plusieurs types (polymorphisme⁶), on définit les règles permettant de construire les structures de la langue (règles de grammaire) avec comme opération de base l'opération d'*application fonctionnelle*. Un opérateur pouvant attendre son argument soit à sa gauche, soit à sa droite, on a deux types d'applications : l'*application-avant* (notée >) et l'*application-arrière* (notée <).

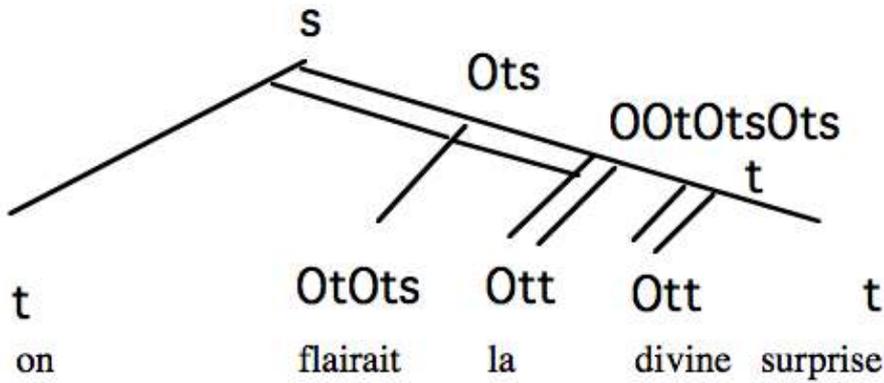
Application-avant (>) : Oxy x -----> y

Application-arrière (<) : x Oxy -----> y

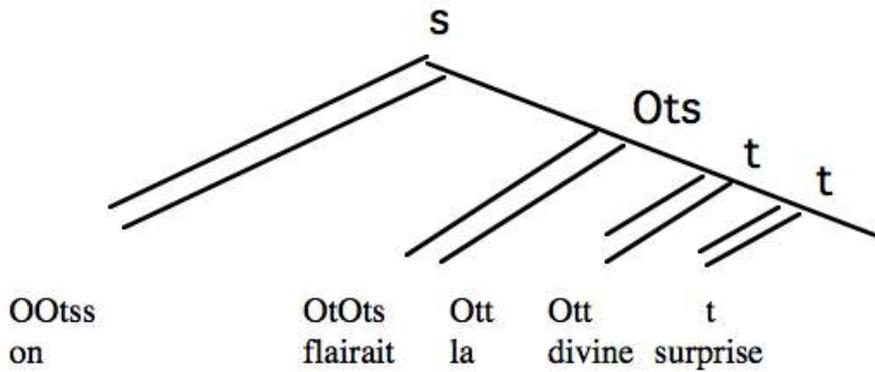


- Figure 2 -

- 9 Ainsi défini, le modèle de la grammaire catégorielle est contraignant en ce sens qu'il n'autorise de simplification (application des règles d'application-avant ou d'application-arrière) qu'entre syntagmes contigus dans la phrase. Or chacun sait que les corpus abondent de ce qu'on appelle les phénomènes périphériques tels que l'ellipse, l'inversion, l'incise, les constructions discontinues etc. L'exception est le lot quotidien de ceux qui s'intéressent à la description d'un corpus⁷. Une théorie grammaticale doit être en mesure de traiter, en particulier, des phénomènes discontinus.
- 10 Quiconque utilise le modèle des grammaires catégorielles est donc amené, tôt ou tard, à proposer des extensions. Nous ne ferons pas ici l'exposé détaillé de toutes les extensions qui ont été proposées au modèle initial⁸. Nous nous bornerons à exposer l'opération de montée du type⁹ afin de la comparer à l'opération de superposition, initialement définie par Shaumyan, et que nous avons utilisée pour le traitement de certaines formes verbales telles que les infinitifs et les participes présents.
- 11 L'opération de montée du type consiste à faire en sorte qu'un mot-argument devienne un mot-opérateur qui agit sur son opérateur. On inverse les rôles opérateur/argument en modifiant le type de l'argument initial. Par exemple, un syntagme nominal en position d'objet (de type t) va pouvoir changer de type et devenir un opérateur qui agit sur le verbe transitif pour former un prédicat (de type OOtOtsOts). C'est ce que l'on appelle la *montée de l'objet* (Figure 3). Formellement il existe quatre opérations de montée du type¹⁰ mais en pratique deux seulement sont utilisées : la *montée de l'objet* et la *montée du sujet* (Figure 4).

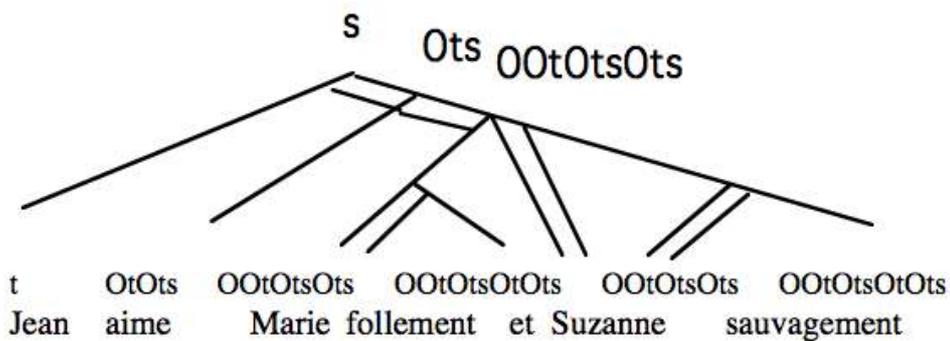


- Figure 3-



- Figure 4-

- 12 Cette opération de montée du type combinée avec l'utilisation de certains combinateurs empruntés à la logique combinatoire (essentiellement le combinateur de composition $B : B(Oxy Ozx \rightarrow Oyz)$) permet de traiter de phénomènes discontinus comme, par exemple, certains cas de coordination. L'exemple classique de coordination à constituants discontinus est donné Figure 5.

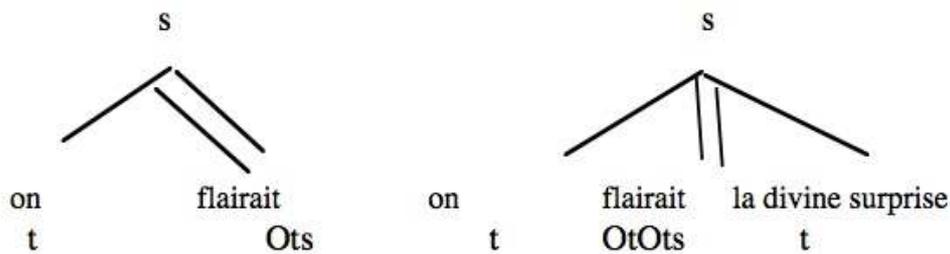


- Figure 5-

- 13 Cependant il apparaît aussi très rapidement que la montée du type est une opération trop puissante qu'il va falloir contrôler si l'on veut éviter l'explosion combinatoire. En effet, si l'on ne contrôle pas, d'une manière ou d'une autre, l'opération de montée du type, on obtient pour une phrase aussi simple que *Jean aime Marie* quatre structures

dictionnaire de machine d'environ 50.000 lemmes¹⁵. Ce dictionnaire comporte une composante morphologique qui définit les formes fléchies et les catégories syntaxiques ainsi qu'une composante morpho-syntaxique qui définit un certain nombre d'attributs (en particulier les sous-catégorisations de près de 6000 verbes).

- 18 Comme nous n'avions pas à notre disposition un dictionnaire de types, la première tâche a été de traduire ce dictionnaire en un dictionnaire de types. La stratégie adoptée a consisté à transformer toutes les parties du discours associées aux mots du dictionnaire en expressions typées. Cette traduction est faite sous forme de règles que nous appelons règles lexicales.
- 19 Une fois chacun des mots associé à un ou plusieurs types catégoriels, les règles d'application-avant et d'application-arrière agissent entre expressions typées. Par exemple, le mot *manger* qui est associé dans le dictionnaire de départ à la catégorie verbe devient une expression qui a pour forme de surface *manger* et pour type catégoriel l'attribut OtOts.
- 20 Les verbes sont en outre associés, dans le dictionnaire initial, à un ensemble d'attributs décrivant leur sous-catégorisation. Ainsi des attributs spécifiques distinguent les verbes copulatifs, les verbes à double compléments, les verbes transitifs, intransitifs, se construisant avec telle ou telle préposition. Ces distinctions ont été utilisées par l'analyseur catégoriel.
- 21 Comme nous l'avons décrit plus haut les verbes transitifs sont des opérateurs qui s'appliquent à deux arguments, le complément d'objet direct d'une part, et le sujet d'autre part. Ils reçoivent donc le type OtOts. De la même manière les verbes intransitifs sont des opérateurs qui s'appliquent à un argument (le sujet) pour former une phrase. Ils reçoivent donc le type Ots.
- 22 Dans notre travail cette distinction verbe transitif / verbe intransitif n'était pas importante, dans la mesure où nous avons pris le parti de poser que tout verbe transitif pouvait se comporter comme un verbe intransitif. Ce n'est bien évidemment pas vrai de tous les verbes transitifs. Cette hypothèse par défaut se justifie pourtant du point de vue du traitement automatique dans la mesure où de nombreux verbes répertoriés comme transitifs peuvent s'employer avec une ellipse du complément d'objet direct : emploi générique du verbe, coordination, etc. Tous les verbes transitifs reçoivent les deux types.



- Figure 7 -

- 23 Considérons maintenant le cas des verbes prédicatifs. Les verbes prédicatifs ont ceci de particulier qu'ils se construisent avec un attribut du sujet. Cet attribut du sujet peut être soit un adjectif, soit un syntagme nominal, soit un participe passé¹⁶ :

Jean est professeur

Jean est un bon professeur
 Paul devient loquace
 L'enfant est agacé

- 24 Les verbes prédicatifs sont donc des opérateurs qui s'appliquent à des adjectifs ou à des termes pour former des prédicats. Ils reçoivent donc les types : OOtOtS et OtOtS.
- 25 Remarquons au passage que ce type de solution, somme toute classique, ne va pas sans poser un problème de surgénération dans le cas de verbes se construisant avec l'auxiliaire *être* et qui sont, par ailleurs des noms, comme c'est le cas du verbe *mourir* au participe passé. En effet *il est mort* aura trois analyses : être + nom, être + pp, être (auxiliaire) + verbe. Ce problème est général à tous les types d'analyseurs, mais reconnaissons qu'il est plus difficile à gérer dans un système de type catégoriel étant donné qu'il impose de définir des choix, des préférences qui tiennent compte du contexte et donc d'alourdir considérablement l'écriture des règles, ce qui fait perdre le bénéfice du formalisme des types.
- 26 Nous avons également traité du cas des verbes se construisant avec une préposition comme c'est le cas de *commencer à*. Différentes solutions s'offrent à nous dans le cadre catégoriel : soit coder la préposition comme un opérateur agissant sur le verbe, soit écrire pour ce cas particulier des « règles lexicales » permettant de rattacher directement le verbe à la préposition. Dans le premier cas, soit on pratique le polymorphisme et on ajoute à la liste des types associés aux prépositions les types OtOtOtS ou OtOtS, soit on enrichit la notion de type.
- 27 La solution du polymorphisme revient à ne pas utiliser l'information dictionnaire concernant la sous-catégorisation des verbes. Tous les rattachements prépositionnels vont donc être produits, y compris ceux qui ne sont pas justifiés linguistiquement. Cette absence de choix préalable garantit de produire la bonne analyse parmi toutes celles construites, au prix d'une éventuelle explosion combinatoire. Ainsi on obtient, pour la phrase, *je dîne à huit heures*, les deux analyses suivantes :
- (Je (dîne à) huit heures))
 (Je dîne (à huit heures))
- 28 La deuxième solution consiste à ne plus considérer un type comme la représentation fonctionnelle simple d'une partie du discours mais comme une structure plus complexe contenant, en plus du type syntaxique, des attributs.
- 29 Pour l'exemple qui nous intéresse on préciserait donc que le verbe-argument de la préposition doit avoir un trait précis dans le dictionnaire.
- 30 Ainsi au cours de l'opération d'application on ne vérifierait plus uniquement la compatibilité entre les types mais aussi une compatibilité entre les attributs d'une liste plus ou moins longue associée au type catégoriel. C'est ce que l'on fait par exemple, grâce au mécanisme d'unification, pour vérifier les règles d'accord morphologiques et c'est ce qu'ont choisi de faire un certain nombre de chercheurs¹⁷. Cependant cette classification entraîne l'écriture de règles de plus en plus élaborées et, rapidement, on perd l'intérêt du formalisme des types pour revenir, en fin de compte, à un système « caché » de règles de réécriture classique. Nous avons donc, pour notre part, opté pour des règles lexicales séparées.
- 31 Reprenons le cas du verbe *commencer* qui se construit avec les prépositions *à* et *de*. Nous avons une règle telle que :

une expression de type OtOts ou Ots ayant le trait <+prep à de> dans le dictionnaire combinée avec la préposition à ou de devient une expression ayant pour type le type de l'expression initiale.

- 32 Remarquons au passage que cela nous permet de rendre également compte de l'insertion d'adverbes. Penchons-nous sur la construction de l'expression *commencer soudain* à. L'opérateur *soudain* s'applique à l'argument *commencer* pour former la suite lexicale *commencer soudain* qui est une expression de type OtOts, porteuse du trait <+prep, à, de> qu'elle a hérité de *commencer*. Elle peut donc se composer avec la préposition à pour former l'expression *commencer soudain* à. Cette solution est une extension du formalisme des types.
- 33 Le même type de solution a été adopté pour les verbes se construisant avec un complément indirect introduit par une préposition (paradigme de *donner*). Supposons qu'au cours du processus d'analyse une expression de type X porteuse du trait <ditrans> ait été créée. Si cette expression est suivie dans la chaîne linéaire de la phrase par la préposition adéquate (à dans le cas de *donner*) alors on crée une nouvelle expression dont le type est OtX. Autrement dit on ajoute un argument à l'expression porteuse du trait <ditrans>.
- 34 Notons que cette solution nous permet de traiter aussi bien de la suite *x donne quelque chose à y* que de la suite *x donne à y quelque chose* ainsi que des insertions adverbiales. En effet considérons la phrase *Eve donne une pomme à Adam*. L'expression *donne une pomme* a été formée, elle est porteuse du trait <ditrans>, qu'elle a hérité du verbe *donner* et elle a pour type Ots. La règle énoncée plus haut nous permet de former *donne une pomme à* qui a pour type OtOts. Nous pouvons alors former la suite *Eve donne une pomme à Adam*. De la même façon on forme *Eve donne à Adam une pomme*. Par application de la règle à *donner* on forme *donne à* qui est une expression de type OtOtOts. On récupère alors les trois arguments (*Eve, pomme, Adam*).
- 35 Remarquons enfin que cette solution ajoutée à celle proposée pour les prépositions nous permet de fournir les trois rattachements, donc trois analyses, pour la phrase *Il a donné un rendez-vous à midi* :
- ((Il a donné un rendez-vous) (à midi))
 (Il (a donné un rendez-vous à midi))
 (Il a donné (un rendez-vous à midi))
- 36 À ce stade de l'analyse nous ne faisons pas de choix. En effet, en l'absence de critères sémantiques ou pragmatiques, les trois rattachements sont justifiés : une analyse distributionnelle nous permet de substituer *pomme* à *rendez-vous* et *Paul* à *Midi* pour obtenir la deuxième analyse. D'autre part, si l'on substitue *chapeau* à *rendez-vous* et *plume* à *midi* on obtient le dernier rattachement.
- 37 Ce même type de règles lexicales a été utilisé pour traiter des verbes se construisant avec une que-Phrase, comme c'est le cas du verbe *dire*. En supposant que ces verbes soient porteurs du trait <+que-P> dans le dictionnaire, l'expression porteuse du trait <+que-P> suivie de la conjonction *que* forme une nouvelle expression qui a pour type OsOts.
- Il (dit que) le Beaujolais nouveau est arrivé.
 t OsOts

4. La superposition

- 38 Nous abordons maintenant le traitement de certaines formes verbales telles que les infinitifs et les participes présents. La solution que nous avons choisie pour traiter de ces formes se rattache à la notion de superposition que l'on trouve chez Shaumyan¹⁸. Cette notion décrit un phénomène linguistique très général : les mots, les syntagmes, ne sont pas attachés de façon rigide à des catégories syntaxiques données. Ainsi, selon le contexte, ils changent de comportement syntaxique, et pour le modèle qui nous intéresse, de type catégoriel. Ce phénomène de superposition est très général en linguistique et apparaît aussi bien au niveau syntaxique qu'au niveau morphologique, phonologique et fonctionnel. Nous ne présenterons pas ici la superposition dans le détail¹⁹ mais plutôt ses principes de base et son application aux formes verbales citées ci-dessus.
- 39 On peut présenter la théorie de la superposition comme une extension de la notion de translation que l'on trouve chez Tesnière :
- « la translation consiste à transférer un mot plein d'une catégorie grammaticale dans une autre catégorie grammaticale, c'est-à-dire à transformer une espèce de mot en une autre espèce de mot »²⁰.
- 40 Considérons les syntagmes nominaux anglais *gold watch*, *silk dress* et *wide open*. Dans ces syntagmes, les noms *gold* et *silk* se comportent comme des adjectifs, au sens où ils modifient des noms. De la même manière l'adjectif *wide* se comporte comme un adverbe au sens où il modifie un autre adjectif.
- 41 Dans la littérature catégorielle ces problèmes ne sont généralement pas abordés car considérés comme relevant du polymorphisme : les deux types nom et adjectif d'une part, adjectif et adverbe d'autre part, vont être associés respectivement à *gold*, *silk* et *wide* dans le dictionnaire. Pourtant, de notre point de vue, ce phénomène n'est pas du polymorphisme. En effet, pour qu'il y ait polymorphisme il faut que les différents types syntaxiques associés à un mot donné aient le même poids. Ainsi les adverbes sont des opérateurs polymorphes, de type Oxx où x peut être un adjectif, un verbe, une phrase ou un autre adverbe. Les conjonctions de coordination sont également des opérateurs polymorphes au sens où ils peuvent agir entre deux phrases, deux noms, deux adjectifs, etc. et, en règle générale, entre deux syntagmes de même type.
- 42 En revanche, il est malheureux d'associer les deux types nom et adjectif aux mots *gold* et *silk*. En effet, si nous avons un cas de polymorphisme, les mots *gold* et *silk* devraient pouvoir se comporter comme des adjectifs dans tous les contextes où un adjectif peut apparaître. Ainsi on devrait pouvoir écrire **the watch is gold* et **the dress is silk* de la même manière que l'on peut écrire *the blue watch* et *the watch is blue*. Or ce n'est pas le cas. Il en va de même pour l'adjectif *wide* qui ne se comporte comme un adverbe que dans un contexte bien défini. Une fois écartée la solution du polymorphisme, on est confronté à un paradoxe : les mots *silk* et *dress* ne peuvent être combinés du point de vue des types catégoriels (ils ont tous les deux le type t) et pourtant d'un point de vue syntaxique le syntagme *silk dress* est bien formé.
- 43 Nous proposons de résoudre ce paradoxe en posant que, dans des contextes déterminés, les mots acquièrent un nouveau type, un type secondaire, qui se superpose à leur type premier, le type qui leur est inhérent. Ce phénomène reçoit le nom de superposition.

- 44 La deuxième propriété importante de la superposition est qu'au cours de l'opération de superposition le mot qui acquiert un type secondaire conserve le comportement syntaxique attaché à son type inhérent.
- 45 Une bonne illustration de cette propriété est donnée par les infinitifs en français. Le type inhérent aux infinitifs est celui des verbes (OtOts ou Ots) et leur type secondaire est celui des termes (t). L'infinitif conserve les propriétés d'un verbe (il admet un sujet, un complément d'objet direct etc. pour former une proposition infinitive) jusqu'à un certain point où il change de type pour devenir un terme. Par exemple, considérons la phrase suivante :
- Boire du vin blanc est un devoir national
- 46 *Boire* est codé OtOts dans le dictionnaire. Dans la phrase précédente il prend le syntagme *du vin blanc* comme argument, formant alors l'expression *Boire du vin blanc* qui a pour type Ots. Cette expression ne trouve pas son deuxième argument dans la phrase et donc elle change de type pour prendre le type t et devenir l'argument du verbe *être*. En d'autres termes les infinitifs se comportent comme des verbes tant qu'ils peuvent opérer sur des arguments dans la phrase. Lorsqu'il n'y a plus aucun candidat pour être argument, ils changent de type pour devenir des termes, eux-mêmes arguments d'autres opérateurs dans la phrase.
- 47 La même chose est vraie des participes présents. Ils se comportent comme des verbes jusqu'à ce qu'ils changent de type pour se comporter comme des adjectifs ou comme des adverbes de phrases (reçoivent les types Ott ou Oss).
- 48 La notion de superposition est formalisée de la façon suivante :
- Soit E une expression de type x. On dit que y est superposé à x s'il existe un contexte dans lequel l'expression E change de type pour prendre le type y. Le type de E est représenté par la formule <y :x>.
- 49 Un opérateur de type Ox<y :x> est appelé superposeur

$$\text{R\`egle de superposition : } \quad \text{Ox<y:z> A} \quad \text{x B}$$

$$\text{<y:z>}$$

- 50 Le lecteur aura donc compris que l'opération de superposition est plus générale que la translation de Tesnière qui, elle, s'attache uniquement à la syntaxe.
- 51 La superposition est différente du polymorphisme comme l'a montré la discussion précédente.
- 52 Enfin la superposition n'est pas de la montée du type. En effet, la superposition décrit, en particulier, un changement de comportement syntaxique, et marque le passage, dans un contexte déterminé, d'une catégorie répertoriée à une autre catégorie répertoriée. La montée du type, partant de l'impossibilité de combiner deux types entre eux (par exemple t et t), fait émerger, par un procédé purement calculatoire, un type compatible. La catégorie qui émerge n'est pas forcément une catégorie répertoriée et n'a donc pas toujours d'interprétation syntaxique²¹. Même si, au bout du compte, on peut être tenté de comparer la superposition à de la montée du type contrôlée, il faut bien garder à l'esprit que les deux démarches sont très différentes.

- 53 Du point de vue de l'implémentation, nous avons défini pour les infinitifs et pour les participes présents des métarègles de changement de type. Lorsqu'un verbe à l'infinitif apparaît dans une phrase il est naturellement associé dans le dictionnaire à la catégorie verbe et à ce titre devient une expression de type Ots ou OtOts. Tant qu'il y a, dans la phrase, des constituants susceptibles d'être arguments du verbe à la forme infinitive, les règles d'application-avant et d'application-arrière s'appliquent librement. Cette métarègle consiste à spécifier à la machine qu'une expression de type Ots, OtOts, ou bien de type s porteuse du trait infinitif, devient une expression de type t.
- 54 Une métarègle analogue traite de la superposition pour les participes présents. Elle spécifie qu'une expression de type Ots ou s porteuse du trait participe présent devient une expression de type Oss ou Ott.
- 55 Nous terminerons sur une remarque concernant le traitement des participes passés. Considérons les phrases suivantes :
- Les yeux rougis par les larmes
Les cheveux décolorés par le soleil
- 56 Du point de vue sémantique, il est clair que *par les larmes* et *par le soleil* se rattachent respectivement à *rougis* et *décolorés*. Cette remarque plaide en faveur de la superposition pour les participes passés. Ils se comportent comme des verbes (admettent un complément) puis changent de type pour prendre le type Ott.
- 57 Mais cette solution pose des problèmes dans le cas d'un verbe transitif suivi d'un groupe nominal et remet en cause les choix faits pour les prépositions. On a donc choisi de coder les participes passés Ott et d'avoir des règles lexicales pour ce qui est des temps composés ; ceci nous permettant, à l'aide des traits du dictionnaire, d'associer chaque verbe avec son auxiliaire et d'éviter ainsi un bon nombre d'analyses parasites.

5. Conclusion

- 58 Dans cet article nous nous sommes attachée à décrire le traitement des verbes dans le cadre d'un formalisme de type catégoriel. Pas à pas nous avons exposé les raisons linguistiques et informatiques qui nous ont conduite à ces choix, dégageant par là même les problèmes attachés à ce type de formalisme.
- 59 Enfin nous avons présenté un concept linguistique général, la superposition, rarement abordé dans la littérature contemporaine²², qui, dans le formalisme des types, fournit une solution élégante à toute une classe de problèmes.
- 60 Nous concluons maintenant par quelques mots sur la couverture de l'analyseur GraCE ainsi que par des exemples d'analyses obtenues.
- 61 Il est, et c'est une chose qui est rarement mentionnée, assez difficile de rendre compte, de façon exacte, de la couverture d'un analyseur syntaxique. En effet, passé un certain stade de description, comment savoir que l'on n'a pas oublié de décrire telle situation (si on ne l'a pas rencontrée dans un texte), comment évaluer l'impact d'une règle sur l'ensemble d'un corpus, etc. Ainsi, dire que l'analyseur est en mesure de traiter, par exemple, des relatives, n'est pas très évocateur du type de couverture obtenu. Même lorsque le concepteur du système pense avoir décrit, dans le menu, ce type de constructions, les tests sur des phrases d'un corpus révèlent souvent des problèmes en suspens. L'enchevêtrement de tous les phénomènes périphériques dont nous avons signalé l'existence au début de cet article apparaît dans les corpus. Les articles de

journaux sont, comme chacun sait, fort instructifs, au moins pour ce qui est de tester des analyseurs syntaxiques.

- 62 Mais sans aller si loin, que le lecteur tente simplement la petite expérience que nous avons faite. Nous avons recherché dans un corpus d'un million de mots, une phrase telle que < sujet, verbe, objet > où le sujet et l'objet sont des groupes nominaux simples (au plus déterminant, nom, adjectif, peu importe l'ordre des constituants). Ce type de phrase *idéale* n'apparaît pratiquement jamais. On s'aperçoit très vite qu'en analyse syntaxique la règle est l'exception et donc l'exception la règle.
- 63 Il est certainement plus intéressant, pour donner une idée de la couverture d'un analyseur, d'indiquer le nombre, la taille et le type de textes que l'on a analysé (avec bien entendu le taux de réussite, et pourquoi pas, une description des problèmes en suspens).
- 64 En ce qui concerne l'analyseur GraCE commençons par énumérer les phénomènes que nous n'avons, délibérément, pas traités. Il s'agit des formes interrogatives, de certaines formes négatives, des cas de coordination entre constituants de nature différente, des dépendances illimitées, des phrases contenant une inversion du sujet (sauf infinitives et relatives) ou une inversion du groupe prépositionnel.
- 65 Notre corpus se compose d'une centaine de phrases extraites du *Monde*, de *Libération* et de 7 à Paris²³.
- 66 Toutes les phrases sont analysées. Le nombre des analyses obtenues pour une phrase donnée varie (le maximum étant 10 analyses pour une phrase). Parmi toutes les analyses on a toujours la bonne. La principale raison aux analyses multiples est la présence de prépositions en fin ou en début de phrase (plusieurs rattachements possibles). Sinon, les problèmes rencontrés sont ceux attachés au formalisme des types ; nous en avons décrit, plus haut, un certain nombre.

BIBLIOGRAPHIE

- Ajdkiewicz K. (1978) *The world-perspective and other essays*. 1931-1933, Dordrecht, Riedel publishing company.
- Bach E. (1987) « Some generalizations of Categorical grammars », in *The formal Complexity of Natural Language*, Savitch at al. ed., Reidel Publishing company.
- Bar-Hillel, Y. (1953) « A quasi arithmetical notation for syntactic description », *Language*, n° 29, 47-58.
- Baschung K. (1991) « Un analyseur UCG du français », *TA Informations*, 2, 71-88.
- (1990) *Grammaires d'Unification à traits et contrôle des infinitifs*, Nouveau doctorat, Clermont-Ferrand.
- Bès G., Gardent C. (1989) « French order without order », *Actes du colloque Association for Computational Linguistics*, Manchester, 249-255.

- Bouma G., (1989) « Efficient processing of flexible categorial grammar », Actes du colloque Association for Computational Linguistics, Manchester, 12-26.
- Buszkowski W., Marciszewski W. and Van Benthem J. (eds.) (1988) *Categorial grammar*, Amsterdam, John Benjamins Publishing company.
- Chanod J.-P. (1993) « Problèmes de robustesse en analyse syntaxique », Actes du colloque « Informatique et Langue Naturelle », Université de Nantes, 223-245.
- Corblin F. 1993 « Catégories et translation en syntaxe structurale », colloque international Lucien Tesnière aujourd'hui, Université de Rouen, Mont Saint-Aignan, Novembre 1992 (à paraître).
- Curry H.B., Fey R. (1958) *Combinatory Logic*, Amsterdam, North Holland.
- Desclés J.-P. (1990) *Langages applicatifs, langue naturelle et cognition*, Hermes.
- Desclés J.-P., Segond F. (1990) « Topicalization : Categorial analysis and Applicative Grammar », Actes du colloque Ordre des mots dans le cadre des Grammaires catégorielles, Clermont-Ferrand, France, Lecomte A. ed., Adosa, 13-37.
- Dowty D.(1988) « Type raising », in Oehrle et al. (ed.), 153-197.
- Heidorn G.E. (1972) Natural Language inputs to a simulation programming system, Technical report from the Naval Postgraduate school.
- Jensen K., Heidorn G.E., Richardson S. (1992) *Natural language Processing : the PLNLP Approach*, Kluwer Academic Publishers, Boston, Massachusetts, USA.
- Karttunen L. (1989) « Radical lexicalism » in Baltin & Kroch (eds.) *Alternative conception of phrases structures*, Chicago, University Press.
- Lambek J. (1958) « The mathematics of Sentence Structure », *American Mathematical Monthly*, n° 65, 154-165.
- Lesniewski S. (1929) « Grundzuge eines neuen Systems der Grundlagen der Mathematik », *Fundamenta Mathematicae*, 14 1-81.
- Montague R. (1970) « English as a formal language », B. Visentini et al. (eds), *Linguaggi nella società e nella tecnica*, Milan, Edizioni di Comunità, 189-224. Repris dans R.H. Thomason (1974) *Formal Philosophy : selected papers of Richard Montague*, New Haven : Yale University Press, 188-221.
- Moortgat M. (1988) *Categorial investigations : Logical and linguistic aspects of the Lambek calculus*, Groningen-Amsterdam Studies in Semantics, Foris publications.
- Oehrle R.T., Bach E., et Wheeler D. (eds) (1988) *Categorial grammars and natural language structures*, Amsterdam., D. Reidel publishing compagny.
- Segond F. (1990) « Approches des grammaires catégorielles », *Mathématiques, Informatique et Sciences Humaines*, n° 110, 47-60.
- Segond F. (1990) *Grammaire catégorielle du français. Etude théorique et implémentation. Le système GraCE (Grammaire Catégorielle Etendue)* Nouveau doctorat, EHESS-Paris.
- Segond F. (1991) « Un analyseur catégoriel pour le français. Une stratégie computationnelle pour les analyses multiples dans le cas des attachements de groupes prépositionnels », *TA Information*, 1, 13-26.
- Shaumyan S. (1965) *Strukturnaja lingvistika*, Moskva, Nauka.

Shaumyan S. (1977) *Applicative Grammar as a Semantic Theory of Natural Language*, University of Chicago Press.

Shaumyan S. (1987) *A semiotic theory of language*, Indiana University Press.

Shaumyan S., Segond F.

Shaumyan S., Segond F. (1992) « Discontinuous constructions and Applicative Universal grammar », Actes du colloque Mathematics of language, Austin, Texas, USA.

Shaumyan S., Segond F. (1993) « The theory of superposition of Applicative Universal Grammar » Actes du colloque « Informatique et Langue Naturelle, Université de Nantes, France, 179-206.

Shaumyan S., Segond F. (1994) « Long-Distance Dependencies and Applicative Universal Grammar », COLING, Kyoto, Japon.

Steedman M. (1990) : « Gapping as constituent coordination », *Linguistics and philosophy* n° 13, 207-263.

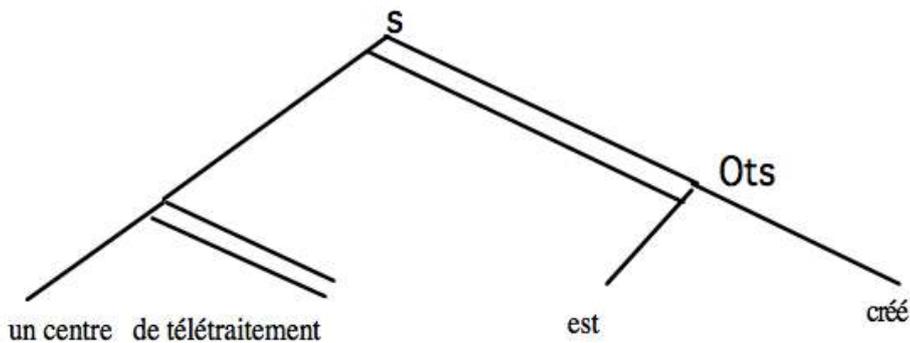
(1988) : « Combinators and Grammars » in Oehrle et al., 417-442.

Tesnière L. (1966) *Éléments de syntaxe structurale*, Paris, Klincksieck.

Uszkoreit H. (1986) « Categorical unification grammars » COLING, Bonn, 187-194.

ANNEXES

Un centre de télétraitement est créé.



Affichage obtenu avec l'analyseur GraCE :

```

SENT T      OTT      "Un"
           T T      "centre"
           OTT OTOTT "de"
           T      "télétraitement"
      OTS      OTS "est"
           VERB      "créé"
PUNC      "."
    
```

Pourtant ses contraintes de fabrication sont sévères.

SENT OXX "Pourtant"
 S T OTT "ses"
 T T "contraintes"
 OTT OTOTT "de"
 T "fabrication"
 OTS OOTTOTS "sont"
 OTT "sévères"
 PUNC "."

On superpose ces couleurs pour obtenir la gravure.

SENT S T "On"
 OTS OTOTS "superpose"
 T OTT "ces"
 T "couleurs"
 OSS OTOSS "pour"
 T OTOTS "obtenir"
 T OTT "la"
 T "gravure"
 PUNC "."

Il m'a répondu qu'il l'avait vue en vitrine

SENT T "Il"
 OTS OOTSOTS "m"
 OTS OSOTS OTS OTOTS "a"
 VERB "répondu"
 CONJ "qu"
 S S T "il"
 OTS OOTOTSOTS "l"
 OTOTS OTOTS "avait"
 VERB "vu"
 OSS OTOSS "en"
 T "vitrine"
 PUNC "."

La première Foire du Trône eut donc lieu à la date indiquée

```

SENT S      T OTT "La"
              T OTT "première"
                T T "Foire"
                  OTT OTOTT "du"
                    T "Trône"
              OTS OTOTS OTOTS "eut"
                OXX "donc"
              T "lieu"
    OSS OTOSS "à"
      T OTT "la"
      T T "date"
        OTT "indiquée"
PUNC "."
    
```

Dehors, un vent violent soufflait sur la terre gelée, mais dans la cuisine, il faisait bon.

```

SENT OXX ADV "Dehors"
      PUNC ","
    S S S T OTT "un"
              T T "vent"
                OTT "violent"
              OTS "soufflait"
    OSS OTOSS "sur"
      T OTT "la"
      T T "terre"
        OTT VERB "gelée"
          PUNC ","
    OXX OXOXX "mais"
      S OSS OTOSS "dans"
        T OTT "la"
          T NOUN "cuisine"
            PUNC ","
    S T "il"
      OTS OOTTOTS "faisait"
        OTT "bon"
PUNC "."
    
```

NOTES

1. Je remercie particulièrement Jean-Pierre Chanod, Gregory Grefenstette, Rachel Panckhurst, Sebastian Shaumyan et Annie Zaenen pour leurs commentaires et leurs critiques.
2. Voir, par exemple, (Uzbek, 1986) et (Karttunen, 1989).
3. Le système GraCE (Segond, 1990).

4. Dans cette notation, empruntée à Shaumyan, la lettre (ou le groupe commençant par O) suivant immédiatement le symbole O représente l'argument et la suivante (ou le groupe commençant par O) le résultant. Dans une notation à la Steedman Ots s'écrit s \ t.
5. Dans toutes les figures qui suivent les opérateurs sont représentés par des lignes doubles et les arguments par des lignes simples.
6. Par exemple *danse* est polymorphe puisqu'il est à la fois nom et verbe.
7. Pour une bonne description des phénomènes linguistiques apparaissant dans les textes français et sur la façon de traiter automatiquement de gros corpus le lecteur pourra se reporter aux travaux de Jean-Pierre Chanod (Chanod, 1993).
8. Pour un tel exposé le lecteur pourra se reporter, entre autre, à l'ouvrage de M. Moortgat (Moortgat, 1988).
9. En anglais *type raising*.
10. Voir Lambek (Lambek, 1958).
11. Ce phénomène est largement discuté dans la littérature sous le nom de *spurious ambiguity* (ambiguïté fallacieuse).
12. X' renvoie à la sémantique du mot X.
13. Le lecteur pourra se reporter à (Segond, 1990, 1991).
14. *Programming Language for Natural Language Processing* développé par G. Heidorn (Heidorn, 1972) et également utilisé pour l'écriture de PEG (PLNLP English Grammar) (Jensen et al., 1992) grammaire sous-jacente au logiciel CRITIQUE.
15. Dictionnaire développé au Centre scientifique d'IBM-France, Paris.
16. Nous reviendrons sur le traitement des participes passés et considérons, dans ce cas, qu'ils se comportent comme des adjectifs et sont codés comme tels (Ott).
17. Le lecteur pourra consulter, en particulier, les travaux de Gabriel Bès et Claire Gardent ainsi que ceux de Karine Baschung (Bès, Gardent, 1989) et (Baschung, 1990, 1991).
18. Cette notion apparaît avec celle de combinateur (Shaumyan, 1977), elle est développée plus en profondeur dans (Shaumyan, 1987).
19. Pour un exposé plus détaillé le lecteur pourra se reporter à (Shaumyan, 1987) et (Shaumyan & Segond, 1992, 1993).
20. (Tesnière, 1966 :364).
21. Pour une justification linguistique de l'opération de montée du type voir Steedman (Steedman, 1990). Pour une critique de cette justification voir Shaumyan et Segond (Shaumyan & Segond, 1992).
22. Voir les travaux de F. Corblin (Corblin, 1993) et J. P Chanod (Chanod, 1993).
23. En annexe, ci-après, quelques-unes des analyses obtenues avec le système GraCE.

RÉSUMÉS

Après une description succincte du modèle des grammaires catégorielles nous nous arrêtons sur le traitement des verbes dans le cadre du système GraCE. Ceci nous conduit à développer le concept linguistique de superposition.

Ce concept rend compte du fait qu'un mot n'est pas attaché de façon rigide à un type donné. Nous montrons comment, dans le cadre catégoriel, cette notion nous permet de traiter des cas où il est impossible de combiner deux types entre eux.

This paper presents a French categorial morpho-syntactic parser : the system GraCE. It focuses on the treatment of verbs. In the first section we give an overview on the categorial formalism. Then while presenting the treatment of verbs, we introduce the notion of superposition. This important linguistic phenomenon points out that words are not strictly attached to a given type. This enables us to resolve the problem of type incompatibility within the categorial framework.

AUTEUR

FRÉDÉRIQUE SEGOND

Rank Xerox Research Centre Meylan

frederique.segond@xerox.fr