



HAL
open science

Approche discursive et approche syntaxique des circonstants en corpus

Lydia-Mai Ho-Dac, Cécile Frérot

► **To cite this version:**

Lydia-Mai Ho-Dac, Cécile Frérot. Approche discursive et approche syntaxique des circonstants en corpus. Jun 2004. sic_00001226

HAL Id: sic_00001226

https://archivesic.ccsd.cnrs.fr/sic_00001226

Submitted on 7 Dec 2004

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.

Approche discursive et approche syntaxique des circonstants en corpus

Mai Ho-Dac, Cécile Frérot

Équipe de Recherche en Syntaxe et Sémantique, UMR 5610
Maison de la Recherche, Université Toulouse le Mirail, France
{hodac,frerot}@univ-tlse2.fr

Dans cette collaboration entre syntaxe et discours, nous souhaitons croiser les dimensions syntaxique et discursive pour étudier en corpus le fonctionnement des syntagmes prépositionnels (SP) circonstants. La notion de complément circonstanciel a fait l'objet de nombreux travaux et définitions (cf. Rémi-Giraud, 1998) ; on peut toutefois la caractériser de manière relativement consensuelle selon quatre composantes conjointes. D'un point de vue sémantique, le complément circonstanciel « précise l'idée du verbe en marquant la connexion de l'action avec un repère (temps, lieu, etc.) situé *autour* d'elle dans le monde des phénomènes » (Grévisse, 1993 : 477). La composition morphologique des circonstants peut couvrir une grande variété de configurations (adverbe, syntagme prépositionnel ou nominal, proposition subordonnée, etc.). Sur un plan syntaxique, les circonstants peuvent être intra ou extra prédicatifs, ces caractères distincts constituant les deux pôles d'un continuum plutôt qu'une opposition binaire (Vigier, 2003). Enfin, dans une approche discursive, les circonstants sont des expressions potentiellement cadratives en ce sens qu'elles posent un critère d'interprétation (le *repère* selon Grévisse) qui peut jouer sur un ensemble de propositions (Charolles, 1997) ; elles participent de ce fait à l'organisation du texte.

C'est à travers la problématique du rattachement prépositionnel que nous nous sommes intéressées au repérage automatique de SP intra ou extra prédicatifs dans un analyseur syntaxique de corpus, SYNTAX (Bourigault & Fabre, 2000)¹. Nous avons cherché à identifier différents types de compléments (compléments essentiels, circonstanciels) et à déterminer la nature de leur rattachement (au verbe ou à la phrase). Cette identification s'appuie sur deux indices : la productivité et la mobilité syntaxique. L'indice de productivité (Fabre & Frérot, 2002) est calculé à partir des informations présentes dans le corpus ; il s'appuie sur deux valeurs : la productivité du verbe vis-à-vis de la préposition (désormais *prod1*, qui évalue le degré de cohésion entre le verbe et son SP) et la productivité du nom constitutif du SP vis-à-vis de la préposition (désormais *prod2*, qui évalue le degré de cohésion interne du SP). L'hypothèse, validée en corpus dans (Fabre & Frérot, op. cit.), est que le croisement de ces deux valeurs permet de mettre au jour différents types de compléments (une valeur *prod1* élevée et une valeur *prod2* nulle tendent à révéler des compléments de nature argumentale ; à l'inverse, une valeur *prod1* nulle et une valeur *prod2* élevée tendent à révéler des compléments de nature circonstancielle). L'indice de mobilité syntaxique (Merlo & Leybold, 2001) est un critère linguistique largement mis en œuvre dans les tests linguistiques (Miller, 1997 et Bonami, 1999) et correspond au fait qu'un élément peut apparaître dans différentes positions dans la phrase. Pour les SP circonstants, nous nous sommes concentrées sur les

¹ Le rattachement automatique des groupes prépositionnels est effectué dans notre analyseur de manière endogène, sur la base d'indices recueillis à partir de contextes non ambigus dans le corpus. Parmi ces indices, une mesure permet d'évaluer le degré de cohésion du GP avec son recteur : la productivité du couple recteur-préposition (prodRecteur). Une prodRecteur élevée est l'indice que le recteur se construit régulièrement avec la préposition, dans la mesure où les mots régis par cette relation sont diversifiés. Par exemple, le couple (*alterner, avec*) est productif, avec une prodRecteur de 7, parce qu'il gouverne sept régis différents : *alterner avec+det (cendre|crue|lit|masse|période|section|surface)*.

positions détachées, marquées à gauche et à droite par une ponctuation. L'hypothèse ici est qu'un SP apparaissant dans différentes positions en corpus tend à être extra prédicatif.

Ces deux indices sont mis en œuvre dans notre analyseur qui exploite les phrases indépendamment les unes des autres, sans s'appuyer sur le niveau extra-phrastique. C'est précisément à ce niveau que nous nous intéressons dans ce travail, en cherchant à explorer un terrain dépassant le cadre strict de la phrase pour prendre en compte l'organisation textuelle, en d'autres termes la dimension discursive. Nous posons ici l'hypothèse qu'il existe des portions de texte dont le contenu est fortement organisé par le biais des circonstants, notamment de lieu et de temps (Ho-Dac *et al.*, 2001 et Virtanen, 1992). Ces portions font apparaître des « chaînes » de SP circonstants qui, dans ce contexte d'apparition, tendent à être plutôt extra prédicatifs, comme l'illustre l'exemple suivant².

En Bretagne, les Gaullistes ne présentent pas, en 1958, une candidature partout, ils sont absents de 10 circonscriptions. Ils totalisent 16,5 % des suffrages et font élire 6 de leurs députés.

Au sein de la Gauche, le PC recule très fortement, de 7 points par rapport à 1956, recueillant 11 % **sur l'ensemble Grand Ouest** pour 18,9 % en moyenne française. Il conserve des positions au-dessus de 20 % **dans six circonscriptions en Bretagne** : Loudéac, Guingamp et Lannion **dans les Côtes-du-Nord**, et Quimper, Douarnenez et Quimperlé **dans le Finistère**. Partout ailleurs, il chute fortement. **En Basse-Normandie**, où en dehors de quelques isolats (Dives-Sur-Mer **dans le Calvados**, quelques villages miniers), il n'a jamais réussi à s'implanter électoralement, il réalise de très faibles scores.

La Gauche non communiste représentée par la SFIO, les radicaux-socialistes et quelques divers gauche, recule elle aussi fortement, plus fortement qu'au niveau national et que la composante communiste **en Bretagne**. Ainsi la Gauche non communiste enregistre 12 points de moins **en Bretagne** que la moyenne nationale.

Dans cet exemple, nous avons une chaîne de SP circonstants de lieu qui organisent le contenu selon l'expression de la localisation spatiale. En général, ce rôle structurant est attribué uniquement aux SP détachés en initiale. Ces SP qualifiés d'introducteurs de cadre (Charolles, 1997) ont un statut particulier puisqu'ils peuvent avoir une portée dépassant leur phrase d'accueil, comme l'illustre le premier *En Bretagne* de l'exemple. Cependant, lorsque nous avons affaire à des chaînes constituées de circonstants d'un même type sémantique et particulièrement de lieu ou de temps, ce sont tous les circonstants qui structurent le contenu et pas seulement les introducteurs de cadres. En effet, si l'on ne s'en tenait qu'à ces derniers, l'exemple ne serait structuré que selon trois localisations spatiales (En Bretagne, . Partout ailleurs, . En Basse-Normandie,), ce qui n'est véritablement pas le cas. L'enjeu est alors de délimiter ces chaînes, particulièrement lorsque l'on étudie les introducteurs de cadres et leur portée, amoindrie dans ces contextes d'apparition. Il y a donc un besoin de repérer l'ensemble des circonstants, et pour cela, les mesures de productivité et de mobilité constituent des indices formels utilisables.

Afin de mesurer l'amélioration qu'apporte la connaissance du contexte d'apparition à l'analyse syntaxique et, conjointement, les indices syntaxiques au repérage des circonstants, nous nous basons sur une analyse en corpus des SP circonstants de lieu. Cette focalisation sur les expressions spatiales se justifie par le fait que ces circonstants prédominent dans notre corpus constitué de textes géographiques³. Nous confrontons ensuite les indices syntaxiques de productivité et de mobilité à l'indice discursif de contexte d'apparition.

Ce travail constitue une première étape vers l'intégration d'informations discursives à Syntax, intégration bénéfique aux deux domaines : au niveau syntaxique, pour étendre la couverture du logiciel et au niveau discursif, pour disposer d'un outil permettant d'analyser des phénomènes dont la réalité dépasse les limites de la phrase.

² Exemple issu de notre corpus.

³ Corpus étiqueté au niveau morphosyntaxique (233925 mots).

Références bibliographiques

- Bonami O. (1999), Les constructions du verbe : le cas des groupes prépositionnels argumentaux, Thèse de doctorat, Université Paris 7.
- Bourigault D., Fabre C. (2000), Approche linguistique pour l'analyse syntaxique de corpus. *Cahiers de grammaire*, 25, pp.131-151.
- Charolles M. (1997). L'encadrement du discours; univers champs domaines et espaces, in *Cahier de Recherche Linguistique*, 6, LanDisCo université Nancy2.
- Fabre C., Frérot C. (2002), Groupes prépositionnels arguments ou circonstants : vers un repérage automatique en corpus. Actes de la conférence *TALN*, 215-224.
- Grévisse M. (1993). Le Bon usage. Paris-Gembloux, Duculot.
- Ho-Dac L.-M., Le Draoulec A., Péry-Woodley M.-P. (2001). Cohabitation des dimensions temps, espace et phénomènes. *Cahiers de Grammaire*, 26, pp. 125-142
- Miller P. (1997), Compléments et circonstants : une distinction syntaxique ou sémantique ?, Actes du 37^{ème} Congrès de la SAES, Nice.
- Rémi-Giraud S. (1998) : Le complément circonstanciel : problèmes de définition, in J.Lyons (eds) *Autour du Circonstant*, pp 65-113.
- Vigier D. (2003). Les syntagmes prépositionnels en « en N » détachés en tête de phrase référant à des domaines d'activité. *Linguisticae Investigaciones*, 26 :1, pp. 97-122.
- Virtanen T. (1992). *Discourse Functions of Adverbial Placement in English : Clause-Initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions*, Abo Akademi University Press.