



**HAL**  
open science

## Construction dynamique du sens : application à la prédication verbale

Guillaume Jacquet, Jean-Luc Manguin, Fabienne Venant, Bernard Victorri

► **To cite this version:**

Guillaume Jacquet, Jean-Luc Manguin, Fabienne Venant, Bernard Victorri. Construction dynamique du sens : application à la prédication verbale. Rochebrune 2010, Jan 2010, Rochebrune, France. hal-00526804

**HAL Id: hal-00526804**

**<https://hal.science/hal-00526804v1>**

Submitted on 15 Oct 2010

**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.

# Construction dynamique du sens : application à la prédication verbale

Jacquet G., Manguin J.-L ; Venant F. et Victorri B.

## Introduction

Le travail présenté s'inscrit dans le cadre théorique de la construction dynamique du sens, initialement proposé par Victorri et Fuchs (1996). Ce modèle s'intéresse à la construction du sens d'une unité lexicale au sein d'un énoncé<sup>1</sup>. Partant du constat initial que toute tentative d'approche de la construction du sens se heurte au fait qu'une unité linguistique prend fréquemment des sens variés selon les énoncés dans lesquels elle est insérée, ce modèle donne une place centrale au phénomène de la polysémie. La polysémie, bien que facile à comprendre intuitivement, est très difficile à formaliser, et soulève de très importantes questions sémantiques (cf. Kleiber 2000, Soutet dir. 2005). De fait, pouvoir jouer sur plusieurs sens constitue bien plus qu'une source d'amusements et de calembours. Richesse, souplesse, efficacité et créativité de la langue reposent sur ces glissements de sens permanents. De ce point de vue, les unités polysémiques sont donc particulièrement intéressantes car c'est sur elles que repose le dynamisme d'une langue.

Le but de la présente étude est d'appliquer ce modèle à l'étude de la polysémie des prédications verbales. D'une manière générale, la prédication fait depuis quelques années l'objet d'une attention toute particulière en linguistique. Cette opération est en effet très intéressante car tout en étant une structure liée à l'acte de langage, à l'énoncé, elle pose de façon cruciale la question de ses constituants et des relations qu'elle entretient avec son environnement textuel.

Nous adoptons ici la définition de la prédication proposée par François (2003)

« Elle (la prédication) est constituée du prédicat verbal et de son cadre prédicatif ('cadres des rôles actanciels, 'schéma de valence sémantique' ou dans le modèle des Principes et Paramètres 'structure argumentale') et elle est le domaine de rattachement des satellites et d'opérateurs spécifiques (par exemple des satellites de localisation spatiale et temporelle et les opérateurs de temps et de modalité objective.) »

Ainsi si on reprend cet exemple proposé par François : *Jean a manifestement mal compris Marie*

On peut distinguer :

- le prédicat verbal *comprendre*, qui entre dans la structure argumentale *GN comprendre GN*
- les deux arguments du prédicat. Le premier argument *Jean*, est identifié comme Agent et sujet grammatical, le second argument *Marie* est identifié comme objet et Patient.

Ces éléments définissent la prédication étroite.

- deux types de satellites :
  - o l'adverbe *mal*, satellite du prédicat.
  - o l'adverbe finalement qui agit à un niveau de stratification supérieur, celui de la phrase (ou, selon les auteurs, clause ou complexe propositionnel), et qui porte sur la valeur de la force illocutoire.

Il s'agit donc ici d'étudier l'interaction des différentes strates dans la construction du sens d'une prédication, afin de déterminer ce qui relève de l'apport propre de la prédication (son sens « lexical », hors contexte) et ce qui relève de l'action du contexte. Ce travail est

---

<sup>1</sup> Un énoncé est une séquence d'unités linguistiques constituant un tout, qu'elle soit prononcée ou écrite.

essentiellement exploratoire cependant les sections 4 et 6 tentent de montrer la faisabilité d'une implémentation de ce modèle.

### ***La construction du sens : un système dynamique***

Dans un énoncé, le sens de chaque unité lexicale dépend de celui pris par les autres unités présentes, ou même parfois de la situation. Il est quasiment impossible de se livrer à un calcul compositionnel classique, où le sens du tout (l'énoncé) se calcule à partir de celui des parties (les unités linguistiques). Nous pouvons illustrer ce problème avec l'exemple suivant :

«On s'entend généralement à dire qu'un système est complexe si on ne peut pas retirer un composant du système sans que le système change de nature. »

Pour un chercheur s'intéressant aux systèmes complexes, cet énoncé ne présente aucune ambiguïté. Et pourtant les unités lexicales qui le composent sont presque toutes polysémiques : s'entendre (se mettre d'accord ? s'écouter ?...), dire (paroles ? propos ? affirmer ? raconter ? parler ?...), système (appareil ? astuce ? doctrine ? assemblage ?...), complexe (infrastructure ? obscur ? composé ? difficile ?...), retirer (abriter ? soutirer ? enlever ?...), nature (âme ? condition ? campagne ? manière ?...)

Et encore, on n'évoque pas ici la polysémie des unités grammaticales. Si on cherche à appliquer le principe de compositionnalité, c'est-à-dire à assigner d'abord un sens à chacune de ces unités, en fonctionnant par essais et erreurs, on tombe sur plusieurs centaines d'interprétations possibles. En fait, il est vain de vouloir mener un calcul ascendant à partir du sens des unités lexicales, puisque chacune ne prend son sens qu'en fonction de la présence des autres. Le sens de *nature* dépend par exemple de celui de *système*, qui dépend lui-même de celui de *complexe*, qui dépend à son tour de celui de *système*...

On se trouve ici en présence d'une complexité non-computationnelle, obéissant aux principes de base de la Gestalttheorie : « Local et global, structure et procès, se déterminent réciproquement et dynamiquement » (Visetti 2004). Le sens d'un énoncé peut être vu comme un tout, résultant de l'interaction de toutes ses parties, c'est-à-dire les unités linguistiques (syntaxiques, lexicales...) qui le composent. Ce tout est cependant plus que la somme de ses parties, puisque qu'un simple calcul ascendant ne suffit pas à le déterminer. Réciproquement, le sens final de chacune des unités constituant l'énoncé est déterminé par le sens global de l'énoncé. L'interaction entre les unités présentes dans l'énoncé et les éléments de la situation d'énonciation conduit à stabiliser dans un même mouvement le sens des unités et le sens global de l'énoncé. Ce double mouvement sémantique, que Victorri qualifie de compositionnalité gestaltiste, car « le tout influence les parties, en même temps que les parties construisent le tout », nous a amené à travailler dans le cadre théorique des systèmes dynamiques. Un système est défini comme un processus quelconque, aussi complexe soit-il. Un système est dit dynamique si son état évolue au cours du temps, et s'il est possible de modifier son comportement. (cf. Nicolis & Prigogine 1992 pour une première approche). Modéliser un système dynamique implique de définir un espace où les processus ont lieu, et dans lequel on puisse définir les états stables. Dans l'étude de la construction du sens, il s'agit donc de définir un espace sémantique, sur lequel s'exerceront les processus de construction du sens. Les états stables contenus dans cet espace sémantique sont les sens pris par un mot ou un énoncé dans une situation donnée (cf. figure 1).

Nous traduisons ensuite l'évolution du système par la donnée d'une dynamique sur l'espace des états (l'espace sémantique dans notre cas). Nous suivons en cela la théorie des catastrophes de Thom (1972). Le rôle de cette dynamique est de spécifier les contraintes qui s'exercent en chaque point, et d'obtenir les points de stabilisation (attracteurs) qui correspondent aux solutions du problème (ici le sens pris par une unité dans un énoncé, ou le sens global pris par un énoncé.) Définir cette dynamique demande de formaliser les contraintes exercées sur une unité donnée par les éléments contextuels (syntaxiques,

sémantiques, discursives et/ou extralinguistique). Les attracteurs de cette dynamique permettent de déterminer le sens pris par l'unité étudiée dans le contexte étudié.

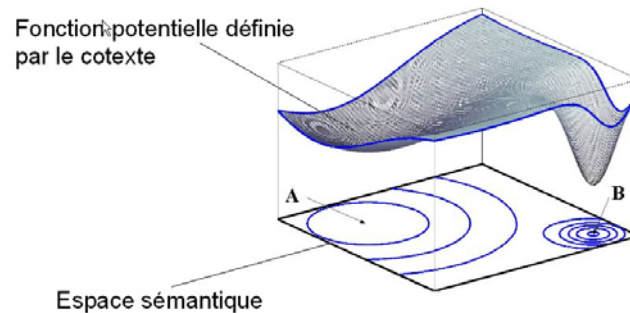


Figure 1. Modèle de la construction dynamique du sens

On peut rapprocher le modèle présenté ici des travaux en morphodynamique (Petitot 1985), développés à la suite des propositions de Thom (1972). Ce cadre théorique nous place aussi dans la lignée des grammaires cognitives américaines (Langacker 1986, Talmy 1988...) qui étudient de façon privilégiée l'ancrage perceptif des structures linguistiques

### **La représentation du sens au cœur du modèle**

Pour modéliser la construction du sens comme un processus dynamique, on va dissocier plusieurs niveaux. Au niveau lexical, on étudie comment une unité lexicale donnée prend son sens au sein d'un énoncé donné. A un niveau plus global, on étudie le sens pris par un énoncé au sein d'un texte ou d'un discours. Nous nous plaçons ici à un niveau intermédiaire, celui de la prédication. Intermédiaire dans le sens où nous ne sommes plus au niveau lexical puisqu'une prédication implique plusieurs unités lexicales et des relations entre ces unités lexicales mais nous ne sommes pas encore au niveau de l'énoncé puisqu'un énoncé est, en tous cas peut être, une combinaison de prédications. Nous voulons rendre compte, d'une part, de la façon dont les constituants d'une prédication construisent son sens hors contexte, son sens lexical, et d'autre part comment ce sens lexical entre en interaction avec d'autres éléments pour l'élaboration d'un sens en contexte. Nous nous restreignons ici à la première étape, c'est-à-dire à la définition de l'espace où les processus ont lieu, et dans lequel on peut définir les états stables. Il s'agit donc de définir un espace sémantique de prédications. Cet espace doit rendre compte des différents sens que peuvent prendre les unités étudiées, mais aussi de la topologie sémantique définie par ces différents sens. Mathématiquement parlant, cet espace est muni d'une structure de variété différentiable. L'intérêt principal de ces espaces est qu'ils sont continus. Ils peuvent ainsi rendre compte de la continuité dans les passages d'un sens à l'autre d'une unité donnée. Ploux et Victorri (1998) ont mis au point une méthode de construction automatique des espaces sémantiques. Cette méthode utilise la relation de synonymie comme accès aux informations lexico-sémantiques. Reposant sur l'analyse d'un graphe de synonymie, elle propose une réponse géométrique à la question de la représentation du sens qui s'inscrit parfaitement dans le modèle que nous venons de décrire. L'étude approfondie des relations de synonymie permet en effet de mettre en évidence à la fois le fonctionnement des unités polysémiques prises individuellement, et leur place dans l'organisation globale du lexique.

La construction des espaces sémantiques repose sur un constat : un synonyme ne suffit pas en général pour définir un sens lexical. Par exemple, dans sa synonymie avec *abandonner*, *laisser* est à la fois synonyme de *quitter* et de *confier*, qui correspondent à deux sens

différents de *abandonner*. L'idée est donc, ici aussi, de caractériser un sens par un ensemble de synonymes. Plus précisément, l'algorithme utilise les cliques du graphe de synonymie. Une clique est un ensemble de mots deux à deux synonymes le plus grand possible. L'idée sous-jacente à la construction des espaces sémantiques est qu'une clique correspond, en première approximation, à une nuance de sens possible pour le mot considéré. Nous ne détaillons pas ici la technique de construction de l'espace sémantique. Disons simplement que les points de cet espace sont les cliques du graphe de synonymie qui relie *abandonner* et tous ses synonymes. Cet espace est muni de la distance du  $\chi^2$ , bien connue en analyse des données, de façon à rendre compte des proximités sémantiques réelles entre les différents sens du mot étudié. La figure 2 montre l'espace sémantique associé à *abandonne* obtenu à partir du *dictionnaire des synonymes du CRISCO*<sup>2</sup>. Il s'agit d'une vision partielle puisque l'espace étant de grande dimension, il faut le projeter pour pouvoir en obtenir une vision 2D. Nous voyons ici la projection selon les deux premiers axes de l'analyse en composantes principales. De telles visualisations permettent une visualisation globale de la sémantique d'une unité. Leur intérêt principal est de rendre compte de la continuité du sens, de la difficulté à tracer des frontières nettes entre les sens, de la « ressemblance de famille » existant entre les différents sens d'une unité. Sur la Figure 2 on voit ainsi se dégager 4 pôles de sens, et la façon dont on peut passer continûment d'un de ces pôles à un autre, en passant par des nuances de sens intermédiaires.

Comparons ces pôles aux 7 sens principaux que l'on peut extraire du TLFi<sup>3</sup> :

1. Renoncer à un pouvoir, à des droits, à la possession d'un bien ou à l'utilisation d'une chose  
*La surcharge rendant la possession des terres onéreuse, l'humble propriétaire abandonna son champ*  
*Je lui sacrifie tout, j'abandonne tout, je renonce à tout pour qu'il me rende l'honneur*
2. Quitter un lieu, ne plus l'occuper  
*Abandonner le domicile conjugal*  
*Abandonner son poste*  
*J'abandonne un pays où je suis oublié*
3. Cesser de défendre une cause, renoncer à des principes, à une idée en la rejetant ou simplement en s'en séparant  
*Je m'aperçois qu'on abandonne les premiers principes du bon sens et de la liberté [...]*
4. Renoncer à poursuivre une action, une recherche, renoncer à une entreprise, un projet  
*Ne la trouvant pas, ils abandonnent cette recherche et n'y pensent plus*
5. Quitter quelqu'un, s'en séparer, le laisser à lui-même  
*Ah! Ce cœur [...] faut-il l'abandonner, le laisser à lui-même?*  
*Mais vous ne m'abandonnez pas, vous ne me laisserez pas seul [...]?*
6. Laisser à qq'n la possession ou le soin d'un bien (ou d'une personne),  
*Abandonner à qq'n le soin de faire qqc.*  
*C'est de maman que je dépendais; papa lui avais abandonné sans réserve le soin de veiller sur ma vie [...]*
7. laisser qqc à l'entière disposition de quelqu'un / ou en proie à quelque chose  
*Il est mieux vu d'abandonner un enfant au bain qu'à l'assistance*  
*Scipion se retira [...] abandonnant aux ravages les terres des Gaulois [...]*

Toute la partie supérieure de l'espace sémantique rassemble des synonymes comme *s'avouer vaincu, se rendre, caler, céder, interrompre, repousser* qui correspondent au sens 4 :

<sup>2</sup> <http://www.crisco.unicaen.fr/Presentation-du-dictionnaire.html>

<sup>3</sup> Trésor de la Langue Française Informatisé ([atilf.atilf.fr/tlf.htm](http://atilf.atilf.fr/tlf.htm))

« Renoncer à poursuivre une action, une recherche, renoncer à une entreprise, un projet : « abandonner des recherches ». En descendant vers la droite de l'espace sémantique, on arrive, via la notion de renoncement et de dépossession, à des synonymes comme *jeter*, *lâcher*, *larguer* qui correspondent (selon que l'objet est humain ou non) aux sens 1 : « renoncer à un pouvoir, à des droits, à la possession d'un bien ou à l'utilisation d'une chose » et 5 « Quitter qqn, s'en séparer ; laisser qqn à lui-même, le laisser seul ». La partie inférieure gauche de l'espace sémantique rassemble des synonymes comme *déguerpir*, *fuir* qui correspondent plutôt au sens 2 : « Quitter un lieu ne plus l'occuper : abandonner un pays, un domicile »

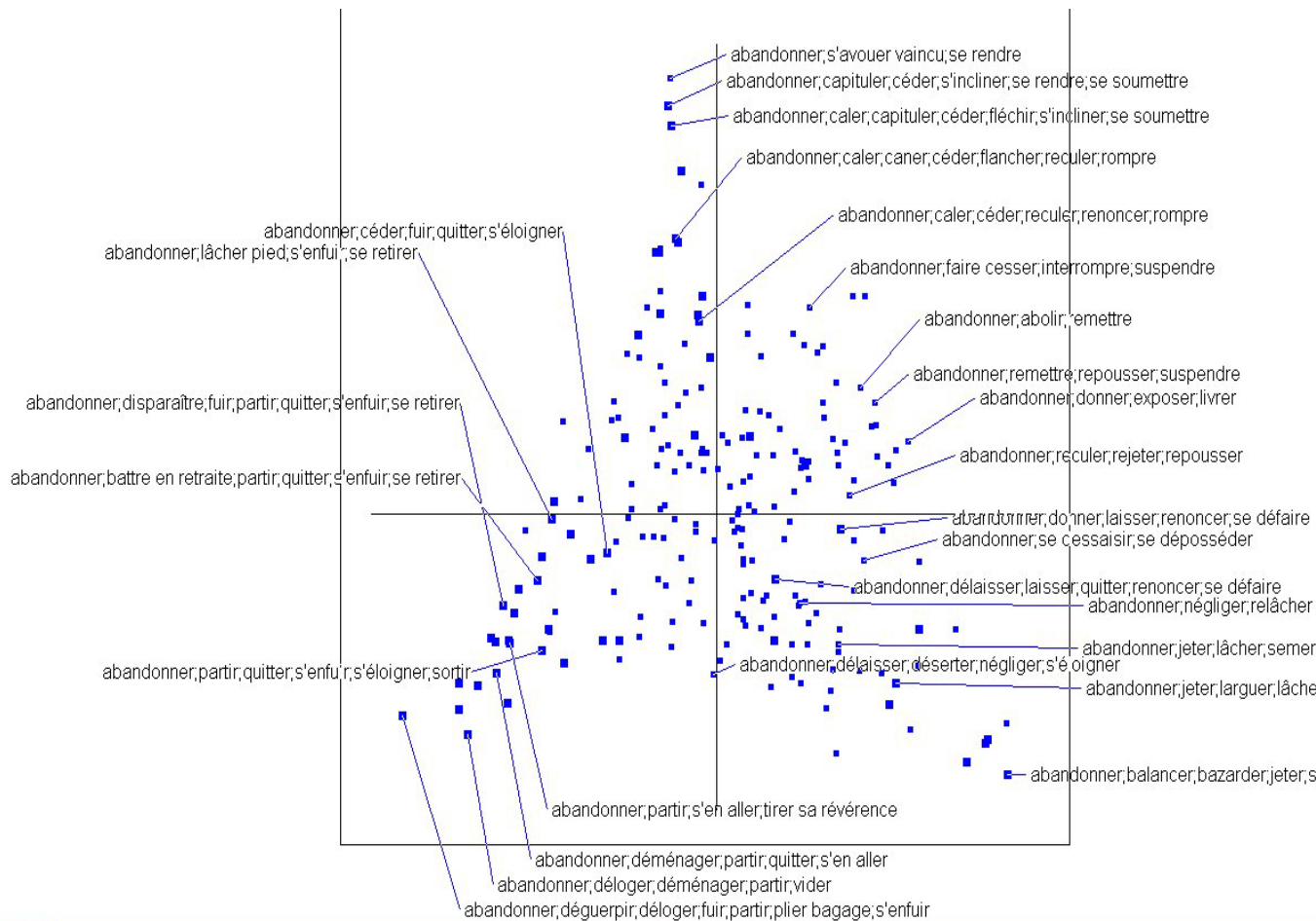


Figure 2. Espace sémantique du verbe *abandonner*

### Un espace sémantique de prédication

Nous cherchons maintenant à étendre cette approche, et à construire non plus l'espace sémantique d'une unité lexicale isolée, mais un espace sémantique de prédication. Il s'agit donc de construire un espace continu au sein duquel la distance géométrique entre deux prédications données soit inversement proportionnelle à leur proximité sémantique. Toute la difficulté de la tâche est d'établir les paramètres pertinents pour la définition d'une telle distance :

- synonymie / proximité sémantique des prédicats verbaux
- structure argumentale identiques
- nature sémantique des arguments
- influence des satellites ?

La dernière question est particulièrement épineuse. L'espace sémantique se doit de ne rendre compte que de ce qui relève de la sémantique lexicale, l'apport propre d'une prédication, ce

qui va entrer en interaction avec le contexte au moment de l'élaboration du sens dans un énoncé ou un texte. Toute la difficulté réside à rendre compte, pour chaque prédication de ce qui relève uniquement du cœur de prédication, et de ce qui relève d'opérateurs satellites contextuels. Autrement dit, il s'agit de savoir où l'on place la frontière entre cadre prédicatif et contexte. Déterminer ce qui relève purement du sens hors contexte (et doit par exemple apparaître dans une définition dictionnaire) dépend grandement du prédicat étudié. Faut-il ou non, à ce niveau de modélisation, prendre en compte la nature sémantique des arguments ? Comment gérer le fait que dans certains cas la présence d'un adverbe semble un élément déterminant dans la discrimination du sens, alors que dans d'autres cas il s'agit d'un opérateur périphérique qui ne caractérise pas l'apport propre de la prédication dans l'énoncé.

Nous appuyons notre réflexion sur des travaux comme ceux de Lacheret et François (2005), consacrés à l'étude des compléments circonstanciels au sein de la phrase verbale. Les auteurs montrent qu'il est bien difficile d'établir une frontière nette entre éléments du cœur et éléments périphérique. Dans bien des contextes, ce type de compléments, loin d'être accessoire, peut entretenir des relations étroites avec le noyau verbal. L'hypothèse d'un continuum entre arguments et circonstanciels est d'ailleurs couramment admise. Fabre et Bourigault (2008) l'ont ainsi testée en développant une méthode permettant de mesurer le degré d'autonomie des compléments prépositionnels vis-à-vis du verbe dont ils dépendent, grâce à des tests élaborés sur un grand corpus analysé syntaxiquement. Ils concluent aussi à un continuum, leur approche leur donnant les moyens de dépasser l'étude des seuls exemplaires prototypiques clairement argumentaux ou circonstanciels, pour mettre au jour des situations intermédiaires.

Les espaces sémantiques que nous cherchons à construire sont particulièrement adaptés pour rendre compte d'un tel continuum. Nous adoptons la position suivante : chaque unité linguistique rattachée au verbe joue un rôle dans la construction du sens de la prédication. Cependant certaines unités sont plus déterminantes que d'autres. Nous ne chercherons donc pas à distinguer compléments essentiels et non-essentiels, mais bien à rendre compte, d'une part, d'un continuum dans les différents types de compléments et, d'autre part de l'influence du corpus et l'influence du genre dans le calcul de similarité entre prédications. Nous verrons dans le paragraphe 6 que ce dernier point se fera au travers de l'exploitation de classes sémantiques construites en corpus.

## ***Prédication et représentation du sens***

Nous allons ici illustrer sur deux exemples, les prédications mettant en jeu le verbe *abandonner* et celles mettant en jeu le verbe *comprendre*, le type de proximités syntaxico-sémantiques que nous cherchons à établir

### **1.1. *abandonner***

#### **Classe sémantique des arguments**

Comme nous l'avons vu au paragraphe 3, on peut dégager sept sens principaux pour *abandonner*. Ces différents sens mettent en jeu deux structures argumentales :

- GN *abandonne* GN
- GN *abandonne* à GN

Les informations fournies par la seule structure argumentale ne sont pas suffisantes pour la discrimination des sens au sein d'un espace sémantique. Ici, par exemple, elles ne permettent pas de discriminer les sens 6 et 7. Il semblerait qu'on ait besoin aussi de caractériser sémantiquement le deuxième argument du verbe. Ainsi si l'objet est un lieu, on a le sens 2, s'il est vivant on a le sens 5, s'il est plutôt du type action on a le sens 3 (ou 4)... Il sera peut-être difficile d'obtenir une caractérisation sémantique de l'objet suffisamment précise pour

discriminer complètement tous les sens. Mais ce genre d'informations semble quand même être nécessaire pour rendre compte de l'apport propre de la prédication.

Il semblerait ainsi que pour caractériser sémantiquement l'apport propre des prédications mettant en jeu le verbe *abandonner* les paramètres structure argumentale + classe sémantique du deuxième argument soit suffisants. L'étude du prédicat *comprendre* va nous montrer que cela peut ne pas être suffisant pour toutes les prédications

## Synonymie des prédicats

Une autre question à laquelle nous devons répondre pour construire notre espace sémantique de prédication, c'est le calcul de la distance sémantique entre deux prédications mettant en jeu non plus le même prédicat, mais des prédicats synonymes. Prenons par exemple *quitter* dans sa synonymie avec *abandonner* :

Toujours d'après le TLFi, pour que *quitter* soit synonyme de *abandonner*, il faut qu'il y ait une notion de rupture définitive :

« quitter sa femme le matin », ce n'est pas l'abandonner, mais « quitter sa femme », oui c'est l'abandonner.

On voit ici à quel point les opérateurs dits « périphériques » peuvent jouer un rôle important au niveau même de la détermination du sens lexical. Dans ce cas, la présence d'un adverbe (sans retour, définitivement, sans regret, à jamais) ou d'un complément comme « le matin » ou « pour une autre » fait toute la différence.

Par ailleurs, les extraits de définitions du verbe *quitter* ci-dessous (toujours d'après le TLFi) montrent que *quitter* et *abandonner*, dans leur synonymie, partagent les mêmes structures argumentales et que l'on retrouve les mêmes discriminations de sens en fonction de la nature sémantique du deuxième argument, ce qui nous conforte dans l'idée qu'il s'agit là d'un paramètre que nous devons absolument prendre en compte dans la construction d'un espace sémantique.

1) [Un inanimé plus ou moins abstr.] Abandonner. *Quitter le passé; quitter une doctrine, la religion; quitter un point de vue...*

2) [Une expression du visage, le ton de la voix, etc.] Abandonner, renoncer à. *Quitter un air (triste, sévère); quitter une mine (attristée); quitter son incognito.*

3) [Une habitude, un défaut, etc. ou encore son propre nom] à quitter *l'oisiveté, quitter son nom*

4) [Une activité professionnelle ou autre, un genre de vie, etc.] Laisser, abandonner, cesser. *Quitter son métier; quitter une charge, une profession; quitter ses fonctions; quitter le barreau, le droit, le journalisme (pour la politique)*

## 1.2. comprendre

L'article du TLFi permet de distinguer quatre sens pour *comprendre* :

1. Inclure dans sa nature propre ou dans un système telle chose ou personnes. Syn : comporter, compter, contenir, embrasser, enclore, impliquer, englober, envelopper, inclure, incorporer, intégrer, mêler...

1.1. Sujet=chose : en soi tel et tel éléments constitutifs.

Le sujet comprend l'attribut

La rétine comprend trois couches superposées

1.2. Sujet= personne + construction comprendre dans/ sous Faire entrer dans une catégorie, dans un ensemble d'éléments nombrables

Les savants ont omis de comprendre les puissances élémentaires dans leur système

Comprendre sous tel terme



2. Elaborer, recevoir dans son esprit la représentation nette d'une chose, d'une personne, Saisir intellectuellement le rapport de signification qui existe entre tel signe et la chose signifiée, se faire une idée claire des causes, des conséquences qui se rattachent à telle chose et qui l'expliquent, Syn : appréhender, pénétrer, imaginer, se figurer, interpréter, se représenter.

*L'intelligence (...) cette faculté de comprendre les relations des choses*

*il est vain d'espérer « **comprendre** » le fonctionnement du système nerveux*

*La vie, je la **comprends** de moins en moins,*

3. Se rendre compte de la réalité de telle chose, prendre conscience de son importance Syn : apercevoir, discerner, réaliser, remarquer, se représenter, saisir, voir  
*elle **comprenait** enfin la gravité de son mal.*

*il me fallait bien distinguer quelque chose de nouveau, de très triste et de très angoissant. (...). Je **compris** tout cela soudain...*

4. Appréhender quelqu'un ou quelque chose dans toute la vérité de sa nature profonde.

Syn : se mettre à la place de, participer, saisir sentir

*Comprendre profondément, comprendre l'Homme. Comprendre et approuver ce qu'est une existence vouée à la galanterie*

Ces différents sens mettent en jeu trois structures argumentales :

Intrans : GN comprend

Trans : GN comprend GN

Prep : GN comprend dans/ sous GN

La prise en compte de la structure argumentale permet ici de distinguer le sens 1.2 des autres sens. Là encore, on voit que si ce paramètre est nécessaire, il n'est pas suffisant. La classe sémantique des arguments va permettre une discrimination supplémentaire :

- si le sujet est une chose, on sélectionne d'emblée le sens 1.1
- si l'objet appartient à une classe du genre [gravité, importance, intérêt, nécessité, portée, prix, utilité, valeur...] → sens 3
- s'il relève d'une classe du genre [attitude, caractère, esprit, inquiétudes, intentions, motifs, position, scrupules, sentiments] → sens 4
- s'il relève d'une classe d'un des classes suivantes ;[un document, texte, discours, explication], [langue, une idée, le langage de, une langue, un mot, paroles], [pensée, question, raisonnement, explication], [sens de, un signe, signification] → sens 2

Cependant, il apparaît ici encore que d'autres éléments, habituellement considérés comme « périphériques » s'avèrent déterminants dans la discrimination du sens. Parmi eux, on peut citer :

- La présence/ absence d'un adverbe:
  - *Comprendre environ, essentiellement, généralement, notamment, ...* → sens 1.1
  - *Comprendre à demi-mot, de travers* → sens 2
  - *Comprendre aisément, bien, difficilement, facilement, mal* → sens 2, 3 voire 4
  - *Comprendre immédiatement, rapidement, sur-le-champ, tout de suite* → sens 3
  - *Comprendre à peine, à merveille, à moitié, sans peine* → sens 2
  - *Comprendre un (certain) nombre de, l'ensemble/ la totalité, tant de* → sens 1.1
- L'intégration dans une construction plus complexe :
  - Verbe + *comprendre* :
    - *(se faire, paraître, vouloir) comprendre* → sens 2
    - *(devoir, pouvoir, savoir) comprendre* → sens 4
  - Verbe + *comprendre* + de
    - *(Essayer, feindre, s'efforcer, permettre) de comprendre* → sens 4 (2 ?)

## **Des classes sémantiques en contexte**

L'analyse de ces deux exemples, nous permet de dégager des paramètres incontournables pour la construction d'un espace sémantique de prédications :

- l'ensemble des sens possibles d'un prédicat indépendamment de tout contexte
- les différentes classes sémantiques qui peuvent caractériser les arguments rattachés à ce prédicat
- le type de relation unissant classe sémantique des arguments et le prédicat.

Il nous faut enfin contraindre ces deux premiers phénomènes par le corpus/domaine dans lequel il évolue, puis déterminer la façon la plus efficace de les combiner dans le calcul d'une distance sémantique

La méthode d'exploration de graphes de synonymie que nous avons présentée nous permettra de réaliser la première étape en nous donnant l'accès à l'ensemble des sens possibles pour un prédicat. La deuxième étape est la caractérisation sémantique des arguments de la prédication. Cette deuxième étape revient non plus à identifier le sens des arguments sous forme de synonymes mais plutôt à identifier la classe sémantique de ces arguments. Cette distinction peut être illustrée par le fait que pour rendre compte du sens de la prédication se trouvant dans l'énoncé « Cette boîte comprend douze bouteilles », la liste des synonymes possibles de *boîte* importe moins que de savoir que *boîte* fait partie de la classe sémantique « objets/entité pouvant contenir d'autres objets/entités » au même titre que *train* (« *Ce train comprend 4 wagons* ») ou *bâtiment* (« *Ce bâtiment comprend 10 appartements* »). Nous comptons pour cela étendre une méthode de construction de classe distributionnelle que nous avons mise au point. Ces classes sont déterminées automatiquement à partir d'un corpus. Leur singularité est qu'elles sont dépendantes de l'énoncé étudié. Ces classes distributionnelles avec pertinence d'emploi sont donc particulièrement bien adaptées à notre propos ici. Elle ne cherche pas, en effet, à créer des classes de mots ayant le même sens mais des classes de mots dont le comportement sémantique influence de la même façon un contexte donné. Autrement dit si nous voulons trouver la classe de *pouvoir* (*pouvoir, puissance, mandat...*) ce n'est pas pour caractériser le sens de *pouvoir* mais pour caractériser le sens de *abandonner* dans *abandonner le pouvoir*.

La technique que nous utilisons s'inscrit dans le cadre bien connu de l'analyse distributionnelle « à la Harris ». Elle est exploitée depuis longtemps dans la communauté du TAL pour la construction de bases de connaissances ou de ressources terminologiques à partir de textes (Habert et Nazarenko, 1996 ; Aussenac-Gilles et al. 2000, Lin et Pantel 2001 ; ...). Notre méthode est entièrement automatique. Elle ne fait appel à aucune modélisation préalable de connaissances sémantiques sur le corpus et utilise des rapprochements de mots sur la base de contextes syntaxiques partagés. Elle utilise les rapports de dépendance syntaxique élémentaires entre un contexte et les mots pleins qu'il régit ou qui le régissent. En tout cela elle se rapproche des travaux de Greffentette (1994). Les contextes nous sont fournis par un analyseur syntaxique robuste. Lors de nos précédentes expérimentations nous avons utilisé les analyseurs Syntex (Bourigault et Fabre, 2000) et XIP (Ait et al. 2002).

Nous ne détaillerons pas ici la méthode de construction des classes distributionnelle (Jacquet, Venant 2005), mais nous présentons simplement quelques exemples de classes obtenues lors de précédentes expérimentations:

- Expérimentations sur le français (Jacquet et Venant 2005)<sup>4</sup>

Nous cherchons ici à caractériser un des arguments d'une des quatre prédications suivantes : GN *descendre* GN (deuxième argument), GN *jouer à* GN (deuxième argument), GN *regarder* GN (deuxième argument), GN *décider* GN (premier argument)

---

<sup>4</sup> Pour cette expérimentation, nous avons utilisé le corpus LM10 qui correspond à l'ensemble des articles du journal *Le Monde* entre 1991 et 2000

<b>GN descendre GN – étude du deuxième argument</b>	
<b>Réalisation lexicale</b>	<b>Classe distributionnelle obtenue</b>
Seine	Fleuve, rivière, seine
Gange	Fleuve, rivière
Chirac	Homme, personne
Pdg	Homme, personne
Mont Blanc	Montagne, piste

<b>GN jouer à GN – étude du deuxième argument</b>	
<b>Réalisation lexicale</b>	<b>Classe distributionnelle obtenue</b>
Monopoly	Billard, bridge, cache-cache, domino, ping-pong, poker, pétanque, souris, yo-yo, monopoly
Lego	Billard, bridge, cache-cache, domino, ping-pong, poker, pétanque, souris, yo-yo, monopoly
Paris	New York, Avignon, Londres, Marseille, Paris
Washington	New York, Londres, Marseille, Paris
Wimbledon	basket, basket-ball, football, loterie, rugby, tennis, jeu vidéo

<b>GN décider GN – étude du premier argument</b>	
<b>Réalisation lexicale</b>	<b>Classe distributionnelle obtenue</b>
Washington	Etas-Unis, Washington
Wimbledon	Europe, France, Italie
président	Pdg, directeur général, Eltsine, Jospin, Mitterand, chef de l'état, premier ministre, président de la république
Otan	Armée, force, police

<b>GN regarder GN – étude du second argument</b>	
<b>Réalisation lexicale</b>	<b>Classe distributionnelle obtenue</b>
Chirac	Enfant, femme, gens, homme, personne, public, Chirac
Seine	mer
Arte	Arte, Chaîne de télévision
Alpes	Montagne

**Tableau 1: exemples de classes sémantiques construites automatiquement à partir du corpus LM10**

- Expérimentation sur l'anglais (Ah-Pine et Jacquet 2009)<sup>5</sup>.

Des expérimentations comparables ont été faites en utilisant une méthode qui essaye de généraliser cette notion de prédication. Autrement dit, cette méthode cherche à caractériser un des arguments en fonction non plus d'une prédication précise mais en fonction d'un ensemble de prédications similaires. Ces expérimentations ont été faites pour le français et pour l'anglais. Ci-dessous une illustration de classes sémantiques pour l'anglais.

<sup>5</sup> Pour cette expérimentation, nous avons utilisé le corpus BNC ([www.natcorp.ox.ac.uk/](http://www.natcorp.ox.ac.uk/))

n° classes	Classe sémantiques	prédications représentatives
classe 1	teacher ; member ; people ; student ; worker ; manager	recruit GN ; enable GN ; employ GN
classe 2	dog ; bird ; cat ; fish ; horse ; rabbit ; boy ; animal ; <b>man</b>	GN eat ; hunt GN ; feed GN ; kill GN ; GN swim ; breed GN ; ride GN ; rear GN ; GN jump
classe 3	mother ; wife ; parent ; woman ; daughter ; dad ; friend ; sister ; <b>man</b>	GN live ; marry GN ; live with GN ; stay with GN ; GN die ; <b>understand GN</b> ; talk to GN ; say to GN
classe 4	paper ; <b>book</b> ; letter ; document ; article ; note ; text	speak to GN ; read GN ; write GN ; publish GN ; edit GN ; print GN ; write in GN
classe 5	paint ; picture ; piece ; <b>book</b> ; photograph ; portrait ; draw ; print ; film	paint GN ; exhibit GN ; GN depict
classe 6	fear ; feel ; anxiety ; anger ; concern ; excitement ; emotion ; frustration ; enthusiasm	arouse GN ; express GN ; voice GN ; heighten GN ; shake with GN ; ease GN ; dispel GN ; experience GN ; relieve GN ; <b>understand GN</b>

**Tableau 2: exemples de classes sémantiques construites automatiquement à partir du corpus BNC**

Ces classes nous montrent qu'un même argument peut appartenir à plusieurs classes sémantiques. Par exemple « *man* » fait partie de la classe « êtres vivants »<sup>6</sup> (classe 2) mais aussi de la classe « être humains » (classe 3). De la même manière, « *book* » fait partie de la classe « documents textuels » (classe 4) mais aussi de la classe « œuvres artistiques » (classe 5). Et pour chacune de ces classes, nous avons la possibilité d'identifier les prédications représentatives de chaque classe. Dans notre cas, ces « prédications représentatives » correspondraient aux prédications dont le sens peut être discriminé en fonction de la classe sémantique à laquelle appartient son argument. C'est le cas de la prédication « understand GN » dont le sens changera selon que son argument fait partie de la classe 3 (classe « êtres humains ») ou qu'il fait partie de la classe 6 (classe « émotions »). Dans le premier cas le sens de « understand GN » sera plutôt l'équivalent du sens 2 de *comprendre* décrit précédemment, dans le second cas, ce sera plutôt l'équivalent du sens 4 de *comprendre*.

Ces classes sémantiques ne constituent pas le cœur du travail que nous présentons, mais l'existence de ces méthodes de construction de classes sémantiques et la qualité des résultats obtenus rendent envisageable l'exploitation du modèle pour une étude approfondie de la prédication verbale.

## Conclusion

Nous avons présenté ici nos premières réflexions pour l'étude de la construction du sens des prédications verbales. Le cadre théorique dans lequel nous sommes placés, la construction dynamique du sens, nous fournit des outils aussi bien théoriques que pratiques permettant d'appréhender les relations particulières que les prédications entretiennent avec leurs environnements textuels.

L'étude que nous voulons mener passe avant tout par la construction d'un espace sémantique de prédication. Nous avons défini nos objets d'étude : chaque prédication est composée d'un prédicat, d'une structure argumentale et pour chaque argument de cette structure argumentale, sa classe sémantique. Nous avons détaillé ici les pistes que nous comptons suivre pour la définition d'une distance entre ces prédications. Cette distance résultera de la combinaison de trois principaux paramètres :

<sup>6</sup> Les noms de classes proposés tels que « êtres vivants » sont simplement des interprétations afin de faciliter la lecture de ces classes.

- Le degré de synonymie entre les deux prédicats
- La proximité des classes sémantiques des arguments entrant en jeu dans les deux prédications
- Le type de relation unissant la classe sémantique des arguments aux prédicats ;

Cette dernière étape est peut-être la moins définie. Elle est néanmoins importante car elle est déterminante pour l'étude de la structure argumentale d'une prédication. Nous avons cependant évoqué dans le paragraphe 4 la difficulté d'identifier les rôles des arguments rattachés au prédicat. Cette question reste ouverte et dans un premier temps nous nous contenterons d'approximer ces relations par des fonctions syntaxiques simples, telles que sujet et objet ainsi que des relations syntaxiques simples, comme le type de relation prépositionnelle (par exemple, GN rattaché au prédicat par la préposition *dans*). Ce type de relation, bien que très réducteur, présente l'avantage d'être appréhendable à l'aide d'analyseurs syntaxiques robustes.

Enfin, notre méthode de construction des classes sémantiques nous permettra aussi de prendre en compte l'influence du corpus. Ces classes sémantiques sont en effet construites à partir du corpus et permettront donc de « contextualiser » la distance entre prédications en fonction du genre du corpus. Il nous sera aussi possible d'appréhender le genre du corpus en contraignant la distance par la représentativité de chaque prédication dans le corpus étudié : si une prédication n'est jamais employée dans le corpus elle aura moins de poids dans la distance que celle qui sont employées.

## **Bibliographie**

- Ah-Pine J. et Jacquet G. (2009), « Clique-Based Clustering for improving Named Entity Recognition Systems », EACL 2009.
- Aït S., Chanod J.P., Roux C. (2002), « Robustness beyond shallowness: incremental dependency parsing », NLE Journal, 2002.
- Aussenac-Gilles N., Biébow B., Szulman N. (2000), Revisiting Ontology Design: a method based on corpus analysis, Actes de 12th International Conference on Knowledge Engineering and Knowledge Management. Juan-Les-Pins
- Bourigault D. et Fabre C. (2000), Approche linguistique pour l'analyse syntaxique de corpus, Cahiers de Grammaires, n° 25, Université Toulouse - Le Mirail, pp. 131-151.
- Fabre C. et Bourigault D. (2008), « Exploiter des corpus annotés syntaxiquement pour observer le continuum entre arguments et circonstants ». Journal of french studies.
- François J. (2003), La prédication verbale et les cadres prédictifs, Bibliothèque de l'Information Grammaticale n° 54, Louvain, Peeters.
- Grefentstette (1994)., Explorations in Automatic Thesaurus Discovery, London, Kluwer Academic Publishers.
- Habert B. and Nazarenko A. (1996)., La syntaxe comme marche-pied de l'acquisition des connaissances : bilan critique d'une expérience, Journées sur l'acquisition des connaissances, AFIA, Sète
- Jacquet G. et Venant F. 2005. Construction automatique de classes de sélection distributionnelle, TALN 2005.
- Kleiber, G. (2000), Problèmes de sémantique, la polysémie en question. Villeneuve d'Ascq, Septentrion.
- Lacheret-Dujour A., François J. (2005) : « Circonstance et prédication en français parlé : contraintes sémantico-pragmatiques et filtrage prosodique », Revue de syntaxe et sémantique : la prédication verbale et averbale dans les langues, n°6, PUC.
- Langacker R.(1986), Foundation of Cognitive Grammar, Standford, Standford University Press

- Lin D. and Pantel P. (2001). Induction of Semantic Classes from Natural Language Text. In Proceedings of ACM Conference on Knowledge Discovery and Data Mining (KDD-01).
- Nicolis, G. & Prigogine, I. (1992), A la rencontre du complexe. Collection Philosophie d'aujourd'hui, Presses Universitaires de France.
- Petitot, J. (1985), Les catastrophes de la parole. Paris, PUF.
- Ploux S., Victorri B., « Construction d'espaces sémantiques à l'aide de dictionnaires de synonymes », TAL, n°39-1, 1998.
- Soutet, O. dir (2005), La polysémie. PU Paris-Sorbonne.
- Talmy, L. (1988), Force Dynamics in Language and thought. Cognitive Sciences.
- Thom R. (1972), Stabilité structurelle et morphogénèse, Paris, Ediscience
- Victorri B. et Fuchs C. (1996), La polysémie, construction dynamique du sens, Paris, Hermès.
- Visetti Y.M. (2004), Constructivismes, émergences : une analyse sémantique et thématique., Intellectica.