

## INTRODUCTION

La cause est une relation logique qui est utilisée pour expliquer tous les événements qui caractérisent la vie de l'univers, du monde, de l'homme. La formation de l'univers, le mécanisme qui fait allumer un moteur d'une voiture, le fait de parcourir une distance en allant en vélo, le processus de photosynthèse chlorophyllienne, tout dans la nature peut être rapporté à un mécanisme où il existe quelqu'un ou quelque chose dont le mouvement, les caractéristiques, les actions agissent sur un autre objet (qu'il soit humain, minéral, animal) en déterminant un changement dans l'état de cet objet.

Puisque la langue doit nous fournir les instruments pour pouvoir exprimer les expériences que nous vivons dans le monde, nous trouvons toujours présente dans les grammaires de langue, en ce qui nous concerne de langue française, l'expression de la relation logique de la cause.

En effet, on peut considérer la causalité comme une notion à la fois logique et linguistique : d'une part elle se présente essentiellement comme une notion conceptuelle, selon laquelle il existe un mécanisme où une cause produit un effet, d'autre part nous la définissons toujours par le biais linguistique en répondant à la question introduite par un mot du type « *pourquoi* ».

Cependant, comme le fait remarquer Patrick Charaudeau dans sa *Grammaire* (Charaudeau 1992 : 525) il n'existe aucune tentative d'expliquer cette

relation d'une façon complète et satisfaisante, en adoptant comme point de départ le plan sémantique.

Du point de vue linguistique, la relation générale de causalité unit deux assertions  $A_1$  et  $A_2$  en les rendant dépendantes l'une de l'autre (plus précisément l'assertion  $A_2$  doit son existence à l'assertion  $A_1$ ) :

1. Il faut avoir 10 € pour aller voir ce film<sup>1</sup>

où l'assertion  $A_1$  est représentée par "avoir 10 €" et l'assertion  $A_2$  est représentée par "voir ce film" et elles sont liées par la conjonction « pour ».

La langue nous offre plusieurs moyens pour exprimer une relation de ce genre et dans un texte ou un discours quelconque le locuteur peut, plus ou moins aisément, localiser quelle est la cause d'un fait, et vice-versa, quelle est la conséquence qui a comme point de départ une cause définie. Selon Nazarenko (2000) les moyens qui nous permettent de repérer des informations causales dans un discours sont de quatre ordres différents :

1. les connecteurs causaux : ce sont généralement des mots syntaxiques qui relient d'autres mots ou des phrases ; ils peuvent être des conjonctions ou locutions conjonctives (telles que *parce que, puisque, car*), des adverbes ou locutions adverbiales (telles que *effectivement, en effet, par conséquent, c'est pourquoi*) et des préposition ou locutions prépositives (telle que *à cause de, grâce à*) ;
2. les tournures syntaxiques de la causalité : ce sont des constructions syntaxiques qui unissent deux propositions qui représentent respectivement la cause et l'effet; la forme la plus classique de construction syntaxique employée pour exprimer la cause est la subordination, comme dans l'exemple 2.

2. Mon frère qui est souvent en retard m'agace<sup>2</sup>.

---

<sup>1</sup> Cfr. Charaudeau 2000: 525.

<sup>2</sup> Cfr. *ibidem*, p. 101.

Il existe aussi d'autres tournures syntaxiques qui expriment la relation de cause à effet telles que la coordination (dans l'exemple 3) et la juxtaposition ( dans l'exemple 4).

3. Il est en retard et il m'agace<sup>3</sup>

4. Plus il est en retard, plus il m'agace<sup>4</sup>.

3. le lexique de la cause qui permet de caractériser plus en détail la nature du lien causal qui unit les deux événements étant donné leur nature sémantique;
4. l'interprétation causale : il arrive parfois que l'information causale soit implicite car qu'il n'existe aucune marque d'elle dans l'énoncé ; elle apparaît donc être véhiculée par de nombreuses autres relations comme la conséquence, le but, la concession, les relations temporelles.

Dans notre parcours de recherche nous avons choisi d'analyser ce concept dans le domaine du lexique, en nous plaçant donc dans le troisième ordre décrit par Nazarenko ; nous avons décidé de prendre en considération parmi tous les verbes qui expriment la cause, ceux qui présentent de façon explicite dans leur sémantisme, les composants de la causalité.

Une fois établi le domaine de notre recherche, à savoir l'étude du lexique verbal de la cause, nous avons procédé parallèlement sur deux axes : d'une part nous avons pris en considération différentes approches théoriques pour l'étude du sens des mots, de l'autre nous avons sélectionné un corpus de textes de façon à pouvoir confirmer ou infirmer nos hypothèses de départ.

### *Etude des différentes théories sémantiques*

Notre travail centré sur l'étude de différentes théories d'analyse du sens s'est développé en plusieurs phases.

Tout d'abord nous avons pris en considération la théorie structuraliste de l'analyse componentielle de la signification des mots, retenant que la

---

<sup>3</sup> Cfr. *ibidem*, p. 101.

<sup>4</sup> Cfr. *ibidem*, p. 101.

subdivision en traits sémantiques pertinents et distinctifs pouvait nous donner un critère précis de désambiguïsation de nos verbes synonymes. A travers un parcours onomasiologique (Baldinger 1984), notamment en partant d'un concept indépendant de la structure pour arriver aux désignations, nous voulions trouver les caractéristiques propres à chaque verbe qui nous permettraient de les différencier. En étudiant le concept de base de « être la cause de » nous avons tout de suite remarqué qu'il faisait intrinsèquement partie du sémantisme des verbes que nous avons choisis. En fait, l'analyse des sèmes nous a montré qu'il existait des différences seulement pour certains verbes, notamment le sème /violence/ pour le verbe *déchaîner*. De plus ce genre d'analyse nous a semblé trop lié à notre intuition dans la mesure où nous ne pouvions pas analyser des caractéristiques concrètes, même après avoir constitué notre corpus.

Nous avons donc abordé une deuxième méthode d'analyse des verbes avec la grammaire des cas proposée par Fillmore aux alentours des années 1970 ; en effet, il s'agit d'une théorie qui adopte un point de vue plutôt syntaxique que sémantique dans l'exploration de la signification et nous avons pris en considération cette méthodologie dans la mesure où les objets que nous avons analysés sont des verbes.

La grammaire des cas fournit la description de la valence sémantique d'un verbe, à savoir le fait que le prédicat (mot principal) a le pouvoir de choisir des *noun phrases* qui sont requis de par sa signification. Fillmore a avancé l'hypothèse de l'existence de neuf cas qui sont choisis au niveau de la structure profonde. Il n'y a toutefois pas de critère pour classer les cas de façon définitive tant qu'ils sont appelés de façon différente par l'un ou l'autre chercheur appliquant la grammaire des cas. Cette méthode ne satisfaisait donc pas non plus notre exigence de clarté et de précision.

Nous avons alors pris en considération une autre théorie d'analyse de la signification, cette fois-ci relevant du domaine psychologique, dans la mesure où les résultats des investigations dépendaient des expérimentations sur des locuteurs de langue. Il s'agit de la sémantique du prototype qui nous a permis

de nous confronter avec des notions telles que le meilleur représentant de catégorie et le degré d'appartenance. Notre but ultérieur était d'établir s'il existait un ou plusieurs représentants pour cette catégorie verbale. Au départ nous pensions mener des expérimentations avec des locuteurs de langue française, comme Rosch et ses disciples l'avaient fait pour la catégorisation de quelques catégories naturelles ; toutefois ce genre d'expérimentation était difficile à réaliser dans la mesure où les catégories verbales ne dénotent pas des objets réels et donc répondre à des questions du type « Quel est le meilleur représentant de la causation, *provoquer* ou *susciter* ? » présupposait d'un côté une réflexion sur une catégorie très abstraite de la langue, de l'autre côté une connaissance approfondie de la notion de causalité. En plus, en analysant un corpus oral<sup>5</sup> (composé malheureusement seulement de 77000 mots) nous avons remarqué l'absence absolue de cette catégorie verbale dans la langue parlée. Nous avons donc déduit que l'expression de la cause dans ce type de corpus était véhiculée par d'autres formes que le lexique verbal.

En raison de ce fait nous nous sommes tournée vers une discipline « nouvelle », la linguistique de corpus qui nous propose une approche concrète à l'analyse de la langue. En effet, la linguistique de corpus nous fournit un modèle d'analyse basé sur l'étude de caractéristiques concrètes des verbes, relevant de leur emploi en contextes linguistique et situationnel. Les notions de collocation, colligation, prosodie sémantique et préférence sémantique relevant de la linguistique de corpus nous ont aidée à analyser le lexique verbal de la cause et à établir de façon précise le degré de synonymie entre les verbes.

Loin de nous d'exclure les apports des théories de l'étude du sens que nous venons de présenter et sur lesquelles nous reviendrons. Nous avons adopté en fait certaines notions fondamentales proposées par la sémantique structurale, comme la subdivision en sèmes de la signification des verbes, car tous ont deux composants nécessaires, à savoir le causateur et la chose causée (ou effet) ; de plus, comme nous le verrons, nous avons établi un

---

<sup>5</sup> Le corpus collecté par Claire Blanche Benveniste et son équipe à l'université de Aix-en-Provence.

rapprochement entre l'échelle de synonymie (proposée par Cruse) et la notion de degré d'appartenance à une catégorie (proposée par Rosch) dans la mesure où il existe des mots qui sont considérés comme plus ou moins synonymiques.

### *Collection du corpus*

Les ressources écrites qui sont disponibles pour les chercheurs qui s'intéressent à la langue française, sont représentées par deux bases de données qui ne sont pas, à proprement parler, de vrais corpus. D'une part nous avons *Frantext*, un recueil d'œuvres principalement littéraires, disponibles en ligne sur abonnement, dont les textes s'échelonnent du XVI<sup>e</sup> au XX<sup>e</sup> siècle. Sur l'intégralité du corpus, il est possible d'effectuer des recherches simples ou complexes (base non-catégorisée). Sur un sous-ensemble comportant des œuvres en prose des XIX<sup>e</sup> et XX<sup>e</sup> siècles, les recherches peuvent en outre répondre à des critères syntaxiques (base catégorisée). Par ailleurs, nous avons le recueil de 10 années de publication des articles du journal *Le Monde*.

En raison de la dominante littéraire de la première base de données et du fait qu'on ne disposait que d'un journal comme échantillon dans le deuxième cas, nous avons choisi de créer notre propre corpus. Au départ notre idée était de collecter des articles du *Monde Diplomatique* et leurs traductions respectives en italien (publiées par le quotidien italien *Il Manifesto*) dans la mesure où, au début, nous voulions mener une analyse contrastive du lexique verbal de la cause. A partir de quelques considérations sur la nature d'un corpus de ce genre, notamment le fait qu'il s'agit d'un corpus parallèle dans la mesure où l'on a des textes originaux et des traductions de ces textes, nous avons pensé que cela aurait entraîné des résultats faussés car l'emploi des termes en italien dépendait des choix du traducteur. En raison de ce fait, nous avons changé d'avis et décidé d'analyser seulement l'une des deux langues, le français, adoptant ainsi un point de vue intralinguistique qui nous permettait de mieux cerner notre objet d'analyse.

Nous avons alors ajouté aux articles du *Monde Diplomatique*, des articles tirés d'autres journaux français différents de par leur public, les sujets traités, les modalités de parution (quotidiens, mensuels, hebdomadaires).

Au terme de notre travail de collecte, nous avons recueilli un corpus d'environ un million de mots, en plusieurs phases qui se sont succédé au fur et à mesure que les résultats obtenus après des analyses partielles nous ont poussée à étendre certains composants et à en ajouter d'autres. Pour nous expliquer : le verbe *causer* que nous considérons comme le verbe causatif par excellence, apparaissait d'une façon très marginale par rapport à d'autres verbes, notamment *entraîner* ou *engendrer*. Nous avons pensé alors étendre notre corpus avec des journaux qui relevaient d'autres domaines, tels que la science et la médecine. Les résultats que nous avons obtenus en analysant ce corpus élargi nous ont confirmé l'hypothèse que l'emploi de tel ou tel autre verbe dépend des sujets dont on parle.

A partir des réflexions concernant le concept de synonymie, réflexions menées par les linguistes, nous avons décidé de vérifier les hypothèses suivantes : (i) la difficulté à trouver des synonymes absolus dans la langue ; (ii) la présence d'un élément, le contexte d'emploi de ces verbes qui, seul, nous permet de désambiguïser les sens des différents synonymes ; (iii) la détermination d'une échelle de la synonymie dans le sens de Cruse 2001 pour établir quels sont dans notre groupe les verbes qui sont plus synonymiques que d'autres.

Notre thèse s'articule donc en 5 chapitres :

- dans le premier chapitre nous avons mis en évidence les caractéristiques de quelques théories sémantiques nées et développées au XX<sup>e</sup> siècle ;
- dans le deuxième chapitre nous avons pris en considération les bases théoriques et méthodologiques de la linguistique de corpus, les critères à suivre pour la constitution d'un corpus et pour l'analyse de la langue et nous avons décrit notre travail de collecte des données pour la composition de notre corpus ;

- dans le troisième chapitre nous avons abordé le problème de la synonymie en tant que notion liée aussi bien à la philosophie du langage qu'à la linguistique, en mettant en évidence les considérations sur cette notion et les problèmes qu'elle soulève dans l'étude de la langue ;
- dans le quatrième chapitre nous avons analysé les différents aspects de la causalité, ses conditions de réalisation, les participants à une action causative, la différence entre agentivité et causativité et nous avons défini le groupe de verbes que nous avons analysé ;
- dans le cinquième chapitre nous avons illustré l'étude sur les occurrences des verbes que nous avons obtenues à partir du traitement de notre corpus avec un concordancier: chaque verbe a été analysé selon ses collocations, ses colligations, ses prosodies sémantiques et ses préférences sémantiques ; cette analyse nous a permis d'établir le degré de synonymie entre les verbes par rapport à leur domaine d'emploi.