

## **Liens conceptuels et relations sémantiques : proposition de représentation des connaissances en sciences du langage**

Laurence Kister, Evelyne Jacquy, Bertrand Gaiffe

Atilf UMR 7118 CNRS/Nancy-Université – 54000 Nancy  
Laurence.Kister@univ-nancy2.fr  
Evelyne.Jacquy@atilf.fr  
Bertrand.Gaiffe@atilf.fr

**Résumé :** Le travail que nous présentons s'intègre dans un projet de production d'une ressource destinée à représenter les connaissances en sciences du langage et à annoter sémantiquement des documents en texte intégral. Il consiste à ce stade de démarrage du projet en un examen des relations entre les liens conceptuels et les relations sémantiques qui organisent un thesaurus afin de prévoir une amélioration de sa structure et conduire à une représentation plus cohérente du domaine.

**Mots-clés :** thesaurus, terminologie, relations sémantiques, liens conceptuels, représentation des connaissances

Le travail consiste à examiner le fonctionnement de Thesaulangue<sup>1</sup> en termes de liens conceptuels et de relations sémantiques. L'analyse est destinée à proposer des modifications afin de le faire évoluer progressivement vers une onto-terminologie au sens de (Roche, 2009) : une représentation qui « permet une construction du sens autour d'une sémantique référentielle et justifie l'intérêt de termes normés en regard des termes d'usage ». L'outil remanié est destiné à faciliter le travail des documentalistes, à offrir une image aussi précise que possible du domaine mais aussi à permettre une annotation sémantique et conceptuelle de documents en texte intégral. Elle fait émerger les différences masquées par la structure et propose des corrections pour renforcer la cohérence de la ressource. Notre démarche rejoint différentes stratégies de constitution de

---

<sup>1</sup> Thesaulangue est le thesaurus des sciences du langage du centre de documentation de l'Atilf. Il présente l'avantage d'être structuré et l'inconvénient d'offrir une image encore incomplète de la linguistique actuelle. Il a été amorcé par l'Inalf (1975). Jusqu'à aujourd'hui, il n'a cessé d'être complété et maintenu. Il a, par ailleurs, été intégré au portail Termosciences conçu par l'Inist (<http://www.termosciences.fr>).

terminologies et d'onto-terminologies : (Nazarenko & Hamon, 2002), (Depecker, 2003), (L'Homme, 2002 et 2005), (Mazuel et Sabouret, 2007) et (Mhiri & *al.*, 2006).

**Les liens conceptuels.** Nos observations et des entretiens avec les documentalistes révèlent une structure de surface hiérarchique et une structure profonde de type treillis (relations transversales parfois implicites). Notre hypothèse est que la représentation repose sur deux relations : (1) des liens conceptuels avec des domaines inférables à partir des termes, (2) des relations d'inclusion entre les domaines inférés. Les liens conceptuels s'établissent entre les termes et les domaines inférés et peuvent être verbalisés par *est identique à* (12% - 110 sur 897) et *étude de* (86% - 770 sur 897) à l'exception de 2% de liens non catégorisés (17 sur 897). La répartition des liens entre les *employés pour* et les termes est assez proche de celle qui existent entre les termes : *étude de*, 80% (274 sur 348) et *identiques à*, 15.5% (54 sur 348).

**Les relations sémantiques entre termes et terme-employés pour.** L'analyse des relations sémantiques vise à rendre compte du travail réalisé par les documentalistes lors de la confection de Thesaulangue et de leur activité d'indexation. Sur ce point, nos travaux se rapprochent de ceux de Christment et *al.* (2008) qui cherchent à « capturer la sémantique que les documentalistes exploitent de manière implicite lorsqu'ils utilisent le thesaurus ». Les relations entre termes sont des relations de généralité (*générique*, 60% - 526/877) et méronymie (*élément de*, 32% - 281/877). L'examen des relations termes-employé pour fait apparaître les mêmes relations qu'entre les termes (*élément de*, 24% - 84/348 et *générique*, 11,5% - 40/348) auxquelles s'ajoute une relation de synonymie (*synonyme*, 58% - 203/348). La relation *synonyme* correspond parfaitement au rôle des *employés pour* en assurant une fonction d'homogénéisation et de contrôle du vocabulaire d'indexation.

Les relations *générique* et *élément de* entre *employés pour* et termes posent questions. La relation *générique* suppose une relation de dominance du terme sur l'*employé pour*. Or, ils ne sont ni synonymes, ni quasi-synonymes. Face à ce type de distorsion nous proposons une extraction de la liste des *employés pour* et une intégration aux termes.

La relation *élément de* est considérée comme un type de *synonymie partielle*. Nous choisissons de les conserver en *employés pour* : être un *élément de* n'entraîne pas une réelle inclusion de domaines. Un élément constitutif possède des propriétés propres et n'hérite pas des propriétés du tout auquel il appartient.

**Correspondance liens conceptuels/ relations sémantiques.** Si on s'intéresse simultanément aux liens conceptuels et aux relations sémantiques en tenant compte des propositions de modifications que nous avons faites nous obtenons la répartition présentée en Tab.1.

		terme-terme		terme-EMP	
générique	identique à	7% (62)	62% (565)	-	-
	étude de	55 % (503)		-	
élément de	identique à	3% (24)	30% (274)	7% (21)	24% (74)
	étude de	27% (250)		17% (53)	
synonyme	identique à	-	-	8% (25)	63% (194)
	étude de	-		55% (169)	

TAB. 1 – Relations sémantiques en fonction des liens conceptuels

Lors de travaux antérieurs, nous avons envisagé une série de verbalisation des relations pour distinguer plus précisément certaines relations entre termes. Pour confirmer la pertinence de ces relations et préciser la représentation nous avons entrepris une analyse des définitions des termes et des *employés pour* de la terminologie que nous avons extraite à partir du TLFi (Trésor de la Langue Française Informatisé).

## Références

- CHRISTMENT C., HAEMMERLÉ C., HERNANDEZ N. & MOTHE J. (2008). Méthodologie de transformation d'un thésaurus en une ontologie. *RIA*, Paris : Hermès-Lavoisier, n°22-1.
- INALF. (1975). Trésor de la langue française. *Le français moderne*, (supplément), nouvelle série, fascicule 2.
- JACQUEY E., KISTER L., GRZESITCHAK M., GAIFFE B., REUTENAUER C., VALETTE M. & SANDRINE O. (2010). Thésaurus et corpus de spécialité en Sciences du Langage : une approche lexicométrique appliquée à l'analyse de termes en corpus. *Taln'10*, Montréal, juillet.
- KISTER L., JACQUEY E. & GAIFFE B. (2009). Fusion d'un thésaurus et d'une terminologie : utilisation de ressources existantes pour amorcer une ontologie. *TIA'09*, Toulouse, novembre.
- L'HOMME M.C. (2005). Sur la notion de terme. *Méta*, n°50-4, 1112-1132.
- L'HOMME M.C. (2002). Fonction lexicales pour représenter les relations sémantiques entre termes. *Structuration de terminologie*, A. Nazarenko et T. Hamon (eds.), *TAL*, n°4-1, p. 19-41.
- MAZUEL L. & SABOURET N. (2007). Degré de relation sémantique dans une ontologie pour la commande en langue naturelle. *18e Journées Francophones d'Ingénierie des Connaissances*, Grenoble, 4-6 juillet.
- MHIRI M., GARGOURI F. & BENSLIMANE D. (2006). Détermination automatique des relations sémantiques entre les concepts d'une ontologie. *INFORSID*, Hammamet, Tunisie
- NAZARENKO A. & HAMON T. (2002). Structuration de terminologie, *TAL*, n°43-1, 174 p.
- ROCHE C. (2009). Le terme et le concept : fondements d'une ontoterminologie, *Toth'09*, 4-5 juin.