Corela

Cognition, représentation, langage

9-1 | 2011

Vol. 9, n° 1

Du thesaurus à l'onto-terminologie : relations sémantiques vs relations ontologiques

Laurence Kister, Evelyne Jacquey et Bertrand Gaiffe

Édition électronique

URL : <https://journals.openedition.org/corela/1962>

DOI : 10.4000/corela.1962

ISSN : 1638-573X

Éditeur

Cercle linguistique du Centre et de l'Ouest - CerLICO

Référence électronique

Laurence Kister, Evelyne Jacquey et Bertrand Gaiffe, « Du thesaurus à l'onto-terminologie : relations sémantiques vs relations ontologiques », *Corela* [En ligne], 9-1 | 2011, mis en ligne le 06 juin 2011, consulté le 18 juin 2021. URL : <http://journals.openedition.org/corela/1962> ; DOI : <https://doi.org/10.4000/corela.1962>

Ce document a été généré automatiquement le 18 juin 2021.

Corela – cognition, représentation, langage est mis à disposition selon les termes de la licence Creative Commons Attribution - Pas d'Utilisation Commerciale - Partage dans les Mêmes Conditions 4.0 International.

Du thesaurus à l'onto-terminologie : relations sémantiques vs relations ontologiques

Laurence Kister, Evelyne Jacquey et Bertrand Gaiffe

Introduction

- 1 La nécessité de proposer des représentations des connaissances dans des domaines de spécialité clairement délimités est liée à la multiplication, au partage et au transfert des informations dans les différents domaines scientifiques et techniques : la diffusion de l'information demande une inter-pénétration toujours plus précise des domaines de spécialité en raison des approches multi- ou interdisciplinaires dont les différents domaines font l'objet. L'information se doit d'être diffusée en fonction du contexte dans lequel évoluent émetteur et récepteur : il ne s'agit pas de transmettre des informations isolées mais d'apporter des informations inscrites dans un réseau de connaissances, de savoirs partagés ou en voie d'acquisition. Dans ce contexte, les ressources terminologiques et ontologiques (Bourigault et Aussenac-Gilles, 2003), (L'Homme, 2004), (Bourigault et al., 2001), (Nazarenko et Hamon, 2002) ainsi que l'extraction et la structuration de ces ressources constituent une contribution majeure pour de nombreuses approches dans différentes disciplines : information et la documentation (Roche, 2004), édition (El Mekki et Nazarenko, 2002), recherche d'informations, classification de documents (Sanjuan et Ibekwe-Sanjuan, 2002) ou détection de documents à caractère raciste sur la Toile (Valette et Grabar, 2004). Les travaux évoqués ci-dessus partagent un point méthodologique : les ressources terminologiques sont extraites à partir de textes relevant d'un domaine de spécialité (Dell'Ortella et al., 2008) d'où la place primordiale des corpus de spécialité lors du processus de constitution de ces ressources (Calberg-Challot et al., 2009) et cela tant en extraction automatique, que manuelle ou assistée. En revanche, l'utilisation de ressources externes, comme des dictionnaires, reste seconde (L'Homme, 2004).

- 2 Dans le contexte de partage des connaissances, l'ontologie constitue un outil de structuration des connaissances et nécessite la prise en compte de relations sémantiques. Le lien de type *est un* n'est pas suffisant lorsqu'on cherche à établir la *carte sémantique* d'un domaine pour produire une représentation de celui-ci (Nazarenko et Hamon, 2002), (Mazuel et Sabouret, 2007) et (Mhiri *et al.*, 2006).
- 3 Le travail que nous proposons vise à établir la liste des relations qu'il nous paraît nécessaire d'introduire pour interpréter les liens sémantiques, parfois complexes, qui s'établissent entre les différents niveaux de l'arborescence d'un thesaurus et prévoir l'évolution de l'outil vers une onto-terminologie. En effet, le thesaurus, l'ontologie et la terminologie sont trois types de ressources complémentaires qu'il nous paraît important de prendre en compte pour notre futur outil d'annotation en texte intégral. A l'heure actuelle, nous disposons de deux de ces ressources terminologiques pour les sciences du langage qui nous servent de domaine d'investigation et d'expérimentation avant d'envisager une généralisation de notre approche à d'autres domaines : un thesaurus – Thesaulangue¹ et une terminologie – TermTLF². Les deux ressources constituent les outils d'annotation de départ et bénéficieront de l'analyse en texte intégral pour leur amélioration en termes de couverture du domaine des sciences du langage.
- 4 Thesaulangue est un langage artificiel normalisé puisqu'il s'agit d'un langage documentaire constitué de termes spécifiques au domaine de connaissances des sciences du langage qui entretiennent des relations sémantiques (synonymie, hiérarchie, associations). Le principal défaut de ce type de langage normalisé se retrouve dans Thesaulangue : il ne propose pas de définitions.
- 5 TermTLF est un ensemble de termes des sciences du langage définis dans le TLFi. Comme bon nombre de terminologies, elle constitue une liste représentative du choix et de l'utilisation des dénominations (ou désignations) des objets et des concepts du domaine considéré.
- 6 Le thesaurus évoqué ci-avant est assez proche de l'ontologie car, comme cette dernière, il consiste en un ensemble structuré de termes désignant, au moyen de termes, les concepts propres au domaine en tenant compte de, relations entre les concepts bien que la représentation par des formes linguistiques diffère de la représentation conceptuelle. Par ailleurs, une ontologie modélise les connaissances relatives à un domaine en en décrivant les objets, les propriétés fonctionnelles, les caractéristiques communes, les relations entretenues par les objets. L'association des propriétés de ces trois types de ressources devraient permettre de rendre compte des relations entre les termes qui désignent les concepts, des liens entre les concepts mais aussi de proposer des définitions qui permettront de mieux comprendre l'organisation du domaine et de permettre la représentation des connaissances.
- 7 Pour la phase de travail que nous présentons dans cet article, nous proposons d'examiner les relations sémantiques présentes dans Thesaulangue qu'il est nécessaire de prendre en compte pour ébaucher la construction d'une onto-terminologie offrant une image satisfaisante et actuelle du domaine des sciences du langage. En effet, Thesaulangue est structuré mais il offre une image incomplète de la linguistique actuelle : sa structure reste fortement marquée par l'époque à laquelle sa constitution a débuté (Inalf, 1975) dans le contexte de la constitution du Trésor de la Langue Française. Face aux nombreuses disciplines possibles en sciences du langage nous nous intéressons, en priorité, à des parties fortement développées, c'est-à-dire des micro-

thesaurus qui comportent suffisamment de niveaux hiérarchiques et qui sont fréquemment utilisés par le centre de ressources de l'Atilf.

- 8 Notre approche sur ce point vise à consolider nos premières observations (Kister *et al.*, 2009), (Gaiffe *et al.*, 2009), (Jacquey *et al.*, 2010), à ajuster certaines représentations et à produire une onto-terminologie, au sens où l'entend (Roche, 2009) quand il précise qu'une ontologie « permet une construction du sens autour d'une sémantique référentielle et justifie l'intérêt de termes normés en regard des termes d'usage ». L'onto-terminologie sera destinée à faciliter le travail des documentalistes et à donner une image aussi précise que possible des sciences du langage à un néophyte.
- 9 Le second objectif est de déterminer, grâce aux définitions associées aux termes, les faisceaux sémiques devant permettre l'annotation sémantique de documents en texte intégral. Différents travaux, dont entre autres (Bourigault *et al.*, 2004), (Aussenac-Gilles et Bourigault, 2000), (Drouin, 2003) et (Drouin 2005) ont montré la faisabilité de l'indexation de textes en utilisant des ressources terminologiques et documentaires, démarche à laquelle s'apparente notre approche.
- 10 A moyen terme notre objectif est d'utiliser les ressources terminologiques pour indexer les textes mais aussi d'annoter sémantiquement les textes. Cette annotation sera quant à elle utilisée pour déterminer, par proximités sémiques, les entités linguistiques aptes à constituer des candidats termes qui viendront enrichir la ressource en cours de production.

1. Thesaulangue : fondement de la futur onto-terminologie

- 11 Thesaulangue est un thesaurus vivant, régulièrement enrichi et mis à jour par le centre de documentation de l'Atilf. Il comporte actuellement 1224³ termes structurés en 20 micro-thesaurus. Les micro-thesaurus les plus développés de Thesaulangue constituent pour notre projet une base solide pour la construction d'une onto-terminologie destinée à faciliter l'indexation du contenu des documents, à révéler les usages qui sont fait des termes du domaine de spécialité dans les textes, à faire apparaître les évolutions sémantiques subies par certains termes et à mettre en évidence la répartition des termes en fonction du niveaux de scientificité ou de vulgarisation du document (Kister et Jacquey, 2007a et b) et (Kister *et al.*, 2008). L'utilisation de ressources préexistantes est une pratique courante en ontologie : bases de données (Benslimane *et al.*, 2003), terminologie structurée (Bachimont *et al.*, 2005) ou glossaire (Navigli R. et Velardi P., 2008).
- 12 La richesse sémantique des micro-thesaurus de Thesaulangue traduit en partie le contenu du fonds du centre de documentation puisque les domaines et sous-domaines des sciences du langage les plus représentés ont conduit à un besoin de précision en relation avec le nombre et le degré de spécialisation des documents disponibles au centre de documentation. La figure 1 (Fig. 1 – Nombre de termes par micro-thesaurus) montre la répartition des termes qui figure dans la structure arborescente sans prendre en compte les 335 termes qui correspondent aux *employés pour* et aux *équivalents*. Pour examiner les relations qui s'établissent entre les différents niveaux hiérarchiques, nous avons choisi de nous intéresser en priorité aux micro-thesaurus qui comportent au moins trois de ces niveaux et qui correspondent à des sous-domaines fortement

représentés dans le fonds documentaire du centre de documentation. Il s'agit des micro-thesaurus de la *morphologie*, de la *syntaxe* et de la *phonétique* (cf. Annexe – Extraits de Thesaulangue).

Fig. 1 – Nombre de termes par micro-thesaurus⁴

Micro-thesaurus	niveau +1	niveau +2	niveau +3	niveau +4	niveau +5	Total
Méthodologie linguistique	18	3	1			22
Histoire de la langue	5	11	7			23
Etudes d'ensemble	7	63				70
Phonétique	7	33	5			45
Orthographe	8	6				14
Grammaire	9					9
Syntaxe	20	59	68	37	2	186
Morphologie	2	17	18	4	8	49
Sémantique	7	26	13			46
Terminologie	5	34				39
Sociolinguistique	4	13	3	19	30	69
Onomastique	6	4	11			21
Etymologie	4	3	4			11
Lexicologie	4	18	4	2		28
Sémiotique	5	10	186			201
Statistique linguistique	1					1
Linguistique appliquée	2	8	3			13
Approches interdisciplinaires	8	11				19
Psycholinguistique	1	2				3
Bibliographie						0
	123	321	323	62	40	869

- 13 Un rapide examen de la figure 1 (Fig. 1 – Nombre de termes par micro-thesaurus) montre qu'il existe un micro-thesaurus qui correspond aux critères que nous avons fixés – trois niveaux hiérarchiques et richesse du micro-thesaurus – mais que nous n'avons pas retenu pour notre analyse. Nous avons écarté le sous-domaine de la *sémiotique* en raison de la répartition des termes qu'il présente sur les différents niveaux et de la nature des termes qu'il contient. En effet, le micro-thesaurus de la *sémiotique* comporte 201 termes dont 186 termes de troisième niveau : cette répartition est due au fait que le micro-thesaurus regroupe essentiellement les termes qui désignent les figures rhétoriques (70 termes soit environ 35 % des termes du sous-domaine) et les genres littéraires (113 termes soit environ 56 % des termes du sous-domaine). De plus, d'un point de vue strictement formel, la structure de ce micro-thesaurus est assez proche de celle du thesaurus de la *phonétique*, retenu pour sa part, qui présente une répartition du même type avec un nombre important de terme au second niveau.

2. Thesaurus, terminologie et onto-terminologie : les relations sémantiques en jeu

- 14 Les travaux sur les relations sémantiques dans les terminologies font partie des problématiques courantes – (L'Homme, 2002), (L'Homme, 2005) et (Bourigault et Lame, 2002), parmi d'autres. Le type d'analyse que nous envisageons consiste à expliciter les

relations sémantiques qui s'établissent entre les termes des différents niveaux afin de structurer l'onto-terminologie de la manière la plus efficace possible en établissant un lien entre la place des termes dans la représentation arborescente et la représentation des connaissances fondée sur les concepts et les liens qu'ils entretiennent. Avant d'examiner les relations sémantiques présentes dans Thesaulangue, nous proposons d'examiner une partie de la littérature portant sur le sujet, tant en linguistique qu'en intelligence artificielle, afin de confronter la théorie et la mise en pratique des relations qui régissent les liens entre les concepts et les termes. Dans ce paragraphe nous passons en revue un certain nombre d'approches dont nous nous inspirerons pour proposer une première ébauche de liste des relations nécessaires à la mise au point de l'onto-terminologie (cf. 3.2.2. Relations proposées par la littérature vs observations sur Thesaulangue).

- 15 (L'Homme, 2002) s'attache à une problématique similaire à celle qui nous intéresse quand elle analyse « les relations sémantiques entre le terme décrit et d'autres unités lexicales apparaissant (...) dans l'article » dans le cadre du travail d'informatisation d'un dictionnaire. Les relations qu'elle propose sont (1) des relations de sémantique lexicale⁵ – taxonomiques comme l'hyperonymie, l'hyponymie, la co-hyponymie et partitives comme la méronymie et l'holonymie, (2) dérivationnelles fondées sur des liens morphologiques réguliers exploitant des relations sémantiques et des relations syntagmatiques qui reposent sur des liens privilégiés entre certains termes et certaines unités lexicales (lien entre un terme et une unité *significative* non reconnue comme un terme spécifique du domaine pour produire un nouveau terme). Ce dernier type de relations s'apparente à des combinaisons lexicales spécialisées ou des cooccurrences qui, par l'usage dans un domaine spécifique, acquièrent un statut de terme. Les relations dérivationnelles présentées par (L'homme, 2002) sont à rapprocher de l'approche lexématique développée par (Corbin, 1991, 1997 et 2000). Cette approche propose un modèle morphologique théorique qui repose sur l'usage de contraintes sémantiques. La description des contraintes sémantiques est alors indispensable à l'explicitation des procédés de construction morphologique. C'est le cas pour *coiffeur* et *coiffer* où le N-dérivé est l'agent du processus décrit par le prédicat verbal qui constitue la base de la construction morphologique. En linguistique, *catégoriseur* et *catégoriser*, par exemple, correspondent au cas où le N-dérivé est un agent non volitionnel, un outil, un instrument utilisé dans le déroulement du processus décrit par le prédicat verbal, base de la construction morphologique qui permet d'obtenir un nom déverbal par suffixation en -eur.
- 16 (Mhiri et al., 2006) définissent différentes relations qu'ils associent à la symétrie et à la transitivité pour traduire la cohérence qui existe entre les concepts et formaliser les liens qu'ils entretiennent. Les relations retenues reposent sur des relations sémantiques (1) d'identité⁶ – les concepts ont une syntaxe, des attributs et des opérations identiques, (2) de synonymie symétrique et transitive de type *sorte de* – le concept correspond à un cas particulier d'un autre concept *canari* et *oiseau*, (3) d'homonymie symétrique⁷ – le concept peut avoir deux sens différents *fil* => le *fil* et *fil* => le *fil* au pluriel, (4) d'équivalence symétrique – des concepts proches jouent des rôles différents selon le contexte *usine* et *fournisseur*, (5) d'antonymie⁸ symétrique – les concepts sont nécessairement disjoints *célibataire* et *marié*.
- 17 (Mazuel et Sabouret, 2007) s'intéressent au degré de relations sémantiques dans une ontologie. Ils évaluent le degré de relations sémantiques des liens non contraints que

sont la méronymie et l'hyponymie mais aussi de liens spécifiques au domaine, comme *estPlusgrandQue* qu'ils présentent comme indispensables à la description du domaine et à sa représentation.

- 18 (Rastier, 2004) évoque des liens verticaux pour la catégorisation (*est un* et *est une partie de*) et de liens horizontaux pour tout ce qui relève de l'actance (synonymie, antonymie, hyperonymie). Les relations sémantiques sont alors de deux ordres – hiérarchiques et non hiérarchiques et permettent de concevoir des *graphes cycliques* en complément des *graphes arborés*.
- 19 (Depecker, 2003) propose une répartition en relations logiques (rapports d'abstraction entre les concepts : *orange, agrume, fruit*) et en relations ontologiques (rapports entre concepts dont les objets entretiennent une relation de présence ou de contiguïté : *aile, avion*). Les relations logiques regroupent les relations (1) génériques qui se subdivisent en espèce-genre – *orange, agrume, fruit* et type-produit – *Airbus A320, Airbus, avion, aéronef*, (2) spécifiques – *courrier* inclut *courrier postal, courrier électronique*, (3) de coordination quand il s'agit de l'ensemble des parties immédiatement constitutives du tout – *pied, chapeau, champignon*. Les relations ontologiques sont présentées comme des relations partitives et associatives. Les partitives se répartissent en relations (1) composant-composé – *fer, acier*, (2) contenu-contenant – *paquet de cigarettes, cartouche*, (3) membre-collection – *footballeur, équipe*. Les relations associatives s'établissent entre concepts, par vertu ou par expérience – *onde, télédétection*. Les relations associatives les plus répandues sont séquentielles et topologiques. Les relations séquentielles sont relativement diverses : (1) cause-effet – *détonateur, explosion*, (2) agent-action-résultat – *dessinateur, dessiner, dessin*, (3) producteur-produit – *déchetterie, produits recyclés*, (4) outil-outil – *vis, tournevis*, (5) procédé-processus – *unité de fabrication initiale, unité de fabrication finale*. La relation topologique situe un objet par rapport à un autre – *molaire, prémolaire*.
- 20 (Roche, 2008) distingue deux grands types de relations entre concepts : les relations hiérarchiques (génériques et partitives) qui permettent d'ordonner le système notionnel, de l'appréhender et d'en maîtriser la complexité et les relations associatives qui correspondent à des liens non hiérarchiques fondés sur l'expérience (cause-effet, producteur-produit, etc.).
- 21 (Le Pesant, 2000) montre que la relation de méronymie demande à faire l'objet de subdivisions afin de rendre compte des nuances sémantiques. Elle peut prendre différentes formes (1) partie fonctionnelle-tout – *lame, couteau*, (2) partie non-fonctionnelle-tout – *pièce, appartement*, (3) élément-collection – *soldat, régiment*, (4) pièce d'un assemblage-tout – *vis, carburateur*, (5) composante-organisation – *locaux, organisme* ou *personnel, organisme*, (6) processus-processus – *flirt, parade amoureuse*, (7) morceau-totalité – *part, gâteau*, (8) localisation externe-site – *berge, fleuve* ou *banlieue, ville*, (9) contenu-contenant – *soupe, assiette*, (10) hyperonyme-hyponyme – *mammifère, chat*.
- 22 (Vieu, 1991), reprise par (Jacquey et Namer, 2003), propose un découpage en cinq sous-classes : (1) composant-assemblage où la partie à une fonction clairement déterminé dans le tout – *roue, vélo*, (2) substance-tout pour laquelle la partie est distribuée de manière homogène dans le tout – *beurre, gâteau*, (3) morceau-tout où le tout ne contient pas d'autre partie similaire au morceau – *cime, arbre*, (4) portion-tout où le tout comporte des parties similaires à la partie – *part de gâteau, gâteau*, (5) élément-collection où la somme des éléments similaires forme le tout – *fil d'un tissu, tissu*.

- 23 Cette brève revue de littérature confirme notre intuition selon laquelle les relations méronymiques et holonymiques sont des constantes dans les terminologies, les ontologies et les thesaurus. Cependant ces notions paraissent souvent trop vagues par rapport aux nuances sémantiques perçues par l'humain. Etablir une classification des relations parties-tout constitue une piste pour l'explicitation d'une partie des relations présentes dans Thesaulangue. L'examen des relations proposées dans la littérature fait apparaître que certaines des relations se fondent à la fois sur des liens conceptuels et des liens sémantiques ce qui nous paraît difficilement intégrable pour la conception de l'onto-terminologie que nous envisageons. La dissociation des liens conceptuels et des liens sémantiques qui nous paraît nécessaire est présentée ci-dessous.

3. Dissociation des relations conceptuelles et des relations sémantiques

- 24 Dans cette partie nous proposons d'examiner les relations qui peuvent être inférées dans Thesaulangue au regard des différentes approches proposées dans le paragraphe précédent. Notre analyse porte sur les relations que nous considérons comme suffisamment représentées. Ce que nous cherchons à faire se rapproche des travaux de (Christment et al., 2008) qui précisent qu'il est nécessaire de « capturer la sémantique que les documentalistes exploitent de manière implicite lorsqu'ils utilisent le thesaurus ». Notre démarche se rapproche aussi de la leur dans la mesure où ils favorisent l'utilisation de relations taxinomiques et de relations associatives sans exclure d'autres formes de liens. Notre approche consiste à déterminer et à verbaliser les relations sémantiques qui existent entre les termes qui désignent les concepts en prenant en compte les types de liens qui s'établissent entre les concepts. En effet, même si certains liens conceptuels laissent présager certaines relations sémantiques, liens conceptuels et relations sémantiques peuvent être confondus. Les premiers concernent la réalité conceptuelle extra-linguistique et la représentation des connaissances et les seconds les formes linguistiques qui permettent de les exprimer au moyen de la langue.

3.1. Les relations d'ordre conceptuel

- 25 Nos hypothèses sont que la structure de Thesaulangue repose sur une relation d'inclusion entre domaines et qu'il existe une relation entre termes et domaines. L'inclusion de domaine entraîne une dominance entre termes qui permet d'établir une forme de hiérarchie et fonde la représentation arborescente du thesaurus. Cependant, d'autres relations s'établissent de manière non hiérarchique. La structure hiérarchique apparente est due à la facilité de représentation qu'elle offre : les relations transversales ne sont pas matérialisées mais elles existent et correspondent, pour partie, à ce qu'on considère comme les inférences implicites que font les documentalistes lorsqu'ils utilisent un thesaurus. Il nous paraît nécessaire, pour comprendre l'organisation de la structure de Thesaulangue, de nous intéresser à trois types de relations : entre termes, entre domaines et entre termes et domaines.
- 26 Un examen des relations (R_i) entre les termes (T_i) et les domaines (D_i) de Thesaulangue fait apparaître que lorsqu'il y a identité entre l'étiquette du terme et l'étiquette du domaine, on peut représenter la relation de la manière suivante :

Si T_1 identique à D_1

Si T_2 identique à D_2

=> Alors D_2 est inclus dans D_1 et T_1 domine T_2

- 27 Ainsi, dans le micro-thesaurus de la *phonétique*, le terme *phonétique* est un nom de domaine ($R_1 = \text{identique à}$) qui domine le terme *phonétique descriptive* qui est lui aussi un nom de domaine ($R_2 = \text{identique à}$) donc le domaine de la *phonétique* inclut celui de la *phonétique descriptive* et les termes entrent dans une relation verbalisable en T_2 est un type (une sous-catégorie) de T_1 . Il y a inclusion de domaine et spécification de terme. Le terme T_2 hérite des contenus sémantiques et conceptuels de T_1 .

Si $T1$ (*phonétique*) identique à $D1$ (*phonétique*)

Si $T2$ (*phonétique descriptive*) identique à $D2$ (*phonétique descriptive*)

=> Alors $D2$ (*phonétique descriptive*) est incluse dans $D1$ (*phonétique*)

et $T1$ (*phonétique*) domine $T2$ (*phonétique descriptive*)

- 28 Pour les trois micro-thesaurus auxquels nous nous intéressons, une deuxième « règle » permet d'établir les relations entre termes et domaines. Il s'agit de la relation *étude de* :

Si D_1 est l'étude de T_1

Si D_2 est l'étude de T_2

=> Alors D_2 est inclus dans D_1 et T_1 domine T_2

- 29 Ainsi, pour *affixe* et *infixe* qui apparaissent dans le micro-thesaurus de la *morphologie*, nous inférons que l'étude de l'*infixe* fait partie de l'étude de l'*affixe* et nous obtenons :

Si $D1$ (*affixe*) est l'étude de $T1$ (*affixe*)

Si $D2$ (*infixe*) est l'étude de $T2$ (*infixe*)

=> Alors $D2$ (*infixe*) est inclus dans $D1$ (*affixe*) et $T1$ (*affixe*) domine $T2$ (*infixe*)

- 30 Les relations *identique à* et *étude de* concernent de manière certaines 276 (99 %) des 278 termes des micro-thesaurus qui nous ont servi de test : 26 (9 %) entrent dans une relation T_i identique à D_i et 250 (90 %) dans une relation D_i est l'étude de T_i .

- 31 Quand on s'intéresse aux relations qui s'établissent entre les domaines, on s'aperçoit que la relation d'inclusion peut suffire à expliquer la structure hiérarchique : la relation d'inclusion D_2 est inclus dans D_1 entraîne la relation hiérarchique D_2 est dépendant de D_1 . Les relations entre termes, qui découlent de ce que nous venons de proposer, sont les relations perçues comme des relations de généricité et de spécificité dans le thesaurus : T_1 est générique de T_2 et T_2 est spécifique de T_1 .

- 32 L'examen de la répartition des relations conceptuelles avec inférence qui s'établissent entre les termes à partir de celles qui existent entre les domaines révèle une majorité de relations *étude de*. Pour ce type de relations, nous avons les configurations suivantes :

- 23 termes liés à leur domaine par une relation *identique à* soit 8.5 %
- 64 termes dominés liés à leur domaine par une relation *étude de* avec des termes dominants liés par une relation *identique à* soit 63.5 %
- 189 termes dominants et dominés liés à leurs domaines par *étude de* soit 68 %.

- 33 Si on observe la répartition en fonction des micro-thesaurus, on se rend compte que, pour les trois micro-thesaurus que nous avons sélectionnés, celui de la syntaxe n'admet pas de combinaisons de relations *identique à-identique à* (Fig. 2 – Répartition des relations en fonction des micro-thesaurus).

Fig. 2 – Répartition des relations en fonction des micro-thesaurus

	Phonétique	Syntaxe	Morphologie
Identique à-identique à	11 (26%)		12 (21.5%)
Identique à-étude de	29 (66.5%)	19 (11%)	17 (30.5%)
Étude de-étude de	3 (7.5%)	157 (89%)	28 (48%)
Total	43	176	57

- 34 La figure 2 (Fig. 2 – Répartition des relations en fonction des micro-thesaurus) montre que le micro-thesaurus *phonétique* privilégie les relations *identique à-étude de* tandis que ceux de la *syntaxe* et de la *morphologie* se fondent principalement sur *étude de* : la combinaison la plus représentée est de loin celle qui utilise deux liens *études de*.
- 35 En ce qui concerne les relations lexicales, nous neutralisons pour cette étude la relation de synonymie puisque nous n'avons observé qu'un cas (*orthoépée* = *orthophonie*).
- 36 Les relations hiérarchiques, comme on pouvait s'y attendre sont nombreuses puisque *est un* est utilisé pour 160 relations soit 58.5 %. Ces relations prennent la forme de combinaisons conceptuelles et se répartissent en relations *identique à* ou *étude de*. La relation de spécificité unit essentiellement des termes reliés à leurs domaines par une relation *identique à* (83 %) ou *étude de* (76 %) : elle est majoritaire dans le cas de termes liés à leur domaine par une relation *identique à*.
- 37 Les relations partitives sont assez peu utilisées (39 soit 14 %). Elles prennent la forme de combinaisons *identique à-identique à*, *identique à-étude de* et *étude de-étude de*. Elles sont assez peu fréquentes entre deux termes liés par des relations différentes (terme dominant lié par *identique à* et terme dominé lié par *étude de*).
- 38 Les relations associatives sont au nombre de 74 soit 17 %. Dans la majorité des cas, elles correspondent à des propriétés étudiées sur les objets, des procédés étudiés dans le domaine. Elles apparaissent lorsque les termes sont liés à leurs domaines par des relations différentes. Dans la majorité des configurations, le terme dominant est lié à son domaine par *identique à* et le dominé par *étude de*.
- 39 Cette représentation des relations internes au thesaurus n'est pas totalement satisfaisante car elle ne permet pas de rendre compte de l'ensemble des liens sémantiques présents dans le thesaurus. Les exemples suivants illustrent des cas pour lesquels d'autres relations sont mises en œuvre. Si nous prenons l'exemple de *phrase* et de *transformation* dans le micro-thesaurus de la *syntaxe*, nous avons $D_1 = \text{étude de la phrase}$ et $D_2 = \text{étude de la transformation}$ avec une ellipse (*transformation de la phrase* quand on restitue l'ellipse) ce qui ne conduit pas à *la transformation (de la phrase) est une phrase*. En effet, *transformation* n'est pas un spécifique de *phrase* tout comme *phrase* n'est pas un générique de *transformation*.
- 40 Ce type d'ellipse n'est pas sans lien avec celui que l'on a pour *transcription (phonétique)* dans le micro-thesaurus de la *phonétique*. La relation qui s'établit entre *phonétique* et *transcription* peut être verbalisée par *technique de* (cf. Fig. 3 – Relations sémantiques repérées dans les micro-thesaurus phonétique, syntaxe et morphologie). Ce type de relation est à rapprocher des relations associatives telles que les définit (Depecker, 2003) : *phonétique* et *transcription* entrent dans une relation *discipline-technique/activité de la discipline* : *phonétique* et *transcription* sont respectivement une discipline et une technique qui portent sur les sons dans les langues naturelles. De plus, il traduit un des phénomènes d'inférence que nous avons évoqué et dont la mise en pratique masque parfois les liens d'héritage qui existent entre les différents niveaux – nous sommes en

présence d'une structure en forme de treillis et non d'une arborescence à héritage simple car il existe des relations transversales.

- 41 Un autre cas correspond à celui des termes complexes de Thesaulangue qui comportent un substantif générique, par exemple *phonétique*, complété par un adjectif pour former un nouveau terme : *phonétique auditive*, *phonétique expérimentale*, *phonétique acoustique*, *phonétique articulatoire*, etc. Ce type de relation relève à la fois de ce que (L'Homme, 2002) classe dans les relations dérivationnelles et des relations de type méronymique : *phonétique* est un holonyme de *phonétique expérimentale* qui en est lui-même un pour *phonétique acoustique* et *phonétique articulatoire*.
- 42 Si on s'intéresse aux travaux de (Mhiri et al., 2006), nous pouvons retenir les relations *identique* à et *synonymie* pour rendre compte des liens *employé pour* présents dans le thesaurus. C'est notamment le cas de *affixe* qui est défini, par Thesaulangue, aux moyens des *employés pour*, comme le seul terme possible pour *affixation*, *élément formant* et *élément formateur*. Le cas d'homonymie, qui existe pour le terme *syntagme nominal* et le terme *syntagme verbal*, tous deux présents dans les micro-thesaurus de la *morphologie* et de la *syntaxe*, ne peut être traduit par aucune des relations proposées. Ils correspondent à des réalités, des phénomènes et des approches différentes dans les deux domaines : le traitement, l'importance et le rôle de ces notions ne sont pas conceptualisés de manière similaire en *syntaxe* et en *morphologie*. Thesaulangue souligne et résout partiellement le problème en précisant le terme par le domaine auquel il appartient : *syntagme nominal (morphologie)*, *syntagme nominal (syntaxe)* et *syntagme verbal (morphologie)*, *syntagme verbal (syntaxe)*.

3.2. Les relations sémantiques mises en œuvre dans Thesaulangue

- 43 Comme nous l'avons évoqué (3.1. Les relations d'ordre conceptuel) la proposition de relations que nous avons faite ne couvre pas l'ensemble des relations sémantiques mises en œuvre lors de la constitution du thesaurus. En effet, certains liens sémantiques peuvent être verbalisés avec un niveau de précision ou des nuances sémantiques propres aux liens qui s'établissent entre les termes dans la représentation que nous avons du domaine. Aussi, il nous paraît intéressant de confronter la verbalisation des liens que nous avons reconstituée avec les informations que nous avons extraites de la littérature sur le sujet (2. Thesaurus, terminologie et onto-terminologie : les relations sémantiques en jeu).

3.2.1. Verbalisation des liens sémantiques présents dans Thesaulangue

- 44 Le tableau proposé ci-dessous dresse l'inventaire des verbalisations des relations que nous percevons entre les termes de Thesaulangue. Il nous paraît nécessaire de les prendre en compte pour rester fidèle à la représentation qu'il propose pour les trois micro-thesaurus sur lesquels nous travaillons (cf. Annexe – Extraits de Thesaulangue).

Fig. 3 – Relations sémantiques repérées dans les micro-thesaurus de la *phonétique*, la *syntaxe* et la *morphologie*.

Relation	Micro-thesaurus	Exemple
technique de	phonétique	phonétique ← transcription
type de	phonétique	son ← consonne
	syntaxe	adjectif ← comparatif
	morphologie	morphologie ← morphologie grammaticale
est analysé en	phonétique	phonétique historique ← amuïssement
	syntaxe	syntaxe ← syntagme
	morphologie	morphologie lexicale ← composition
manifestation de	syntaxe	syntaxe ← ordre des mots
condition nécessaire	syntaxe	phrase ← construction
construction apparentée à	syntaxe	fonction ← apposition
contrainte syntaxique	syntaxe	phrase ← accord
élément de	syntaxe	phrase ← préposition
est le référent	syntaxe	pronom relatif ← antécédent
est le cumul de plusieurs	syntaxe	affixe ← (verbe) parasynthétique

- 45 La diversité des verbalisations possibles nous conduit à formuler et à utiliser plusieurs types de relations, aussi nous nous inspirons largement des relations présentées au paragraphe 2 (2. Thesaurus, terminologie et onto-terminologie : le rôle des relations sémantiques) et mettons à profit nos propres observations sur Thesaulangue pour proposer une première sélection de relations destinée à évaluer la démarche que nous sommes en train d'établir.

3.2.2. Relations proposées par la littérature vs observations sur Thesaulangue

- 46 La confrontation des relations proposées par la littérature et des relations repérées dans Thesaulangue nous a permis de déterminer un ensemble de relations qu'il nous paraît nécessaire de considérer pour amorcer une onto-terminologie. En effet, nous souhaitons prendre en compte des aspects de sémantique lexicale, des relations syntagmatiques ainsi que des relations morphologiques que nous soumettons aux contraintes du courant lexématique.

Fig. 4 – Relations sémantiques retenues pour la confection de l'onto-terminologie

Relations de sémantique lexicale	
	Hierarchiques : <i>sorte de</i>
	Partitives
	Méronymes
	Holonymes
	Associatives
	Procédé de
	Propriété de
	Fonction identique à
	Résultat de
Relations syntagmatiques inspirées de (L'homme, 2002) et enrichies des observations sur Thesaulangue	
	Collocation – modification adjectivale (phonétique, phonétique descriptive)
	Ellipse (phonétique, transcription (phonétique))
	Complémentation du nom (propositions, ordre des propositions)
Relations morphologiques proposées par (L'Homme, 2002) contraintes par les conditions imposées par le courant lexématique (Corbin, 1991, 1997 et 2000)	
	Conversion (phonétique _{substantif} transcription phonétique _{adjectif})
	Suffixation (verbe _{substantif} périphrase verbale _{adjectif relationnel})

- 47 Le détail du décompte des relations présentes dans les trois micro-thesaurus retenus est proposé en figure 5 (Fig. 5 – Répartition des relations retenues dans Thesaulangue).

Fig. 5 – Répartition des relations retenues dans Thesaulangue⁹

		Si T ₂ est dominé par T ₁ , relations de T ₂ vers T ₁									
		Est un	Partitive		Associative						
Combinaison conceptuelle	Micro-thesaurus	Générique	Méronyme	Holonyme	Propriété	Procédé	Fonctionne comme	Caractérisé par	Résultat de	Total micro-thesaurisations de concepts	Total combinaisons de concepts
Id, Id	Phon	9	1	1						11	22
	Morp	10	1							11	
Id, Etud	Phon		5		6	18				29	66
	Synt		13		6					19	
	Morp		9		2	7				18	
Etud, Etud	Phon	3								3	188
	Synt	121	3	3	24	1	5	1		158	
	Morp	18	2	1					6	27	
Total relation sémantique		161	34	5	38	26	5	1	6	273	
Total type de rel. sémantique		161	39		76						276

- 48 Si on tente de synthétiser le fonctionnement des relations de sémantique lexicale en regard de celui des combinaisons conceptuelles, nous pouvons dire que :

- deux relations *identique* à produisent
 - 83 % de relations générique-spécifique
 - 13 % de méro-holonymie
 - 4 % de relations associatives
- deux relations *étude de* produisent
 - 76 % de relations générique-spécifique
 - 19 % de relations associatives
 - 5 % de méro-holonymie
- une relation *identique à* et une relation *étude de* produisent
 - 64 % de relations associatives
 - 42 % de méro-holonymie

- 49 L'utilisation des relations termes/domaines que nous proposons permet d'établir une corrélation entre le niveau conceptuel (combinaisons de relations termes/domaines) et le niveau sémantique (tendances nettes dans le choix des relations de sémantique lexicale). Nous pouvons conclure cette étape du travail en proposant une sorte de présélection des relations ontologiques en fonction des relations sémantiques. L'analyse que nous avons faite permet d'établir des règles prévisionnelles de mise en place des relations utilisées dans Thesaulangue : ces règles sont en quelques sortes des prédictions de correspondances entre les liens conceptuels et des relations sémantiques que nous résumons dans le tableau proposé ci-dessous 6 (Fig. 6 – Corrélation niveau conceptuel/niveau sémantique).

Fig. 6 – Corrélation niveau conceptuel/niveau sémantique

T ₂ dominé par T ₁ Relation ordonnée (T ₂ , T ₁)	T ₁ identique T ₂	D ₁ étude de T ₁
T ₂ identique D ₂	Spécifique-générique Méro-holonymie Relations associatives	
D ₂ étude de T ₂	Relations associatives Méro-holonymie	Spécifique-générique Relations associatives Méro-holonymie

Perspectives

- 50 L'état d'avancement de nos analyses nous laisse entrevoir le travail de structuration, de formalisation qu'il reste à effectuer pour produire une onto-terminologie offrant une image satisfaisante des sciences du langage. La représentation sémantique du domaine et les informations sémantiques codées et héritées doivent être utilisées pour annoter des documents en texte intégral. Ce travail d'analyse doit permettre :
- l'insertion de candidats termes dans le thesaurus pour améliorer la représentation du domaine
 - l'adjonction des définitions des termes du thesaurus afin de constituer les faisceaux sémantiques nécessaires à l'annotation sémantique
 - l'enrichissement du thesaurus en traits sémantiques pour permettre l'annotation sémantique automatique ou semi-assistée
 - la validation de la ressource produite pour nous assurer que l'image qu'elle donnera du domaine des sciences du langage est satisfaisante. Cette étape revient à introduire de nouveaux termes ou supprimer des termes du thesaurus actuel et à harmoniser les relations que les termes entretiennent (nature des relations, hiérarchie des relations, etc.)
- 51 A plus long terme l'onto-terminologie peut constituer un outil pédagogique facilitant la connaissance et l'utilisation des termes et du thesaurus des sciences du langage par des documentalistes novices dans ce domaine scientifique ou par tout utilisateur qui le découvre.

BIBLIOGRAPHIE

- Aussenac-Gilles N. et Bourigault D. (2000). The Th[IC]2 Initiative : Corpus-Based Thesaurus Construction for Indexing WWW Documents ». Proceedings of the EKAW'2000 Workshop Ontologies and texts, Juan-Les-Pins, Université Paul Sabatier, Toulouse, p. 71-78, octobre.
- Bachimont B., Baneyx A. Malaisé V., Charlet J. et Zweigenbaum P. (2005). Synergie entre analyse distributionnelle et patrons lexico-syntaxiques pour la construction d'ontologies différentielles, Terminologie et Intelligence artificielle, Rouen, 4-5 avril.
- Benslimane D, Arara A., Yetongnon K. et Ben Abdallah H. (2003). Two approches for ontologies building : from-scratch and from existing data sources, Interntional Conférence on Information System and Engineering ISE, Montréal, Canada, 20-24 juin.
- Bourigault D., Jacquemin C. et L'Homme M.C. (2001). Recent Advances in Computational Terminology. Amsterdam/Philadelphie : John Benjamins.
- Bourigault D. et Lame G. (2002). Analyse distributionnelle et structuration de terminologie. Application à la construction d'une ontologie documentaire du Droit, Structuration de terminologie, A. Nazarenko et T. Hamon (eds.), Traitement automatique des langues, vol. 43, n° 1, 129-150.

- Bourigault D. et Aussenac-Gilles N. (2003). Construction d'ontologies à partir de textes. Actes de la conférence TALN 2003, Batz-sur-Mer.
- Bourigault D., Aussenac-Gilles N. et Charlet J. (2004). Construction de ressources terminologiques ou ontologiques à partir de textes : un cadre unificateur pour trois études de cas. M. Slodzian (Ed.), *Revue d'Intelligence Artificielle (RIA)*, Numéro spécial sur les techniques informatiques de structuration de terminologies, Hermès, Paris, vol. 18, N. 1/2004, p. 87-110, 2004.
- Calberg-Challot M., Lerat P. et Roche C. (2009), Quelle place accorder aux corpus dans la constitution d'une terminologie ?, Actes de la conférence Toth'09, Annecy, 4-5 juin.
- Corbin D. (1991). Introduction : la formation des mots, structures et interprétations. *Lexique* 10, Villeneuve d'Ascq, PUL, 7-30.
- Corbin D. (1997). Décrire un affixe dans un dictionnaire. Les formes du sens. Etudes de linguistique française, médiévale et générale offertes à Robert Martin à l'occasion de ses 60 ans, Kleiber G. et Riegel M. (eds), Louvain la Neuve, Duculot, 79-94
- Corbin D. (2000). Pour en finir avec la parasyntèse. Mélanges de grammaire et de linguistique française en hommage à M. Riegel, Kleiber G. Pellat J.C. et Buridant C. (eds.)
- Dell'Ortella F., Lenci A., Marchi S., Montemagni S., Pirelli V. et Venturi G. (2008). Dal testo alla conoscenza e ritorno : estrazione terminologica e annotazione semantica di basi documentali di dominio. *Analisi Testuale e Documentazione nella città digitale*, Convegno nazionale dell'Associazione Italiana per la Terminologia. I-TerAnDo, Università di Calabria, Rende, 5-7 juin, AIDAinformazioni, 26, 1-2, pp. 185-206.
- Dendien J. et Pierrel J.M. (2002). Le trésor de la langue informatisé : un exemple d'informatisation d'un dictionnaire de langue de référence. *TAL*.
- Depecker L. (2003). Entre signe et concept : éléments de terminologie générale, Presse de la Sorbonne Nouvelle.
- Drouin P. (2004). Acquisition des termes simples fondée sur les pivots lexicaux spécialisés, TIA, Strasbourg.
- Drouin P. (2003) Terme Extraction Using non-technical corpora as a point of leverage, *Terminology*, vol. 9, no 1, 99-117.
- El Mekki T. et Nazarenko A. (2002). Comment aider un auteur à construire l'index d'ouvrage. Actes de la conférence CIFT 2002. pp. 141 – 157. Tunisie.
- Gaiffe B., Jacquy E. et Kister L. (2009). Approche lexico-sémantique de l'extraction terminologique : utilisation de ressources lexicographiques et validation sur corpus. Actes de la conférence Toth'09, Annecy, 4-5 juin.
- Inalf. (1975). Trésor de la langue française. Le français moderne, (supplément), nouvelle série, fascicule 2.
- Jacquy E., Kister L., Grzesitchak M., Gaiffe B., Reutenauer C., Valette M. et Sandrine O. (2010), *Thesaurus et corpus de spécialité en Sciences du Langage : une approche lexicométrique appliquée à l'analyse de termes en corpus*, Taln'10, Montréal, juillet.
- Jacquy E. et Namer F. (2003) Morphosémantique et modélisation : les verbes dénominaux préfixés par é-, Actes du Colloque Le sens en linguistique 25-27 mai, Montréal.
- Kister L., Jacquy E. et Gaiffe B. (2009). Fusion d'un thesaurus et d'une terminologie : utilisation de ressources existante pour amorcer une onto-terminologie, TIA'09, Toulouse, novembre.

- Kister L., Jacquey E. et Gaiffe B. (2008). Repérage de la référence à partir du thesaurus, de la terminologie et de la sémantique lexicale. *Analisi Testuale e Documentazione nella città digitale. Convegno nazionale dell'Associazione Italiana per la Terminologia, I-TerAnDo, Università di Calabria, Rende, 5-7 juin, AIDAinformazioni*, 26, 1-2, pp. 25-36.
- Kister L. et Jacquey E. (2007a). Comparaison des structures thématiques de textes spécialisés et de thésaurus ou de terminologies. *Terminologia e mediazione linguistica : approcci e metodi a confronto. ASS.I.term et Università di Bologna, sede di Forli, Bertinoro, 8 juin, Realiter, en ligne.*
- Kister L. et Jacquey E. (2007b). Acquisition sémantique à partir de données lexicographiques au service de la comparaison entre des structures thématiques de textes spécialisés et de thésaurus. *Terminologie : approches transdisciplinaires*, Gatineau (Québec), 2-4 mai, en ligne.
- L'Homme M.C. (2005). Sur la notion de terme, *Méta*, 50-4, 1112-1132.
- L'Homme M.C. (2004). *La terminologie : Principes et Techniques*. Presses Universitaires de Montréal.
- L'Homme M.C. (2002). Fonction lexicales pour représenter les relations sémantiques entre termes. *Structuration de terminologie*, A. Nazarenko et T. Hamon (eds.), *Traitement automatique des langues*, vol. 43, n° 1, 19-41.
- Le Pesant D. (2000). Suggestions méthodologiques pour une typologie des classes deméronymes. *Scolia*. Strasbourg : Publications de l'Université Marc Bloch, n° 12, 161-179.
- Mazuel L. et Sabouret N. (2007). Degré de relation sémantique dans une ontologie pour la commande en langue naturelle, *Actes de la conférence IC 2007 – 18e Journées Francophones d'Ingénierie des Connaissances*, Grenoble, 4-6 juillet.
- Mhiri M., Gargouri, F. et Benslimane D. (2006). Détermination automatique des relations sémantiques entre les concepts d'une ontologie. *Conférence INFORSID*. Hammamet. Tunisie
- Nazarenko A. et Hamon T. (2002). Structuration de terminologie. *TAL*, vol 43, n° 1, 174 pages.
- Navigli R. et Velardi P. (2008). From glossaries to ontologies : extracting semantics structures from textual definitions. *Conferences on Ontology Learning and Prpulation : Bridgin the gap between Text And Knowledge*, IOS Press Amsterdam, The Netherlands.
- Rastier F. (2004). Ontologie(s). *Structuration de terminologie*. J.M. Pierrel et M. Slodzian (eds.), *Revue des sciences technologies de l'information, Revue d'intelligence artificielle*, vol 18, n° 1 15-40.
- Roche M. (2004). *Intégration de la construction de la terminologie de domaines spécialisés dans un processus global de fouille de textes*. PhD Thesis, Université Paris 11.
- Roche C. (2008), Faut-il revisiter les Principes terminologiques ?, *Actes de la conférence Toth'08*, 5-6 juin.
- Roche C. (2009). Le terme et le concept : fondements d'une ontoterminologie. *Actes de la conférence Toth'09*, 4-5 juin.
- Sanjuan E. et Ibekwe-Sanjuan F. (2002). Terminologie et classification automatique des textes, *Actes de la conférence JADT 2002*, pp. 677-688.
- Valette M. et Grabar N. (2004). Caractérisation de textes à contenu idéologique : statistique textuelle ou extraction de syntagme ? L'exemple du projet PRINCIP. *Actes de la conférence JADT 2004*.

NOTES

1. Thesaulangue est le thesaurus mis au point et maintenu par le centre de documentation de l'Atilf (initié par l'Inalf) qui a été intégré au portail Termsciences par l'Inist après avoir été mis au format TMF.
 2. TermTLF est une terminologie qu'on peut définir comme le résultat de l'extraction des définitions du dictionnaire de langue générale, le TLFi (Trésor de la Langue Française informatisé), lorsque celles-ci dépendent directement ou par héritage de l'un des domaines suivants : grammaire, lexicographie, lexicologie, linguistique, philologie, phonétique, phonologie, rhétorique, sémiologie, sémiotique, stylistique, toponymie. (Gaiffe et al. 2009)
 3. Les termes correspondants aux intitulés des 20 micro-thesaurus sont inclus dans les 1224 termes que compte le thesaurus (mai 2011).
 4. Détail du nombre de termes par micro-thesaurus en mai 2011.
 5. L'Homme parle de relations pragmatiques.
 6. Selon nous, cette relation est incluse dans la relation de synonymie. La relation ne porte pas sur le concept mais plutôt sur les entités linguistiques qui le désigne.
 7. Pour cette relation, il s'agit, selon nous, de concepts totalement distincts. L'ambiguïté est une due à la polysémie de la forme *fil* et n'atteint en rien les concepts.
 8. L'auteur parle d'antononie.
 9. La combinaison *identique à-identique* n'existe pas en syntaxe. Pour alléger le tableau nous avons choisi de ne pas y faire figurer de ligne sans effectifs.
-

RÉSUMÉS

Cet article s'intéresse aux relations sémantiques et aux liens conceptuels qui s'établissent entre les termes et les concepts du domaine des sciences du langage. Il vise à déterminer les relations qu'il nous paraît nécessaire d'introduire pour interpréter les relations sémantiques, parfois complexes, qui s'établissent entre les différents niveaux de l'arborescence d'un thesaurus et à prévoir l'évolution de celui-ci vers une onto-terminologie. La future ressource présentera l'avantage de cumuler les apports des trois approches développées par le thesaurus, la terminologie et l'ontologie et pourra être utilisée comme outil d'annotation en texte intégral. La ressource terminologique ainsi constituée bénéficiera des analyses sémantique et conceptuelle réalisées à partir du texte intégral : ces analyses contribueront à son enrichissement.

This article deals with the semantic relations and the abstract links which become established between the terms and the concepts of the linguistic's domain. It aims at determining the relations usefull to interpret the semantic links, which become established between the levels of the thesaurus arborescence and to transform the thesaurus into an onto-terminology. The future resource will present the advantage to accumulate the contributions of three approaches – thesaurus, terminology and ontology – and can be used to annotate the complete text. The terminological resource constituted will benefit from semantic and abstract analyses realized from the complete text : these analyses will enrich it.

INDEX

Keywords : domain, ontological relation, ontology, semantic relation, term, terminology, thesaurus

Mots-clés : domaine, ontologie, relation ontologique, relation sémantique, terme, terminologie, thesaurus

AUTEURS

LAURENCE KISTER

Atilf UMR 7118 CNRS-Nancy-Université

EVELYNE JACQUEY

Atilf UMR 7118 CNRS-Nancy-Université

BERTRAND GAIFFE

Atilf UMR 7118 CNRS-Nancy-Université