
Etude des marqueurs de méronymie dans un corpus aligné français-italien

Anne Condamines^{*1}, Marie-Pierre Escoubas^{*†2}, and Silvia Federzoni^{*‡3}

¹CNRS et Université Jean Jaurès (CLLE) – CNRS : UMR51532 – France

²Université La Sapienza (MEMOTEF) – Italie

³Université Jean Jaurès (CLLE) – Université Jean Jaurès, Toulouse 2 – France

Résumé

Des marqueurs lexico-syntaxiques sont souvent utilisés pour repérer, en corpus spécialisés, des relations conceptuelles afin de construire des réseaux terminologiques (Meyer, 2001). Dans ce cadre général, nous utilisons un corpus aligné, traduit du français vers l'italien (Le Monde diplomatique) afin de tirer parti des phénomènes qui se mettent en place lors de la traduction : ajout d'éléments, choix différent dans la structuration des informations (Baker, 1996), (Pearson, 2000)... À la suite d'une étude sur l'hyperonymie (Condamines et al., à paraître), nous nous focalisons ici, sur la méronymie. Des listes de marqueurs de méronymie en français et en italien (essentiellement des verbes comme *réunir* et *décomposer en* ou *dividere in* et *organizzare in*) sont projetés sur le corpus (d'un peu plus d'un million de mots) via l'outil TXM, préalablement traité avec Alinéa pour aligner les textes français et italien. Nous obtenons ainsi des portions de texte qui comprennent les marqueurs équivalents dans les deux langues. Les données obtenues sont annotées pour ce qui concerne : la présence ou pas de la méronymie, la présence d'une autre relation, la traduction par un autre type de marqueur (verbe à pas de marqueur ; verbe à autre élément syntaxique (par exemple, la partie du discours ou l'ordre des éléments variant) ; verbe à verbe "synonyme" dans l'autre langue).

Deux phénomènes particulièrement intéressants apparaissent lors de l'analyse. Tout d'abord, un phénomène qui concerne le fait que, lors de la traduction, de nouveaux marqueurs apparaissent, non recensés dans les listes initiales. Ainsi le verbe italien *comprendere* (comprendre) a permis d'accéder à un marqueur non recensé dans la liste initiale française : *coprire*. Dans une méthode récursive bien connue en ingénierie des connaissances, les marqueurs ainsi identifiés sont à leur tour projetés afin de détecter des marqueurs nouveaux, ceci jusqu'à ce que tous les marqueurs du corpus aient été identifiés. La présentation se focalisera sur ce premier aspect.

Un autre aspect intéressant est en lien avec le fait qu'il existe des marqueurs qui sont beaucoup plus polysémiques dans une langue que dans l'autre. Ainsi le verbe italien *comportare* peut marquer en italien soit une méronymie soit une conséquence (comme dans *la riorganizzazione del personale comporta disoccupazione tecnica / la réorganisation du personnel implique le chômage technique*) alors que le plus souvent, c'est la méronymie qui s'exprime dans le corpus en français à travers l'usage de *comporter*. Des exemples de ce phénomène

*Intervenant

†Auteur correspondant: marie-pierre.escoubas@uniroma1.it

‡Auteur correspondant: silvia.federzoni@univ-tlse2.fr

seront présentés.

Baker, M. (1996). Corpus-based translational studies : the challenges that lie ahead , In Harold Somers (Éd) *Terminology, LSP and Translation: Studies in language engineering* in honour of Juan C. Sager, 175–186.

Condamines, A., Escoubas, M.-P., & Federzoni, S. (sous presse). Apport de la traduction dans l'analyse des marqueurs de relations conceptuelles. Une étude en corpus aligné français-italien. ”. Dans Cécile Frérot & M. Pecman (Éds), *Des corpus numériques à l'analyse linguistique en langues de spécialité*. Grenoble : Presses de l'UGA.

Meyer, I. (2001). Extracting knowledge-rich contexts for terminography: A conceptual and methodological framework. In Didier Bourigault, Jacquemin, C. & L'Homme, M.-C (Éds), *Recent Advances in Computational Terminology*, Amsterdam : John Benjamins Publishing Company, 279-302.

Pearson, J. (2000). Une tentative d'exploitation bi-directionnelle d'un corpus bilingue. *Cahiers de Grammaire, Sémantique et Corpus*, (25), 53-69.

Mots-Clés: corpus aligné, marqueurs de relation, méronymie, traduction