

Patrons morphosyntaxiques nominaux et repérage des termes dans un corpus comparable de communication financière en italien et en français

M.-P. Escoubas-Benveniste¹, A. Condamines^{2,3}, F. Montermini^{2,3}

1. Université La Sapienza, Rome
2. Université Jean-Jaurès, Toulouse
3. CNRS

Mots-clés : syntagme nominal marqueurs morphologiques, terminologie financière, corpus comparables en italien et en français

On sait que les termes se manifestent de manière prédominante, quoique non exclusive, sous la forme de noms et de syntagmes ayant pour base un nom (Rey 1975, 1987 ; Lerat, 1995). En français et en italien, il existe plusieurs moyens, syntaxiques ou morphologiques, de construire une lexie complexe où deux noms sont mis en relation. Le nom tête (N1) peut par exemple être modifié par un syntagme prépositionnel (1a), par un adjectif morphologiquement construit sur un nom (1b) ou par un nom juxtaposé dans un composé morphologique (1c) :

- | | | |
|----|------------------------------|-------------------------|
| 1) | italien | français |
| | a. <i>chiusura per ferie</i> | <i>course à pied</i> |
| | b. <i>chiusura serale</i> | <i>course landaise</i> |
| | c. <i>chiusura lampo</i> | <i>course poursuite</i> |

Dans les deux langues, il existe, cependant, d'autres modèles de construction de noms qui sont productifs, comme le montrent les exemples suivants :

- 2) it: *calciomercato, fantacalcio*¹, *tiro sottoporta*²
fr: *éco-innovation, écoredevance, éco-PTZ (éco-prêt à taux zéro)*³

L'objectif de notre contribution est d'examiner et de comparer les patrons morphologiques et syntaxiques qui sont attestés dans un genre textuel de la langue économique en français et en italien. Nous nous proposons d'identifier, sur la base d'une description linguistique, lesquels de ces patrons pourraient correspondre à des termes dans les deux langues. Comme les textes examinés sont des versions linguistiques originales, l'attribution du statut de terme possible à une « portion de texte » ne peut se fonder sur l'hypothèse d'une équivalence de traduction instaurée par le contexte – postulat qui sous-tend généralement l'extraction automatique des termes bilingues à partir de corpus alignés. De plus, même si le genre textuel considéré est pragmatiquement comparable, il n'est cependant pas soumis à de fortes contraintes rédactionnelles, ce qui empêche d'opérer un alignement fondé sur la structure du texte. C'est

¹ Exemples illustrant la langue du foot cités et traduits par F. Montermini (2012) qui signifient respectivement "marché des joueurs de football" et "jeu de fantasia lié au football".

² Cf. (Montermini 2008). En français "tir près du but".

³ Exemples extraits d'un corpus de presse. (Escoubas-Benveniste, 2012). En ce qui concerne le statut terminologique du formant *éco-*, nous renvoyons le lecteur à l'étude de P. Dury (2007) sur sa "déterminologisation".

pourquoi nous envisageons, dans un premier temps, d'extraire et de décrire, séparément pour chaque sous-corpus, les patrons récurrents dans le genre de texte choisi, qui émane de deux réalités sociolinguistiques distinctes. L'intérêt d'utiliser un corpus bilingue comparable réside dans la possibilité de détecter des portions de texte ou contextes susceptibles d'être mis en correspondance après l'analyse.

La ressource que nous utilisons est un corpus du domaine économique. Ce dernier se compose d'un échantillon de communiqués de presse publiés par des entreprises italiennes et françaises cotées en bourse pour répondre à une obligation d'information concernant leur situation. Ce genre textuel relève de la communication financière (Léger 2010) une branche de la communication institutionnelle de l'entreprise. Bien que le corpus que nous analysons soit de dimensions modestes (env. 147 000 mots pour le sous-corpus italien ; environ 112 000 mots pour le sous-corpus français) il va nous permettre, à partir de faits linguistiques réels, d'amorcer une analyse comparée qui, à notre connaissance, est inédite pour ces deux langues dans le domaine économique et financier et le genre textuel considérés.

Pour analyser les séquences à tête nominale dans le but d'évaluer si des processus de composition sont à l'œuvre, la nature syntaxique ou morphologique de l'expansion du nom offre une première ressource. On sait en effet que, si le nom N1 est déterminé par un adjectif dénominal, le syntagme qui en résulte donne lieu à différentes relations possibles entre le nom base dont dérive l'adjectif et la tête du syntagme N1 :

- it : *sistema muscolare (sistema dei muscoli, sistema costituito dai muscoli)*
*strappo muscolare (strappo del muscolo, il muscolo si strappa)*⁴
- fr : *cheminée volcanique (cheminée du volcan, i.e. partie du volcan)*
éruption volcanique (éruption du volcan, le volcan entre en éruption)

Pour certains linguistes, ces "pseudo-adjectifs relationnels" (Monceaux 1997) auraient un sens lexical réductible à celui de leur base, et leur suffixe aurait un rôle purement fonctionnel, consistant à signaler que le radical et le nom recteur du syntagme sont liés par une relation de détermination (Mélis-Puchulu 1991, Wandruska 2004, Rainer 2013). D'autres linguistes considèrent au contraire que les affixes sont polysémiques, à l'instar des lexèmes (Lehrer, 2000). Selon cette interprétation l'adjectif dénominal pourrait acquérir un sens partiellement différent de celui du nom. La relation sémantique qui s'établit est ainsi "marquée" par le suffixe, dont la présence est susceptible d'indiquer la présence d'un "contexte riche en connaissances" (Meyer 2001). Le contenu conceptuel de la relation sous-jacente à ce contexte doit être construit. La construction du sens s'opère à plusieurs niveaux : celui des concepts du domaine spécialisé associés aux noms du syntagme – ainsi *muscolo* et *sistema*, *cheminée* et *volcan* sont dans une relation de la partie au tout, tandis que *muscolo* et *strappo*, *volcan* et *éruption* sont un des participants du prédicat auquel réfère le nom d'action – mais aussi celui du genre textuel, issu de contraintes situationnelles précises, dont on sait qu'il peut créer ses propres marqueurs, ou des fonctionnements de marqueurs connus propres à ce genre (Condamines 2007).

⁴ Cf. Wandruska 2004, in Grossmann, M., Rainer, F. (eds) p. 383.

Les adjectifs dénominaux relationnels sont d'un usage très fréquent dans les langues spécialisées (Grossman *et al.* 2004) : dans son étude de 2009 sur un corpus d'articles de médecine (750 000 mots) Maniez fait apparaître que les cinq suffixes du français les plus productifs pour la construction d'adjectifs à base nominale sont à l'origine de plus de 400 adjectifs différents dans le corpus médical. Par ailleurs, les études en traitement automatique de la langue naturelle ont mis en évidence l'intérêt méthodologique que présentent les adjectifs relationnels pour l'extraction de termes et de connaissances à partir de corpus (Daille 1999, Cartoni 2009)⁵.

Les adjectifs suffixés seront donc une première voie d'accès privilégiée aux syntagmes nominaux dans la méthode d'exploration du corpus que nous utiliserons. Pour ce qui concerne les syntagmes qui ne contiennent pas d'adjectif relationnel, N prép N (1a), et plus particulièrement N + N (1c), l'analyse tirera parti d'autres types de marqueurs, pour établir leur statut terminologique potentiel sur la base d'autres relations que les mots du syntagme instaurent avec d'autres éléments de leur co-texte. Nous nous posons la question de savoir quelles similarités et quelles divergences les patrons morphologiques susceptibles d'être des termes présentent, d'un sous-corpus à l'autre, d'une langue à l'autre, dans le genre textuel examiné.

Bibliographie

Cartoni Bruno (2009) "Les adjectifs relationnels dans les lexiques informatisés : formalisation et exploitation dans un contexte multilingue", TALN 2009, Senlis, 24-26 juin 2009.

Condamines, Anne (2007) « L'interprétation en sémantique de corpus : le cas de la construction de terminologies », Revue française de linguistique appliquée 2007/1 (Vol. XII), p. 39-52.

Daille Béatrice (1999) "Identification des adjectifs relationnels en corpus", Conférence TALN 1999, Cargèse, 12-17 juillet 1999

Dury, P. 2007. « La déterminologisation du formant *eco-* et la terminologisation de l'adjectif *carbon neutral* : un aperçu diachronique de la migration des termes » Actes de la 7ème Conférence Terminologie et Intelligence Artificielle (TIA). Sophia-Antipolis, France, 8-10 octobre

Escoubas-Benveniste, (2012) "Le Monde et le « Grenelle de l'Environnement » : pistes pour l'analyse sémantique assistée par ordinateur d'un corpus de presse" In Proceedings of 10th International Conference Journées d'Analyse statistique des Données Textuelles, JADT 2010, a cura di Sergio Bolasco *et al.*, Milano: Edizioni universitarie di Lettere Economia Diritto, vol. 1/2010, p. 409-422.

Grossmann, Maria, Rainer, Franz (eds) (2004) *La formazione delle parole in italiano*. Tübingen: Niemeyer

Léger, Jean-Yves (2010) *La communication financière*, Paris : Dunod

Lerat, Pierre (1995) *Les langues spécialisées*, coll. Linguistique nouvelle, Paris : PUF

Maniez, François (2009) L'adjectif dénominal en langue de spécialité : étude du domaine de la médecine. Revue française de linguistique appliquée, 2009, XIV-2 (117-130)

⁵ Dans l'application décrite par B. Daille, *produit laitier* est ainsi associé à d'autres termes *produit au lait*, *produit du lait*, *produit issu du lait* qui sont l'expression de relations conceptuelles différentes.

- Meyer I., (2001). « Extracting Knowledge-rich Contexts for Terminography : A Conceptual and methodological Framework ». D. Bourigault, M.C. L'Homme, C. Jacquemin (eds) : *Recent Advances in Computational Terminology*, John Benjamins. pp. 279-302.
- Mélis-Puchulu Agnès. (1991). "Les adjectifs dénominatifs : des adjectifs de "relation"." *Lexique*, 10, La formation des mots, structures et interprétations, Presses Universitaires de Lille. 33–60
- Monceaux Anne (1997) "Adjectifs de relation, complémentation et sous-classification". In: *Langages*, 31^e année, n°126, 1997.
- Montermini, Fabio (2012). *L'italien. La vie d'une langue*, Toulouse : Éditale éditions.
- Montermini, Fabio (2008). *Il lato sinistro della morfologia. La prefissazione in italiano e nelle lingue del mondo*, Milano, FrancoAngeli
- Rainer, Franz (2013) "Can relational adjectives really express any relation? An onomasiological perspective, *International Journal of Social Sciences and Education* 10(1):12-41
- Rey, Alain (1979, 1992²) *La terminologie : noms et notions*, Que sais-je ? Paris: PUF.
- Rey, Alain (1987) « Les terminologies, un défi pour le lexicologue », in I. Tamba (ed.) *Etudes de lexicologie, lexicographie et stylistique offertes en hommage à Georges Matoré*, Paris : Société pour l'information grammaticale
- Wandruska, Ulrich (2004) "Aggettivi di relazione". In Grossmann, M., Rainer, F. (eds) *La formazione delle parole in italiano*. Tübingen: Niemeyer, 382-401