

Une nouvelle ressource pour la description des collocations : le DicoLiLex, dictionnaire de liens lexicaux

Sébastien Marengo et Benoît Robichaud, Université de Montréal

Parmi les contraintes pesant sur l'expression du sens linguistique, celles impliquées par les collocations occupent une place prépondérante. Ainsi, lorsqu'un locuteur souhaite exprimer un sens donné auprès d'une unité lexicale, la *base*, il doit savoir s'il dispose d'une autre unité lexicale, le *collocatif*, dont le choix est contraint par la première (Mel'čuk 2013). Comme les collocations sont nombreuses au sein d'une langue et mettent en jeu un large éventail de relations sémantico-lexicales, leur description représente un défi de taille: il faut non seulement distinguer clairement les bases, mais aussi caractériser de manière précise les relations qui les unissent aux collocatifs. Dès lors, le dictionnaire de collocations apparaît comme un outil indispensable (Tutin 2005), véritable passerelle entre le linguiste qui décrit les collocations et l'utilisateur qui souhaite les maîtriser ou les étudier.

Pour produire un tel dictionnaire, la Lexicologie explicative et combinatoire (Mel'čuk, Clas & Polguère 1995) est particulièrement adaptée. D'une part, elle dispose d'un ensemble de critères pour distinguer et décrire sémantiquement les unités lexicales, ou *lexies* (acceptions). D'autre part, elle exploite un puissant formalisme, celui des *fonctions lexicales*, pour caractériser les liens lexicaux, notamment syntagmatiques. Ceci permet de consigner les collocations sous forme de triplet: la lexie vedette (la base), la fonction lexicale en jeu et la lexie ramenée par la fonction (le collocatif).

Cependant, les fonctions lexicales sont souvent difficiles d'accès. Cela tient à ce qu'elles sont abstraites et/ou nécessitent une bonne connaissance du cadre théorique dont elles sont issues. Prenons comme exemple la collocation *se faire ami*, comme dans la phrase *Il s'est fait ami avec sa petite voisine*. La base est ici la lexie AMII.1, dont la forme propositionnelle est *l'individu X est un ami de l'individu Y*, tandis que le collocatif est la lexie SE FAIRE. La fonction lexicale en jeu se nomme Caus₁Oper₁ :

$$\text{Caus}_1\text{Oper}_1(\text{AMII.1}) = \text{SE FAIRE} [\sim \text{avec } N_Y]$$

Cela signifie: 'SE FAIRE est un verbe exprimant que le 1^{er} actant sémantique de AMII.1 cause le commencement de la situation décrite par un verbe support dont le 1^{er} actant syntaxique profond correspond au 1^{er} actant sémantique de AMII.1 et dont le II^e actant syntaxique profond est AMII.1'. Or, on ne peut pas exiger d'un usager non-spécialiste qu'il soit familier avec les formalismes sous-jacents à une telle description (cf. L'Homme 2014).

C'est pour résoudre ce genre de problème qu'est né le DicoLiLex (*Dictionnaire de liens lexi-caux*), une ressource actuellement développée par l'Observatoire de linguistique Sens-Texte (OLST). L'objectif visé est de décrire les liens lexicaux du français, essentiellement pour les collocations, à l'aide des principes de la Lexicologie explicative et combinatoire, mais de façon à permettre un accès modulable. L'hypothèse posée est la suivante: si les données lui sont présentées adéquatement, l'utilisateur non-spécialiste aura une compréhension intuitive des concepts théoriques qui les sous-tendent.

Concrètement, le DicoLiLex est constitué de deux composantes :

- Le cœur de la ressource est une base de données de liens lexicaux. Les fonctions lexicales y sont accompagnées de gloses en langue naturelle (Mel'čuk & Polguère 2007), ce qui résout le problème de leur opacité. Ainsi, à l'article AMIL.1, on voit l'indication suivante :

'X faire en sorte qu'il devient ~ de Y' : SE FAIRE [~ avec N_Y]

À l'heure actuelle, la base de données contient environ 5000 lexies vedettes et 5900 fonctions lexicales, pour un total d'environ 37 000 liens lexicaux. Chaque lien lexical sera illustré par un exemple.

- La deuxième composante est une interface en ligne offrant deux modes de consultation : un mode classique, grâce auquel l'utilisateur peut choisir une lexie dans la liste alphabétique des vedettes, et un mode recherche, qui permet une interrogation selon les différents éléments des liens lexicaux (bases, fonctions lexicales accompagnées de leurs gloses, collocatifs). Chaque article se présente comme une hiérarchie que l'utilisateur peut déployer ou compacter selon ses besoins. À terme, l'utilisateur pourra basculer entre les gloses des fonctions lexicales et leurs notations techniques.

Dans l'ensemble, le DicoLiLex s'adresse à tous ceux qui s'intéressent à la représentation du sens véhiculé par les collocations : linguistes, terminologues, traducteurs, spécialistes du traitement automatique des langues, enseignants, apprenants du français, étudiants, etc.

Références

DiCouèbe, Observatoire de linguistique Sens-Texte, <http://olst.ling.umontreal.ca/dicouebe/>.

L'HOMME, Marie-Claude (2014). « Why Lexical Semantics is Important for E-Lexicography and Why it is Equally Important to Hide its Formal Representations from Users of Dictionaries », *International Journal of Lexicography*, vol. 27, n° 4, p. 360-377.

L'HOMME, Marie-Claude, Benoît ROBICHAUD & Patrick LEROYER (2012). « Encoding collocations in DiCoInfo: From formal to user-friendly representations », dans Sylviane Granger & Magali Paquot (éd.), *Electronic Lexicography*, Oxford, Oxford University Press, p. 211-236.

MEL'ČUK, Igor (2013). « Tout ce que nous voulions savoir sur les phrasèmes, mais... », *Cahiers de lexicologie*, n° 102, p. 129-149.

MEL'ČUK, Igor, André CLAS & Alain POLGUÈRE (1995). *Introduction à la lexicologie explicative et combinatoire*, Louvain-la-Neuve, Duculot.

MEL'ČUK, Igor A. & Alain POLGUÈRE (2007). *Lexique actif du français*, Bruxelles, De Boeck.

MEL'ČUK, Igor et al. (1984-1999). *Dictionnaire explicatif et combinatoire du français contemporain*, Montréal, Presses de l'Université de Montréal, 4 volumes.

TUTIN, Agnès (2005). « Le dictionnaire de collocations est-il indispensable ? », *Revue française de linguistique appliquée*, vol. X, n° 2, p. 31-48.