



PHRASÉOLOGIE ET TRADUCTION : BASES DE DONNÉES ET APPLICATIONS

A. GREZKA

CNRS - Université Paris 13 Sorbonne Paris Cité

Le traitement automatique des langues quel que soit son objectif doit s'appuyer sur des ressources linguistiques. Ces ressources sont de natures diverses (dictionnaires, lexiques, textes, etc.). L'élaboration de telles ressources est un axe de recherche fondamentale. Les dictionnaires de « mots simples » sont les ressources les plus faciles à constituer et les plus importantes puisqu'elles sont exploitées lors de l'opération d'analyse lexicale, ce qui correspond à la première étape de n'importe quel type de traitement. Le résultat de cette analyse est ensuite utilisé au cours de traitements ultérieurs, comme l'analyse syntaxique puis sémantique. Mais la langue n'est pas constituée uniquement de mots simples. On y trouve en grand nombre des suites figées, c'est-à-dire dont les éléments constitutifs n'ont pas d'autonomie. Le figement pose des problèmes spécifiques en analyse automatique des textes en langues naturelles. Les systèmes existants ne le prennent pas en compte de manière satisfaisante car des méthodes précises pour le reconnaître manquent.

Le figement, souvent négligé dans l'enseignement des langues et dans la traduction, occupe une place importante dans les préoccupations de la recherche actuelle, et ce en raison de son rôle déterminant dans les systèmes linguistiques de reconnaissance. Les études récentes montrent qu'il s'agit d'une donnée de base incontournable dans la description des langues et d'un fait économique pour le fonctionnement du système. C'est en partant d'un bilan général des études portant sur ce phénomène que nous allons expliciter ces caractéristiques et préciser les perspectives que l'étude du figement ouvre devant la recherche linguistique en général. L'objectif de cette communication est tout d'abord de présenter une méthode d'extraction de ces locutions figées en corpus. Nous proposons un moyen d'organisation des concordances, destiné à aider les terminologues, linguistes, informaticiens, etc. qui souhaitent créer de la ressource linguistique. Nous souhaitons proposer un modèle capable de produire toutes les formes possibles d'une expression donnée. Il faut en effet fournir une ressource susceptible d'être projetée sur un document écrit au cours du processus d'analyse. Ce type d'élément lexical est complexe à représenter car il se présente sous les aspects les plus divers et ne respecte souvent pas les règles standards de morphologie ou de syntaxe. Nous présentons ensuite les applications possibles de cette ressource, par le biais de la base FixISS, réunissant actuellement des locutions françaises, espagnoles et polonaise. L'utilisation des locutions figées dans les systèmes de traitement automatique est fondamentale quel que soit le type d'application envisagé. Cependant les systèmes où l'inclusion de ces locutions est la plus intéressante sont les applications où la langue est envisagée par rapport à une autre langue : les outils de traduction automatique et les plateformes d'apprentissage des langues. Ces unités lexicales tiennent une place importante dans la maîtrise d'une langue. En effet, l'utilisation de ces unités en contexte s'avère difficile pour des locuteurs non-natifs tandis que les locuteurs natifs y ont souvent recours. L'emploi par un non-natif d'expressions figées est un bon indice de ses compétences langagières. En ce qui concerne les outils de traduction il suffit d'utiliser les produits à large couverture, comme ceux proposés en libre accès sur le web, pour se rendre compte rapidement que leur performances sont peu satisfaisantes, tant sur le plan syntaxique que lexical et sémantique.



Bibliographie

- ANSCOMBRE, J.-C. et MEJRI, S. (2011), *Le figement linguistique : La parole entravée*, Paris, H. Champion.
- BISKRI, I., MEUNIER, J.G. et JOYAL, S. (2004), « L'extraction des termes complexes : une approche modulaire semi-automatique », *Le Poids des mots*, PURNELLE G., FAIRON C. & DISTER A. (éds), Actes des 7èmes Journées Internationales d'Analyse Statistique des Données Textuelles, Presses Universitaires de Louvain, Louvain-La-Neuve, Belgique, Volume 1, pp.192-201.
- GONZÁLES REY, I. (dir.) (2014), *Outils et méthodes d'apprentissage en phraséodidactique*, Bruxelles, E.M.E., Proximités « Didactique », 321 pages.
- GREZKA, A. (2009), *La polysémie des verbes de perception visuelle*. Collection *Sémantiques*, L'Harmattan, Paris, 298 pages.
- GREZKA, A. (2011), « La base de données Figement ». In HAJOK A. et S. MEJRI (eds), *Le figement linguistique et les 3 fonctions primaires (prédicats, arguments, actualisateurs)*, *Neophilologica*, 23, Université de Silésie, Katowice, Pologne, pp. 15-28.
- GREZKA, A. (2013), « Le figement absolu : les locutions adverbiales ». In MOGORRON HUERTA P., GALLEGO HERNANDEZ D., MASSEAU P. et M. TOLOSA IGUALADA (eds), *Fraseología, Opacidad y Traducción*, Collection *Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation*, vol. 86, Peter Lang GmbH, Frankfurt, pp. 67-81.
- GREZKA, A. (2014), "Elaboration d'outils méthodologiques pour décrire les séquences figées". In P. MOGORRON (ed), *Figement, traduction, variantes-défigement*, 5, Temática.
- GREZKA, A. (2015), "Problèmes théoriques et méthodologiques sur la constitution d'outils en phraséologie". In P. MOGORRON et F. NAVARRO DOMINGUEZ (eds), *Fraseología, didáctica y traducción*, Collection *Studien zur romanischen Sprachwissenschaft und interkulturellen Kommunikation*, vol.101, Peter Lang GmbH, Frankfurt, pp. 93-104.
- GREZKA, A. (2016). "La base FixISS : traitement du figement dans les locutions adverbiales". In KRAIF O. et A. TUTIN (eds), *Phraséologie et linguistique appliquée*, Cahier de lexicologie, 108, Classiques Garnier, Paris, pp.75-90.
- GREZKA, A. (2016) (Traduction Takada H. et A. Kijima). "凝結表現の記述のための方法論的ツールの開発". In TAKADA H. (ed), *Gengo no fuhensei to kobetusei, 言語の普遍性と個別性*, 7, Maison Doctorale de la société et de la culture modernes de l'Université de Niigata, Niigata (Japon), pp. 21-33. http://dspace.lib.niigata-u.ac.jp/dspace/bitstream/10191/40561/1/7_21-33.pdf
- GREZKA, A. (à paraître). "Présentation du projet FixISS", In GREZKA A. et M.-H. VIGUIER (eds), *Figement et mouvement : critères et outils*, Revue ELA (Études de linguistique appliquée), Klincksieck.
- GREZKA, A. et L. MENESES LERIN (2013). "La fijación absoluta en la enseñanza de los idiomas y en la traducción". In I. GONZÁLES REY (ed), *Didáctica y traducción de las unidades fraseológicas*, Publicación de la Universidade de Santiago de Compostela, pp. 261-274.
- GREZKA, A. et M.-H. VIGUIER (eds) (à paraître). "Figement et mouvement : critères et outils", *Actes de la Journée d'étude du 17 juin 2015*, Revue ELA (Études de linguistique appliquée), Klincksieck.
- GROSS, G. (1996), *Les expressions figées en français. Noms composés et autres locutions*, Paris, Ophrys.
- GROSS, M. (1988), *Grammaire transformationnelle du français. Syntaxe de l'adverbe*, Paris, Cantilène.



GROSSMANN, F. et TUTIN, A. (2003), *Les collocations : analyse et traitement*, De Werelt, Amsterdam.

JESPERSEN, O. (1971), *La philosophie de la grammaire*, Les Editions de Minuit, Paris (Traduction de A.- M. Léonard).

LAMIROY, B. (coord.), KLEIN J.-R., LABELLE J., LECLÈRE C., MEUNIER A., et C. ROSSARI (2009), *Les expressions verbales figées de la francophonie. Belgique, France, Québec et suisse*, Paris, Ophrys.

SVENSON, M.-H. (2004), *Critères de figement. L'identification des expressions figées en français contemporain*, Umeå, Skrifter från moderna språk.