



**HAL**  
open science

## Une entrée lexicale spécifique : les collocations

Cristelle Cavalla

### ► To cite this version:

Cristelle Cavalla. Une entrée lexicale spécifique : les collocations. María Laura Perassi; Martín Tapia Kwiecien. Palabras como puentes : estudios lexicológicos, lexicográficos, y terminológicos desde el Cono Sur, Buena Vista Editora, pp.58-67, 2019, 978-987-4984-09-8. <hal-02533954>

**HAL Id: hal-02533954**

**<https://hal.science/hal-02533954v1>**

Submitted on 6 Apr 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

# UNE ENTRÉE LEXICALE SPÉCIFIQUE: LES COLLOCATIONS

Cristelle Cavalla  
(Université Sorbonne Nouvelle-Paris 3, France)

## Introduction

Comme les autres langues, la langue française utilise de nombreuses collocations – le polonais et le français (Grobelak, 1990), l’anglais (Lewis, 2000), l’espagnol (Solano Rodriguez, 2012)– malheureusement l’enseignement de ces éléments phraséologiques n’est que peu pris en compte en Français Langue Étrangère (FLE). Le choix des collocations comme objet d’étude tant en linguistique que pour l’enseignement découle du constat récurrent de leur absence dans les écrits des étudiants allophones en France (Cavalla, 2008). Ajoutons que de nombreux auteurs s’accordent sur la citation de Hoey (2005) traduite par Edmonds :

Un des traits qui différencient le langage d’un apprenant avancé de celui d’un locuteur natif, c’est que les énoncés prononcés sont souvent corrects d’un point de vue grammatical, mais inappropriés d’un point de vue collocationnel (2013a, p. 131).

Les écrits d’étudiants allophones en Master ont révélé ce point de vue de Hoey. De fait, il paraît nécessaire d’enseigner ces éléments aux étudiants allophones afin qu’ils puissent rendre des écrits universitaires répondant aux critères rédactionnels attendus dans les universités françaises.

Ce sont d’abord les obstacles rencontrés par les étudiants allophones face aux collocations que nous allons décrire. Leurs erreurs ont permis de comprendre les difficultés qu’engendrent ces éléments quasi-figés dans le cadre de leur écriture académique. Ensuite, nous verrons les définitions des collocations pour l’enseignement des langues afin de cerner le phénomène dans un cadre pédagogique. Enfin, nous présenterons une séquence didactique développée pour l’aide à l’écriture à l’aide du corpus numérique en ligne Scientext<sup>14</sup>.

---

<sup>14</sup> ANR Scientext, dirigé par A. Tutin et F. Grossmann au Lidilem de 2007 à 2011. En ligne sur <<http://scientext.msh-alpes.fr/scientext-site/spip.php?article1>>.

## Les collocations : difficultés pour les apprenants

Trois grandes difficultés ont été observées au cours des différentes études menées sur des corpus d'apprenants, notamment des mémoires de master et quelques thèses de doctorat. Il s'agit de la traduction, du repérage incertain de figements et de leur fréquence élevée.

Ce sont d'abord les essais de traduction de ces éléments à l'écrit et à l'oral, qui nous ont permis de repérer les difficultés. L'enseignant repère que l'apprenant ne les connaît pas car la phrase contient des éléments qui ne font pas tout à fait sens. Voici quelques exemples :

- A l'écrit : « nous nous *demandons* la question suivante » pour « nous *posons* la question suivante ; « prendre *recours à* » pour « prendre *appui sur* ».
- A l'oral, seulement depuis quelques années : « visiter quelqu'un<sup>15</sup> ».

Ces deux types d'erreurs lexicales sont d'autant plus intéressants qu'ils se généralisent auprès d'allophones non anglophones natifs. Ces erreurs sont désormais fréquentes chez des étudiants chinois ou latino-américains dont le français est la L3 après l'anglais qu'ils ont étudié comme L2.

Le deuxième critère de reconnaissance des difficultés est la présence d'erreurs chez les apprenants qui révèlent une connaissance partielle de la collocation. Peut-être par absence d'un enseignement de ces éléments. Voici quelques exemples :

- Il peut y avoir un intérêt à *continuer à investiguer* cette *piste de recherche* :
  - intérêt à *poursuivre les investigations* ; mais « piste de recherche » est connu par l'apprenant.
- C'est un problème qui concerne aussi les *petits natifs*... :
  - les *jeunes locuteurs natifs*.
- Ils faisaient un *vacarme de bruit* pour m'empêcher de... :
  - Ils faisaient un *vacarme assourdissant* (qui m'empêchait de...).

Ainsi, on s'aperçoit que les apprenants ont repéré une certaine fréquence dans des figements, mais ne les ayant pas appris de façon systématique, ils ne les connaissent que partiellement. Ces deux entrées d'erreurs pour les apprenants, mettent au jour la fréquence des collocations à l'écrit d'abord, à l'oral ensuite.

---

15 De l'anglais : *to visit someone*.

Plusieurs auteurs avaient attesté la présence fréquente des collocations en linguistique de corpus, notamment Sinclair qui en 1991 pour l'anglais :

Les collocations sont apparues partout dans les concordances générées à partir du corpus. L'idiomaticité, loin d'être un aspect marginal de la langue, semblait être omniprésente et au moins aussi importante que la syntaxe dans la construction du sens (p. 112)<sup>16</sup>.

Pour Lewis (2000), ces « chunks » représenteraient plus de la moitié de nos discours (oraux et écrits) et les collocations seraient alors la trame de fond du lexique mental d'un locuteur natif. De fait, il faudrait les enseigner au même titre que le reste du lexique :

Nous savons désormais qu'une grande part de notre 'vocabulaire' est constituée de différentes structures préfabriquées. Les plus simples d'entre elles sont les collocations. En conséquence, nous pensons que l'enseignement des collocations devrait être une priorité dans tous les cours de langues étrangères (Lewis, 2000)<sup>17</sup>.

Pour Mel'cuk (2003), il s'agit clairement d'un trait de reconnaissance du natif :

En fait, ce sont la fréquence et la qualité de leur usage [les phrasèmes] qui déterminent la différence entre un locuteur natif et un étranger qui a bien appris la langue : UN NATIF PARLE EN PHRASÈMES<sup>18</sup> (p. 26).

González-Rey (2008) décrit des « prêt-à-parler » (p. 7), de petits figements de deux mots (*travailleur indépendant*) que les locuteurs natifs utilisent abondamment sans en être réellement conscients. Une image est alors intéressante à noter, celle de Calaque :

La collocation c'est comme la prose de M. Jourdain, on en fait tous les jours sans même le savoir (Calaque, 2006).

---

<sup>16</sup> « Collocational patterns appeared everywhere in the lexical concordances generated from the corpus. Idiomaticity, far from being a marginal aspect of language, seemed to be ubiquitous and at least as significant as syntax in the construction of meaning » (Sinclair, 1991, p. 112 ; traduction personnelle).

<sup>17</sup> « We now recognise that much of our 'vocabulary' consists of prefabricated chunks of different kinds. The single most important kind of chunk is collocation. Self-evidently, then, teaching collocation should be a top priority in every language course » (Lewis, 2000 ; traduction personnelle).

<sup>18</sup> Les majuscules sont de l'auteur.

Il est donc communément admis qu'un locuteur non natif d'une langue – quelle qu'elle soit – est inconsciemment évalué positivement par les natifs quand il utilise les collocations. Ainsi, la fréquence représente une difficulté supplémentaire pour les apprenants dès l'instant où elle est repérée par les natifs.

Pour résumer, en matière de collocations, les linguistes savent qu'il existe dans la langue une attirance entre certaines lexies (Hausmann *et al.*, 2006). On retrouve ce constat chez Hoey (2005) dans la théorie du *lexical priming* (« amorçage lexical ») qui rejoint finalement le point de vue de Rastier (1987). En effet, ces deux chercheurs expliquent que l'utilisation des lexies dans les phrases est fortement influencée par leurs usages contextuels précédents (comme l'environnement lexical, sémantique, syntaxique, discursif, etcétera). Enfin, la fréquence de la collocation dans un corpus doit être considérée comme un facteur important. Comme le mentionne Kamber (2014), les grands corpus numériques notamment, fournissent des exemples prototypiques d'usage des lexies, mais pas décisifs pour certaines collocations qui, malgré une fréquence peu élevée, restent « disponibles » à l'usage<sup>19</sup>. Par conséquent, ces collocations peu fréquentes sont pertinentes du point de vue didactique et entrent dans la conception des *collocations statistiquement significatives* de Williams (2003) et de *la disponibilité lexicale des collocations* d'Hausmann (1985), reprise par Tutin *et al.* (2002). Bak Sienkiewicz (2016) en conclut que, pour les noms d'émotions, les lexies prises isolément sont moins fréquentes qu'en association lexicale (en collocations). Mais qu'est-ce qu'une collocation ? Pourquoi l'a-t-on si longtemps ignorée malgré sa fréquence ?

### Les collocations : définitions

La définition de la collocation est ici envisagée pour son enseignement. De quelle définition a-t-on besoin pour ensuite enseigner cet élément linguistique à des apprenants ou de futurs enseignants de langue ? Une réflexion autour de la transposition didactique de la définition des collocations s'est mise en place dans le cadre de la formation de futurs enseignants de langue. Comme souvent, quelques exemples suffisent aux étudiants et aux enseignants non linguistes (dans le cadre de formations continues en FLE) pour comprendre ce phénomène et le repérer. Ces mêmes

---

<sup>19</sup> Nous constatons ainsi que les faibles résultats statistiques de certaines combinaisons peuvent être liés au type de corpus choisi, aux préférences linguistiques de tel ou tel locuteur ou même de tel ou tel support.

étudiants – futurs enseignants et enseignants en formation continue – apprécient ensuite d’avoir les critères définitoires du phénomène, certainement pour le fixer dans leur mémoire. De fait, une définition quelque peu vulgarisatrice mais apparemment efficace pour ces publics d’enseignants est apparue intéressante à plusieurs niveaux pour la découverte des particularités des collocations (Bard, 2007 ; Montagnon, 2007). Par exemple, comprendre que les deux éléments de certaines d’entre elles sont inséparables (« avoir une peur bleue ») tandis que d’autres moins. En effet, nombre de collocatifs et de bases (terminologie de Hausmann [1979]) peuvent être séparés par des éléments (« prendre *de longues* vacances »). Cependant, reste la question des éléments qui ne fonctionnent pas réellement *en base et collocatif* tel que le décrit Hausmann mais qui sont considérés comme des collocations de par leur haute fréquence associative (« partir en vacances »). Ainsi, pour l’enseignement de ce phénomène fréquent, notamment auprès de futurs enseignants de FLE non linguistes, il existe deux grandes entrées définitoires pour les collocations, de deux traditions linguistiques un peu différentes telles que l’explique Gledhill (2011).

D’un côté une vision contextualiste lexicale, où des linguistes extraient des corpus numériques les associations lexicales fréquentes. Halliday, McIntosh et Strevens (1964) et Sinclair (1987, 1991, 2004a) sont de cette tradition. L’idée de Firth (1957a) est que, outre le fait que le sens d’un mot apparait en contexte large (le texte, l’inférence élargie), il se révèle surtout en contexte restreint, c’est-à-dire entre les éléments associés proches. Le sens d’une lexie s’élabore dans son usage avec d’autres lexies, ce que Firth a résumé ainsi : « *You shall know a word by the company it keeps* » (Firth, 1957b, p. 11). Le sens se crée dans le contexte, comme l’ont décrit Rastier ou Hoey après Firth, et s’alimente de l’entour large (le global) puis proche (le local). Rappelons que le lexème de Rastier ne devient sémène, donc porteur de sens, qu’une fois actualisé en contexte. Hanks (2008) propose que le mot contienne des sens potentiels qui attendraient d’être activés en contexte par les mots qui l’entourent. Le local est donc privilégié. Il est alors intéressant d’analyser les constructions (*patterns*) lexico-syntaxiques associées aux mots car elles révèlent un sens potentiel du mot et de ceux qui l’entourent (Yan, 2016). Trouver ces constructions est rendu possible par les corpus numériques, outil d’étude de ces chercheurs qui contribuent depuis des décennies au développement de la linguistique de corpus.

Ce contextualisme prend pour principale source de connaissances l’usage et pour principal objet d’analyse les textes. Pour exemple de l’apparition du sens en contexte restreint, prenons le verbe « poser » associé à

des substantifs (Figure 1). La requête, dans le corpus Scientext, était la suivante : [poser + N]. Les 100 premières occurrences font apparaître deux substantifs : *question* (33 occ.) et *problème* (37 occ.). Ces lexies sont donc attirées les unes vers les autres, il s'agit de collocations. Reste la question du pourquoi cette attirance privilégiée ? Ceci est une question davantage sémantico-discursive et nous avons tenté d'y répondre avec Sorba (Cavalla *et al.*, à paraître).

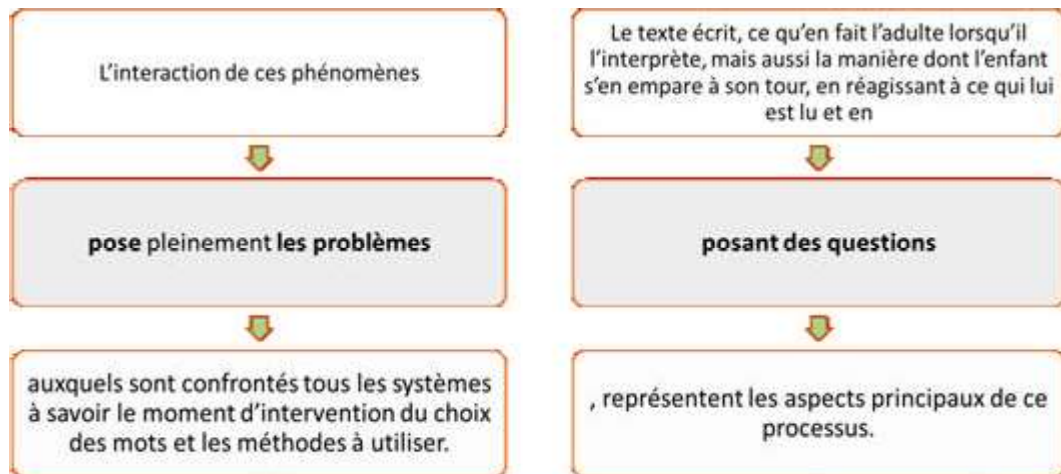


Figure 1 : Extraction de Scientext de la requête du lemme [poser + N]

D'un autre côté les linguistes qui privilégient les éléments syntaxiques et sémantiques de la collocation tout en prenant en compte l'aspect fréquentiel de l'association des éléments en présence. C'est la définition de Tutin et Grossmann (2002), se référant à Hausmann (1989), que nous retenons : une collocation est le résultat d'une cooccurrence privilégiée de plusieurs éléments lexicaux. L'un de ces éléments est la base et conserve son sens premier (ou le plus fréquent), l'autre est le collocatif et acquiert un sens autre que celui généralement présent. Dans « émettre une hypothèse », il s'agit bien d'une *hypothèse*, et le verbe *émettre* perd la partie de son sens qui a trait au son.

Ces deux points de vue, complémentaires pour le FLE, sont résumés dans la figure 3 ci-dessous, avec deux exemples : « prendre un risque » dans le corpus journalistique de Emolex et « résultats similaires » dans le corpus de Scientext en Biologie.

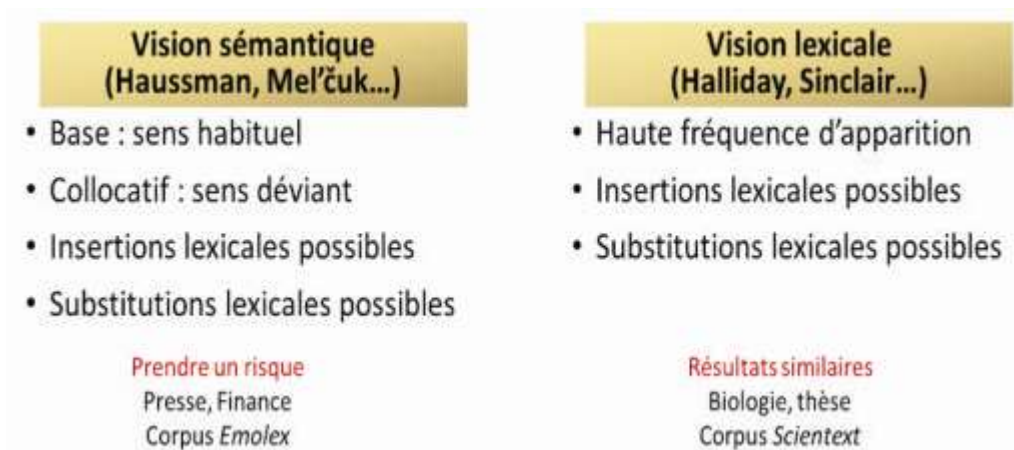


Figure 3 : Résumé de la définition de la collocation selon deux entrées linguistiques

Les aspects sémantiques permettent de comprendre l'utilisation de certains verbes dont le sens déviant peut paraître incongru pour des apprenants. Par exemple le sens de *prendre* dans « prendre un bain » fait rire plus d'un étudiant allophone. Les aspects lexicaux liés à la fréquence permettent aux apprenants de se constituer des listes toutes prêtes d'associations lexicales et de ne plus buter sur le choix de l'adjectif à utiliser. Dans « résultats intéressants » par exemple, les apprenants retiennent que *résultats* s'associe souvent avec *intéressant* dans les conclusions d'articles scientifiques en économie par exemple (travail mené avec Tutin dans Cavalla *et al.*, 2008). L'enseignant peut leur donner deux ou trois adjectifs à retenir dans un premier temps : cela rassure et permet d'avancer dans l'écriture.

Ainsi, au plan didactique, ces deux entrées sont indispensables. Que le sens se réalise en contexte large ou restreint importe peu aux apprenants dès l'instant où ils ont accès simultanément au sens et à la lexie appropriée. Toutefois, au plan des choix didactiques cela me paraît important : si le sens se crée en contexte (large et restreint), alors il est indispensable de présenter ces lexies en contexte. Le genre reste donc une entrée privilégiée me semble-t-il. Reste alors à extraire et rassembler les collocations – puis les autres éléments semi-figés – afin de les avoir à disposition dans l'enseignement. En l'absence de listes il est difficile d'enseigner systématiquement ces éléments. On pourrait les classer sémantiquement et les présenter dans les genres qui les incluent de façon privilégiée. Dès l'instant où de tels éléments apparaîtront dans les méthodes selon une logique onomasiologique, alors l'enseignant pourra plus facilement les traiter et l'apprenant pourra y avoir accès.

## Conclusion

Pour conclure, nous retiendrons que les collocations sont des éléments très fréquents de la langue et qu'il est important de les enseigner à des publics allophones. Dans le cadre de la phraséologie, les collocations peuvent être envisagées sous deux entrées indispensables pour l'enseignement en langue étrangère : une entrée qui retient leur fréquence d'utilisation par les natifs grâce notamment aux extractions possibles dans les corpus numériques et une entrée fondée sur leur combinatoire syntaxique et les choix lexico-sémantiques opérés dans ces conditions. Des recherches récentes tentent alors à séparer les deux types d'éléments en deux classes distinctes : d'un côté les collocations dans la vision sémantique selon la définition de Hausmann et de l'autre, les « routines langagières » de la vision de Firth et Sinclair selon la fréquence d'apparition des éléments dans les corpus numériques.

Pour l'enseignement en FLE nous prenons en compte les deux points de vue et décrivons depuis quelques années les phraséologismes de ce type dans les écrits scientifiques de natifs afin d'aider ensuite les étudiants étrangers à rédiger leurs écrits académiques en français.

La phraséologie est un élément reconnu dans l'enseignement du lexique des langues étrangères mais n'apparaît pas encore systématiquement dans les manuels d'enseignement. La multiplication des descriptions linguistiques mettant en lumière la présence abondante des collocations, notamment à l'écrit, a permis aux didacticiens de mieux comprendre leur pertinence pour l'apprenant dans l'apprentissage d'une langue. Nous pouvons retenir que l'acquisition d'un mot se fait à l'aide des éléments qui l'entourent mais surtout que retenir un élément lexical, c'est mémoriser deux types d'informations : le sens et la forme. Ceci peut paraître banal. Toutefois, l'absence de l'un (la forme des collocations) ou de l'autre (le sens des figements que finalement l'apprenant doit deviner) voire des deux dans les manuels d'enseignement du FLE est un défaut auquel il faudra remédier dans un avenir proche.

---

## Références bibliographiques

Bak Sienkiewicz, M. (2016). *Les constructions V causatif + N d'émotion. Aspects linguistiques et pistes didactiques*. (Doctorat). Grenoble : Université Grenoble Alpes.

- Bard, A. (2007). *Constitution d'un corpus d'écrits universitaires d'étudiants étrangers*. (Master). Grenoble : Université Grenoble Alpes.
- Calaque, E. (2006). Collocations et image de l'organisation lexicale. *EduFLE – Didactique du lexique, de la sémantique FLE*. Récupéré de <[www.edufle.net/Collocations-et-image-de-l](http://www.edufle.net/Collocations-et-image-de-l)>.
- Cavalla, C. (2008). Les collocations dans les écrits universitaires : un français spécifique pour les apprenants étrangers. Dans O. Bertrand et I. Schaffner (éd.), *Apprendre une langue de spécialité : enjeux culturels et linguistiques* (pp. 93-104). Paris : Editions Ecole Polytechnique.
- Cavalla, C. et Tutin, A. (2008, 13-16 août). Etude des collocations évaluatives dans les écrits scientifiques. *Europhras 2008*, Helsinki, Finlande.
- Cavalla, C. et Sorba, J. (à paraître). Etude diachronique du figement : collocations verbo-nominales. Dans S. Mejri et I. Sfar (éd.), *Europhras*. Paris : Champion.
- Edmonds, A. (2013a). 'Parfaitement sensé' – La préférence sémantique et l'enseignement des collocations en L2. Dans C. Garcia-Debanc, C. Masseron et C. Ronveaux (éd.), *Enseigner le lexique* (pp. 131-151). Namur : Presses Universitaires de Namur.
- Firth, J. (1957a). Modes of Meaning. *Papers in Linguistics 1934-1951* (pp. 190-215).
- Firth, J. (1957b). A Synopsis of Linguistic Theory 1930–1955. Dans J. R. Firth (Ed.), *Studies in Linguistic Analysis* (Special Volume of the Philosophical Society, pp. 1-32). Oxford : Basil Blackwell.
- Gledhill, C. J. (2011). The 'lexicogrammar' approach to analysing phraseology and collocation in ESP texts. *ASp - anglais et Français de spécialité*, 59, 5-23. Récupéré de <<http://asp.revues.org/2169>>.
- Gonzalez-Rey, M-I. (2008). *La didactique du français idiomatique*. Fernelmont, Belgique : E. M. E.
- Grobelak, L. (1990). *Dictionnaire collocationnel du français général*. Warszawa : Państwowe Wydawnictwo Naukowe.
- Halliday, M., McIntosh, A. y Strevens, P. (1964). *The Linguistic Sciences and Language Teaching*. London : Longman.
- Hanks, P. (2008). Lexical Patterns: from Hornby to Hunston and beyond. Dans E. Bernal et J. De Cesaris (éd.), *Proceedings of Euralex 2008*. Barcelona: Universitat Pompeu Fabra. Récupéré de <[http://www.patrickhanks.com/uploads/5/1/4/9/5149363/hornby\\_to\\_hunston.pdf](http://www.patrickhanks.com/uploads/5/1/4/9/5149363/hornby_to_hunston.pdf)>.
- Hausmann, F. J. (1979). Un dictionnaire des collocations est-il possible ? *Travaux de littérature et de linguistique de l'Université de Strasbourg*, XVII:1, 187-195.

- Hausmann, F. J. (1985). Trois paysages dictionnaires : la Grande-Bretagne, la France et l'Allemagne. Comparaisons et connexions. *Lexicographica*, 1, 24-50.
- Hausmann, F. J. et Blumenthal, P. (2006). Collocations, corpus, dictionnaires. *Langue Française*, 150, 119-126.
- Hoey, M. (2005). *Lexical Priming: A New Theory of Words and Language*. London/New York : Routledge.
- Kamber, A. (2014). Prendre, un verbe support dans l'enseignement du FLE : une analyse sur corpus. *Revue Mosaïques, Au cœur du verbe. Discours, syntaxe et didactique*, 3-16.
- Lewis, M. (dir.). (2000). *Teaching collocation : Further developments in the lexical approach*. Hove : Language teaching publications LTP.
- Mel'cuk, I. (2003). Les collocations: définition, rôle et utilité. *Revue française de linguistique appliquée, thématique*, 23-31.
- Montagnon, G. (2007). Développement d'une séquence didactique pour l'utilisation des collocations dans les écrits universitaires. (Master). Grenoble : Université Grenoble Alpes.
- Rastier, F. (1987). *Sémantique interprétative*. Paris : Presses Universitaires de France.
- Sinclair, J. (1987). *The nature of the evidence. Looking Up: an account of the COBUILD Project in lexical computing*. London/Glasgow : Collins.
- Sinclair, J. (1991). *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.
- Sinclair, J. (2004a). *How to use corpora in language teaching*. Amsterdam : Benjamins.
- Solano Rodriguez, M. A. (2012). Las unidades fraseológicas del francés y del español: tipología y clasificación = Phraseological units in French and Spanish: typology and classification. *Paremia*, 21, 117-128.
- Tutin, A. et Grossmann, F. (2002). Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue Française de Linguistique Appliquée, VII*, 7-25.
- Williams, G. (2003). Les collocations et l'école contextualiste britannique. Dans A. Tutin et F. Grossmann (éd.), *Les collocations: analyse et traitement* (pp. 33-44). Paris : De Werelt.
- Yan, Rui. (2016). *Étude des constructions verbales scientifiques dans une perspective didactique : utilisation des corpus dans l'enseignement du FLE à l'aide des techniques du TAL*. (Thèse). Grenoble : Université Grenoble-Alpes.