



HAL
open science

Exemple d'enseignement de la phraséologie transdisciplinaire à l'aide de corpus numériques en FLE

Cristelle Cavalla

► **To cite this version:**

Cristelle Cavalla. Exemple d'enseignement de la phraséologie transdisciplinaire à l'aide de corpus numériques en FLE. La lettre de l'AIRDF, 2018, La didactique du lexique, 64, pp.43-47. hal-02087791

HAL Id: hal-02087791

<https://hal.science/hal-02087791>

Submitted on 23 Apr 2019

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.

EXEMPLE D'ENSEIGNEMENT DE LA PHRASÉOLOGIE TRANSDISCIPLINAIRE À L'AIDE DE CORPUS NUMÉRIQUES EN FLE

Cristelle Cavalla, Université Sorbonne Nouvelle-Paris 3

Depuis quelques années désormais, nous¹⁶ développons des formations pour l'aide à l'écriture académique à l'université auprès d'étudiants dont le français est une L2, L3 ou L4. Ces formations répondent à un besoin d'aide à l'écriture académique que ces étudiants ont parfois du mal à combler seuls avec les outils à leur disposition : dictionnaires et traducteurs en ligne. Dans un premier temps, c'est une rapide étude des écrits de certains de ces étudiants qui nous a permis de constater que leurs lacunes étaient souvent d'ordre phraséologique. En effet, ceux-ci connaissent le lexique scientifique approprié à leur discipline, mais connaissent mal, voire méconnaissent, le lexique transdisciplinaire qui contribue à la cohérence d'un écrit scientifique. En outre, parmi les éléments de ce lexique transdisciplinaire, la phraséologie transdisciplinaire est peu étudiée et peu enseignée malgré sa fréquence élevée. Nous entendons par phraséologie, l'ensemble des unités de type figé dans lequel une variété d'éléments coexiste, telles les collocations et les locutions (C. Gledhill & Frath, 2007 ; Polguère, 2011 ; Tutin & Grossmann, 2013). Les éléments sur lesquels nous travaillons ici sont des collocations (Hausmann, 1979) et des routines langagières (Firth, 1957 ; Sinclair, 1991). Dans cette recherche, la phraséologie transdisciplinaire s'attache à voir les types de figements rencontrés dans le discours scientifique et plus précisément ici les collocations et les routines langagières du discours scientifique. Nous reviendrons plus précisément sur les définitions de ces éléments ci-dessous. L'un des premiers exemples erronés rencontrés dans les écrits des étudiants est « L'hypothèse suivante est déclarée... » pour « L'hypothèse suivante est émise/avancée... » (type collocation). Nous avons donc à faire à des éléments discursifs qui contribuent à la cohérence de l'écrit à produire et à son insertion dans un genre spécifique qu'est l'écrit académique.

Afin de répondre aux attentes des étudiants et de mieux cibler des interventions de remédiation, plusieurs projets de recherche ont émergé avec le double objectif de constituer un corpus d'écrits scientifiques pour l'étude des spécificités linguistiques de ce genre et l'élaboration de séquences didactiques. Le premier projet fut l'ANR Scientext (Tutin, 2013) et le deuxième un PPF (Plan Pluri-Formations) nommé FULS¹⁷ pour l'élaboration de séquences. L'équipe de Scientext a constitué un corpus d'écrits académiques en français et en anglais, dans plusieurs disciplines scientifiques. Nous avons alors développé des séquences didactiques à partir des extractions du corpus pour l'aide à l'écriture scientifique auprès d'étudiants en formation de Master ou de Doctorat. Cette présentation va permettre de montrer comment nous avons fait entrer deux éléments de formation dans ces classes d'écriture scientifique : la phraséologie transdisciplinaire et les corpus numériques.

Les éléments phraséologiques représentent un pan non négligeable du lexique. Des chercheurs estiment qu'ils représentent 70 % du lexique anglais (Lewis, 2000) et certainement presque autant dans les langues, de façon générale. Les collocations (*prendre un bain* [Cavalla & Sorba, 2017]) et les routines langagières (*principaux résultats* [Tutin, 2010]), appartiennent à ce grand ensemble et sont décrites selon deux perspectives : la perspective lexicale (Firth, 1957 ; Sinclair, 1987) et la perspective sémantique (Hausmann, 1979 ; Mel'čuk, 2003).

¹⁶ Cette brève présentation rassemble quelques éléments du travail de l'équipe de Grenoble au LIDILEM et de Paris au DILTEC.

¹⁷ Formes et Usages des Lexiques Spécialisés : <http://webtek-66.iut2.upmf-grenoble.fr/>

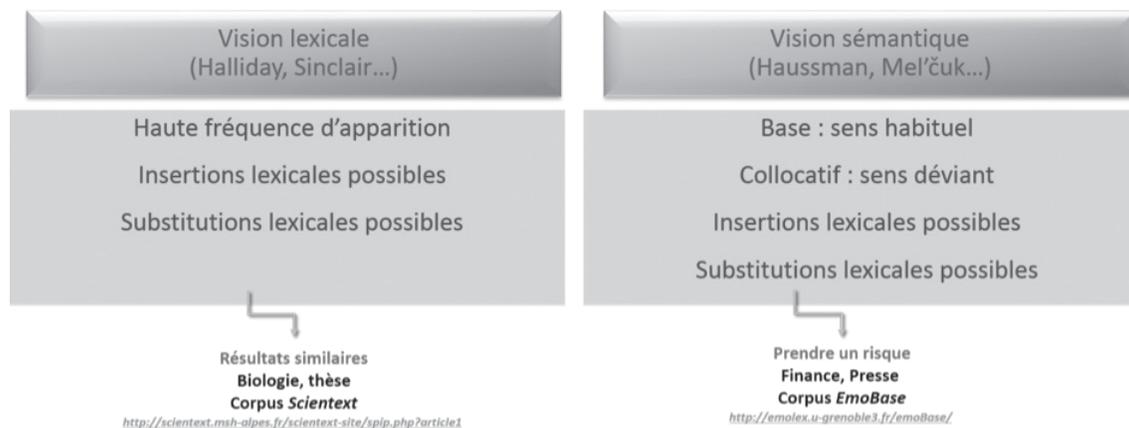


Figure 1. Deux types de figements : collocation et routine langagière

La vision lexicale se fonde d'abord sur les statistiques possibles grâce à la linguistique de corpus. Deux éléments qui apparaissent souvent ensemble dans un corpus, seront généralement reconnus des locuteurs qui les utilisent habituellement de façon associée (C. J. Gledhill, 2011). Les éléments combinés conservent leur sens, on trouve souvent une compositionnalité sémantique, des insertions lexicales sont possibles entre les deux éléments (« des résultats *très* similaires »), et des substitutions quasi-synonymiques sont envisageables : « des résultats *identiques* ».

La vision sémantique permet de distinguer les collocations en tant que sous-type d'élément phraséologique. Les collocations sont des éléments fréquents et présentent la caractéristique de posséder un élément dont le sens est modifié. Ainsi, la collocation ne permet pas la compositionnalité sémantique des deux éléments associés. Il est question de « base », c'est-à-dire l'élément porteur du sens de l'unité phraséologique, l'élément sur lequel le locuteur s'appuie pour choisir le collocatif qui voit son sens modifié, de façon plus ou moins intense. L'association des deux éléments (base + collocatif) crée alors une combinaison lexico-sémantique non prédictible pour un locuteur non averti : « peur bleue » est un exemple prototypique de collocation. La lexie « peur » conserve son sens habituel, il s'agit de la base, l'élément dont il est question ; tandis que la lexie « bleue » est le collocatif, l'élément modifié sémantiquement. Une telle combinaison lexicale n'est pas prévisible par un apprenant (Binon & Verlinde, 2003) du français L2 puisque rien ne permet de prévoir que « bleue » signifie « fort, très, intense » dans ce contexte. L'autre exemple est courant et familier, mais très représentatif de notre point de vue. Il s'agit de « se faire une toile » qui signifie « aller au cinéma voir un film projeté sur un écran qui est en toile ». Il n'est pas possible pour un locuteur de France d'utiliser le synonyme immédiat qui serait « écran » (*se faire un écran). Le seul quasi-synonyme est « ciné » dans « se faire un ciné ». Nous sommes face à une collocation non prédictible par des apprenants, mais très fréquente et donc connue des locuteurs.

Dans ces conditions, pour l'enseignement de la phraséologie transdisciplinaire à des étudiants de Master et de Doctorat, nous avons opté pour une approche globale du phénomène phraséologique. En effet, les apprenants visés ne sont pas des spécialistes linguistes, mais des économistes, des chimistes et des juristes entre autres. De fait, nous ne sommes entrées dans ces précisions linguistiques qu'en cas de requête explicite de la part des étudiants. Ainsi, pour les étudiants, les éléments phraséologiques abordés sont des éléments « qui apparaissent fréquemment ensemble » dans des corpus numériques. Ce critère est valable dans les deux perspectives descriptives des éléments et donc convient pour des non-spécialistes. Notons toutefois que certains étudiants remarquent la différence de non-compositionnalité sémantique et utilisent ce critère pour s'appropriier le sens des collocations notamment¹⁸. Ceci fera l'objet d'une étude ultérieure autour de l'appropriation du sens par des apprenants de L2.

Concrètement, plusieurs séquences didactiques ont été testées au fil du temps et des niveaux des apprenants. Nous avons abouti à un dispositif de formation qui introduit simultanément les corpus

¹⁸ Il s'agit ici du phénomène de représentation concrète de l'image développée dans la collocation et qui devient une aide mnémotechnique. L'idée de « prendre un bain » fait rire les étudiants qui voient les Français « attrapant » leur baignoire.

numériques dans la classe de langue et les éléments phraséologiques (émettre une hypothèse, *avoir des résultats intéressants...*) quel que soit le niveau des apprenants. Pour davantage de détails à propos de ces séquences, je me permets de renvoyer le lecteur à des articles publiés par l'équipe (Cavalla, 2015 ; Cavalla, Loiseau, Lascombe, & Socha, 2014 ; Tran, 2014 ; Tran, Tutin, & Cavalla, 2016 ; Yan, 2017).

La séquence la plus aboutie et la plus diffusée tend vers l'aide à la mémorisation en s'appuyant sur les acquis antérieurs des apprenants, l'évitement de la traduction (sauf si deux éléments phraséologiques sont identiques en L1 et L2), l'appropriation simultanée de l'outil « corpus numérique » et des éléments linguistiques. La séquence débute par la découverte de l'outil avec lequel nous demandons aux étudiants de chercher des éléments phraséologiques. L'approche est celle du FOU¹⁹ (Mangiante & Parpette, 2011) dans le sens où l'analyse des besoins a été effectuée en amont comme précédemment indiqué, la collecte des données est recensée dans les corpus et leur traitement a été effectué au fil des études linguistiques (Tutin & Grossmann, 2013). Pour résumer, la séquence se déroule en trois temps décrits précisément par ailleurs (Cavalla, à paraître) et présentée sous forme de tableau en annexe 1 :

- la découverte de l'outil corpus,
 - possibilités de requêtes
 - limites pour les besoins en écriture académique
- son appropriation d'abord libre, puis avec des contraintes linguistiques (recherche d'éléments phraséologiques),
 - une entrée lexicale (chercher le mot « hypothèse » et les éléments associés)
 - puis une entrée sémantique (chercher comment énoncer une opinion dans un écrit académique)
- enfin, la réutilisation des corpus et des éléments phraséologiques pour la présentation d'un sujet de mémoire avec un diaporama.

Cette séquence a été mise en place auprès d'étudiants de différentes disciplines et de différents niveaux à la fois en langue – avec des L1 variées aussi – et dans leur cursus universitaire (Master et Doctorat). Dans tous les cas rencontrés, un écrit long de type mémoire de master, article de recherche ou chapitre de doctorat était à produire.

Pour conclure, brièvement, sur ce premier aperçu de nos travaux, nous avons constaté que les apprenants n'avaient plus systématiquement recours à la traduction dès l'instant où l'outil les aidait à voir les éléments phraséologiques et qu'ils avaient compris que la traduction littérale n'était que rarement possible. La séquence répond en partie à l'aide à l'apprentissage à la fois grâce aux deux outils que sont les corpus numériques et les éléments phraséologiques. Il faut peu de temps pour qu'ils comprennent que l'entrée sémantique permet de chercher « comment rédiger une hypothèse ? » ou « comment introduire une citation ? » sans obligatoirement utiliser les lexies « hypothèse » et « citation ». Après plusieurs exemples de consultation et de recherche avec eux, leur capacité à rechercher ce dont ils ont besoin personnellement pour leur rédaction est alors impressionnante. En somme, une fois que l'apprenant sait où chercher l'information utile à sa rédaction personnelle, une fois qu'il a pris conscience de l'existence de la phraséologie transdisciplinaire, alors son autonomie scripturale dans la production de textes scientifiques s'en trouve accrue.

Références bibliographiques

- Binon, J., & Verlinde, S.. (2003). Les collocations : clef de voûte de l'enseignement et de l'apprentissage du vocabulaire d'une langue étrangère ou seconde. *La lettre de l'AIRDF*, 33, 31-36.
- Cavalla, C. (2015). Collocations transdisciplinaires dans les écrits de doctorants FLS/FLE. *Linx*, 72, 95-110.
- Cavalla, C. (à paraître). Exemples d'utilisations de corpus pour l'enseignement des collocations en FLE. In M.-P. Jacques & A. Tutin (Eds.), *D'une discipline à l'autre : lexique transversal et formules discursives des sciences humaines*. London : ISTE editions.

¹⁹ Français sur Objectif Universitaire.

- Cavalla, C., Loiseau, M., Lascombe, V., & Socha, J. (2014). Corpus, base de données, cartes mentales pour l'enseignement. In P. Blumenthal, I. Novakova & D. Siepman (Eds.), *Les émotions dans le discours. Emotions in discourse* (p. 327-341). Francfort : Peter Lang.
- Cavalla, C., & Sorba, J. (2017). Prendre un bain, des risques ou la fuite : étude diachronique du figement. In F. Grossmann & S. Mejri (Eds.), *La phraséologie : Sémantique, Syntaxe, Discours* (p. 187-201). Paris : Honoré Champion.
- Firth, J. R. (1957). A Synopsis of Linguistic Theory 1930–1955. In J. R. Firth (Ed.), *Studies in Linguistic Analysis* (Vol. Special Volume of the Philosophical Society, p. 1-32). Oxford : Basil Blackwell.
- Gledhill, C. J., & Frath, P. (2007). Collocation, phrasème, dénomination : vers une théorie de la créativité phraséologique. *La linguistique*, 43(1), 63-88. doi : 10.3917/ling.431.0063
- Gledhill, C. J. (2011). The 'lexicogrammar' approach to analysing phraseology and collocation in ESP texts. *ASp - anglais et Français de spécialité*, 59, 5-23. Repéré à : <http://asp.revues.org/2169>
- Hausmann, F. J. (1979). Un dictionnaire des collocations est-il possible ? *Travaux de littérature et de linguistique de l'Université de Strasbourg*, 17(1), 187-195.
- Lewis, M. (2000). *Teaching collocation : Further developments in the lexical approach*. Hove: Language teaching publications LTP.
- Mangiante, J.M., & Parpette, C. (2011). *Le français sur objectif universitaire*. Grenoble : Presses universitaires de Grenoble.
- Mel'čuk, I. (2003). Les collocations : définition, rôle et utilité. *Revue française de linguistique appliquée, thématique*, 23-31.
- Polguère, A. (2011). Figement et ellipse dans une perspective lexicographique : le cad de « dé à jouer » et « dé à coudre ». In J.-C. Anscombe & S. Mejri (Eds.), *Le figement linguistique : la parole entravée* (pp. 363-373). Paris : Champion.
- Sinclair, J. M. (1987). *The nature of the evidence. Looking Up: an account of the COBUILD Project in lexical computing*. London and Glasgow: Collins.
- Sinclair, J. M. (1991). *Corpus, Concordance, Collocation*. Oxford : Oxford University Press.
- Tran, T. T. H. (2014). Les séquences lexicalisées à fonction discursive comme outil d'aide à l'écriture auprès des étudiants étrangers. In P. Blumenthal, I. Novakova & D. Siepman (Eds.), *Les émotions dans le discours. Emotions in discourse* (p. 355-364). Francfort : Peter Lang.
- Tran, T. T. H., Tutin, A., & Cavalla, C. (2016). Pour un enseignement systématique des marqueurs discursifs à l'aide de corpus en classe de FLE : l'exemple des marqueurs de reformulation. *Linguistik Online*, 78(4), 113-128. Repéré à : <https://bop.unibe.ch/linguistik-online/article/view/2953>
- Tutin, A. (2010). Evaluative adjectives in academic writing in the humanities and social sciences. In R. Lores-Sanz, P. Mur-Duenas & E. Lafuente-Millan (Eds.), *Constructing Interpersonality: Multiple Perspectives on Written Academic Genres* (p. 219-240). Cambridge : Cambridge Scholars Publishing.
- Tutin, A. (2013). La phraséologie transdisciplinaire des écrits scientifiques : des collocations aux routines sémantico-rhétoriques. In A. Tutin & F. Grossmann (Eds.), *L'écrit scientifique : du lexique au discours. Autour de Scientext* (p. 27-43). Rennes : Presses Universitaires de Rennes.
- Tutin, A., & Grossmann, F. (2013). *L'écrit scientifique : du lexique au discours. Autour de Scientext*. Rennes : Presses Universitaires de Rennes.
- Yan, R. (2017). *Étude des constructions verbales scientifiques dans une perspective didactique : utilisation des corpus dans le diagnostic des besoins langagiers du fle à l'aide des techniques de tal*. (Thèse de doctorat, Université Grenoble Alpes, Grenoble).

Annexe 1 : Synopsis de la séquence

Objectifs généraux	Objectifs spécifiques	Déroulement	Temps – Matériel
Apprendre à rédiger des écrits académiques à l'aide 1. de corpus numériques 2. de la phraséologie transdisciplinaire	Découverte de l'outil corpus. Comprendre les apports des corpus numériques par rapport à l'offre du net en général.	Échange et discussion sur leurs représentations de l'écrit académique. Présentation d'outils numériques à partir d'énigmes à résoudre pour en comprendre les limites. Jeu d'intuition linguistique.	4h : dépend des prérequis des apprenants. Ngram Viewer ²⁰ Linguee ²¹ et/ou Tradooit ²² Logiciel d'interrogation du net.
	Appropriation des corpus numériques. Prise de conscience du rôle de la phraséologie transdisciplinaire dans les discours académiques.	Manipulation de corpus à l'aide de requêtes guidées. 1. Entrée lexicale : chercher le mot « hypothèse » et les éléments associés. 2. Entrée sémantique : chercher comment énoncer une opinion dans un écrit académique. Élargissement à d'autres corpus. Réflexion métalinguistique sur les extractions de concordances.	6h : dépend des profils disciplinaires des apprenants. Le Corpus Français de Leipzig ²³ Lextutor ²⁴ Lexicoscope ²⁵ ScienQuest
	Préparation, à l'aide des corpus et des éléments phraséologiques, d'une présentation de leur sujet de mémoire.	Sélection et manipulation d'éléments de la phraséologie transdisciplinaire : 1. à partir de questions générales (rédiger une introduction, des hypothèses...) 2. à partir de leurs interrogations personnelles	10h : dépend des attentes pour le produit final. Corpus au choix Diaporama

Pour résumer, la séquence se déroule en trois temps consécutifs complémentaires. D'abord la découverte de la phraséologie transdisciplinaire et des corpus numériques puis la manipulation guidée à la fois des outils numériques et des éléments linguistiques. Enfin, la 3^e et dernière partie du cours est consacrée à leur rédaction personnelle. Les trois phases contiennent des moments de travail individuel et collectif, l'échange et la discussion entre pairs sont au cœur du dispositif.

²⁰ URL : <https://books.google.com/ngrams>

²¹ URL : <https://www.linguee.fr/>

²² URL : <https://www.tradooit.com/>

²³ URL : http://corpora.uni-leipzig.de/fr?corpusId=fra_mixed_2012

²⁴ URL : <https://lextutor.ca/conc/fr/>

²⁵ URL : <http://dip01.u-grenoble3.fr/~kraifo/lexicoscope/lexicoscope.php>