

VITALIJA KAZLAUSKIENĖ
Université de Vilnius, Lituanie
vitalija.kazlauskiene@flf.vu.lt
ORCID : 0000-0003-0505-7770

Combinatoire collocationnelle et linguistique de corpus dans l'apprentissage lexical du FLE

Collocational Combinatorics and Corpus Linguistics in Lexical Learning in French as a Foreign Language

Abstract

The importance of stereotypical uses of language, especially in the area of collocational combinatorics, is decisive in language teaching and learning. This type of lexical relationship is difficult for non-native learners to acquire because of its complexity, not only in terms of lexical use, but also in terms of particular linguistic awareness. Learners' collective corpora can be revealing in describing their transitional competence. The diagnostic of interlanguage-specific difficulties makes it possible to evaluate the progression of a target language, to describe it, to identify its hegemonic variety and to create the most effective activities. In this article, we will discuss the issue of interlanguage in the learner corpora and language corpora for use in lexical learning in French as a foreign language lessons. We address questions that a learner of L2 French has, which are sometimes difficult to find answers to in scholarly grammars or L2 French workbooks.

Keywords: collocation; lexical competence; L2 French; language corpora; learner' corpora; interlanguage

Mots-clés : collocation ; compétence lexicale ; français L2 ; corpus de langues ; corpus d'apprenants; interlangue

Introduction

L'acquisition de la compétence lexicale est complexe. Souvent, un mot dont la signification semble transparente dans un contexte prend une signification inattendue dans un autre. Et ce n'est que dans la combinaison des mots, dans leur interaction soit à l'intérieur d'un groupe de mots, soit dans un contexte

langagier plus large, parfois même purement socioculturel, qu'il est possible de déchiffrer correctement son sens.

Les usages stéréotypés revêtent une importance déterminante dans l'enseignement et l'apprentissage d'une langue étrangère. Ce type de relation lexicale est apparemment difficile à maîtriser pour les apprenants non natifs (Grossmann 2012 ; Kazlauskienė 2018, 2021) en raison de sa complexité non seulement lexicale, d'usage, mais aussi à cause de la conscience linguistique particulière qu'elle mobilise (Kazlauskienė 2018, 2021).

L'analyse des corpus collectifs d'apprenants peut être révélatrice de la compétence transitoire de ces derniers. Le diagnostic des difficultés propres à l'interlangue permet d'évaluer la progression d'une langue cible, de la décrire, de relever sa variété hégémonique et de créer des activités les plus efficaces possibles.

Dans cet article, nous allons évoquer la question de l'interlangue dans un corpus de productions écrites (essais et lettres) des années 2011–2021¹ d'apprenants lituaniens de FLE de niveau B1. Nous allons aussi décrire la possibilité de recours aux corpus de français pendant les cours de langue en tant qu'outil d'apprentissage lexical, notamment des particularités collocationnelles. Dans notre étude, nous examinons les difficultés lexicales auxquelles les apprenants de français L2 peuvent être confrontés, qui se traduisent par les erreurs relevées dans le corpus analysé ainsi que par des questions et des hésitations sur les usages stéréotypés de la langue. Ces interrogations peuvent s'avérer difficiles à résoudre à l'aide des grammaires, des manuels de FLE ou des dictionnaires disponibles.

Importance des usages stéréotypés de la langue

La compétence lexicale² dépasse la simple connaissance du lexique, elle englobe une compréhension plus complexe du sens. En effet, le sens d'un mot ne peut être pleinement compris qu'en tenant compte de son contexte linguistique, culturel, et parfois même social ou politique, autrement dit, du langage, de l'environnement dans lequel il est utilisé (Kazlauskienė et Dryjańska 2022). L'apprentissage des usages stéréotypés de la langue est important, car chaque unité de ces usages, avec leur propre structure, leur forme et leur contenu, présentent des caractéristiques particulières de la langue considérée. La plupart des études didactiques sur le lexique soulignent aussi cet aspect (Granger et Paquot 2008 ; Martinez 2018 ; Kazlauskienė 2021). Les apprenants non natifs peuvent avoir des difficultés à maîtriser les relations lexicales de ce type, car elles nécessitent souvent plus qu'une simple manipulation de la forme : il est indispensable de connaître le cadre conceptuel d'un lexème pour l'intégrer correctement à un contexte donné. La compréhension et la production linguistiques sont influencées non seulement par les connaissances et les compétences des locuteurs, mais aussi par leur capacité à interpréter le message codé dans une situation donnée.

1 La longueur des textes narratifs varie entre 22 et 449 mots, la longueur moyenne est de 222,6 mots ; cependant, pour une longueur des lettres de 10 à 236 mots, la taille moyenne est de 113,4 mots. Le corpus entier contient 163 609 mots.

2 Le CECRL (2001 : 87) décrit la compétence lexicale comme la connaissance et la capacité à utiliser le vocabulaire d'une langue qui se compose d'éléments lexicaux et d'éléments grammaticaux. Les éléments lexicaux sont des expressions toutes faites, des locutions figées et des mots isolés. Les éléments grammaticaux appartiennent à des classes fermées de mots.

Notion de collocation

Dans cet article, nous nous pencherons sur la question des usages stéréotypés en mettant l'accent sur la notion de collocation. Les collocations constituent une unité sémantique et elles sont soumises à des contraintes syntaxiques, sémantiques et pragmatiques. Ces unités linguistiques se caractérisent par un certain degré de figement et de récurrence. Une cooccurrence lexicale binaire (Tutin 2013 : 49) semi-figée ou « semi-idiomatique » (Tremblay 2014) se compose d'une base et d'un collocatif : la base conserve son propre sens, tandis que le collocatif complète ce sens d'une manière particulière, contrainte. La première unité lexicale (la base, le noyau) est choisie librement. La seconde unité lexicale (le collocatif) est choisie en fonction de la base (Heid et Raab 1989 ; Mel'čuk 1998 ; Mel'čuk et Polguère 2007) : *la fleur fanée, le vent souffle, un maigre salaire, le bon sens*, etc. La spécialisation du sens en contexte et la restriction combinatoire sont des propriétés qui donnent à ces séquences un caractère préfabriqué et prêt à l'emploi (Tutin 2013 : 47). Il convient de souligner que deux lexèmes peuvent se succéder ou être séparés par d'autres mots (Kazlauskienė 2021 : 144). Plusieurs typologies de collocations sont à distinguer. Ainsi Cowie (1981 : 230) distingue des collocations transparentes et des collocations restreintes. Les premières se combinent librement au niveau sémantique, tandis que les secondes se fondent plutôt sur les capacités du noyau. Mel'čuk (1993, 1995 ; Mel'čuk et Polguère 2008) parle aussi de catégorisation sémantique. Il distingue ainsi, fondamentalement, quatre types de phrasèmes : les pragmatèmes, expressions idiomatiques, collocations et quasi-phrasèmes. Dans le contexte de l'apprentissage de la langue, la catégorisation des collocations en collocation grammaticale et collocation lexicale est la plus pertinente (Xiao et Girani 2018 : 233), vu la possibilité de développement lexical multidimensionnel. Conformément à cette catégorisation, on pourrait travailler la compétence lexicale, et plus particulièrement, les collocations, au niveau non seulement lexical mais aussi grammatical, en relevant les particularités de la syntaxe et de la morphologie associées à chaque collocation. Selon Tutin et Grossmann (2002 : 7), les collocations sont souvent caractérisées par leur sens transparent lors de la réception (le sens se « devine »), tandis que les locuteurs non natifs ont du mal à produire les lexèmes appropriés, principalement en raison d'une maîtrise insuffisante de la morphosyntaxe.

Présence des collocations dans les méthodes de FLE

La question se pose donc de savoir quels sont les aides et les moyens possibles d'acquisition des collocations par les apprenants du français. Dans une certaine mesure, les dictionnaires proposent des solutions, mais ils ont tendance à s'orienter plutôt vers des mots isolés. Les collocations et les cooccurrences lexicales sont généralement présentées dans la section réservée aux exemples, mais toutes les collocations pouvant être contrôlées par un mot n'y figurent pas forcément (Tremblay 2014 : 75). Avec le développement de la linguistique de corpus, les dictionnaires collocationnels se répandent de plus en plus. Le *Dictionnaire des combinaisons de mots* (Le Fur 2007) en est un exemple.

Quoi qu'il en soit, Martinez (2018 : 36) affirme que le développement de la compétence lexicale en matière d'expressions semi-figées est considéré comme une acquisition langagière inachevée dans les programmes d'enseignement et d'apprentissage du FLE à tous les niveaux. Toutefois, si l'on compare les méthodes utilisées avant l'avènement de l'approche communicative et de la perspective actionnelle

à celles d'aujourd'hui, il est évident que ce type de relations lexicales apparaît progressivement dès le niveau faux débutant. Des collocations telles que *faire du shopping*, *être passionné par la lecture*, *j'ai une grande nouvelle*, *elle va bien*, *vous avez l'heure, svp ?*, *les horaires d'ouverture*, par exemple³, sont facilement identifiables à ce niveau. Cependant, les listes de mots et d'expressions figées ainsi que les tableaux de verbes conjugués sont encore couramment utilisés dans les méthodes de FLE. Cela soulève donc deux questions : *comment et à quel niveau enseigner les collocations en FLE ? et quand faut-il commencer à les introduire ?*

D'après le CECRL (2021 : 138), la maîtrise du vocabulaire se manifeste par la capacité des utilisateurs/apprenants à sélectionner l'expression adéquate dans leur répertoire. À mesure que la compétence progresse, cette capacité est de plus en plus caractérisée par l'utilisation de collocations et d'éléments lexicaux préfabriqués. Cependant, il est erroné de penser qu'il faut commencer par enseigner des mots isolés pour ensuite passer aux combinaisons de mots. Au contraire, certaines collocations et expressions très courantes méritent d'être introduites et mises en avant dès le début, tout comme certaines expressions ou formules de communication. L'analyse de la fréquence de ces unités polylexicales dans les corpus montre d'ailleurs qu'elles ont souvent une fréquence supérieure à celle de mots tout à fait courants. Des unités polylexicales telles que *faire l'objet de*, *avoir besoin de*, *jouer un rôle*, *mettre en scène*, qui figurent parmi les 2 000 mots les plus fréquents du français, en sont un exemple (Verlinde et Selva 2001).

Par conséquent, il est recommandé d'introduire les collocations dès le début de l'apprentissage, en travaillant le plus possible le bagage lexical, la diversité des formules, des phrases de routine, des expressions figées. Le CECRL donne une explication de la compétence lexicale en soulignant que « l'adjectif «lexical» n'est pas utilisé ici uniquement pour des mots isolés mais plutôt pour des expressions linguistiques attestées et utilisées comme des ensembles fixes et indivisibles » (CECRL 2001 : 33). Ainsi, l'enseignement du lexique peut être beaucoup plus efficace s'il est structuré et systématique, c'est-à-dire s'il tient compte de la combinatoire lexicale et morphosyntaxique des mots dès le début.

Difficultés d'usage des collocations lexicales en FLE

Comme déjà mentionné, les collocations lexicales sont des unités complexes qui peuvent poser des difficultés aux apprenants de FLE, comme cela ressort du corpus analysé, en particulier lorsque la langue maternelle de l'apprenant ne dispose pas d'un équivalent approprié. Le rôle de la langue maternelle et son influence dans l'acquisition d'une langue étrangère a donné lieu à de nombreuses hypothèses (Klein 1989 : 42). L'interférence, autrement dit le transfert négatif de la langue maternelle, pourrait susciter des erreurs interlinguistiques. Zgusta (1971 : 296) utilise le terme « anisomorphisme » pour désigner ce phénomène d'absence d'équivalence totale entre deux langues. Ainsi, le corpus des apprenants lituaniens de FLE se caractérise par la présence de collocations calquées sur le lituanien, telles que : *demander une question (au lieu de *poser une question*), *recevoir (au lieu de *obtenir*) un diplôme, *visa fini (au lieu de *expiré*), *produit vieilli (au lieu de *périmé*), etc. Ces exemples illustrent non seulement l'interférence du lituanien, mais ils montrent également combien il est nécessaire de prêter attention aux particularités syntaxiques et sémantiques des collocations françaises.

3 Source : les méthodes FLE *Latitudes* (2008) ; *Alter Ego Plus* (2012) ; *Casquette* (2016) ; *Edito* (2018).

L'expérience acquise dans d'autres langues, ainsi que l'expérience vécue à différentes étapes de l'apprentissage de la langue cible influencent inévitablement, de manière plus ou moins prononcée, la production en langue cible. Les apprenants s'approprient l'information et mobilisent les connaissances acquises pour construire et produire des textes en français, selon une logique qui leur appartient. Toute cette expérience pourrait être profitable et produire de bons résultats, liés à un transfert positif. D'un point de vue pédagogique, il est bon de connaître les possibilités de transfert positif pour les favoriser. Cependant, ce qui est aussi inévitable, c'est la présence d'un transfert négatif, avec toute sorte d'erreurs interlinguistiques, et des erreurs intralinguistiques surtout présentes dans les constructions collocationnelles. Les extraits suivants tirés d'un corpus écrit d'apprenants lituaniens illustrent la présence nette de ce type d'erreurs :

- (1) (...) *les technologies moderne ce n'est pas une bien solutions.* (2011E301898)⁴
- (2) *Bonjour, mon bien ami. Ça va ?* (2013L165432)
- (3) *Peut être il commençait beaucoup des médicaments pour un large vie ?* (2011E301931)

Ces phrases extraites du corpus illustrent des exemples d'erreurs intralinguistiques fréquentes. On observe notamment à plusieurs reprises la confusion entre l'adjectif *bon(ne)* et l'adverbe *bien*, ainsi qu'entre les adjectifs *longue* et *large*. En général, l'interlangue des apprenants lituaniens de FLE est principalement caractérisée par des transferts positifs ainsi que des interférences entre leur langue maternelle et la langue cible, et une tendance à la surgénéralisation des règles de la L2. Compte tenu de cette tendance, il est crucial de leur fournir une aide spécifique dans l'apprentissage des collocations, par exemple en proposant des exercices ciblés sur :

- les nuances des synonymes pouvant intervenir dans les collocations, qui ne sont pas évidentes pour les apprenants non natifs, comme dans les exemples suivants tirés du corpus analysé : **la génération dernière*, **un mûr consommateur* ;
- les particularités de la langue française, pour éviter la traduction littérale de leur langue maternelle, comme dans : **futur plan (au lieu de les projets d'avenir)* ;
- la syntaxe, en attirant l'attention sur les erreurs syntaxiques fréquentes, comme dans l'exemple suivant :

- (4) *Les test de français était très difficile et je peur que je ne finiserais ma école avec resultats mal.* (2015L2643D04)

Une autre particularité de l'interlangue des apprenants lituaniens de FLE est que, jusqu'au niveau B1, la catégorie nominale prédomine dans leur production écrite. Les collocations nomino-adjectivales sont fréquentes et la question de la position de l'adjectif demeure importante, étant donné qu'en lituanien, l'adjectif est toujours placé avant le nom (Kazlauskienė 2021). Parmi les collocations verbales, c'est la structure prépositionnelle qui prévaut. Les exemples ci-dessus et une analyse approfondie du corpus de productions écrites d'apprenants lituaniens de FLE au niveau B1 montrent que la combinaison des mots

⁴ Tous les travaux des apprenants ont été introduits manuellement dans des fichiers électroniques sans corrections, c'est-à-dire que la composition originale a été conservée.

n'est pas toujours typique du français standard. Les apprenants mobilisent le lexique qu'ils maîtrisent dans un contexte particulier.

Le diagnostic des productions écrites telles que les essais et les lettres de niveau B1 nous a permis d'identifier des erreurs propres à ce type de texte et de les classer. Nous avons constaté plusieurs occurrences de l'interférence interlinguale : les apprenants recourent notamment à l'analogie, aux emprunts, aux néologismes, à la traduction et aux mots internationaux dans leurs écrits. Par ailleurs, l'emploi de synonymes incorrects ou d'homonymes, le rapprochement sémantique, la généralisation et la simplicité des formes sont autant d'exemples d'interférence intralinguale.

Le processus d'interférence et d'hybridation peut être vu comme une contamination essentielle des collocations, tant au niveau lexical que syntaxique. L'analyse de travaux d'apprenants réalisée par d'autres chercheurs (Xiao et Girani 2018 : 236) met en évidence le même type d'erreurs grammaticales. Les erreurs de collocations se présentent sous la forme d'omissions, ajouts, déplacements, substitutions, pléonasmes, interférences, etc.

Enfin, la dernière grande catégorie d'erreurs est liée à la combinatoire des mots, impliquant des erreurs lexicales directement liées à la notion de collocation. Les extraits du corpus analysé illustrent bien ces erreurs :

- (5) *J'ai beaucoup de plan pour l'avenir.* (2015L2643D21)
- (6) *Personnes de ce type n'avons pas bien connection avec leurs famille, [...].* (2014Epra4)
- (7) *J'ai aussi peur que cette chose soit contemporaine et le jour prochain tout va changer.* (2011E301892)
- (8) *Néanmoins, quelquefois les jeunes sont trop sous une pression des personnes mûres.* (2015E2643D13)

Ces extraits mettent en évidence les tendances générales du corpus, et l'on constate les mêmes erreurs que celles relevées par Xiao et Girani (2018), notamment : le choix d'un mauvais collocatif (collocation inexistante), le choix d'un collocatif approprié dont le sens véhiculé n'est pas le bon, ainsi que le non-respect des contraintes d'emploi du collocatif (Anctil 2010, cité par Xiao et Girani 2018 : 237).

L'analyse des constructions collocationnelles révèle les rapports entre la compétence et la performance des locuteurs. Le diagnostic effectué sur les écrits des apprenants souligne à son tour la nécessité d'introduire les cooccurrences semi-figées dès le début de l'apprentissage et de travailler intensivement le bagage lexical ainsi qu'une variété de collocations, d'expressions et de phrases toutes faites. Il serait nécessaire de systématiser l'enseignement des collocations aux apprenants de FLE. De ce fait, la stratégie de la « compétence collocationnelle » (*collocation competence*) proposée par Brown (1994 : 25) se présente comme un élément essentiel de la compétence lexicale.

Corpus linguistique pour faciliter l'apprentissage des collocations

Dans cette partie, nous décrirons comment l'utilisation de certains corpus linguistiques peut remédier aux erreurs et contribuer à l'évolution de la langue des apprenants. L'enseignement du lexique peut être beaucoup plus efficace s'il est structuré et systématique, c'est-à-dire s'il prend en compte la combinatoire lexicale et syntaxique des mots.

Un corpus linguistique de langue authentique constitue une ressource lexicale intégrant des routines verbales qui permettent de réaliser des activités d'apprentissage qui ont pour objectif d'aider les apprenants dans leurs rédactions. L'approche adaptée est fondée sur la *Corpus Pattern Analysis* (CPA – analyse de modèles basée sur les corpus) (Hanks 2013). La CPA, qui vise à analyser les usages réguliers des lexèmes en contexte, permet d'accéder à leurs sens à travers l'analyse des constructions syntaxiques et des cooccurrences associées à ces lexèmes, c'est-à-dire par l'analyse de leurs emplois usuels. Cette méthode d'analyse combine des techniques de traitement automatique du langage (TAL) à une analyse linguistique manuelle. Dans ce but, il est possible de travailler sur différents types de corpus, à commencer par ceux de Google ou de YouTube, mais aussi sur des corpus de langue authentique (textes littéraires, textes web, textes juridiques, etc.), et même les corpus des apprenants, pour leur autocorrection.

L'analyse des données du corpus permet de travailler la signification d'un lexème, son contexte et son entourage, et de mettre en évidence le processus verbal du langage en temps réel. Toutes sortes d'activités sont possibles pour exploiter cette ressource inépuisable du langage, depuis les exercices traditionnels à trous jusqu'aux tâches adaptées à la classe inversée. Voici quelques propositions d'activités possibles basées sur les données des corpus :

1. L'extraction et la modélisation de routines, séquences linguistiques verbales récurrentes appliquées à une liste de lexèmes établis, conformément aux étapes suivantes :
 - extraction de séquences (par exemple à l'aide de l'interface visuelle proposée par un corpus) ;
 - repérage du sens et structure syntaxique ;
 - classification des séquences et traitement des routines discursives.

Une accessibilité et une adaptation langagière s'imposent pour attirer l'attention des apprenants et faciliter la compréhension. Il s'agirait d'aider les élèves à retenir le sens et l'association lexicale simultanément, ainsi que de fixer les structures qu'ils connaissent déjà partiellement.

2. L'exploitation d'extraits du corpus par l'enseignant peut également se révéler pertinente et productive. En effet, l'utilisation de la méthode de la classe inversée s'adapte parfaitement à l'analyse individuelle préalable avant le travail collectif sur les corpus. La démarche déductive, les exercices habituels à trous, la recherche individuelle de mots et d'expressions dans le cadre du corpus permettent aux apprenants de s'approcher de la langue authentique construite en partie de séquences linguistiques stéréotypées parfois introuvables dans les dictionnaires.
3. L'observation du fonctionnement lexical en contexte et de la façon dont il peut varier et être nuancé en fonction de son environnement linguistique est un élément important dans l'apprentissage du FLE. Par exemple, lors de l'introduction du vocabulaire de l'écologie, l'enseignant peut souligner les différences réelles entre des termes tels que *ville verte* et *espace vert* sans nécessairement détailler la terminologie de la collocation.
4. De petits projets linguistiques peuvent également être réalisés à partir d'une liste de collocations syntaxiques ; pour les apprenants plus avancés, on peut aussi proposer une liste de collocations lexicales, ou même de pragmatèmes. L'utilisation de documents authentiques et de contextes réels est essentielle pour aider les apprenants à identifier les limites sémantiques d'un lexème et à intégrer des expressions souvent utilisées par les locuteurs natifs dans leur propre discours.

Illustration par un exemple concret

Pour illustrer ce point, après avoir diagnostiqué les productions écrites des apprenants de FLE lituaniens de niveau B1, nous avons constaté différentes sortes de difficultés, de la structure syntaxique à la combinatoire lexicale, en passant par la forme morphologique. Entre autres, la difficulté de choix du verbe s'est révélée évidente. Ainsi par exemple, les apprenants hésitent entre les verbes *poser* et *demander* dans la collocation contenant le collocatif *une question*. Nous avons donc décidé de travailler cette collocation, à titre d'exemple démonstratif, à l'aide des corpus Sketchengine (accès institutionnel) et Leipzig Corpora Collection (accès gratuit).

Sketchengine

1. Une activité initiale consisterait à effectuer une analyse des verbes *poser* et *demander* dans le corpus, afin de relever les différents types de collocations syntaxiques associées à ces verbes. Notre recherche, qui applique la méthodologie *corpus-driven* (Tognini-Bonelli 2001), a permis d'extraire une catégorie de motif collocationnel VN (verbe + nom : *poser une question*) ainsi que des patterns prépositionnels caractéristiques du lexème *poser*, ce qui permet de travailler non seulement la dimension lexicale, mais aussi la construction syntaxique d'une collocation et ses variétés.
 2. La seconde étape d'analyse est concentrée sur le texte et le contexte : nous demandons aux apprenants de relever les particularités propres à ces verbes dans des textes authentiques qui illustrent des situations linguistiques réelles.
- L'analyse des lignes de concordance peut aussi fournir des exemples de règles syntaxiques que les apprenants pourraient analyser eux-mêmes afin de tirer des conclusions sur l'emploi du déterminant.

Document ID	Text Snippet	KWIC
#113523	unique européen s'élève environ à 1 milliard d'euros.	pose une première question : la Commission peut-elle se satisfaire d'une telle
#128028	à l'Institut européen de technologie?	poser la question suivante: ces parrains privés imposeront-ils aussi leurs conditi
#174786	s'initier un suivi des conclusions de la commission d'enquête sur la CIA et	poser la question formelle suivante: quelles mesures le Conseil et la Commission
#175432	d'anti-américaine.	posent les questions critiques, qui ont réalisé des évaluations globales ces dernie
#182342	a lettre et qui sont impliqués dans le terrorisme.	poser la question de savoir pourquoi.
#183261	ite.	poser deux questions fondamentales au Conseil.
#217000	e.	poser une question à la Commission: la Commission annoncera-t-elle, lors de la
#230449	unauté internationale, y compris l'UE, est impuissante.	poser une courte question supplémentaire au Conseil, étant donné que la présid
#245060	khé.	poser la question de savoir si les dialogues - au cours desquels nous avons expri
#245923	age et les sanctions spécifiques instaurées par l'UE.	posons la question de savoir pourquoi nos mesures sont inefficaces et pourquoi le
#259345	lèglement Rome I.	posant la question de savoir quel droit sera applicable.
#276645	défi pour tous les pays.	poser la question de la qualité de la surveillance de nos frontières extérieures tar
#315407	écritité d'approvisionnement.	poser les questions que plusieurs intervenants ont déjà soulevées clairement au
#359475	u élevé de plomb dans les peintures utilisées pour ces jouets.	pose la question très sérieuse du type de futur que nous souhaitons pour nos ei
#387800	née européenne contre la peine de mort, il conviendrait de demander, et je	pose la question à la présidence portugaise, dans quelles circonstances un pay
#405283	snir?	poser la question.
#420661	ponses apportées par le Commissaire.	posons nos questions et, à ce jour, nous ne sommes pas encore parvenus à obter
#472909	avons besoin des meilleurs cerveaux du monde.	posent aussi une autre question : " Prenez-vous nos préoccupations au sérieux. "
#533812	je souhaiterais m'exprimer maintenant au nom de M. Navarro, et ensuite	poser ma propre question, ou plutôt réitérer une question.
#533967	endant, je vous donnerai 15 secondes supplémentaires, si vous souhaitez	poser votre question immédiatement car, en vertu du règlement, je ne peux auto

Figure 1. Concordance de *poser + question* sur Sketchengine.

3. L'étape suivante consiste, en suivant la visualisation, à classer les collocatifs dans le tableau selon leur fonction grammaticale (nom, verbe, adjectif, etc.) et à analyser leur fréquence d'apparition. Finalement, des exemples pourront être donnés et comparés à la langue maternelle des apprenants :

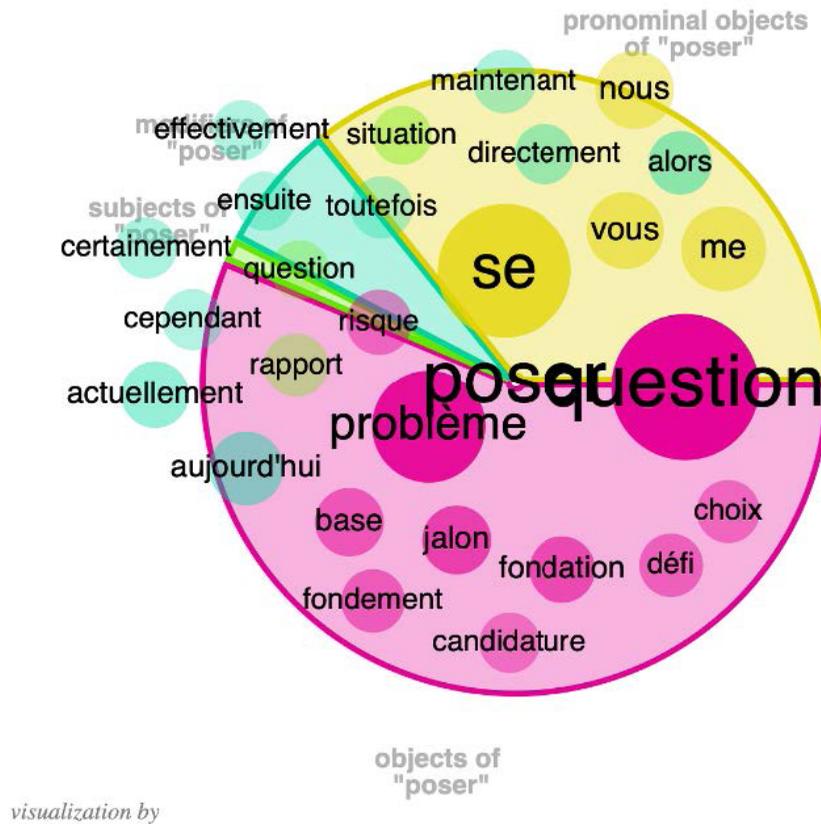


Figure 2. Visualisation graphique du verbe *poser* sur Sketchengine.

4. Pour une analyse plus approfondie, les apprenants sont invités à comparer l'utilisation des deux verbes, *poser* et *demander*, en présence de collocatifs différents, à relever leurs particularités d'utilisation, à identifier des parties communes et finalement, à analyser les contextes de ces cas particuliers.

L'analyse visuelle des graphiques du corpus met en évidence une utilisation commune des lexèmes *poser* et *demander*. Il s'agit du cas de l'expression *poser* ou *demander la main*, ce qui confirme la nécessité d'une analyse détaillée de ce cas particulier.

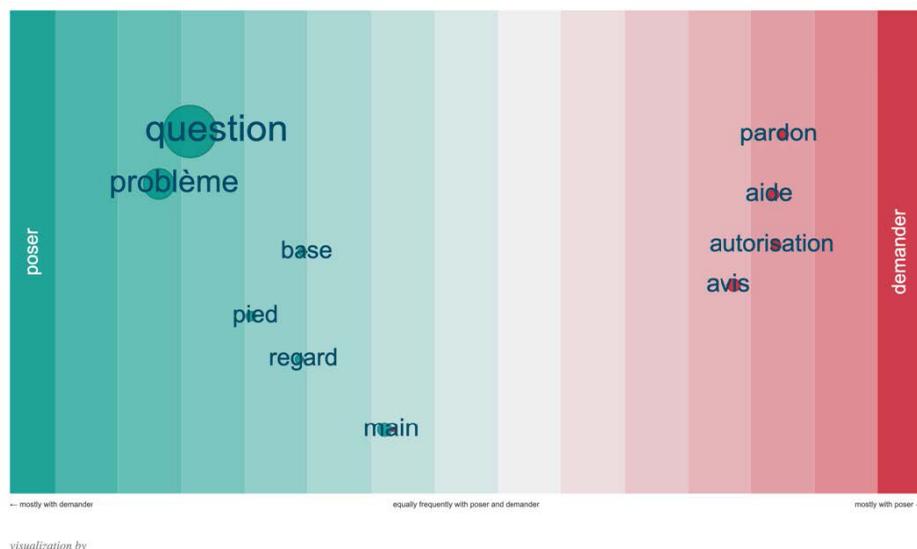


Figure 3. Différence des séquences entre *poser* et *demander* sur Sketchengine.

5. Pour compléter les résultats de leur analyse, les apprenants ont recours aux données du graphique du verbe *poser* présentées dans le corpus de Leipzig Corpora Collection. Ils choisissent des exemples concrets du corpus afin de les comparer et ainsi compléter la liste des particularités d'utilisation de ce verbe.
6. Enfin, des activités pratiques préparées sur la base des corpus sont proposées aux apprenants. En utilisant l'application *learningapps.org*, un exercice est créé pour encourager la réflexion et la prise de décision sur ce qu'il est possible et ce qu'il n'est pas possible de *poser*. De plus, dans le contexte des classes plurilingues, l'activité peut porter sur une comparaison entre ces verbes et les verbes équivalents dans la langue maternelle des apprenants : est-ce que l'on *pose* et *demande* la même chose dans leur langue maternelle ?

Conclusion

L'interlangue du corpus des apprenants et un diagnostic sur cette interlangue pourraient être révélateurs et servir à améliorer l'acquisition du FLE. Ces données pourraient mieux orienter l'enseignant dans son choix du matériel, en fonction des besoins des élèves, mais aussi l'aider à construire une démarche d'enseignement qui tienne compte des particularités linguistiques et langagières de chaque apprenant. Une analyse permettrait entre autres de comparer les travaux des apprenants afin d'évaluer leur progression individuelle aussi bien que celle du groupe. Ce type de données pourrait en outre servir de base à diverses démarches scientifiques relatives à l'apprentissage des langues étrangères, à l'élaboration de matériel didactique et au développement de programmes d'enseignement appropriés. L'observation et l'analyse de l'interlangue et des erreurs rencontrées dans la production écrite sont nécessaires, mais pas

suffisantes. Il faut aller plus loin dans cette démarche et trouver une remédiation tendant vers l'autonomie de l'apprenant.

Les questions fréquemment posées par les apprenants de français L2 et pour lesquelles il est parfois difficile de trouver des explications dans les grammaires, dans les manuels de FLE ou dans les dictionnaires disponibles peuvent recevoir des réponses grâce aux corpus linguistiques de la langue cible, en l'occurrence le français. Ces types de corpus sont des outils précieux pour remédier aux erreurs et favoriser l'évolution de la langue des apprenants.

Le schéma que nous venons de présenter, qui consiste en un diagnostic de l'interlangue, un travail sur un corpus linguistique de la langue cible, une observation de textes authentiques et une mise en pratique, est une approche possible et efficace pour l'apprentissage des collocations, voire du lexique en général, car il évite la fossilisation de l'interlangue et favorise son évolution constante. L'objectif est d'aider les apprenants à retenir à la fois le sens et l'association lexicale, et à consolider les structures qu'ils connaissent déjà partiellement. Il est clair que l'enseignement du lexique peut devenir beaucoup plus performant si l'on tient compte de la combinaison lexicale et syntaxique des mots en les observant dans des textes authentiques.

Bibliographie

- Anctil, Dominic (2010) *L'erreur lexicale au secondaire*. Thèse de doctorat non-publiée. Montréal : Université de Montréal.
- Brown, Phillip R. (1994) « Lexical Collocation : a Strategy for Advanced Learners. » [In :] *Modern English Teacher*. Vol. 3 (2) ; 24–27.
- Conseil de l'Europe (2021) *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer. Volume complémentaire*. Strasbourg : Éditions du Conseil de l'Europe.
- Cowie, Anthony P. (1981) « The treatment of collocations and idioms in learner's dictionaries. » [In :] *Applied Linguistics*. Vol. 2 (3) ; 223–235.
- Granger, Sylviane, Magali Paquot (2008) « Disentangling the phraseological Web. » [In :] Sylviane Granger et Fanny Meunier (éds.) *Phraseology. An interdisciplinary perspective*. Amsterdam/Philadelphia : John Benjamins ; 27–50.
- Grossmann, Francis (2012) « Le rôle de la compétence lexicale dans le processus de lecture et l'interprétation des textes. » [In :] *Forumlecture.ch*. Vol. 1 ; 1–13.
- Hanks, Patrick (2013) *Lexical Analysis : Norms and Exploitations*. Cambridge : MIT Press.
- Heid, Ulrich, Sybille Raab (1989) « Collocations in multilingual generation. » *Proceedings of the Fourth Conference of ACL, European Chapter, Manchester, 10–12 April 1989*, Manchester : Association for Computational Linguistics.
- Kazlauskienė, Vitalija (2018) *Prancūzų K2 daiktavardinis junginys : mokinių tekstyno analizė*. Thèse de doctorat non-publiée. Vilnius : Université de Vilnius.
- Kazlauskienė, Vitalija (2021) « Collocation nomino-adjectivale dans la production écrite en FLE. » [In :] *Taikomoji kalbotyra*. Vol. 15 ; 143–154.
- Kazlauskienė, Vitalija, Agnieszka Dryjańska (2022) « Le sens de fête en polonais, en lituanien, en français et sa (non)coïncidence collocationnelle. » [In :] *Annales Universitatis Paedagogicae Cracoviensis. Studia Linguistica*. Vol. 17 ; 20–42.
- Klein, Wolfgang (1989) *L'acquisition de langue étrangère*. Paris : Armand Colin.

- Le Fur, Dominique (éd.) (2007) *Dictionnaire des combinaisons de mots : les synonymes en contexte*. Paris : Le Robert.
- Martinez, Christine (2018) « Comment faciliter l'apprentissage des collocations aux apprenants non natifs. Une approche didactique et linguistique. » [In :] *Synergies pays riverains de la Baltique*. Vol. 12 ; 35–45.
- Mel'čuk, Igor (1993) « La phraséologie et son rôle dans l'enseignement/apprentissage d'une langue étrangère. » [In :] *Éla. Études de linguistique appliquée*. Vol. 92 ; 82–113.
- Mel'čuk, Igor (1995) « Phrasemes in language and phraseology in linguistics. » [In :] Everaert, Martin, van der Linden, Erik-Jan, Schenk, André, & Schreuder, Robert (éds.) *Idioms : Structural and Psychological Perspectives*. Hillsdale : Lawrence Erlbaum ; 167–232.
- Mel'čuk, Igor (1998) « Collocations and Lexical Functions. Phraseology. » [In :] A.P. Cowie (éds.) *Theory, Analyses and Applications*. Oxford : Clarendon Press ; 23–53.
- Mel'čuk, Igor, Alain Polguère (2007) *Lexique actif du français : l'apprentissage du vocabulaire fondé sur 20 000 dérivations sémantiques et collocations du français*. Bruxelles : De Boeck.
- Mel'čuk, Igor, Alain Polguère (2008) « Prédicats et quasi-prédicats sémantiques dans une perspective lexicographique. » [In :] *Lidil*. Vol. 37, <http://journals.openedition.org/lidil/2691> (consulté le 2/01/2023).
- Tognini-Bonelli, Elene (2001), *Corpus Linguistics at Work*. Amsterdam : John Benjamins.
- Tremblay, Ophélie (2014) « Les collocations : des mots qui font la paire. » [In :] *Québec Linguistique Appliquée*. Vol. VII (1) ; 7–26.
- Trim, John (éd.) (2001) *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer. Guide pour les utilisateurs*. Division des Politiques Linguistiques : Strasbourg.
- Tutin, Agnès (2013) « Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument. » [In :] *Langages*. Vol. I (189) ; 47–63.
- Tutin, Agnès, Francis Grossmann (2002) « Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. » [In :] *Revue française de linguistique appliquée*. Vol. I (7) ; 7–25.
- Verlinde, Serge, Thierry Selva (2001) « Nomenclature de dictionnaire et analyse de corpus. » [In :] *Cahiers de Lexicologie*. Vol. 79 ; 113–139.
- Xiao, Chen, Frédéric Girani (2018) « Typologie d'erreurs des collocations dans les essais argumentatifs et perspectives didactiques de l'enseignement du français langue étrangère dans les universités. » [In :] *Synergies Chine*. Vol. 13 ; 229–246.
- Zgusta, Ladislav (1971) *Manual of Lexicography*. Prague : Academia; The Hague & Paris : Mouton.

Corpus

Leipzig Corpora Collection, https://corpora.uni-leipzig.de/en?corpusId=ind_mixed_2013_
 Sketch Engine, <https://www.sketchengine.eu>.