

## SEQUENCES PREFABRIQUEES DANS LE CORPUS D'APPRENANTS EN FOS/FOU

VITALIJA KAZLAUSKIENĖ

Université de Vilnius, Lituanie

**Résumé.** L'analyse de séquences préfabriquées permet de décrire l'acquisition précoce de séquences lexicales. Ces dernières sont l'un des facteurs les plus importants de l'idiomaticité et de la richesse lexicale de la langue ; ils constituent donc un outil efficace pour développer le FOS. L'objet principal de notre recherche est l'analyse quantitative et qualitative de séquences lexicales dans le corpus écrit d'étudiants lituaniens en philologie française. Ce corpus est constitué de mémoires d'étudiants de troisième année de licence. Ayant pour but de recenser la métalinguistique propre à ce type d'écrits, plus précisément l'utilisation des termes spécifiques au sujet de la linguistique, nous présentons un aperçu des constructions lexicales utilisées dans les séquences préfabriquées (désignées ci-après SP) du corpus à notre disposition ; nous décrivons ensuite les particularités de leur composition. Les données issues de l'analyse du corpus-pilote permettent d'identifier certaines particularités et démontrent la nécessité de ce type d'analyse pour le développement du FOU et du FOS chez les apprenants.

**Mots-clés :** séquence préfabriquée, FOS, FOU, écrit académique, corpus d'apprenants

### INTRODUCTION

Le français sur objectifs spécifiques (FOS) se différencie du français général au niveau lexical. En d'autres termes, les besoins des apprenants en FOS se focalisent plutôt sur la compétence lexicale. Or l'acquisition de cette compétence est complexe, dans la mesure où il ne suffit pas aux apprenants de distinguer la forme et la signification d'un élément lexical pour savoir comment l'intégrer correctement dans le contexte langagier (Kazlauskiene, 2019 ; 2020 ; 2021). Cette compétence suppose en effet de comprendre les relations syntaxiques, les systèmes morphologiques, la valence, les nuances modales et la structuration des modèles grammaticaux intégrant logiquement l'ensemble des variations lexicales dans un texte. Les apprenants, quant à eux, préfèrent certaines combinaisons lexicales à d'autres. En général, ces combinaisons sont fixées par l'usage et par la fréquence d'emploi dans la langue cible.

De Cock a constaté que les séquences préfabriquées (SP) répétitives fournissaient une variété d'informations sur la langue des apprenants du français langue étrangère (De Cock, 2004 : 233). L'analyse de ces séquences permet de décrire

non seulement certaines différences entre la langue produite par des locuteurs natifs et non natifs mais aussi l'acquisition précoce de séquences lexicales : celle-ci est l'un des déterminants les plus importants de l'idiomaticité et de la richesse lexicale de la langue (Juknevičienė, 2011 : 45) et elle représente un levier puissant pour le développement du FOS.

Cet article pourrait avoir un apport pour la formation linguistique en français sur objectifs universitaires (FOU) à destination d'étudiants lituaniens qui intègrent la licence en lettres de l'université de Vilnius. Celle-ci prévoit en effet des éléments relevant du FOS. En l'occurrence, le FOS relève du domaine de la linguistique générale aussi bien qu'appliquée.

## NOTION DE SEQUENCES PREFABRIQUÉES

On sait aujourd'hui que le nombre d'unités polylexicales dans une langue dépasse le nombre d'unités monolexicales (Mejri, 2005). Ces unités polylexicales ou séquences figées (SF) font l'objet d'un intérêt croissant dans la recherche en sciences du langage et en didactique du FLE. Est considérée ici comme séquence figée toute unité polylexicale dont la structure et le fonctionnement font un tout, ayant un certain degré de figement. Cette dénomination, proposée par Mejri dans ses travaux sur le figement couvre l'ensemble des objets de recherche décrits sous des appellations telles que expressions figées, locutions, collocations, clichés, phrasèmes, parémies, pragmatèmes, expressions idiomatiques, idiotismes, etc.

Pour aider les apprenants à développer et améliorer leur FOU/FOS en formation linguistique, nous nous intéressons à un type de co-occurrences de langage préfabriqué nommé 'blocs lexicaux *lexical bundles*' par Biber (Biber *et al.*, 1999 : 989) et par d'autres auteurs. Ces structures sont aussi appelées 'séquences récurrentes *recurrent sequences*' (Howarth, 1998 ; Gledhill, 2000 ; de Cock, 2004), autrement dit des séquences préfabriquées.

Quel que soit le terme choisi, les chercheurs définissent les séquences préfabriquées ou blocs lexicaux comme des séquences répétitives de mots dans le texte (Juknevičienė, 2011 : 24). Biber explique que ce 'sont des blocs composés de trois ou quatre unités, *quelque chose comme, comme tu sais, plus ou moins*, dans des textes avec des registres spécifiques' (Biber *et al.*, cités par PerezBettan, 2015 : 35). De plus, Biber (1999 : 989) remarque que '[l]es blocs lexicaux peuvent être considérés comme des collocations étendues : des blocs de mots qui montrent une tendance statistique à coexister'<sup>1</sup>.

Le nombre, la richesse, la complexité et la variété des SP vont de pair avec la fréquence de leur usage aussi bien dans la langue courante que dans les discours de spécialité. Néanmoins, comme le montrent des analyses précédentes

---

<sup>1</sup> 'Lexical bundles can be regarded as extended collocations: bundles of words that show statistical tendency to co-occur' (notre traduction)

de séquences extraites, les blocs lexicaux possèdent souvent des objectifs fonctionnels tels que 'l'expression du point de vue, l'organisation du discours et l'expression d'un sens référentiel'<sup>2</sup> (Biber *et al.*, cités par Vilkaitė, 2016 : 34) où le sens est défini comme le produit des processus de signification et de référence. Les structures référentielles inscrivent un énoncé à la manière pertinente dans la situation d'énonciation (Béliston, 2017 : 4).

Corollairement, maîtriser les SP dans une langue, c'est aussi pouvoir comprendre les discours de spécialité, quel que soit le domaine. L'avantage d'utiliser des corpus d'apprenants dans des études sur l'acquisition des langues, vu leur taille et les textes inclus, est qu'ils permettent d'observer des régularités d'utilisation de moyens lexicaux mais sont aussi un moyen efficace d'analyser le développement langagier et de l'interlangue en général (Myles, 2005 ; Granger, 2012 ; Kazlauskienė, 2019 ; 2020 ; 2021).

## OBJECTIFS DE LA RECHERCHE

L'objectif principal de notre étude est d'analyser, sur les plans quantitatif et qualitatif, des séquences lexicales dans le corpus écrit d'étudiants lituaniens en philologie française ; ce corpus est constitué de mémoires d'étudiants de troisième année de licence. En prenant en considération notre contexte spécifique directement lié au FOU et au FOS, les informations relatives à l'emploi général de séquences répétitives de deux ou de plusieurs mots (ou d'autres unités linguistiques) dans le corpus sont importantes dans la mesure où elles donnent l'image la plus nette possible de la construction typique du discours académique, des écrits universitaires et de la capacité des étudiants à employer les mots de façon cohérente dans un texte spécifique et orienté.

Les besoins linguistiques des étudiants en FOU sont divers ; ils concernent la capacité à la réflexivité critique dans la rédaction d'un écrit universitaire. Les productions universitaires visées dans l'élaboration d'une formation en FOU sont directement liées au discours scientifique et recèlent de nombreuses formes de communication tant orales qu'écrites : article de recherche, proposition de communication, thèse, compte-rendu, affiche, conférence, mémoire de recherche ou professionnel, rapport de stage problématisé, cours et débats (Mroue, 2014 : 9). En FOS/FOU, le savoir-agir recouvre la nécessité pour l'apprenant de comprendre les différents paramètres contextuels de l'action, qu'elle soit professionnelle, pédagogique ou scientifique et d'acquérir les compétences linguistiques associées aux tâches professionnelles ou académiques exigées : connaissance préalable du protocole actionnel, statut des différents (inter) actants, positionnement dans l'entreprise ou l'institution, objectifs de la tâche, lieu, moment, etc. (Mangiante, 2017 : 24).

---

<sup>2</sup> 'expressing stance, organizing discourse, and expressing referential meaning' (notre traduction)

Dans le présent article, nous nous intéressons plus précisément à l'écrit de recherche universitaire ; nous examinons les composantes et spécificités de la langue académique des étudiants lituaniens en philologie française. Notre but étant de retracer l'opération par la métalinguistique, plus précisément l'utilisation des termes spécifiques au sujet de la linguistique, nous présentons une vue d'ensemble des constructions lexicales observées dans les séquences préfabriquées propres au corpus à notre disposition. Nous décrirons le rapport qu'entretiennent les scripteurs avec cette pratique langagière, et plus particulièrement les difficultés qu'éprouvent les jeunes chercheurs à la maîtriser.

## **METHODOLOGIE**

Notre but étant d'étudier le corpus d'apprenants à des fins didactiques, notamment pour proposer des astuces qui aideront les étudiants à mieux maîtriser le discours académique des écrits universitaires, nous avons constitué un corpus-pilote pour les analyses primaires qui vise à évaluer la nécessité de procéder à ce type d'analyse.

Les écrits universitaires des étudiants lituaniens en philologie française sont constitués de mémoires d'étudiants de troisième année de licence sur la période 2014-2017. Les apprenants sont des étudiants débutants, c.à.d. ayant commencé à apprendre le français à l'université, ce qui pourrait avoir une certaine influence sur les résultats obtenus.

Notre corpus est un corpus universitaire destiné à la conception de ressources de formation. Les sujets (FOS) des mémoires sont directement liés à la spécialité des étudiants, à savoir la linguistique, compte tenu de leur filière philologique.

Le corpus-pilote est composé de 7 717 mots (dont 1 870 mots différents). Toutes les citations de linguistes ont été retirées du corpus pour éviter l'interprétation fautive d'une langue écrite qui ne serait pas celle des étudiants. Les pages de titre et les bibliographies ont aussi été laissées de côté, en prenant en considération leur correspondance aux exigences formelles ce qui ne pourrait pas être interprété en tant qu'un exemple de la langue authentique des étudiants.

## **ANALYSE DES DONNEES**

Nous avons repéré dans le corpus-pilote certaines caractéristiques de séquences préfabriquées (SP) propres aux écrits universitaires d'étudiants lituaniens en philologie française. Notre évaluation démontre l'intérêt de mêler extraction automatique et intervention manuelle pour repérer des SP et permettre ainsi de les évaluer.

Pour commencer, nous avons constitué une liste des mots du corpus. Les premiers résultats font ressortir une faible diversité lexicale : le corpus de 7 717 mots au total se compose de 1 870 mots différents ; la liste des mots fréquemment

utilisés, c.à.d. au moins 5 fois, se compose de 245 mots. Le répertoire des mots du corpus fait aussi ressortir la prévalence de la classe nominale. Statistiquement, un mot sur quatre du corpus est un nom. Les déterminants représentent 22,6% des mots les plus fréquents dans le corpus-pilote ; les prépositions, 10,4% ; les noms 6,3 % et les verbes 0,03%. Le pourcentage prépondérant de déterminants, qui sont des composants caractéristiques des phrases nominales, et la fréquence des noms montre une fois de plus que la présence nominale prévaut, dans le corpus. Le déterminant le plus fréquent dans la construction de SP de FOU est un démonstratif, comme *sur ce point de vue*.

## 1 SP NOMINALES

En ce qui concerne les SP nominales, elles sont étroitement liées à la thématique des écrits académiques comme nous le voyons dans le tableau 1 :

Tableau 1 Exemples de SP illustrant la corrélation avec la thématique des écrits académiques

Progression d'Instagram dans l'enseignement du FLE	Attrition de la langue française du fait que l'on vit en Lituanie	Dialecte picard dans les médias locaux	Parfums exprimés par les mots	Figures de style dans le discours publicitaire
apprentissage des langues	situation de multilinguisme en	langue maternelle	le discours du parfum	un phénomène social
réseaux sociaux en ligne	l'attrition de L2 en environnement de	langue officielle	dénominations du parfum	publicité imprimée
applications d'apprentissage	impact de l'attrition	d'initiation à la langue	le parfum des fleurs	publicité comme une offre de
données visuelles contextualisées	l'interférence inter-langagière	langues régionales appartiennent	senteur parfumée	exprimer l'opinion

Les exemples du tableau 1 montrent comment la composition des SP est influencée par la thématique, c.à.d. Le sujet des écrits des étudiants. Les SP relevées sont de nature simple, correspondent plutôt à la compétence linguistique générale et ne présentent pas de traits caractéristiques du discours académique, les écrits analysés ne sont pas marqués par le lexique transversal ou spécialisé. L'inventaire lexical correspond plutôt au contenu des manuels de FLE. Cela correspond aux résultats des recherches de Cortes (Cortes, citée par Juknevičienė, 2011 : 41), qui a démontré que les écrits académiques d'étudiants contenaient beaucoup moins de blocs lexicaux de langage académique professionnel, qui sont

des blocs de référence nominaux, moins que d'autres types fonctionnels de blocs, qui contribuent à l'expression du contenu propositionnel du texte. Pour réussir leurs écrits, les étudiants doivent donc maîtriser un ensemble d'éléments ; nous supposons que les étudiants se trouvent très souvent amenés à produire des écrits auxquels 'ils ne sont pas formés' (Mroue, 2014 : 40). Mroue (2014 : 27) affirme que le scripteur est amené à 'interagir' – implicitement ou explicitement – dans son processus d'écriture, avec la communauté scientifique omniprésente et aussi avec les sources de ses connaissances (auteurs, courants, publications). Ce qui entraîne la possibilité d'une multiplicité de voix qui traversent son discours (Mroue, 2014 : 27). Ceci est aussi le cas des SP nominales tirées du corpus analysé dans le cadre de cette recherche. Ils présentent, en effet, un contenu lexical directement lié au sujet des mémoires. En revanche, les cas qui pourraient être analysés dans le cadre du FOU sont présents fragmentairement. Tel est le cas de *par exemple* (21 cas), qui introduit des arguments, des illustrations et des faits. Les blocs *la partie pratique*, *la partie théorique* assurent la cohésion et la structuration du discours ; les expressions figées telles qu'*en fonction de* annoncent une démarcation, une argumentation, mais elles ne sont présentes qu'en quantité limitée. Bien entendu, les SP nominales expriment théoriquement moins de nuances du discours académique que les SP verbales.

Les pronoms ne sont pas dans le nombre des SP fréquentes. Ils illustrent seulement la démarche privilégiée par un étudiant pour présenter les étapes de son travail et ses réflexions. *On* et *nous* sont en proportion presque égale : *nous* (58 cas), *on* (56 cas). Ce qui distingue souvent l'emploi de l'une ou de l'autre forme, c'est la prudence des scripteurs dans la rédaction de leurs conclusions. Ils privilégient le *on* (*on constate*, *on peut tirer la conclusion*, *on distingue*, *on suppose*, etc.) pour éviter d'assumer la responsabilité d'une position personnelle. La surutilisation du pronom *on*, dans certains cas, pourrait être influencée par le contenu des manuels de FLE utilisés qui sont orientés à l'approche communicative. Par conséquent, les apprenants du français ignorent souvent le registre familier de ce pronom. Cependant, les étudiants donnent la préférence au *nous* dans la structuration de leur discours écrit (*nous présentons*, *nous décrivons*, *nous voyons*, *nous nous basons*, etc.).

Pour leur part, les SP adjectivales ne présentent pas non plus d'intérêt majeur dans le corpus analysé. Un cas particulier est l'adjectif *linguistique* qui encadre les recherches des étudiants dans le domaine de la linguistique : *un concept linguistique*, *un plan linguistique*, *contexte linguistique*, *pragmatique linguistique*. Un autre cas intéressant est l'adjectif *important*, qui ne trouve pas de synonymes dans les textes des étudiants et qui est employé aussi bien pour relever la particularité, la certitude, l'importance d'un fait que pour justifier la présence de tel ou tel élément dans le discours : *une présence importante*, *défini comme importante*, *il est très important*, *le plus important* etc. Après l'analyse primaire des listes des SP nominales et adjectivales, nous avons constaté que le FOS était moins représenté

que le FOU, la métalinguistique n'était pas clairement développée et que les étudiants passaient par des notions largement utilisées et proposées ou codées dans le sujet de leurs travaux.

## 2 SP VERBALES

En dépit de la capacité des verbes à introduire la parole de l'autre, à argumenter, à proposer des points de vue et à effectuer autres fonctions nécessaires à ce type d'écrit, force est de constater que parmi les 100 mots les plus utilisés, cette catégorie ne se présente que sous la forme de quatre verbes simples – *pouvoir, dire, faire, comprendre* – et recense 73 cas dans tout le corpus-pilote. Toutes sortes des verbes introducteurs ne sont pas en quantité significative. Le tableau 2 ci-dessous montre une fréquence plus grande des verbes basiques comme être et *avoir*. Cependant, même si la quantité de SP verbales (cf. tableau 2) n'est pas remarquable, leur présence est beaucoup plus fonctionnelle du point de vue de l'écrit académique.

Tableau 2 **Panorama des 100 premiers mots du corpus**

Classe grammaticale	Quantité pour les 100 premiers mots les plus utilisés
déterminant	1743
préposition	803
conjonction	508
nom	488
pronom	251
adverbe	182
verbes <i>être</i> et <i>avoir</i> , autres verbes	240 ; 73
adjectif	71
nom de nombre	44

Nous constatons que les étudiants connaissent certains éléments phraséologiques fréquents, mais leur intuition linguistique ne les aide pas à savoir pourquoi insérer certains plutôt que d'autres, à exprimer des nuances ou à choisir les SP qui conviennent le mieux à chaque situation. Les étudiants appliquent à tous les contextes linguistiques les SP bien apprises. Les résultats des recherches précédentes (Cavalla, 2019 : 101) montrent que l'apprenant construit souvent

son discours de façon inverse : il choisit un mot, dont l'utilisation est familière et tente de le placer au bon endroit.

Par ailleurs, l'observation en contexte nous a permis de relever les particularités de la langue analysée des apprenants. Nous avons en effet analysé le positionnement des apprenants dans leur texte à l'aide de SP verbales en nous basant sur la liste des collocations dressée par Tutin (2007 : 7) qui s'est intéressée au lexique et à la phraséologie des écrits scientifiques. Dans notre corpus-pilote, l'inventaire des SP verbales du lexique consacré au positionnement (démarcation et prolongement) (cf. tableau 3) ne se distingue pas par sa variété et richesse.

Tableau 3 Inventaire des SP du lexique du positionnement

Positionnement		
Prolongement	Démarcation	Nouveauté
64%	3%	8%
Défendre, fournir, adapter : 25%		

La catégorie la plus pertinente est celle du prolongement, puisqu'on trouve des SP de forme *confirmer l'hypothèse, suivre une approche, appliquer le principe*, etc. Encore une fois, nous percevons la capacité des étudiants à présenter des faits et les synthétiser, mais ce qui manque ce sont des compétences scripturales de l'écriture académique universitaire. C'est toujours de mettre ces faits en question, de présenter les limites des recherches actuelles, sans parler de proposer une approche ou d'introduire une notion qui posent des difficultés aux étudiants. 'L'écriture universitaire doit permettre de s'approprier des savoirs disciplinaires de référence et d'être capable de les interroger' (Rinck, 2011 : 79). Or les résultats de l'analyse n'en apportent pas une preuve suffisante.

### 3 SP PREPOSITIONNELLES

Le panorama des 100 premiers mots les plus répétitifs du corpus (tableau 2), qui sont à la base des SP les plus fréquentes, se caractérise entre autres par la nécessité de distinguer les prépositions. Même si leur présence pourrait être interprétée comme étant dictée par la valence des verbes, les concordanciers verbaux ne valident pas cette hypothèse, la listes des verbes étant très limitée. Notons la prépondérance de *en, par* et *sur* parmi les prépositions les plus présentes.

*En* est une préposition qui forme le plus des SP directement liées au FOS et au FOU grâce à ses capacités à former le gérondif : *en parlant, en analysant*, etc. Quelques exemples : *en parlant de (ce cas), en cherchant le nombre de*. Parmi les autre SP les plus fréquentes avec *en*, relevons celles qui expriment la quantité : *de plus en plus, de moins en moins, en moyenne, en somme* ; la présentation :



*en ce cas, en ce qui concerne, en tel façon, en particulier ; l'étude : étudier plus en détail ; l'explication : en fait plus, en nous laissant entendre, présenter plus en détail.* Toutes ces prépositions sont employées pour présenter des données.

*Sur* est souvent utilisée pour donner certaines explications de contenu : *un livre sur l'étude, les recherches sur, trouver des données sur ; argumenter, se fonder sur les données : s'accorde sur le fait que, sur le plan de, est basé sur le principe de, se concentre sur la façon, sur un plan sémantique, agir sur le même façon, sur un plan linguistique, sur ce point, nous nous baserons sur le schéma de, s'appuyer sur l'avis de.*

*Par* sert plutôt à la démonstration : *par exemple, par hasard ;* référence à un linguiste *X : le schéma fait par, illustré par, présenté par, proposé autrefois par, utilisé par, crée par chercheurs, l'étude fait par, cité par, distingué par, on voit par une explication de dictionnaire ;* explication : *par son origine, séduire par cette originalité, aussi bien que par ses, finissent par des points, par son origine, par cette raison, on peut nommer X par un autre mot.*

Entre autres SP, il est à relever la conjonction fréquente *comme* qui, *comme par,* exprime la référence à un autre (*des paroles comme, nombreux auteurs comme, tout comme les autres, définit comme, comme écrit X, définit cette fonction comme importante, comme souligne X*), fournit des explications fondées sur les données reçues (*les cas comme ceux-là, comme on voit par, les expressions comme*) ou sur une autre opinion (*considéré comme, c'est comme, définit cette fonction comme importante*), fonde sa propre opinion ou observation (*comme nous avons déjà mentionné, comme nous l'avons évoqué*). Parmi les conjonctions fréquentes dans le corpus, il est aussi à mentionner les formes figées *bien que, afin que,* etc.

Tous les exemples cités démontrent la capacité des SP prépositionnelles et conjonctionnelles d'être fortement orientées sur la fonctionnalité de l'écrit académique.

## CONCLUSION

Alors que nous débutons notre recherche sur l'écriture universitaire d'étudiants lituaniens en philologie française, les données de la première analyse du corpus-pilote nous permettent déjà de relever certaines particularités démontrant la nécessité de ce type d'analyse pour le développement de FOU et de FOS chez les étudiants. Le corpus révèle :

- la prévalence de SP nominales de contenu général, liées à la thématique ;
- une quantité et une diversité limitées des SP verbales ;
- la manipulation fréquente d'expressions figées bien apprises, surtout pour les SP prépositionnelles ;
- une bonne présentation des faits, tandis que les compétences liées à l'agir, à l'argumentation, la discussion, l'expression de doutes et le questionnements ne sont pas bien développées.

A l'instar de Cavalla et Bou (2010), nous avons constaté deux types de difficultés dans les écrits analysés, à savoir des difficultés 'discursives' et 'méthodologiques', typiques exprimant la position du chercheur, le point de vue, la problématisation et la démarche argumentative. Les différents dysfonctionnements rapportés sont observés dans des écrits d'étudiants lituaniens. Entre les difficultés les plus significatives, nous relevons plutôt l'incapacité à bien présenter la problématique et, par conséquent, les difficultés à savoir agir, argumenter et nuancer le discours. Les étudiants mobilisent de la terminologie adéquate mais celle-ci est souvent limitée au cadre donné. Cependant, le corpus relève les SP qui aident les apprenants à démontrer, expliquer, synthétiser l'information, introduire et relier les idées, mettre en relief, citer un auteur, de façon plus ou moins convenable. Ce type d'analyse permet de définir comment les étudiants maîtrisent cette réflexivité, ce qu'ils possèdent mais ne savent pas toujours transposer. Etant donné que le corpus est constitué d'écrits d'étudiant débutants, on peut supposer qu'il s'agit de difficultés de 'littératie académique qui influencent leurs productions écrites' (Barton et Hamilton, 2010 ; Chiss, 2012), voire de lacunes dans le développement de la compétence générale linguistique et de l'influence de leur interlangue.

Puis, à un niveau micro, le champ d'étude de la phraséologie a montré qu'il était indispensable de travailler, d'une part sur le lexique spécialisé car 'le discours universitaire fait usage de phrasèmes qui lui sont spécifiques' (Cavalla, 2018 : 192), d'autre part sur les concordanciers, pour analyser ces éléments lexicaux en contexte (Tran, 2014). Il est envisageable de se focaliser sur certaines situations de communication correspondant aux besoins identifiés des étudiants lituaniens de philologie française, l'objectif étant d'aider les apprenants à utiliser les corpus numériques pour leurs écrits académiques et de les guider dans la découverte de ce lexique et de la phraséologie transdisciplinaire liée à ce type d'écrit.

La rédaction académique répondant à des règles discursives très rigoureuses est l'une des préoccupations majeures des étudiants, d'où l'intérêt à lui accorder en FOU et pas uniquement en FOS. Ce type de données pourrait en outre constituer la base de diverses démarches scientifiques relatives à l'apprentissage pour le développement du FOS aussi bien que du FOU.

## RÉFÉRENCES

- Barton, D. et Hamilton, M. (2010) La littératie: Une pratique sociale. *Langage et société*, 3 (133): 4562.
- Béllisson, C. (2017) le sens des référentiels professionnels et les processus de signification et de référence. *Questions Vives* 28. Disponible en ligne <http://journals.openedition.org/questionsvives/2341> [Consulté le 2 janvier 2021].
- Biber, D., S. Johansson, G. Leech, S. Conrad and Finegan, E. (1999) *Longman Grammar of Spoken and Written English*. Harlow. Harlow: Longman.

- Cavalla, C. (2018) Lexique transdisciplinaire et enseignement aux étudiants allophones. In A. Tutin et M.-P. Jacques *Lexique transversal et formules discursives des sciences humaines*: 191-214. London : ISTE Éditions.
- Cavalla, C. (2019) Une méthodologie sur corpus pour l'écriture en FOU. *Points Communs* 47: 91-108. Disponible en ligne <https://www.lefrancaisdesaffaires.fr/professeurs/ressources/points-communs/> [Consulté le 2 janvier 2021].
- Cavalla, C., BOU, P. (2010) Un référentiel-outil de compétences méthodologiques. In Cadet, L., Goes, J. et Mangiante, J.-M. (Eds). *Langue(s) et intégration: Dimensions institutionnelle, socio-professionnelle et universitaire*. Bruxelles: Peter Lang : 365-381.
- Chiss, J.-L. (2012) De la littératie aux littératies: conceptions et frontières. *Recherches en didactique des langues et des cultures*: 9-2. Disponible en ligne <http://journals.openedition.org/rdlc/3943> [Consulté le 2 janvier 2021].
- Cortes, V. (2004) Lexical bundles in published and student writing in history and biology. *English for Specific Purposes*, 23: 397-423.
- De Cock, S. (2004) Preferred Sequences of Words in NS and NNS Speech. *Belgian Journal of English Language and Literatures (BELL)*, 2: 225-246.
- Gledhill, C. (2000) *Collocations in Science Writing*. Tuebingen: Gunter Narr Verlag.
- Granger, S. (2012) How to use foreign and second language learner corpora. In A. Mackey and S. Gass (éd.). *Research methods in second language acquisition: a practical guide*: 7-29. Oxford: Wiley-Blackwell.
- Howarth, P. (1998) The phraseology of learners' academic writing. In A. Cowie (éd.). *Phraseology: theory, analysis, and applications*: 161-86. Oxford: Oxford University Press.
- Juknevičienė, R. (2011) *Leksinės samplaikos svetimkalbių ir gimtakalbių vartotojų rašytinėje anglų kalboje*. Thèse de doctorat dirigée par J. Grigaliūnienė. Université de Vilnius.
- Kazlauskienė, V. (2019) The Apposition in the Written Production of Lithuanian Learners in French. *Sustainable Multilingualism*, 15 (1): 89-104.
- Kazlauskienė, V. (2020) L'adjectif dans l'interlangue française des apprenants lituaniens. *Language for International communication: Linking Interdisciplinary Perspectives*, 3: 213-228.
- Kazlauskienė, V. (2021) Collocation nomino-adjetivale dans la production écrite en FLE. *Taikomoji kalbotyra*, 150: 143-154.
- Mangiante, J.-M. (2017) Discours et action(s) en milieux professionnel et universitaire: d'une norme d'usage à une contextualisation didactique en FOS et FOU. In: *le français en contextes: Approches didactiques, linguistiques et acquisitionnelles* [en ligne]. Perpignan: Presses universitaires de Perpignan.
- Mejri, S. (2005) Figement, néologie et renouvellement du lexique. *Linx* 52. Disponible en ligne <http://journals.openedition.org/linx/231> [Consulté le 2 janvier 2021].
- Mroue, M. (2014) Écrit de recherche universitaire: éléments pour une sensibilisation au positionnement scientifique à travers la phraséologie transdisciplinaire. *Linguistique*. Université Grenoble Alpes.
- Myles, F. (2005) Interlanguage corpora and second language acquisition research. *Second Language Research*, 21: 373-391.
- Perez-Bettan, A. (2015) *Apprentissage et utilisation du langage préfabriqué chez des apprenants de français langue étrangère*. Thèse de doctorat dirigée par H. Hilton et O. Theophanous. Université de Paris 8.

Rinck, F. (2011) Former à (et par) l'écrit de recherche: quels enjeux, quelles exigences? *le Français Aujourd'hui*, 174 : 79-89. Disponible en ligne: <http://www.cairn.info/revue-le-francais-aujourd-hui-2011-3-page-79.htm> [Consulté le 2 janvier 2021].

Tran, T. T. H. (2014) Description de la phraséologie transdisciplinaire des écrits scientifiques et réflexions didactiques pour l'enseignement à des étudiants non-natifs: application aux marqueurs discursifs. Thèse de doctorat, Université de Grenoble.

Tutin, A. (2007) Modélisation linguistique et annotation des collocations: application au lexique transdisciplinaire des écrits scientifiques. In S. Koeva, D. Maurel, M. Silberztein (Eds), *Formaliser les langues avec l'ordinateur*. Besançon, Presses universitaires de Franche-Comté: 189-215.


Vilkaitė, L. (2016) Formulaic language is not all the same: comparing the frequency of idiomatic phrases, collocations, lexical bundles, and phrasal verbs. *Taikomoji kalbotyra*, 8: 28-54.

## PREFABRICATED SEQUENCES IN THE CORPUS OF LEARNERS IN FSP / FAP

**Abstract.** The analysis of prefabricated sequences makes it possible to describe the early acquisition of lexical sequences being one of the most important factors determining the idiomatic character and the lexical richness of the language which represents a strong springboard for the development of the FSP. The main objective of our analysis focuses on the quantitative and qualitative analysis of lexical sequences in the written corpus of the memoirs of third-year undergraduate Lithuanian students of French philology. The aim of the study being the analysis of the metalinguistic properties of the corpus, this research presents an overview of the lexical constructions of the prefabricated sequences specific to the corpus at our disposal, and provides a description of the peculiarities of their composition. The data obtained from the study of the pilot corpus allows us to identify certain peculiarities which demonstrate the need for this type of analysis for the development of FSP and FAP in students.

**Key words:** prefabricated sequence, FSP, FAP, academic writing, corpus of learners

**Vitalija Kazlauskiene** (Maître de conférences, Dr en linguistique appliquée) travaille actuellement à l'Université de Vilnius, en Lituanie. Ses intérêts de recherche incluent la communication écrite, la grammaire communicative, l'acquisition du langage, le langage et migration.

 <https://orcid.org/0000-0003-0505-7770>

Courriel : [vitalija.kazlauskiene@flf.vu.lt](mailto:vitalija.kazlauskiene@flf.vu.lt)