

Collocation nomino-adjectivale dans la production écrite en FLE

Vitalija Kazlauskienė

Université de Vilnius / Vilnius University

Adresse électronique : vitalija.kazlauskiene@ff.vu.lt

Annotation. D’après Tutin et Grossmann, « on caractérise souvent les collocations par le fait que le sens est transparent en réception, alors que, pour un locuteur non-natif, il est difficile de produire les lexèmes adéquats » (Tutin, Grossmann 2002). Nous tenterons de vérifier cette affirmation et de définir les régularités des combinaisons lexicales dans l’interlangue du FLE des apprenants lituaniens. Nous nous fixons ainsi comme tâche, sans prétendre à l’exhaustivité, de faire le point sur les fréquences et particularités des collocations nomino-adjectivales dans les productions écrites des apprenants. Le corpus révèle que la combinaison des mots n’est pas toujours typique du français standard. Le processus d’interférence et d’hybridation peut être vu comme une contamination essentielle des collocations, tant au niveau lexical que syntaxique. Cependant, les expressions partiellement figées, autrement dit les collocations, sont un moyen important de s’exprimer pour les élèves en atteinte du niveau B1.

Les mots clés: *collocation; cooccurrence; SN; corpus; FLE*

Nomino-Adjectival Collocation in Written Production of L2 French

Abstract. The acquisition of lexical competence is complex, since it is not enough for learners to distinguish the form and the meaning of a lexical element in order to know how to correctly integrate it into the language context. Learners prefer some lexical combinations to others. Generally speaking, these combinations are determined by usage and by frequency of use in the target language. The question of frequency leads to characterizing the particular cases of co-occurrences as collocations (Anctil, Tremblay 2016). According to Tutin and Grossmann, “one often characterizes collocations by the fact that the meaning is transparent in reception (it is “guessed”), whereas, for a non-native speaker, it is difficult to produce the appropriate lexemes” (Tutin, Grossmann 2002). We will try to verify this statement and define the regularities of lexical combinations in L2 French interlanguage of Lithuanian learners. The aims of this presentation is to take stock of the frequencies and particularities of nomino-adjectival collocations in the written productions of non-native French learners, without claiming to be exhaustive. Thanks to the corpus at our disposal, we have extracted certain characteristics of nomino-adjectival collocations specific to interlanguage. Nomino-adjectival collocations are characterized by the structure (Adj) + N + (Adj) in the corpus. Collocations of the N + Adj type are more frequent than that of the Adj + N type. The corpus also reveals that the combination of words is not always characteristic of the general language. The process of interference and hybridization can be seen as an essential contamination of collocations, both lexically and syntactically. However, partially set phrases, in other words collocations, are an important way of expressing themselves for learner with B1 level. The results of this study demonstrate that the analysis of collocational constructions can reveal the relationship between competence and performance of speakers. This type of data could also form the basis of various scientific approaches relating to teaching and learning of foreign languages.

Keywords: *collocation; co-occurrence; SN; corpus; L2*

1. Introduction

L'acquisition de la compétence lexicale est complexe, dans la mesure où il ne suffit pas pour les apprenants de distinguer la forme et la signification d'un élément lexical pour savoir comment l'intégrer correctement dans le contexte langagier. Cette compétence suppose de comprendre les relations syntaxiques, les systèmes morphologiques, les nuances modales, et la structuration des modèles grammaticaux. Les apprenants préfèrent certaines combinaisons lexicales à d'autres. De façon générale, ces combinaisons sont fixées par l'usage et par la fréquence d'emploi dans la langue cible. La question de la fréquence amène à caractériser les cas particuliers de cooccurrences en tant que collocations (Anctil, Tremblay 2016). D'après Tutin et Grossmann, « on caractérise souvent les collocations par le fait que le sens est transparent en réception (il se « devine »), alors que, pour un locuteur non-natif, il est difficile de produire les lexèmes adéquats » (Tutin, Grossmann 2002). Nous tenterons de vérifier cette affirmation et de définir les régularités des combinaisons lexicales dans l'interlangue du FLE des apprenants lituaniens. Nous nous fixons ainsi comme tâche, sans prétendre à l'exhaustivité, de faire le point sur les fréquences et particularités des collocations nomino-adjectivales dans les productions écrites des apprenants non-natifs du français.

Nous comptons sur les exemples fournis par le corpus d'apprenants pour saisir les particularités combinatoires ainsi que les spécificités du champ collocationnel des noms accompagnés d'adjectifs. Les écrits des élèves nous offrent également la possibilité d'évaluer la fréquence générale de ce type de construction. Ils permettent d'orienter les réflexions didactiques vers l'emploi général des collocations nomino-adjectivales dans le corpus d'apprenants, lesquelles sont importantes pour distinguer le plus nettement possible la construction typique et propre à l'interlangue. Ce type de données pourrait en outre constituer la base de diverses démarches scientifiques relatives à l'apprentissage des langues étrangères. L'analyse des spécificités linguistiques des collocations nomino-adjectivales devrait notamment conduire à préciser la réflexion didactique pour leur enseignement.

2. La notion de collocation

La définition de la collocation retenue dans cette étude, largement partagée (Mel'cuk 1998 ; Polguère 2000 ; Grossmann et al. 2003 ; Cavalla et al. 2014), consiste en la rencontre fréquente de deux lexies qui se suivent ou sont séparés par d'autres mots. Hausmann soutient qu'« il ne suffit pas de voir la collocation comme une combinaison sous contrainte, il faut encore comprendre qu'elle est une combinaison orientée. En tenant compte de cette orientation on appelle l'un des éléments de la collocation – base et l'autre – collocatif. En effet, dans la collocation *célibataire endurci*, le signifié de la base (*célibataire*) est autonome. La base n'a pas besoin du collocatif (*endurci*) pour être clairement définie. Il en va tout autrement pour le collocatif qui ne réalise pleinement son signifié qu'en combinaison avec une base (*célibataire*) » (Hausmann 1989 cité par Cavalla 2014 : 153). En d'autres termes, la première unité lexicale (la base ou mot-clé) est choisie librement. La seconde unité lexicale (le collocatif) est choisie en fonction de la base (Heid et Raab 1989 ; Mel'cuk 1998 ; Mel'cuk et Polguère 2007). De cette manière, l'un des éléments conserve son sens habituel tandis que l'autre est quelque peu métaphorisé (les degrés de métaphorisation varient d'une collocation à l'autre) (Grossmann et al. 2003). Les collocations forment une unité de sens, et elles sont soumises à certaines restrictions syntaxiques, sémantiques et pragmatiques. Ces unités linguistiques se caractérisent par un certain degré de figement et de récurrence.

De manière générale, les études menées en phraséologie durant les dernières décennies sont représentatives de deux grandes tendances (Granger et Paquot 2008 ; Bolly 2008 ; 2010 ; 2011) : l'approche

fonctionnelle, qui se centre sur l'étude des unités caractérisées par un degré minimal de figement sémantique (selon des critères linguistiques), et l'approche statistique, qui s'intéresse aux unités polylexicales caractérisées au minimum par une attraction lexico-grammaticale entre unités (selon des critères de récurrence et de cooccurrence). L'orientation fonctionnelle décrite par Mel'cuk (cf. Mel'cuk, Clas et Polguère 1995 ; Mel'cuk et Polguère 2007) porte sur la notion de « fonctions lexicales » (Grossmann et Tutin 2002 ; 2003) et met l'accent sur la présence ou l'absence d'une paire de constituants fonctionnant dans un rapport de base à collocatif. Sinclair (1991) et Biber (1990) exposent quant à eux une orientation statistique (*data driven corpus approach*). Cette approche « procède de l'étude de formes de mots, de leur fréquence, de la fréquence de leur cooccurrence avec d'autres formes de mots et du calcul statistique de la significativité d'une telle cooccurrence » (Clear cité par François et Manguin 2006 : 50).

Les collocations soulèvent des difficultés majeures dans différentes applications de notre domaine d'étude. Ramisch (2012) classe la recherche sur les expressions polylexicales, dont les collocations sont un sous-groupe, en cinq tâches : extraction, identification, désambiguïsation, représentation et application des expressions polylexicales. Plusieurs travaux ont été consacrés à l'extraction de collocations et d'expressions polylexicales à partir de textes. En général, les travaux sur l'extraction de collocations se basent sur des méthodes statistiques. Nous avons nous aussi choisi l'extraction par la base de collocations en utilisant un corpus annoté avec des parties du discours et des informations statistiques, comme l'écart type et la moyenne de la distance entre les mots dans une phrase.

Dans les collocations nomino-adjectivales, les deux éléments constitutifs (c.à.d. un nom, suivi ou précédé par un adjectif, accompagné ou non d'un modificateur sous forme d'un adverbe) forment ensemble une expression plutôt descriptive.

Notre analyse se fonde sur un échantillon de plus de 3000 syntagmes nominaux (SN) de composition (Adj) + N + (Adj) tirés du corpus de l'interlangue des apprenants lituaniens.

3. Les particularités de la collocation nomino-adjectivale dans le corpus de langue française des apprenants lituaniens (résultats de l'analyse empirique)

3.1. Matériel et méthodes

Cette partie expose les résultats de l'analyse empirique des collocations nomino-adjectivales basée sur les données du corpus annoté des apprenants lituaniens (LPRATT). Ce dernier a été constitué dans le cadre de la thèse de doctorat de l'auteure (Kazlauskienė 2018) et complété par les travaux de l'examen de l'année 2018. Il a été élaboré manuellement à partir des travaux écrits des élèves de la classe de terminale pendant leur examen final. Toutes les productions écrites ont été anonymisées et codées. Le corpus entier contient 114 245 mots. Le nombre total de participants s'élève à 325. Le niveau à atteindre est celui de B1-Utilisateur indépendant. Dans cet article les exemples tirés du corpus sont présentés dans leur intégralité en conservant l'originalité des phrases des apprenants, c'est-à-dire sans aucune modification ni correction. Tous les types d'erreurs ont été marqués et étiquetés, dans les phrases les corrections apparaissent entre les signes #...\$. Les constructions de SN spécifiques à ce corpus ont tout d'abord été compilées en faisant l'extrait du concordancier. L'étude de cette liste permet de réaliser une analyse articulée autour de la collocation nomino-adjectivale dans l'interlangue. Sur le plan linguistique, la forme lexicale et syntaxique pour la combinatoire semble primordiale à décrire ; nous nous attacherons à cerner les spécificités du lexique collocationnel propre aux apprenants litua-

niens du FLE. Pour ce qui est du sens, souvent non compositionnel, nous comptons sur les exemples fournis par le corpus.

Une cooccurrence est statistiquement significative lorsque la co-apparition des termes ne semble pas être due au hasard. En d'autres termes, la fréquence observée correspond bien à la fréquence attendue. Pour l'analyse des particularités des collocations, nous avons relevé les cooccurrences de la valeur statistique significative de 5 récurrences au sein du corpus, lequel n'est pas un corpus général de l'interlangue au vu de sa taille et de ses limites lexicales mais est nonobstant représentatif. En d'autres termes, un syntagme répété au moins cinq fois dans le corpus et ayant certaines particularités sémantiques propres à la collocation pourrait fournir des informations sur cette notion caractéristiques au corpus analysé. L'analyse s'appuie sur la liste des SN relevés dans le corpus analysé, dont les cooccurrences adjectif + nom et nom + adjectif. Le nombre de ces deux types de syntagmes est quasiment identique. Les SN étiquetés à la construction adjectif + nom sont au nombre de 1508, et les SN nom + adjectif au nombre de 1573, ce qui montre l'emploi équivalent des deux types de SN dans le corpus d'apprenants. On ne peut relever de préférence pour l'une de ces structures. Cependant, tous les SN ne représentent pas des collocations. Nous pourrions distinguer quelle construction est la plus favorable à la réalisation de collocations au cours de l'analyse des données empiriques.

3.2. La structure adjectif + nom

La structure adjectif + nom est assez répandue dans le corpus, cet ordre étant le plus habituel pour les apprenants dont la langue maternelle est le lituanien. Nous présenterons une vue d'ensemble du lexique des cooccurrences lexicales, la description des particularités de leur composition ainsi que le marquage des difficultés lexicales principales.

Parmi les adjectifs les plus présents dans le corpus, nous pouvons relever les adjectifs tels que : *bon(ne)s*, *nouveau(el)le(s)*, *grand(e)s* etc. Les collocations les plus répandues contiennent l'adjectif *bon(ne)s*, voir le tableau 1.

Tableau 1. Cooccurrences contenant l'adjectif *bon(ne)s*

Cooccurrences	Pourcentage de Adj + N de tous les SN
En bonne forme/en pleine forme	13
Une bonne santé	1,3
Une (très) bonne idée	1
Bonne chance	1
Une bonne chose	0,5
Bon courage	0,42
De bonne qualité	0,36
Un bon choix	0,36
Une bonne nourriture	0,36
Une bon(ne) ami(e)	0,36

Ce tableau présente les SN qui correspondent statistiquement à ce que nous considérons comme des cas représentatifs de ce corpus. En général, l'adjectif *bon* est utilisé, dans la plupart du temps pour intensifier respectivement les qualités, les attitudes positives du nom. L'omission possible de *bon* ne change pas le sens principal, mais la préférence des apprenants se manifeste par la fréquence répétitive de cet adjectif. L'emploi de l'adverbe *bien* au lieu de *bon(ne)s* est également fréquent, ex. :

- (1) <...> *les technologies #modernes\$ moderne #ce n'est pas\$ snapas #une\$ #bonne\$ bien #solution\$ solutions.* (2011E301898)
- (2) *Bonjour, mon #bon\$ bien ami. Ça va?* (2013L165432)

L'exemple 1 illustre l'intrication, significative pour ce corpus, de deux parties du discours sémantiquement proches, à savoir de l'adverbe *bien* à la place de l'adjectif *bon*. L'exemple 2 représente aussi l'utilisation incorrecte de l'adverbe *bien* à la place de *bon*. La situation linguistique prévoit l'emploi d'une collocation possible et bien typique pour ce type d'écriture – *mon cher ami* ou simplement *mon ami*. Cette confusion illustrée par ces exemples est la plus perceptible dans le présent corpus, pour la simple raison que l'adjectif *bon* est le plus révélateur dans les écrits analysés. La seule collocation la plus significative statistiquement est *en bonne forme*, qui est recopiée de la formule du sujet de l'examen « *Que devrions-nous faire pour rester en bonne forme ?* » (2013).

Goes (2018, 592) appelle adjectif « passe-partout » un adjectif pouvant qualifier tous les substantifs ayant un sème [+ intensifiable], comme dans notre cas l'adjectif *bon*. En prenant en considération cette assertion, les cooccurrences du tableau 1 – *une bonne santé, une bonne idée, de bonne qualité, un bon choix, une bonne nourriture* – pourraient être interprétées comme des combinaisons libres. C'est à cause de leur emploi de passe-partout qu'elles se répètent souvent et sont statistiquement représentatives. A ce titre, malgré le nombre d'occurrences, il paraît difficile de ranger toutes les constructions statistiquement significatives du corpus dans la catégorie de la collocation. Tandis que Mejri (2000, 49) écrit que « théoriquement, toute séquence libre est candidate au figement ».

Dans le corpus analysé, nous pouvons relever encore quelques adjectifs passe-partout privilégiés par les élèves, notamment *nouveau* et *grand*. Ainsi, pour parler de nouvelles connaissances, les élèves donnent parfois la préférence à la collocation *les nouveaux amis* (0,6). L'adjectif *nouveau* apparaît également dans le figement de *Nouvel An* (0,4%). *Une nouvelle technologie* est aussi une collocation nomino-adjectivale propre à ce corpus et constitue presque 5 % de tous les SN contenant l'adjectif. Cette présence, relativement assez distincte, s'explique par son occurrence dans le thème de l'examen « *Les technologies modernes apportent-elles des avantages importants à tous ?* » (2011). Dans les syntagmes étudiés, nous avons constaté la présence de deux adjectifs synonymes en parlant de la technologie : *technologie nouvelle* et *technologie moderne* (1,4%). Le choix du bon adjectif et du bon nom au bon endroit est étroitement lié au plan pragmatico-discursif. Le choix ou la place discursive spécifique est lié à la fonction pragmatique recherchée : l'intention de relever la nouveauté de ce qui n'existait pas avant ou la modernité ce qui existe. Cependant, le contexte plus large du corpus nous montre que dans ce cas l'utilisation d'une construction ou une autre se développe de manière inconsciente, plutôt intuitive.

Les collocations avec d'autres adjectifs ne sont pas nombreuses non plus. Parmi les SN les plus variés avec le potentiel de la présence de collocations, nous pourrions encore mentionner les cas avec le déterminant quantitatif (déterminant indéfini) *tout*. Il est à noter que, selon les approches grammaticales dites traditionnelles, il s'agit d'un adjectif indéfini. Mais la classe grammaticale de l'adjectif est problématique dans ce cadre théorique, raison pour laquelle les « adjectifs » possessifs (*ma, ta sa, etc.*), démonstratifs (*ce, cette, etc.*) ou indéfinis de la grammaire traditionnelle ne sont plus considérés comme tels dans les grammaires rénovées actuelles (Riegel, Pellat et Rioul 2009 ; Béguelin 2000 ; Bronckart 2015 ; Chartrand 2016 ; etc.). N'ayant pas la même distribution syntaxique que les adjectifs « standards » (ils ne commutent pas de la même manière), ces types de termes sont rangés dans la catégorie du « déterminant » (déterminants possessifs, indéfinis, quantitatifs, etc.). En dépit de son statut complexe, *tout* forme des combinaisons plus au moins stables avec le nom en le déterminant par une information sur sa qualité et s'accorde avec celui-ci, voir tableau 2 :

Tableau 2. Cooccurrences contenant le déterminant quantitatif tout(e)s

Cooccurrences	Pourcentage de Dét + N de tous les SN
Tout le monde (entier)	1,9
Tout le temps	1,6
Tous les jours	1
Toute la famille	0,8
Toute sa (la, ma) vie	0,6
Toute la journée	0,43
Tous les gens	0,43

Le tableau 2 illustre la fréquence de la construction *tout le monde*, qui est plutôt figement que collocation. Les collocations telles que *toute la famille*, *tout le temps* sont à la frontière du figement. Le reste n'a pas de traits particuliers. En dehors des cooccurrences mentionnées, quelques constructions singulières se distinguent, dont la liste est exposée dans le tableau 3.

Tableau 3. Autres cooccurrences du type Adj (Dét) + N

Cooccurrences	Pourcentage
Une grande liberté	2,7
(En) premier lieu	2
Cher(e)s ami(e)s	2
Un long voyage	1,6
Grosses bises	1,5
Chaque jour	1,1
(Mon/ma) meilleur(e) ami(e)	1,1
Gros bisous	1
À mon avis	0,9
Les jeunes personnes	0,8
Le jeune homme	0,6
Les petits enfants	0,6
Un grand problème	0,5
En même temps	0,5
D'autre part	0,4
Quelque chose	0,4
Grand avantage	0,36

Le tableau ci-dessus représente une liste des cooccurrences qui ne sont ni diverses et ni nombreuses. Certains cas (tableaux 1–3) listés comme collocation semblent figés : *bonne chance*, *bon courage*, *tout le monde*, *toute la journée*, *en premier lieu*, *grosses bises*, *à mon avis*, *en même temps*, *d'autre part*, *quelque chose*, *grand avantage*. Même si ces syntagmes sont stables, nous considérons certains d'entre eux comme des collocations, notamment *bonne chance*, *toute la journée*, *grand avantage* où soit l'adjectif soit le déterminant conserve son rôle de modificateur et accepte la gradation. À l'inverse, *chaque jour*, *en même temps* sont plutôt considérés comme figements. « C'est bien la relation stable entre l'expression linguistique X et l'élément de la réalité X qui impose le figement et non les caracté-

ristiques particulières de l'adjectif » (Goes 2018 : 599). Selon Mejri, « l'action du figement ne se limite pas au lexique ; elle englobe également la grammaticalisation par le biais de laquelle la langue se dote des outils syntaxiques permettant de structurer le discours » (Mejri 2005 : 164). De cette manière, *tout le monde*, *quelque chose* correspondent à des concepts préexistant dans la langue et ne sont en conséquence pas considérés comme collocations.

Parmi les collocations moins fréquentes mais pertinentes pour les écrits des apprenants, nous trouvons *les jeunes gens*, *une simple question*, *l'air frais*, *temps libre*, etc. L'éventail des collocations est prédéterminé par la thématique du sujet et par le niveau linguistique des apprenants. Certaines difficultés d'apprentissage du FLE viennent s'y ajouter. Le contact de la langue maternelle et de la langue étrangère peut constituer un atout ou au contraire un obstacle, ce que dénote la présence de greffe¹ collocationnelle interlinguale et intralinguale. Selon Goes, « la collocation apparaît comme une rencontre préférentielle d'unités lexicales ; elle revêt un caractère attendu et souvent culturellement lié au sein d'une même communauté linguistique » (2018 : 592). C'est notamment pour cette raison que l'on trouve le résultat de cette interférence dans le corpus d'apprenants, c.à.d. les constructions faites par analogie avec la langue lituanienne telles que *tous les gens*, *une autre semaine*, *les petits enfants*, *dans la première place*, *un sérieux problème* qui sont hybrides du point de vue de la norme et dont on trouve cependant sans peine des équivalents dans la langue française : *tout le monde*, *la semaine prochaine*, *les petits enfants*, *en premier lieu*, *un grand problème*, etc. Ou *un grand voyage* (qui a le sens particulier de la mort), au lieu de *long voyage*. Quelques exemples tirés du corpus :

- (3) *Peut #être\$ être il #commençait\$ commençait beaucoup #de\$ des #médicaments\$ médicaments pour #une\$ un #longue\$ large vie ?* (2011E301931)
- (4) *Mais il y a #un\$ une #sérieux\$ sérieuse #problème\$ problème.* (2015L2643D35)

Les exemples 3 et 4 représentent des cas intéressants d'utilisation des collocations et suggèrent que les élèves ont une intuition linguistique en français et qu'un enseignement pourrait les aider à la développer. Il est à noter que non seulement le choix inconvenable d'un adjectif est un obstacle pour établir, relever et analyser une collocation dans le corpus d'interlangue, mais aussi les formes grammaticalement incorrectes des adjectifs, comme dans le cas de l'exemple 4 illustrant l'accord fautif en genre.

Les données de l'analyse nous montrent bien que les collocations dans la position de Adj + N ne sont pas une manière habituelle de s'exprimer pour les apprenants lituaniens de ce niveau. Les mots apparaissent dans le texte sans collaboration particulière. Les apprenants sont peu habitués à se positionner dans leurs écrits et donc peu enclins à l'utilisation du lexique approprié. Les collocations simples, générales, communes, répétitives, influencées par la thématique de l'examen mais toujours peu utilisées sont caractéristiques du corpus analysé. Les collocations les plus nombreuses sont soit recopiées des thèmes de l'examen (*en bonne forme*, *nouvelles technologies*, *une grande liberté*), soit apprises au niveau inférieur, telles que les collocations *cher ami*, *gros bisous* qui représentent une forme d'adresse pour saluer ou dire au revoir, bien naturelle pour les écrits épistolaires. Après avoir compilé la base nominale en ajoutant en préposition le collocatif en forme d'adjectif, on peut constater que les collocations de ce type ne sont pas intrinsèques au corpus. En général, les séquences libres prévalent dans les écrits des apprenants. La précision du sens du nom par un adjectif qualificatif n'est pas marquée par une grande diversité non plus.

¹ Greffe - l'amalgame de deux ou plusieurs unités linguistiques, de sorte que la physionomie de chacune de ces unités soit plus ou moins directement identifiables. L'amalgame repose sur la base partielle d'une homophonie ou/et d'une dimension sémantique (Duprez 1984, cité par Legallois 2013, 105).

3.3. La structure nom + adjectif

Le deuxième type de SN au contenu adjectival se présente sous la forme de N + Adj. Il se caractérise par sa diversité des adjectifs, collocations (tableau 4) et figements. Parmi les figements les plus répandus, on peut citer *le téléphone portable* (3,3%) et *le monde entier* (2%). Le premier est influencé par le sujet « *Les technologies modernes apportent-elles des avantages importants à tous ?* » (2011), tandis que le figement *le monde entier* est une construction générale, habituelle pour ce niveau d'apprentissage. Les collocations les plus fréquemment rencontrées dans le corpus sont présentées dans le tableau 4.

Tableau 4. Collocations du type N + Adj

Collocations	Pourcentage de N + Adj de tous les SN
La langue étrangère	4,4
Les technologies modernes	2,5
Le pays lointain	2,2
Les personnes (plus, peu) âgées	1,1
La vie quotidienne	1
La nourriture saine	0,9
Réseaux sociaux	0,7
L'année prochaine	0,5
Un message électronique	0,5
Le temps libre	0,5
L'année dernière	0,4
Point positif	0,3
La semaine dernière	0,3
La semaine prochaine	0,3
La langue maternelle	0,3
Beau temps	0,3
La communication virtuelle	0,3
Une vie personnelle	0,3

On constate que les collocations les plus répandues sont celles qui sont communes, générales et souvent reflètent le sujet des textes. Ainsi, dans différents textes, les collocations les plus courantes changent en fonction du sujet. Granger (1998), Juknevičienė (2011) ont remarqué que les apprenants emploient intensément certaines constructions de phrases qui ne sont pas difficiles à utiliser. Siepmann (2008 : 194) a observé que les apprenants d'une langue étrangère ont tendance à surutiliser certains lexèmes et collocations qu'ils ont pu rencontrer dans des manuels d'apprentissage du FLE. Nous avons constaté que les apprenants surutilisent en effet certaines collocations adjectivo-nominales, telles que *la langue étrangère*, *les technologies modernes*, *le pays lointain* etc. Comme mentionné précédemment, les collocations de cette construction nomino-adjectivale sont plus variées. On trouve dans le corpus davantage de collocations que celles présentées dans le tableau 4 mais elles sont parfois difficiles à discerner. Parmi les différentes causes déjà évoquées précédemment, on peut citer l'obstacle de l'ordre incorrect des mots. En raison de l'interférence de la langue lituanienne, l'adjectif se trouve souvent dans la position préposée même quand cela n'est pas nécessaire. Par exemple :

- (5) *Premièrement #,\$ je voudrais dire que les nouvelles, #X modernes\$ #technologies modernes\$ modernes technologies sont très #importantes\$ importants pour le #médecin médecin.* (2011E301856)

- (6) *Il #fait\$ faut \$beau temps\$ **temps très belle**. (2011L301937)*
(7) *#Premièrement\$ Première, la langue #X de\$ de #française\$ français est #une\$ #deuxième\$ deuxième #langue officielle\$ **officielle langue** de l'Union \$européenne\$ Europe, <...>. (2012E510262)*

Les exemples ci-dessus montrent la position inappropriée des adjectifs qui repose sur l'interférence. Il en résulte la présence fréquente d'erreurs de production, qu'on pourrait interpréter comme greffe syntaxique. Il est à noter que le continuum entre la formation collocationnelle et les régularités syntaxiques est direct, en d'autres termes l'ordre erroné des mots contrarie certaines fluidités verbales.

Les greffes collocationnelles sont une difficulté distinguée dans le corpus, visibles dans les exemples 8, 9 et 10.

- (8) *J'ai aussi peur que cette chose soit contemporaine et **le jour** #suivant\$ **prochain** tout va changer. (2011E301892)*
(9) *Néanmoins, quelquefois les jeunes sont trop sous une pression **des personnes mûres**. (2015E2643D13)*
(10) *#le\$ Les test de français était #très\$ très #difficile\$ difficile et je #avais\$ peur que je ne #finirais\$ finisserais #mon\$ ma école #aux\$ avec #résultats\$ **resultats** #mauvais résultats\$ **mal**. (2015L2643D04)*

L'hybridation illustrée dans les exemples 5–10, considérée comme la greffe collocationnelle, suppose la notice d'unités préformées chez les apprenants, ce qui pourrait devenir une sorte de fossilisation. La construction **le jour prochain* au lieu de *jour suivant* (ex. 8) est répétée dans le corpus. L'interférence interlinguale est bien présente, ce que prouve l'emploi par l'analogie, la traduction, etc. L'exemple 9, représentant l'analogie de la langue lituanienne, illustre le choix des apprenants en faveur de **personnes mûres*, au lieu de la collocation *gens adultes*. L'exemple 10 représente plutôt l'interférence intralinguale qui se manifeste par le rapprochement sémantique de l'adjectif *mauvais* à l'adverbe *mal*. En général, l'interférence intralinguale au sein des collocations de ce corpus émerge non seulement par le rapprochement sémantique, mais aussi par l'emploi de synonymes incorrects. Ces types de compatibilité des éléments linguistiques constituent une contrariété essentielle de concaténation de fragments linguistiques. En fait, la combinaison de collocations inappropriées est plus explicite dans la combinaison de N + Adj. La plupart des difficultés liées aux collocations sont sémantiques, puisque les collocations font intervenir des acceptions de vocables dont le sens n'est pas prévisible et dont la combinatoire est incertaine. Ainsi, les élèves d'un niveau B1 en français ont besoin d'une aide pour la fixation de structures lexicales qu'ils connaissent partiellement, peut-être pour les avoir entendues ou lues. De façon générale, les collocations nous apprennent beaucoup « sur le plan appliqué, sur l'énonciation des formes préfabriquées, et la compétence/performance des locuteurs » (Legallois 2013 : 107). Il conviendrait d'attirer l'attention des apprenants sur la palette des moyens linguistiques dont ils disposent et d'insister sur la découverte des collocations afin de les aider à manipuler toutes les unités lexicales à leur disposition au cours de leurs productions écrites personnelles.

4. En guise de conclusion

Grâce au corpus à notre disposition, nous avons extrait certaines caractéristiques des collocations nomino-adjectivales propres à l'interlangue. Les conclusions présentent une synthèse des principaux résultats. Les collocations nomino-adjectivales se caractérisent par la structure (Adj) + N + (Adj) dans le corpus. Suite à l'observation de la fréquence lexicale des SN du corpus d'apprenants à l'aide du logiciel WordSmith Tools, nous avons constaté que les collocations nomino-adjectivale du type N + Adj sont plus fréquentes que celle du type Adj + N.

Le corpus révèle aussi que la combinaison des mots inhérente aux productions des apprenants n'est pas toujours typique du français standard. Le processus d'interférence et d'hybridation peut être vu comme une contamination essentielle des collocations, tant au niveau lexical que syntaxique. Certains éléments privilégient la coprésence des uns et l'omission des autres dans une séquence de mots. Les apprenants opèrent le lexique bien appris dans un contexte particulier, voire dans une certaine construction. Ils font leur choix parmi un nombre assez limité d'unités lexicales. Dans le corpus analysé, nous pouvons relever quelques adjectifs passe-partout privilégiés par les élèves, notamment *bon*, *grand*, *nouveau*. La tendance des élèves à recopier des cooccurrences de formules des thèmes donnés à l'examen est aussi marquée. Cependant, les expressions partiellement figées, autrement dit les collocations, sont un moyen important de s'exprimer pour les élèves en atteinte du niveau B1.

Les résultats de cette étude montrent que l'analyse des constructions collocationnelles est capable de révéler les rapports entre compétence et performance des locuteurs. Nous convenons qu'une approche systématique en corpus pourrait apporter des éléments supplémentaires sur la constitution des collocations nomino-adjectivales. Un regard quantitativement plus informé pourrait davantage éclairer la complexité de ce phénomène. Nous retiendrons, en conséquence, qu'il serait nécessaire de systématiser l'enseignement/apprentissage des collocations auprès des apprenants lituaniens en FLE. Il s'agirait d'aider les élèves à retenir le sens et l'association lexicale simultanément en FLE, ainsi que de fixer les structures qu'ils connaissent déjà partiellement.

Abréviations

FLE – français langue étrangère ; Adj – adjectif ; N – nom ; B1 – niveau intermédiaire ; SN – syntagme nominal ; LPRATT – corpus annoté des apprenants lituaniens ; #...\$ – les corrections des erreurs étiquetées ; Dét – déterminant ; * – forme fautive.

Bibliographie

- Anctil, D., O. Tremblay. 2016. Les collocations : des combinaisons de mots privilégiées. Dictionnaire Lexique. *Logiciels d'aide à la rédaction* 21 (3).
- Béguelin, M.-J. 2000. *De la phrase aux énoncés. Grammaire scolaire et descriptions linguistiques*. Bruxelles: De Boeck-Duculot.
- Biber, D. 1990. Methodological issues regarding corpus-based analyses of linguistic variation. *Literary and Linguistic Computing* 5, 257–269.
- Bolly, C. 2008. *Les unités phraséologiques : un phénomène linguistique complexe ?* Thèse de doctorat, Louvain-la-Neuve, Université catholique de Louvain.
- Bolly, C. 2010. Flou phraséologique, quasi-grammaticalisation et pseudo marqueurs de discours : un no man's land entre syntaxe et discours? *Linx* [En ligne], 62–63.
- Bolly, C. 2011. *Phraséologie et collocations. Approche sur corpus en français L1 et L2*, Bruxelles, Bern, Berlin, Frankfurt am Main, New York, Oxford, Wien, P.I.E. Peter Lang (GRAMM-R. Études de linguistique française, Vol. 9).
- Bulea Broncart, E. 2015. À grammaire rénovée, exercices... ? *La Lettre de l'AIRDF* 57, 39–45.
- Cavalla, C. 2014. Collocations transdisciplinaires : réflexion pour l'enseignement. Maria-Isabel Gonzalez-Rey. *Outils et méthode d'apprentissage en phraséodidactique, EME*, 151–169.
- Chartrand, S.-G. (dir.). 2016. *Mieux enseigner la grammaire au primaire et au secondaire. Pistes didactiques et activités pour la classe*. Montréal: Éditions du Renouveau Pédagogique.
- Goes, J. 2018. Collocations et dénominations : quelques réflexions sur le figement adjectival. *Philologia Mediana*, 589–602.
- Granger, S. 1998. *Learner English on Computer*. London & New York: Addison Wesley Longman.

- Granger, S., M. Paquot. 2008. Disentangling the phraseological web. *Phraseology: An interdisciplinary perspective*. S. Granger & F. Meunier (eds.). Amsterdam, Philadelphia, John Benjamins. 27–49.
- Grossmann, F., A. Tutin (ss. dir.). 2003. *Les collocations : analyse et traitement*. De Werelt: Amsterdam.
- Heid, U., S. Raab. 1989. Collocations in multilingual generation. Dans *Proceedings of the Fourth Conference of ACL, European Chapter, Manchester, 10–12 April 1989*. Manchester.
- Jacques, F., J.-L. Manguin. 2006. Dispute théologique, discussion oiseuse et conversation téléphonique : les collocations adjectivo-nominales au cœur du débat. *Langue française* 2 (150), 50–65.
- Juknevičienė, R. 2011. *Leksinės samplaikos svetimkalbių ir gimtakalbių vartotojų rašytinėje anglų kalboje*. [Les cooccurrences lexicales dans les écrits en anglais L2 et des anglophones natifs]. Thèse de doctorat. Université de Vilnius.
- Kazlauskienė, V. 2018. *Syntagme nominal en FLE : analyse d'un corpus d'apprenants lituaniens*. Thèse de doctorat. Université de Vilnius.
- Legallois, D. 2013. Les greffes phraséologiques – ou quand la syntaxe se compromet. *Langages* 1 (189), 103–120.
- Mejri, S. 2000. Figement et renouvellement du lexique : quand le processus détermine la dynamique du système. *Le Français Moderne* 1, 42–61.
- Mejri, S. 2005. Figement, néologie et renouvellement du lexique. *Linx* [En ligne], 52. URL : <http://journals.openedition.org/linx/231> ; DOI : 10.4000/linx.231
- Mel'cuk, I. 1998. Collocations and lexical functions. *Phraseology. Theory, Analysis and Applications*, 23–53.
- Mel'cuk, I., A. Polguère. 2007. *Lexique actif du français - l'apprentissage du vocabulaire fondé sur 20.000 dérivations sémantiques et collocations du français*. Bruxelles: De Boeck.
- Mel'cuk, I., A. Clas et A. Polguère. 1995. *Introduction à la lexicologie explicative et combinatoire*. Louvain: AUPELF-UREF-Duculot/ Coll. Champs linguistiques.
- Polguère, A. 2000. Towards a theoretically-motivated general public dictionary of semantic derivations and collocations for French. *Proceedings of the 9th EURALEX International Congress*. Dans E. L. C. R. Ulrich Heid, Stefan Evert (dir.). Stuttgart, Germany. 517–527.
- Ramisch, C. 2012. *A generic and open framework for multiword expressions treatment: from acquisition to applications. (Un environnement générique et ouvert pour le traitement des expressions polylexicales)*. Thèse de doctorat. Grenoble Alpes University, France.
- Riegel, M., J.-C. Pellat et R. Rioul. [1994] 2009. *Grammaire méthodique du français*. Paris: Presses Universitaires de France.
- Siepmann, D. 2008. Phraseology in learners' dictionaries. What, where and how? *Phraseology in Language Learning and Teaching*. Granger, Sylviane & Fanny Meunier. (dir.). Amsterdam, Philadelphia: John Benjamins. 185–202.
- Sinclair, J. 1991. *Corpus, Concordance, Collocations*. Oxford: Oxford University Press.
- Tutin, A., F. Grossmann. 2002. Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif. *Revue française de linguistique appliquée* VII-1, 7–25. Université de Grenoble.

Daiktavardinės kolokacijos su būdvardžiu prancūzų K2 rašytinėje kalboje

Vitalija Kazlauskienė

Vilniaus universitetas

Santrauka

Leksinės kompetencijos įsisavinimas – sudėtingas procesas: kad mokiniai vieną ar kitą leksinį vienetą teisingai integruotų į kalbinį kontekstą, jiems nepakanka atskirti jo formos ir reikšmės. Besimokantieji turi savitą leksinių samplaikų inventorių. Paprastai šios samplaikos pasirenkamos pagal tai, kaip dažnai jos vartojamos tikslinėje kalboje ir kiek nusistovėjusi jų vartoseną. Dažnumo aspektas leidžia atskirus žodžių junginių atvejus traktuoti kaip kolokacijas (Ancil, Tremblay 2016). Pasak Tutino ir Grossmanno, „kolokacijoms dažnai būdingas reikšmės skaidrumas recepcijoje, nors negimtakalbiam kalbėtojui sunku produkuoti tinkamas leksemas“ (Tutin, Grossmann 2002). Šiame straipsnyje siekiama patikrinti šio teiginio relevantiškumą ir apibrėžti leksinių samplaikų dėsningumus prancūzų kalbos kaip K2 besimokančių lietuvių

tarpukalbėje. Šiuo tikslu keliamas uždavinys pernelyg neišsiplečiant įvertinti daiktavardinių kolokacijų su būdvardžiu dažnumą ir ypatumus prancūzų kalbos besimokančių negimtakalbių rašto darbuose. Tekstyno analizė leido išskirti tam tikras tiriamų tarpukalbės kolokacijų ypatybes. Tekstyne daiktavardinėms kolokacijoms su būdvardžiu būdinga struktūra (Adj) + N + (Adj): N + Adj tipas yra dažnesnis už Adj + N tipą. Tekstynas taip pat parodė, kad žodžių samplaikos ne visuomet atitinka standartinės prancūzų kalbos žodžių junginius. Interferencijos ir hibridizacijos procesai gali būti laikomi esmine kolokacijų kontaminacija tiek leksiniu, tiek sintaksiniu lygmeniu. Vis dėlto iš dalies sustabarėję posakiai, kitaip tariant, kolokacijos, – svarbi B1 kalbos lygio siekiančių mokinių raiškos priemonė. Šio tyrimo rezultatai rodo, kad kolokacinių konstrukcijų analizė atskleidžia kalbėtojų kompetencijų ir atlikties santykį. Be to, šie duomenys galėtų padėti pamatus tolesniems svetimųjų kalbų mokymosi moksliniams tyrimams.

Raktažodžiai: *kolokacija; žodžių junginys; DJ; tekstynas; prancūzų K2*

Submitted 11 November 2020