

AGNIESZKA DRYJAŃSKA
ORCID : 0000-0003-1649-8408
Université de Varsovie
a.dryjanska@uw.edu.pl

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

VERS UNE (R)ÉVOLUTION DU DÉVELOPPEMENT LEXICAL : INTÉGRATION DE LA COMPÉTENCE COLLOCATIONNELLE EN FLE PAR L'ANALYSE DE CORPUS

1. INTRODUCTION

Dans l'enseignement et l'apprentissage du Français Langue Étrangère (FLE), les collocations semblent indispensables pour des raisons communicationnelles et lexico-sémantiques. Tout d'abord, l'emploi de ces « associations lexicales privilégiées » (Tutin 2012 : 47)¹ facilite la communication en améliorant sa fluidité.

¹ La notion de collocation peut être analysée *grosso modo* dans deux perspectives : plus large, fondée sur un critère statistique (Halliday, Hasan 1976 ; Cruse 1986 ; Sinclair 1991) et plus restreinte, prenant sa source dans les études phraséologiques (Hausmann *et al.* 1989 ; Mel'cuk 1998 ; Tutin, Grossmann 2002 ; Śliwa 2011, Tutin 2012). La définition de la collocation citée ci-dessus correspond à la première perspective, qui est celle que nous adoptons, car elle est plus appropriée à l'analyse de corpus et équivalente à la définition du *Cadre européen commun de référence pour les langues* (CdE 2001 : 88). Nous avons exposé de manière plus détaillée la problématique collocationnelle dans les cadres linguistique et didactique dans nos recherches précédentes (Kazlauskienė 2022, 2023 ; Dryjańska, Kazlauskienė 2022, 2023).

té et son efficacité. Comme le souligne Amanda Edmonds, « une séquence dite formulaire, stockée et extraite en bloc, est appréhendée plus rapidement qu'une séquence similaire mais non phraséologique » (2013 : 125). L'acquisition des collocations en tant que sous-catégorie d'unités lexicales se fait dans le cadre du développement des compétences linguistiques (Conseil de l'Europe [CdE] 2001 : 115), notamment de la compétence lexicale (CdE 2001 : 88). L'intégration des collocations à la compétence sémantique (CdE 2001 : 91) s'explique, en premier lieu, par le fait qu'elles sont considérées comme des unités de sens particulières, se situant entre le préconstruit et le libre. Elles sont prototypiquement transparentes, étant donné la compositionnalité de leur sens, mais leur structure est imprédictible. Il convient par ailleurs de souligner les difficultés définitoires de ces « unités aux contours flous » (Tutin, Grossmann 2002), dues au fait que le critère de compositionnalité du sens n'est pas toujours respecté (Lehmann, Martin-Berthet 1998, 2018). En deuxième lieu, les nuances sémantiques liées à la synonymie ou à la polysémie sont traçables grâce à des combinatoires lexicales dans des contextes particuliers (Henriksen 2013 : 31) selon le principe de John R. Firth (1957) que le sens d'une unité sémantique, quelle que soit sa complexité lexico-syntaxique, est déduit de son co-texte.

La continuité des réflexions théoriques sur l'acquisition des collocations est assurée par de nombreux chercheurs (Benigno *et al.* 2015 ; Cavalla *et al.* 2018 ; Dryjańska, Kazlauskiene 2022, 2023 ; Kazlauskiene 2023, et d'autres). Cependant, l'apprentissage de ces expressions semi-figées et, corollairement, le développement de la compétence collocationnelle, englobant la diversité des aspects sémantico-syntaxiques notés ci-dessus, restent souvent sous-exploités dans les programmes d'enseignement du FLE à tous les niveaux (Martinez 2018 ; Dryjańska 2024). C'est pourquoi, dans la présente étude, nous nous sommes fixées pour objectif d'analyser la place et le traitement des collocations dans une méthode de FLE et de concevoir un scénario basé sur des corpus de textes palliant les insuffisances identifiées et permettant d'exposer les avantages de l'intégration des acquis de la linguistique de corpus (désormais LC) pour faire avancer le développement de la compétence lexicale, dont fait partie la compétence collocationnelle, en didactique des langues et cultures (désormais DDLC).

Il est à noter que le recours à la LC en FLE permet d'intégrer un paradigme méthodologique fondé sur les acquis de la science des données (*data science*) afin de sensibiliser les apprenants à l'existence de régularités lexico-sémantiques révélées par de grandes quantités de données, ainsi que de leur apprendre à interpréter la représentation symbolique de ces régularités dans des corpus de textes. Si l'intégration de la LC en DDLC est déjà bien étudiée et développée en anglais, notamment à travers de nombreuses recherches sur l'apprentissage fondé sur les données (*data-driven learning*) (O'Keeffe 2020), elle demeure moins avancée en FLE (André 2018). Par exemple, les corpus de textes sont rarement utilisés en FLE à l'Université de Varsovie et à l'Université de Vilnius.

Toutefois, cette question suscite un intérêt croissant, et de nouvelles approches sont proposées pour exploiter pleinement son potentiel didactique, y compris celles intégrant une dimension plurilingue et pluriculturelle (Cavalla, Pecman 2020). Notre but est donc de mettre en lumière l'utilité des corpus de textes pour renouveler et améliorer le développement de la compétence lexicale chez les apprenants de FLE en mettant en avant sa composante collocationnelle.

2. LA COMPÉTENCE COLLOCATIONNELLE AU SEIN DE LA COMPÉTENCE LEXICALE

La compétence lexicale telle qu'elle est définie dans le *Cadre européen commun de référence pour les langues*, inclut, à côté des mots isolés, la connaissance et l'emploi de différents types de *structures toutes faites* et de *locutions figées*, structures « constituées de plusieurs mots, apprises et utilisées comme des ensembles » (CdE 2001 : 87), parmi lesquelles des « collocations constituées de mots fréquemment utilisés ensemble », comme *faire un discours* ou *faire une faute* (CdE 2001 : 89).

Étant donné que les collocations sont considérées, au sein de la compétence lexicale, comme des unités lexicales au même titre que les mots isolés, et que leur acquisition suit le même processus que celle des mots isolés (Henriksen 2013), la compétence collocationnelle peut être définie comme la connaissance de ces structures semi-figées et la capacité à les utiliser. La connaissance des collocations est interprétée en tant que capacité à les reconnaître et à comprendre leur sens, leur fonction, mais également à maîtriser les contraintes sémantico-syntactiques qui les régissent (*semantically vs formally motivated collocations*) et leur place dans le système des relations interlexicales d'une langue² (Henriksen 2013 : 39–40). De plus, elle est considérée comme un critère essentiel d'évaluation du niveau des connaissances lexicales, parfois même sociales et culturelles, en ce qu'elle permet de « révéler les rapports entre compétences et performances des apprenants » (Kazlauskiene 2022 : 194).

L'acquisition des collocations par un apprenant d'une langue étrangère est un processus cumulatif, ce qui explique que les résultats de l'apprentissage sont souvent visibles à des niveaux plus élevés de maîtrise d'une langue. Cependant, à notre sens, cela ne devrait pas mener à la conclusion que l'apprentissage des collocations doit être commencé à ces niveaux³, d'autant que c'est un processus plus complexe que l'acquisition des mots isolés, étant donné, entre autres, une exposition plus rare à ces structures due à leur fréquence plus basse que celle

² Notons que cette conception est basée sur le fait que ces structures reposent sur des « liens mentaux associatifs entre les mots dans l'esprit des utilisateurs de la langue » (Henriksen 2013 : 32).

³ Selon Philip Brown (1994 : 25), le développement de la compétence collocationnelle relève d'une stratégie active destinée aux apprenants de haut niveau.

de nombreux mots isolés (Henriksen 2013). Le développement de la compétence collocationnelle devrait occuper une place qui lui est propre dans le développement de la compétence lexicale des apprenants de langues étrangères dès le début de leur apprentissage, notamment dans le cadre des philologies.

3. LE RÔLE DE LA COLLOCATION DANS L'ACQUISITION DU VOCABULAIRE ET SES DIFFICULTÉS POUR LES APPRENANTS

Tout comme d'autres éléments de la langue, les collocations sont utilisées de manière instinctive par les locuteurs natifs et font partie intégrante de leurs compétences linguistiques implicites. Cependant, pour un locuteur non natif en cours d'apprentissage d'une nouvelle langue, leur utilisation est problématique. Les problèmes rencontrés sont variés et apparaissent à la fois lors des phases de réception et de production. Ils sont dus à la complexité et aux propriétés des collocations, qui ont déjà été signalées : par exemple, leur diversité, leur degré de transparence sémantique, leur imprédicibilité syntaxique (Tutin, Grossmann 2002), mais aussi à l'influence de la langue maternelle, à la faible présence des collocations dans les dictionnaires ou les manuels de FLE et au manque de pertinence de ces structures dans la perception des apprenants (Henriksen 2013 : 31).

Dans ce contexte, il est essentiel de sensibiliser le locuteur non natif à la distinction entre les collocations facilement reconnaissables en compréhension (décodage) et celles qui sont réutilisables en production (encodage), car ces catégories ne se recouvrent pas nécessairement. Cette distinction concerne à la fois la forme et le sens : une collocation peut être identifiable dans une langue sans être directement réutilisable dans une autre en raison de l'anisomorphisme (Zgusta 1971 : 296), c'est-à-dire de l'absence d'équivalence totale entre les deux langues. Ainsi, dans une perspective contrastive, une même forme en L1 peut ne pas correspondre exactement à son équivalent en L2, ce qui se manifeste dans le phénomène des faux-amis – sans toutefois s'y limiter.

Ensuite, l'enseignement des collocations permet aux apprenants de comprendre « en quoi les différences combinatoires entre des mots appartenant au même paradigme sémantique, synonymes ou antonymes, reflètent différentes visions des choses auxquelles ces mots se réfèrent » (Blumenthal 2006). En effet, les lexèmes synonymiques *bavarder* et *papoter*, ou encore *groupe*, *troupe* et *équipe*, peuvent paraître similaires, mais leurs utilisations spécifiques dans des collocations particulières mettent en lumière des nuances subtiles de sens qui les différencient. Quant aux équivalences apparentes entre les unités lexicales en français et dans la langue maternelle des apprenants, celles-ci peuvent se caractériser par une divergence collocationnelle due à une non-équivalence sémique entre deux mots (sémèmes) dans les deux langues. Ainsi, les collocations du substantif polysémique *parole* comme *prendre la parole*, *donner la parole* et *adresser la parole*

illustrent bien des nuances sémantiques qui peuvent ne pas avoir d'équivalents collocationnels directs, c'est-à-dire dont la base serait l'équivalent du substantif *parole*, dans la langue maternelle des apprenants.

4. L'ENSEIGNEMENT DU NOUVEAU LEXIQUE

Il apparaît ainsi que l'objectif de l'enseignement lexical du FLE, surtout dans le contexte philologique, est double : rendre la communication plus efficace en fournissant des collocations courantes pour une expression plus naturelle et fluide (Henriksen 2013), et sensibiliser les apprenants à la spécificité des relations interlexicales d'une langue à travers les patrons lexico-syntactiques.

Une méthodologie permettant d'acquérir de nouveaux phénomènes linguistiques devrait suivre la procédure standard d'exercisation définie par Puren (2016), qui se décompose en trois étapes principales : observation, conceptualisation et entraînement.

À la première étape, l'observation, les apprenants sont encouragés à repérer les collocations dans des contextes authentiques. Il est préférable d'utiliser des contextes typiques et prototypiques, adaptés au niveau de langue des apprenants, afin de les familiariser avec l'utilisation réelle des collocations et de faciliter ainsi leur conceptualisation.

La deuxième étape, la conceptualisation, consiste à aider les apprenants à appréhender les concepts désignés par les collocations. Cette phase implique un travail de réflexion, où les apprenants sont amenés à saisir les règles et les logiques sous-jacentes aux collocations, c'est-à-dire à reconnaître des structures syntaxiques, des traits sémantiques et des contextes pragmatiques.

Enfin, la troisième étape, l'entraînement, engage les apprenants dans la pratique active des collocations à travers divers exercices d'application, puis de réemploi, jusqu'à une appropriation leur permettant de les utiliser de manière fluide et appropriée dans différents contextes de communication.

En suivant ces trois étapes, les enseignants peuvent mieux guider les apprenants dans l'acquisition des collocations. Dans la suite, nous allons examiner si cette procédure est respectée, sachant que les insuffisances, surtout aux deux premières étapes, peuvent être palliées par une analyse de corpus de textes fournissant de nombreux exemples de contextes d'emploi.

5. LA PRÉSENTATION DE LA DÉMARCHE EMPIRIQUE

Notre recherche a un double objectif : en premier lieu, nous visons à analyser la place et le traitement des collocations en FLE, sur l'exemple de la méthode *Édito B1*, en examinant le nombre de collocations proposées, la typologie des pa-

trons syntaxiques, leurs caractéristiques sémantiques et la prosodie sémantique. Nous nous intéressons également à la manière dont ces collocations sont enseignées et mises en pratique, en analysant les activités lexicales proposées et la pertinence du métalangage utilisé dans ce contexte. En deuxième lieu, à titre d'exemple, nous élaborerons un scénario reposant sur des corpus de textes pour un travail complémentaire en classe. Ce scénario éclaircira les divergences sémantico-syntaxiques entre des unités lexicales synonymiques présentées sous forme de « listes de mots » dans le manuel *Édito B1*, telles que *bavarder, papoter, blablater*.

La méthode a été choisie pour les raisons suivantes :

- sa modernité : *Édito B1* est un manuel récent, notamment dans sa nouvelle version de 2023, ce qui suggère qu'il intègre des techniques pédagogiques modernes ;
- son approche actionnelle, alignée sur les recommandations actuelles en didactique des langues ;
- son utilisation à l'université ; *Édito B1* est couramment utilisé aux Universités de Varsovie et de Vilnius.

Le niveau B1 a été sélectionné étant donné que c'est à ce niveau que l'on introduit de nombreuses collocations comme *faire une faute* ou *faire un discours* (CdE 2001 : 88), qui sont très courantes dans la communication. De plus, au niveau B1, les apprenants ont « une bonne maîtrise du vocabulaire élémentaire » (CdE 2001 : 89). Par conséquent, le rôle du scénario qui sera présenté est de faciliter la réalisation de cet objectif.

5.1. L'ANALYSE LEXICO-COLLOCATIONNELLE DE LA MÉTHODE ÉDITO B1

Le manuel *Édito B1* introduit approximativement 1800 unités lexicales dans des sections consacrées au vocabulaire, sous la forme de « listes de mots », dans lesquelles 25 % des unités lexicales sont des collocations. Celles-ci ont été classées en fonction de leur structure syntaxique, selon la typologie présentée dans le tableau ci-dessous. Les exercices lexicaux se trouvent également dans ces sections et sur les pages *Vocabulaire* et *Grammaire*, et ils ne fournissent en général qu'un exemple pour chaque unité lexicale.

Nous avons relevé environ 550 collocations répondant à des patrons sémantico-syntaxiques différents. Parmi eux, on observe la prédominance de deux patrons : verbaux, de type VN, qui englobent 210 collocations, et nominaux, suivant le modèle (A)N(A), qui comprennent 130 collocations. La prédominance du patron VN est compréhensible étant donné que le substantif et le verbe sont les catégories grammaticales de base de la construction de la phrase, permettant aux locuteurs d'exprimer et d'échanger des informations essentielles, avec une « précision suffisante » pour le niveau B1 (CdE 2001 : 101), sans transmettre toutes les nuances de sens fines requises au niveau C2 (CdE 2001 : 101).

Tableau 1. Typologie des collocations (les collocations opaques sont en italiques)

typologie des collocations (~540/510) (Édito, B1, 2023)				
V+N/SN (~210)	V+Prép+N (~60)	N+A (~130)	N1+de+N2 (~80)	(N1)+à/sur/ en+N2 (~20)
attendre le signal (de départ sur la piste de course)	arriver en tête (des ventes)	vagabondage mental	charge de travail	mise en action
avoir / ne pas avoir le droit (de)	contribuer à la vie sociale	haute résolution	amis d'enfance	personne à part entière
avoir des facilités (en langues étran- gères)	décider dans un cadre collectif	reconnaissance visuelle	travail de terrain	paiement en ligne
avoir du mal (à gé- rer le stress)	être à l'écoute	imaginaire col- lectif	baisse de concen- tration	droit à la décon- nection
avoir le sens (de l'initiative, de la créativité)	entrer en interaction	société inclusive	capacité d'adaptation	mise en scène
avoir le sourire	mettre en relation	expérience pro- fessionnelle	capacité d'analyse	communication sur Internet
avoir un (petit) mo- ment (de détente)	éclater de rire	activité profes- sionnelle	changement de carrière	film à gros budget
avoir une période (d'essai de)	mettre dans des cases	contrainte profes- sionnelle		trajet en corres- pondance
battre les records (de vente)	naitre avec une malformation	effort physique		la garde à vue
		dépense physique		

La pertinence du schéma collocationnel (A)N(A) pour le niveau B1 s'explique par le fait qu'il introduit un nombre important d'adjectifs, augmentant ainsi la précision des énoncés produits par les apprenants. L'analyse de ce type de collocations leur permet de comprendre que l'emploi d'adjectifs comme *collectif*, *mental* ou *inclusif* est soumis à des contraintes syntaxiques parfois plus fortes que celles imposées aux adjectifs très fréquents introduits aux niveaux de maîtrise de la langue A1 ou A2, souvent utilisés dans des constructions libres.

Nous avons également relevé 160 schémas collocationnels – verbaux et nominaux – plus complexes dont le collocatif est constitué d'un syntagme prépositionnel composé d'un groupe nominal précédé, le plus souvent, des prépositions *à*, *de*, *sur* et *en*. Le patron collocationnel N1deN2 s'est avéré particulièrement productif dans notre analyse (80 exemples). Cela est important, car le choix entre l'utilisation (ex : *sens de l'initiative*) ou la non-utilisation d'un déterminant (ex : *capacité d'analyse*) (N1deDétN2 vs N1deN2) dans un complément du nom est souvent problématique pour les apprenants, surtout ceux dont la langue maternelle est flexionnelle.

Les verbes présents dans les syntagmes verbaux analysés sont premièrement des verbes très fréquents : des verbes supports comme *avoir* (*avoir du mal à faire*

quelque chose), faire (faire face à quelque chose), être (être à l'écoute) ou prendre (prendre en charge) et des verbes ayant davantage de valeur sémantique, tels que : *mettre, sentir, entrer, arriver*. Ces verbes, certes, font déjà partie du répertoire lexical au niveau de maîtrise du français A2, mais les apprenants ne connaissent pas les schémas lexicaux fréquents dans lesquels ils sont utilisés, comme *se mettre en valeur, entrer par effraction, arriver en tête*. Dans l'apprentissage lexical, il ne s'agit pas uniquement d'augmenter la quantité et l'étendue du vocabulaire, mais également sa profondeur. Ainsi, en apprenant la combinatoire lexicale, les apprenants approfondissent⁴ leur connaissance des mots. De plus, les verbes mentionnés, fréquemment utilisés dans la communication, sont également très utiles au niveau d'apprentissage visé.

5.2. LES CARACTÉRISTIQUES SÉMANTIQUES DES COLLOCATIONS ANALYSÉES

Une autre particularité de certaines collocations – leur transparence sémantique (Tutin, Grossmann 2002) – a été observée dans la plupart des collocations relevées. Cette caractéristique facilite la réception, cependant, elle ne rend pas la collocation prédictible, ce qui fait que son emploi est toujours problématique pour un apprenant (Tutin, Grossmann 2002 ; Tutin 2012). De plus, la transparence et la compositionnalité de sens peuvent compliquer l'apprentissage des collocations puisque ces structures ne sont pas considérées comme des unités de sens qui doivent être « apprises comme un tout, une unité » (Xiao, Girani 2018).

Parmi les constructions recensées, on en distingue également quelques-unes qui sont sémantiquement opaques⁵. Ces collocations sont « proches des expressions perçues comme figées, mais le sens de la base reste interprétable » (Tutin, Grossmann 2002). Pour l'apprenant, les constructions comme *charge de travail, société inclusive* ou *haute résolution*, sont, d'un côté, plus difficiles à comprendre, mais de l'autre, plus susceptibles de fonctionner comme des unités de sens distinctes, ce qui peut constituer une motivation à les apprendre.

Les observations ci-dessus mettent en évidence la complexité du phénomène collocationnel dans le manuel analysé, qui devrait toutefois être valorisée par le biais d'activités pédagogiques variées et d'une réflexion métalinguistique approfondie pour améliorer la compétence lexicale des apprenants. Les objectifs de cet enseignement seraient les suivants :

⁴ Dans l'apprentissage lexical, on parle des trois critères : la quantité (*size*) et l'étendue (*range*) du vocabulaire, à savoir le nombre de mots et de domaines thématiques dans lesquels ils sont utilisés (CdE 2001), et la profondeur (*depth*), désignant le degré de connaissance d'un mot et incluant des éléments tels que, entre autres, le sens, la fréquence ou les traits syntaxiques et morphologiques (Schmitt 2014 ; Hasan, Shabdin 2016).

⁵ Pour être plus précis, l'opacité est un trait du collocatif. Notons également le statut particulier des collocations opaques qui sont « proches des expressions perçues comme figées, mais [dont] le sens de la base reste interprétable » (Tutin, Grossmann 2002).

- renforcer la compréhension du sens des collocations opaques ;
- proposer des activités permettant l’application, l’entrainement à l’utilisation des collocations et leur réemploi ;
- renforcer la conceptualisation des collocations et les distinguer des constructions libres.

Afin de les atteindre, nous proposons de recourir aux corpus de textes, qui fournissent de nombreux contextes d’emploi regorgeant de différents patrons lexicaux. Cela peut se faire en mode *hands-off*, moins chronophage, ou *hands-on*, favorisant le développement de l’autonomie de réflexion et de l’esprit critique. Dans le premier cas, c’est l’enseignant qui explore des corpus de textes pour y sélectionner des exemples qui seront ensuite analysés par les apprenants afin d’en tirer des conclusions propres à leurs besoins ou à leurs intérêts. Si les exemples fournis par l’enseignant ne dissipent pas tous les doutes des apprenants, il est recommandé de proposer une activité *hands-on* afin que les apprenants puissent se familiariser avec les fonctionnalités des corpus de textes et rechercher par eux-mêmes des exemples appropriés.

6. UNE PROPOSITION DE SCÉNARIO

Dans la méthode *Édito B1*, parmi les mots isolés, on distingue des synonymes comme les verbes *bavarder*, *papoter*, *blablater*, qui ne sont pas présentés ni « exercisés » de manière à assurer leur compréhension ni leur emploi correct.

Selon la procédure d’exercisation de Puren (2016) déjà évoquée, nous avons identifié les étapes suivantes de l’enseignement des verbes *bavarder* et *papoter* (le verbe *blablater* ne figure pas dans les textes) :

– Observation – repérage des verbes *bavarder*, *papoter* dans des textes du manuel :

- (1) Il n’est plus question d’un passage en caisse mais de souffler ou **bavarder** avant de reprendre sa route. [*Édito B1* : 34]
- (2) Ça m’énerve qu’on doive en arriver là pour pouvoir **papoter** un peu avec les caissières. [*Édito B1* : 32]

Ces deux contextes ne permettent pas de distinguer la différence sémantique entre ces verbes.

– Conceptualisation – une meilleure compréhension du sens de ces verbes se fait à travers un même exercice dans lequel une définition des trois verbes est proposée :

- (3) Parler de différentes choses, souvent sans importance. [*Édito B1* : 36]

– Entrainement – à cette étape, on propose un exercice comprenant quelques phrases dans lesquelles on peut utiliser le verbe *papoter*, par exemple :

- (4) Dans ce monde déshumanisé, c’est important de créer du lien / **papoter** / faire connaissance dès que possible. [*Édito B1* : 37]

La démarche présentée ci-dessus, caractéristique de ce manuel et citée à titre d'exemple, nous paraît insuffisante, parce qu'elle ne permet pas d'éclairer les enjeux communicationnels des lexèmes synonymiques. Afin de pallier ces lacunes, nous proposons un scénario à deux volets : consultation des dictionnaires et analyse de corpus de textes afin d'analyser les aspects suivants :

– la différence de fréquence d'occurrence des lexèmes synonymiques, à savoir la corrélation de la fréquence et de la pertinence communicationnelle ;

– les collocations les plus caractéristiques des synonymes analysés et notamment :

- la différence concernant les collocatifs nominaux en position de sujet des verbes *bavarder*, *papoter* et *blablater* ;
- les collocatifs adverbiaux les plus fréquents révélant la prosodie sémantique⁶.

Dans le cas des verbes *bavarder*, *papoter* et *blablater*, l'analyse de trois dictionnaires unilingues : *Le Robert*, *Le Larousse*, *Le Trésor de la langue française informatisé* [TLFi]) et d'un dictionnaire bilingue (Dobrzyński *et al.* 2005, 158, 222) n'a pas permis d'éclaircir les différences sémantiques entre les trois verbes. Le dictionnaire bilingue propose les mêmes équivalents polonais *gadać*, *paplać* pour les verbes *bavarder* et *papoter*. À titre d'exemple, dans *Le Robert*, nous trouvons les définitions suivantes :

- (5) *Bavarder* – Parler beaucoup, de choses et d'autres.
- (6) *Papoter* – Parler beaucoup en disant des choses insignifiantes. *Bavarder*.
- (7) *Blablater* – Faire des blablas ; tenir des propos sans intérêt.

Vu l'insuffisance des ressources fournies par les textes et exercices du manuel *Édito B1* ainsi que par les dictionnaires consultés, nous passons à une exploration des corpus de textes suivants : Sketch Engine frTenTen20 [SE] et la Leipzig Corpora Collection [LCC] French mixed 2012. Les corpus comptent respectivement 15 115 914 647 et 1 468 766 604 mots.

Les deux corpus de textes analysés fournissent les données de fréquence d'occurrence suivantes :

Tableau 2. Fréquence d'occurrence dans les corpus analysés

MOT	LCC French (classe de fréquence) ⁷	frTenTen20 SE fréquence d'occurrence
bavarder	17	25 631
papoter	18	25 440
blablater	20	2 139

⁶ La prosodie sémantique se réfère au sens attitudinal d'un élément lexical façonné par ces collocatifs (Cheng 2013).

⁷ « Les mots de fréquence similaire sont regroupés en classes, l'objectif étant que la classe de fréquence d'un mot ne change que rarement d'un corpus à l'autre. Le mot le plus fréquent d'un corpus a toujours la classe de fréquence 0 [...]. Les mots extrêmement rares peuvent avoir une classe de fréquence de 20 ou plus dans les grands corpus » (LCC).

L'analyse des fréquences permet de constater que le verbe *blablater*, dont la fréquence est plusieurs fois plus basse que celle de ses synonymes et dont la classe de fréquence est de 20, est très rare ; par conséquent, ce verbe est beaucoup moins pertinent pour la communication, notamment au niveau B1. C'est le verbe *bavarder* qui est le plus couramment utilisé dans la LLC. Les cooccurrences les plus importantes des verbes *bavarder* et *papoter* sont représentées dans la LCC French par les graphes⁸ suivants (LCC French) :

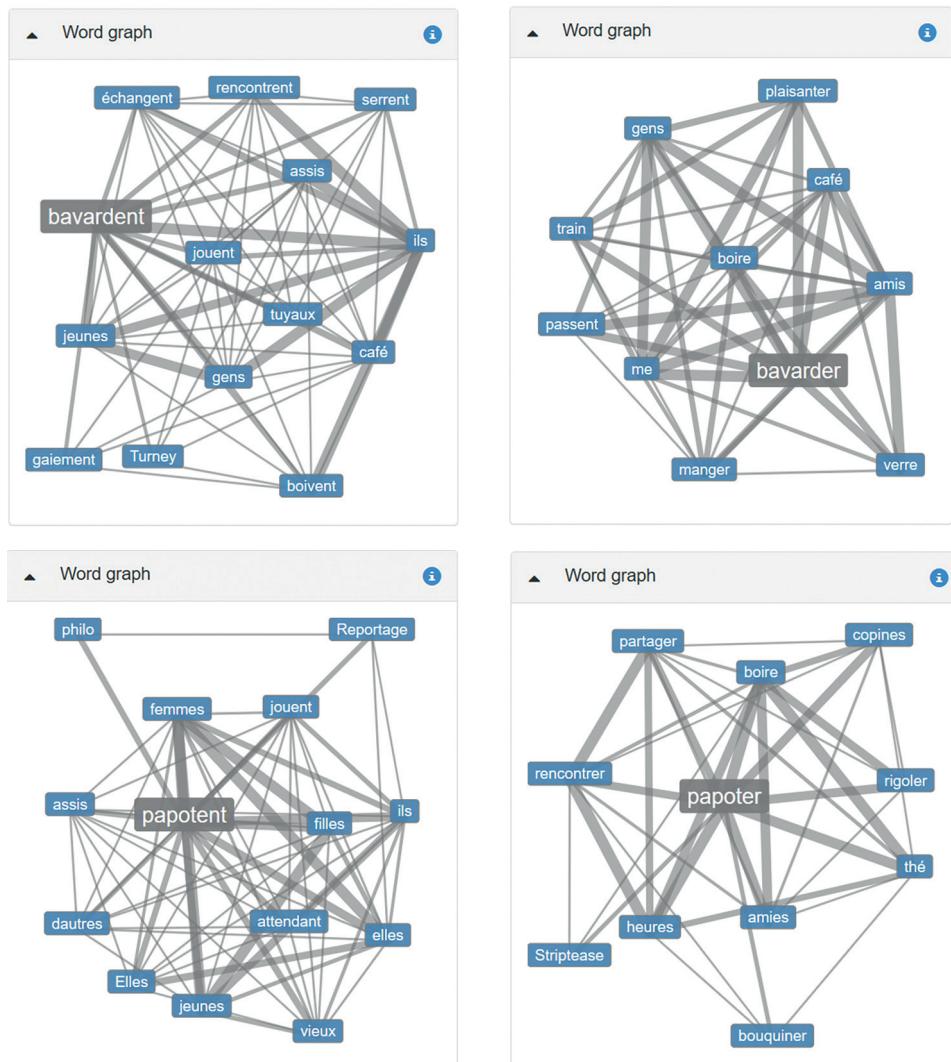


Figure 1. Graphes représentant les cooccurrences principales des mots analysés dans la LCC French

⁸ Les mots cooccurrents les plus significatifs d'un mot d'entrée sont représentés dans des graphiques. La force du lien rattachant les mots cooccurrents est représentée par la largeur de la ligne (LCC).

Tout d'abord, le verbe *papoter* est plus couramment utilisé avec les substantifs féminins : *copines*, *amies*, *filles* et *femmes* et le pronom personnel *elles*, ce qui n'est pas le cas du verbe *bavarder*, ce dernier étant employé dans des contextes moins marqués en genre, où interviennent, par exemple, les mots *amis* et *gens*.

Ensuite, le collocatif adverbial *gaiement* et les cooccurrences avec *verre*, *manger*, *boire*, *café*, suggèrent que le bavardage peut être défini comme une activité permettant à différentes personnes de passer agréablement du temps en mangeant ensemble ou en prenant un verre ou un café. Étant donné que les cooccurrences mentionnées sont statistiquement significatives, on pourrait supposer que le verbe *bavarder* est souvent associé au contexte gastronomique, ce qui n'est pas visible dans le manuel.

subjects of "bavarder/papoter"		
Pie	13	0
pie	16	0
singe	5	0
élève	51	0
classe	7	0
couple	14	0
voisin	6	0
oiseau	6	0
gens	66	0
ami	17	0
homme	73	0
deux	16	0
garçon	6	0
étudiant	8	0
femme	122	79
adulte	11	11
fille	42	56
copine	5	10
dame	12	27
maman	8	25
mec	0	5
parent	0	12
monde	0	49

Figure 2. La fréquence d'occurrence des collocations formées par les mots *bavarder* et *papoter*.
Source: Word Sketch Difference

L'étape suivante est l'analyse des collocatifs nominaux en antéposition (en position de sujet) des verbes *bavarder* et *papoter* à l'aide Word Sketch Difference dans Sketch Engine (figures 2 et 3), fonctionnalité conçue pour comparer deux mots en analysant leurs collocatifs et particulièrement utile pour analyser des différences subtiles de sens, d'usage, de connotations ou de registres entre deux mots. Les couleurs et les ombrages indiquent la préférence des collocatifs pour chacun des deux mots. Les chiffres présentés dans la figure 2 indiquent la fréquence d'occurrence des collocations formées par les mots analysés avec les collocatifs énumérés dans ce tableau. Ces relations peuvent être visualisées d'une autre manière, comme le montre la figure 2. La fréquence de chacune des collocations est représentée par la taille du cercle autour de chaque collocatif. La position horizontale dans cette représentation indique la force de la collocation, en termes de score LogDice, par rapport aux lemmes analysés : à droite (rouge), *papoter*, et à gauche (vert), *bavarder*.

Cette étude élargit l'inventaire des collocatifs nominaux en fonction sujet du verbe *bavarder* par rapport aux résultats que nous avons relevés dans la LCC French. Parmi eux, figurent également les lexèmes *élèves*, *classe*, *étudiants*, qui indiquent l'importance du contexte éducatif associé à ce verbe. Ensuite, nous notons la présence de collocatifs animaliers comme *pies*, *singes* et *oiseaux*, différence majeure concernant l'emploi des verbes *bavarder* et *papoter* (exemples 1–4).

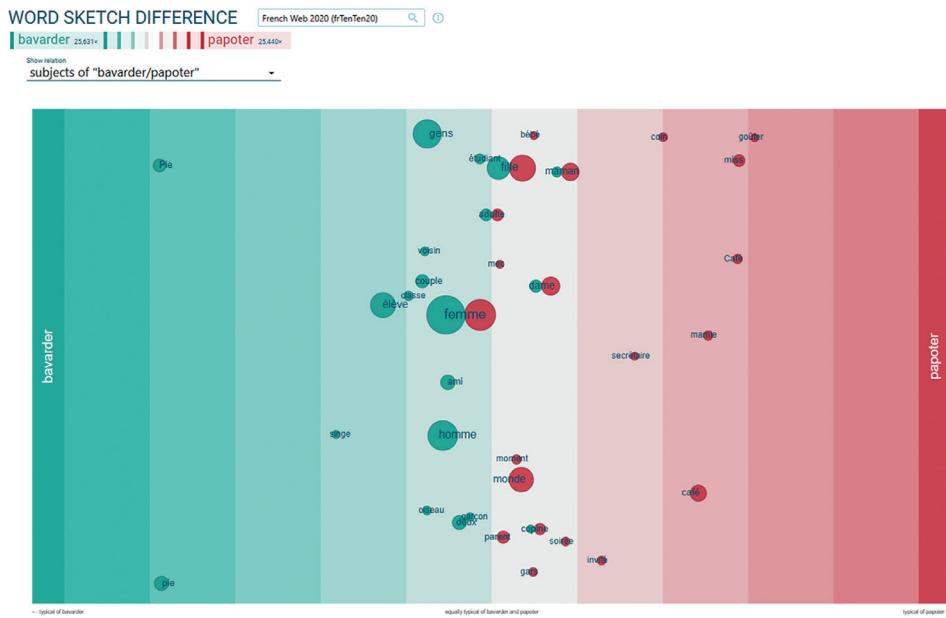


Figure 3. Visualisation d'occurrence des collocations formées par les mots *bavarder* et *papoter*.

Source: Word Sketch Difference

L'analyse permet également l'identification de quelques ambiguïtés liées à l'homophonie et au sens métaphorique, phénomènes assez fréquents en français et problématiques pour les apprenants. À titre d'exemple, observons les phrases suivantes :

(8) Seuls sont concernés le corbeau freux, la corneille noire, la **pie bavarde** jusqu'au 10 juin. [LCC]

(9) Les **pies bavardent** et jacassent dans les bois près du lac de Champfleuri. [LCC]

(10) À cet égard, une femme d'intellectuel a une mission que peut-être il est bon de signaler : si souvent elle l'oublie et, au lieu d'être la Béatrix, ne sait être que la **perruche bavarde** et dissipatrice ! [LCC]

(11) Les **oiseaux bavardent** dans les feuillages. [LCC]

Les exemples cités indiquent très clairement aux apprenants que le contexte est crucial pour la construction du sens. En effet, il permet de distinguer la forme verbale *bavardent* dans les phrases 2 et 4 de l'adjectif homophonique *bavarde* dans le mot composé *pie bavarde*, désignant une espèce de corvidés, mais également dans la cooccurrence *perruche bavarde*, souvent à sens métaphorique.

Ce qui est visible dans la figure 2, c'est que le papotage n'est pas un domaine masculin : les hommes et les garçons ne papotent pas. Ce sont surtout les femmes, filles, copines, dames et mamans qui papotent, ce qui n'exclut pas qu'elles puissent

s’adonner également au bavardage. Notons aussi la collocation intéressante *tout le monde papote*. Ces observations se complètent par des exemples comme ceux-ci.

(12) Pendant que les plus petits faisaient connaissance, les **mamans papotaient** devant un thé. [LCC]

(13) Dans une salle de sport, deux **copines papotent** en boxant dans un sac de frappe. [LCC]

(14) Les retrouvailles sont touchantes, les deux **copines bavardent**, se confient, se marrent, mangent de la colle en pots avec du sucre. [LCC]

(15) Ici les **hommes bavardent** et flânerent autour d’une fontaine publique tout comme ils le feraient à Istanbul ou à Damas. [LCC]

(16) Finalement, y a du monde près des jus de fruits, **tout le monde papote** mais personne n’achète. [LCC]

En analysant ensuite les collocatifs adverbiaux des deux verbes à l’aide des visualisations présentées ci-dessous⁹, nous concluons qu’ils confirment les résultats de la LCC French, à savoir la spécificité du verbe *bavarder*, dont la collocation la plus forte, *bavarder gaiement*, se caractérise par une connotation positive.

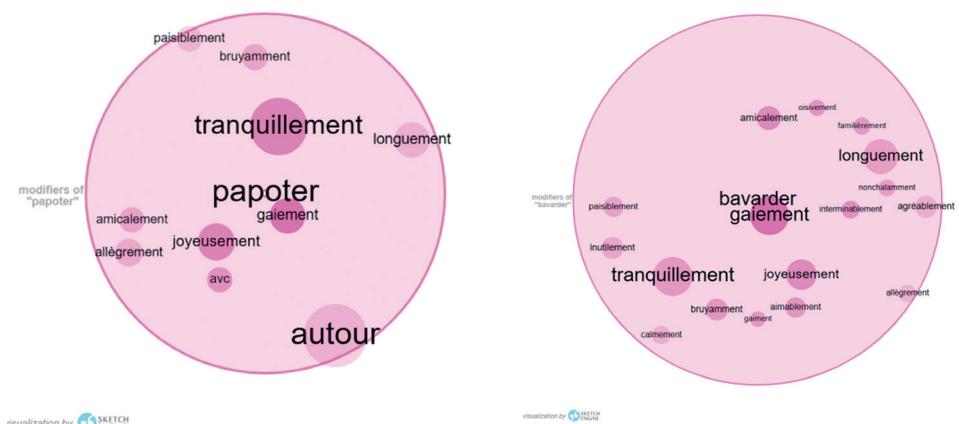


Figure 4. Visualisation des principaux modificateurs adverbiaux des verbes *papoter* et *bavarder*. Source : SE

Cependant, parmi ces collocatifs, l’on distingue également deux adverbes à connotation péjorative, *inutillement* et *oisivement*. La prosodie sémantique du verbe *bavarder* est toutefois dans l’ensemble positive, puisqu’on ne relève que deux cas opposés.

⁹ Les paramètres de cette visualisation sont la taille du cercle près de chaque adverbe et la distance du centre qui correspondent respectivement à la fréquence et à la typicité d’une cooccurrence (force de la collocation).

Même si les observations que nous venons de présenter sont fondées sur les collocations les plus significatives, vu la pertinence didactique de ce type d'analyse, elles permettent de sensibiliser les apprenants à la spécificité du système linguistique français et à ses nuances sémantico-connotatives. Précisons également que ce scénario peut être réalisé en mode *hands-off* ou *hands-on*.

CONCLUSIONS ET RECOMMANDATIONS

L'objectif de cette étude était d'analyser la place et le traitement des collocations dans la méthode *Édito B1* en FLE et de proposer une approche didactique complémentaire fondée sur l'analyse de corpus. L'analyse a révélé que la présentation des collocations dans le manuel repose principalement sur des listes de mots, sans mise en contexte suffisante ni exploitation approfondie de leur valeur sémantique et syntaxique. Cette approche limite le développement de la compétence collocationnelle des apprenants et ne leur permet pas d'appréhender pleinement les régularités lexico-syntaxiques propres au français.

Afin de remédier à ces insuffisances, un scénario didactique, basé sur l'exploitation de corpus de textes et adoptant une approche à la fois inductive et contextualisée des collocations, a été présenté. Cette méthodologie vise non seulement à renforcer la compréhension des relations lexicales et des particularités combinatoires, mais aussi à promouvoir l'autonomie des apprenants en leur fournissant des outils analytiques pour relever et interpréter les données linguistiques.

Même si l'utilisation de la linguistique de corpus en didactique des langues n'est pas récente, notre contribution met en lumière l'intérêt d'une analyse critique d'une méthode moderne intégrant des collocations et la nécessité d'un complément méthodologique fondé sur les corpus. Ce travail s'inscrit ainsi dans une perspective plus large de renouvellement des approches d'enseignement du lexique en FLE intégrant des outils issus du Big Data pour améliorer l'efficacité du développement de la compétence lexicale, notamment en termes de profondeur lexicale.

SOURCES DES EXEMPLES

[Édito B1] Heu-Boulhat É. (2023). *Édito B1 méthode de français*. Lavis : Didier FLE.

[LCC] Leipzig Corpora Collection (French), URL : Leipzig Corpora Collection - French (uni-leipzig.de) (consulté le 26/09/2024).

[SE] Sketch Engine, URL : <https://www.sketchengine.eu/> (consulté le 26/09/2024).

BIBLIOGRAPHIE

DICTIONNAIRES

- Dobrzyński J. *et al.* (2005). *Wielki słownik francusko-polski*. Warszawa : Wiedza Powszechna.
- Le Larousse, URL : Larousse.fr : *encyclopédie et dictionnaires gratuits en ligne* (consulté le 26/09/2024).
- Le Robert, URL : <https://dictionnaire.lerobert.com/> (consulté le 26/09/2024).
- Le Trésor de la langue française informatisé en ligne [TLFi], URL : <https://www.atilf.fr/> (consulté le 26/09/2024).

LIVRES, ARTICLES ET AUTRES

- André V. (2018). « Nouvelles actions didactiques : faire de la sociolinguistique de corpus pour enseigner et apprendre à interagir en français langue étrangère », *Action Didactique* 1, 71–88.
- Benigno V. *et al.* (2015). « Les collocations fondamentales : une piste pour l'apprentissage lexical », *Revue française de linguistique appliquée* 20(1), 81–96.
- Blumenthal P. (2006). « De la logique des mots à l'analyse de la synonymie », *Langue française* 150, 14–31.
- Brown Ph. (1994). « Lexical Collocation : a Strategy for advanced learners », *Modern English Teacher* 3(2), 24–27.
- Cavalla C. *et al.* (2018). « Un corpus pour l'enseignement des collocations en FLE. Études en didactique des langues », *Les corpus / Corpora* 31, 93–107.
- Cavalla Ch., Pecman M. (2020). « Enseignement des expressions préfabriquées », *Action Didactique* 6.
- Cheng W. (2013). « Semantic Prosody », [dans :] C. A. Chapelle (dir.), *The encyclopedia of applied linguistics*, 1–7. Oxford : Wiley-Blackwell Publishing Ltd.
- Conseil de l'Europe [CdE] (2001). *Cadre européen commun de référence pour les langues : apprendre, enseigner, évaluer*. Strasbourg : Didier.
- Cruse A. (1986). *Lexical Semantics*. Cambridge : Cambridge University Press.
- Dryjańska A. (2024). « Elementy językowego obrazu świata w nauczaniu języka francuskiego w kontekście filologicznym », *Neofilolog* 62(2), 409–426.
- Dryjańska A., Kazlauskienė V. (2023). « Jour, Journée au sens de ‘fête’ et leurs équivalents collocationnels en polonais et lituanien : défis pour le FLE », *Białostockie Archiwum Językowe* 23, 101–128.
- Edmonds A. (2013). « Une approche psycholinguistique des phénomènes phraséologiques : le cas des expressions conventionnelles », *Langages* 189, 121–138.
- Firth J. R. (1957). *Studies in linguistic analysis*. Oxford : Basil Blackwell.
- Halliday M. A. K., Hasan R. (1976). *Cohesion in English*. London : Longman.
- Hasan K., Shabdin A. A. (2016). « Conceptualization of Depth of Vocabulary Knowledge with Academic Reading Comprehension », *PASAA* 51, 235–268.
- Hausmann F. J. (1989). « Le dictionnaire de collocations », [dans :] F. J. Hausmann *et al.* (dir.), *Wörterbücher : ein internationales Handbuch zur Lexicographie. Dictionaries*. Berlin–New-York : De Gruyter, 1010–1019.
- Henriksen B. (2013). « Research on L2 learners' collocational competence and development – a progress report », [dans :] C. Bardel, C. Lindquist, B. Laufer (dir.), *L2 Vocabulary knowledge and use. New perspectives on assessment and corpus analysis*. [S. l.] : EuroSla, 29–59.
- Kazlauskienė V. (2022). « Nature et acquisition des collocations lexicales en FLE », *Darnioji daugiaikalbystė* 20, 194–209.
- Kazlauskienė V. (2023). « Combinatoire collocationnelle et linguistique de corpus dans l'apprentissage lexical du FLE », *Academic Journal of Modern Philology* 20, 95–106.

- Kazlauskienė V., Dryjańska A. (2022). « Le sens de *fête* en polonais, en lituanien, en français et sa (non)coïncidence collocationnelle », *Annales Universitatis Paedagogicae Cracoviensis. Studia Linguistica* 17, 20–42.
- Lehmann A., Martin-Berthet F. (1998). *Introduction à la lexicologie : sémantique et morphologie*. Paris : Dunod.
- Lehmann A., Martin-Berthet F. (2018). *Lexicologie. Sémantique, morphologie et lexicographie*. Paris : Armand Colin.
- Martinez Ch. (2018). « Comment faciliter l'apprentissage des collocations aux apprenants non natifs. Une approche didactique et linguistique », *Synergies pays riverains de la Baltique* 12, 35–45.
- Mel'čuk I. (1998). « Collocations and Lexical Functions. Phraseology », [dans :] A. P. Cowie (dir.), *Theory, Analyses and Applications*. Oxford : Clarendon Press, 23–53.
- O'Keeffe A. (2020). « Data-driven learning – A call for a broader research gaze. », *Language Teaching* 54(2), 1–14.
- Puren Ch. (2016). « La procédure standard d'exercisation en langue », URL : www.christianpuren.com/mes-travaux/2016c/ (consulté le 1/08/2024).
- Schmitt N. (2014). « Size and Depth of Vocabulary Knowledge? What the Research Shows », *Language Learning* 64, 913–951.
- Sinclair J. (1991). *Corpus, Concordance, Collocations*. Oxford : Oxford University Press.
- Śliwa D. (2011). « Cœur et sentiments : une étude cognitive des collocations françaises et polonoises », [dans :] A. Krzyżanowska, R. Jakubczuk (dir.), *Parler des émotions : entre langue et littérature*. Lublin : Wyd. UMCS, 191–199.
- Tutin A. (2012). « Les collocations lexicales : une relation essentiellement binaire définie par la relation prédicat-argument », *Langages* 189, 47–63.
- Tutin A., Grossmann F. (2002). « Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif », *Revue française de linguistique appliquée* 1(7), 7–25.
- Xiao Ch., Girani F. (2018). « Typologie d'erreurs des collocations dans les essais argumentatifs et perspectives didactiques de l'enseignement du français langue étrangère dans les universités », *Synergies Chine* 13, 229–246.
- Zgusta L. (1971). *Manual of Lexicography*. Prague : Academia ; The Hague–Paris : Mouton.

TOWARDS A (R)EVOLUTION OF LEXICAL DEVELOPMENT: INTEGRATING COLLOCATIONAL COMPETENCE IN TEACHING FRENCH AS A FOREIGN LANGUAGE THROUGH CORPUS ANALYSIS

Abstract

The study has a dual objective. Firstly, it aims to analyse the place and pedagogical treatment of collocations in French language textbooks, specifically using the *Édito B1* textbook as a case study. The findings indicate that the current pedagogical approach is insufficient, primarily due to the lack of pertinent usage examples that would facilitate learners' comprehension of the collocation phenomenon and its proper application. Consequently, the second objective is to leverage corpus linguistics to propose a didactic scenario that enhances lexical acquisition. Within this framework, both educators and learners engage with extensive textual data to extract lexico-syntactic patterns, thereby identifying the most relevant collocations and their contextual usage.

Keywords: lexical competence, collocation, FLE, didactic applications.

Mots-clés : compétence lexicale, collocation, FLE, applications pédagogiques.