



**HAL**  
open science

## Analyse des erreurs de collocations dans les écrits universitaires : étude comparative chez quelques étudiants allophones et francophones

Azadeh Pirooz

### ► To cite this version:

Azadeh Pirooz. Analyse des erreurs de collocations dans les écrits universitaires : étude comparative chez quelques étudiants allophones et francophones. *ILCEA: Revue de l'Institut des langues et cultures d'Europe, Amérique, Afrique, Asie et Australie*, 2025, 57, 10.4000/137v1. hal-04927914

**HAL Id: hal-04927914**

**<https://hal.science/hal-04927914v1>**

Submitted on 4 Feb 2025

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



**ILCEA**

Revue de l'Institut des langues et cultures d'Europe,  
Amérique, Afrique, Asie et Australie

**57 | 2025**

**Langues de spécialité, corpus et genres : réflexions  
méthodologiques, caractérisation et transposition  
didactique**

---

## Analyse des erreurs de collocations dans les écrits universitaires : étude comparative chez quelques étudiants allophones et francophones

*Analysis of Collocation's Errors in Academic Writing: A Comparative Study of  
Some Allophone and Francophone Students*

**Azadeh Pirooz**

---



### Édition électronique

URL : <https://journals.openedition.org/ilcea/21738>

DOI : 10.4000/137vl

ISSN : 2101-0609

### Éditeur

UGA Éditions/Université Grenoble Alpes

### Édition imprimée

ISBN : 978-2-37747-527-8

ISSN : 1639-6073

### Référence électronique

Azadeh Pirooz, « Analyse des erreurs de collocations dans les écrits universitaires : étude comparative chez quelques étudiants allophones et francophones », *ILCEA* [En ligne], 57 | 2025, mis en ligne le 31 janvier 2025, consulté le 31 janvier 2025. URL : <http://journals.openedition.org/ilcea/21738> ; DOI : <https://doi.org/10.4000/137vl>

---

Ce document a été généré automatiquement le 31 janvier 2025.



Le texte seul est utilisable sous licence CC BY-SA 4.0. Les autres éléments (illustrations, fichiers annexes importés) sont « Tous droits réservés », sauf mention contraire.

---

# Analyse des erreurs de collocations dans les écrits universitaires : étude comparative chez quelques étudiants allophones et francophones

*Analysis of Collocation's Errors in Academic Writing: A Comparative Study of Some Allophone and Francophone Students*

Azadeh Pirooz

---

## 1. Introduction

- 1 Les compétences linguistiques sont essentielles à la réussite académique des étudiants, en particulier dans la production d'écrits universitaires de qualité. Parmi les différents aspects de la compétence linguistique, la compétence lexicale revêt une importance particulière, notamment en ce qui concerne les collocations, qui sont des cooccurrences privilégiées de deux unités lexicales (la base et le collocatif) liées par une relation syntaxique. Les collocations jouent un rôle important dans la fluidité, la clarté et la cohérence du discours écrit, et leur utilisation correcte est un indicateur clé de la maîtrise de la langue.
- 2 Les recherches sur l'acquisition des collocations en particulier en L2 confirment que les collocations représentent des freins à l'apprentissage surtout en productions, quel que soit le niveau d'apprentissage (Cavalla, 2008 ; González-Rey, 2015) et que les apprenants même au niveau avancé en L2 produisent moins de collocations par rapport aux natifs (Forsberg, 2006 ; Nesselhauf, 2005). Afin de mieux utiliser les collocations, tant le locuteur natif que l'apprenant d'une langue doivent acquérir une connaissance approfondie des collocations. Les travaux portant sur l'évaluation et la maîtrise des

compétences collocationnelles révèlent également la véritable difficulté lexicale des étudiants dans leurs écrits universitaires (Boch & Rinck, 2015 ; Cavalla, 2008).

- 3 Dans cette étude de corpus, nous nous intéressons aux collocations transdisciplinaires, telles qu'*effectuer une analyse*. Ces unités lexicales ne sont pas spécifiques à une discipline (il ne s'agit pas de terminologie), mais peuvent être utilisées dans le discours scientifique (Jacques & Tutin, 2018). Elles renvoient, selon Jacques et Tutin, « non seulement aux procédures, démarches, objets scientifiques, mais aussi aux éléments d'argumentation, d'évaluation et de structuration du discours » (2018 : 6).
- 4 Notre objectif est double : d'une part, étudier les traits phraséologiques qui caractérisent l'interlangue des non-natifs, et d'autre part, analyser les difficultés liées à la combinatoire lexicale des étudiants non natifs dans leur utilisation des collocations transdisciplinaires, en les comparant avec les productions des étudiants natifs. Nous étudions les associations de type verbe-nom, qui sont plus fréquentes dans les productions, mais représentent également un obstacle majeur dans la production des non-natifs (Nesselhauf, 2005). Nous observons ce type d'association dans les productions écrites des étudiants, car « le décodage de la collocation ne pose pas de problème. Les problèmes se situent au niveau de l'encodage » (Hausmann, 1997 : 282).
- 5 Cet article explore les collocations, en particulier les collocations transdisciplinaires dans les écrits universitaires. Nous réservons ici le terme « écrit universitaire » aux travaux produits dans le cadre du cursus des étudiants, à l'exclusion des mémoires et des thèses. Ces derniers sont considérés comme des écrits de recherche, car ils représentent une activité visant à construire et diffuser des connaissances, en s'appuyant sur une exposition raisonnée et une discussion de théories, conformément aux normes de la communauté scientifique et légitimée par un cadre établi (Pollet, 2014). Nous mettons en lumière l'importance des collocations dans les écrits universitaires ainsi que les défis qu'elles posent aux étudiants natifs et non natifs, comme le montrent les recherches existantes sur l'acquisition de la phraséologie collocationnelle. Nous soulignons les approches méthodologiques prédominantes et les lacunes actuelles dans ce domaine. La méthodologie de notre étude sera détaillée, incluant les participants, la collecte de données, la grille d'erreurs utilisée, et le traitement des données. Les résultats offriront un aperçu des erreurs collocationnelles identifiées, avec une analyse approfondie des erreurs lexicales, grammaticales et contextuelles. Enfin, la conclusion résumera les principales découvertes de notre étude.

## 2. Cadre théorique

### 2.1. Les collocations et les collocations transdisciplinaires définitions

- 6 La définition de la collocation que nous adoptons dans cette étude s'inscrit dans une approche fonctionnelle (Hausmann, 1989 ; Mel'čuk, 1998 ; Tutin & Grossmann, 2002). Elle se définit comme suit : des expressions binaires composées de deux unités lexicales, généralement deux unités lexicales, qui apparaissent fréquemment ensemble et sont reliées par une relation syntaxique. L'un des deux éléments lexicaux, la *base*, conserve sa fonction sémantique habituelle, tandis que l'autre, le *collocatif*, est moins prédictible et son sens dépend de la base. Par exemple, dans la collocation *prendre la décision*, la base est *décision* et le collocatif est le verbe *prendre*.

- 7 Les collocations, également appelées combinaisons semi-figées, se situent entre deux extrêmes : les *combinaisons libres* et les *combinaisons figées*. Contrairement aux collocations, où la base conserve son sens habituel alors que le collocatif adopte un sens métaphorique, comme dans *peur bleue*, une combinaison libre se caractérise par une association où les constituants maintiennent leur sens usuel. Ce type d'association est compositionnel et transparent sur le plan sémantique, facilitant ainsi la prédictibilité du sens pour les apprenants, comme dans *armoire bleue*.
- 8 Nous étudions ici les collocations transdisciplinaires telles que *mener une analyse*. Il s'agit d'un sous-ensemble du lexique scientifique transdisciplinaire (LST)<sup>1</sup>. Le LST se caractérise par trois propriétés principales : sa transdisciplinarité, car il n'est pas spécifique à une discipline, mais à un genre ; sa fréquence, car il inclut des éléments lexicaux fréquemment utilisés dans le discours scientifique ; et enfin sa spécificité, en tant que propre au genre de l'écrit scientifique, assurant la fonction de désigner les procédures et outils de l'activité scientifique (Hatier, 2016). Le LST est classé selon différents critères : le classement selon l'ordre alphabétique de Phal (1971), le classement par fréquence de Coxhead (2002), ainsi que le classement fonctionnel de Pecman (2004), sémantique de Tutin (2007) et notionnel d'Hirsh (2010).

## 2.2. Les collocations dans les productions des natifs et des non-natifs

- 9 Notre attention s'est particulièrement focalisée sur les collocations verbo-nominales, car ce patron syntaxique est non seulement fréquent, mais aussi source de difficulté dans les productions chez les locuteurs natifs (LN) et non natifs (LNN) par rapport aux autres patrons syntaxiques (Nesselhauf, 2005). Ce type de collocations véhicule des informations essentielles, indispensables dans la communication (Gyllstad, 2007). Ce patron représente 41 % des collocations dans le discours académique Howarth (2013).
- 10 D'un point de vue quantitatif, les LN ont tendance à produire un nombre plus élevé de collocations d'usage conventionnel que les LNN (Howarth, 1998). Parallèlement, d'un point de vue qualitatif, les productions des LN intègrent les collocations de manière plus fluide et naturelle, tandis que celles des LNN présentent souvent des erreurs de sélection ou d'emploi même au niveau avancé (Granger & Paquet, 2008 ; Laufer & Waldman, 2011 ; Nesselhauf, 2005).
- 11 Les résultats de l'étude de Webb, Newton et Chang (2012) sur l'apprentissage des collocations en L2 montrent que les collocations sont souvent apprises de manière fortuite, par des rencontres répétées dans des situations de communication réelles plutôt que dans le cadre d'un apprentissage formel. En rencontrant fréquemment une collocation dans divers contextes linguistiques, les apprenants renforcent leur compréhension de son sens et de son utilisation. Cette exposition répétée favorise une meilleure mémorisation et une intégration plus solide des collocations dans le lexique des apprenants, ce qui facilite leur utilisation future dans des contextes appropriés.
- 12 Les études sur l'acquisition des collocations soulignent plusieurs sources d'erreurs dans l'utilisation des collocations. Tout d'abord, l'influence de la langue première (L1) est une cause majeure, notamment en ce qui concerne la congruence, c'est-à-dire présence ou absence d'un équivalent littéral en L1 (Nesselhauf, 2005 ; Peters, 2016). De plus, le manque de connaissances sur les propriétés sémantiques et syntaxiques des

collocations contribue également aux erreurs. Cela se traduit parfois par un suremploi, un sous-emploi ou un mésusage de certaines constructions, ainsi qu'un manque de compréhension du registre attendu dans le contexte linguistique approprié (Bolly, 2011). Les travaux de Bolly (*ibid.*) révèlent que les apprenants anglophones et néerlandophones de français langue seconde (L2) suremployent les associations libres du type V+N avec deux verbes à haute fréquence (*prendre* et *donner*). De plus, les résultats de l'étude d'Anctil (2012) sur les erreurs lexicales d'élèves de 3<sup>e</sup> année du cycle secondaire québécois montrent que la majorité des erreurs liées à la combinatoire lexicale atteint 78,7 %. Il constate que plus de la moitié de ces erreurs sont des collocations verbales, qu'il s'agisse de verbes supports sémantiques vides tels que *faire*, *porter*, etc., ou de verbes véhiculant un sens général (Anctil, 2012 : 21). Par ailleurs, le suremploi de certaines collocations peut être attribué à leur similitude avec les équivalents dans la langue première ou à leur fréquence dans la langue cible, reflétant ainsi le processus de développement de l'interlangue chez les apprenants. Une étude récente sur l'utilisation des collocations de type V+N (*make/take* + noun) des apprenants selon le CECRL révèle que les débutants utilisent des collocations concrètes et quotidiennes, tandis que les niveaux intermédiaires et avancés emploient des collocations plus complexes. Les résultats montrent également que les apprenants avancés maîtrisent mieux les collocations avec des noms plus difficiles et longs (Du, Afzaal & Al Fadda, 2022). La nature arbitrairement restreinte des collocations rend leur maîtrise difficile, même pour les apprenants avancés, qui ont du mal à distinguer entre des verbes sémantiquement proches, mais spécifiques dans leur usage. Les erreurs sont particulièrement fréquentes avec les verbes de création, de réalisation, d'obtention et de mise en place. Cette acquisition partielle des verbes conduit à des erreurs fréquentes, ce qui souligne l'importance de comprendre les nuances sémantiques complètes pour réussir l'apprentissage des collocations (Men, 2018).

- 13 Howarth (1998) explore trois aspects majeurs de l'utilisation des collocations, à savoir le *figuratif*, le *délexical* et la *technique*. Les LN intègrent intuitivement les aspects figuratifs dans leur discours, tandis que les LNN peuvent rencontrer des difficultés à les comprendre et à les utiliser correctement. De même, les collocations délexicales courantes peuvent poser problème aux LNN, tout comme les termes spécialisés, pour lesquels les LNN peuvent manquer de connaissances dans des contextes professionnels ou techniques spécifiques.
- 14 Enfin, les LN montrent peu d'erreurs syntaxiques, indiquant une bonne maîtrise des collocations et une familiarité avec leurs spécificités syntaxiques. Les erreurs des LNN se répartissent généralement en deux catégories principales : les écarts phraséologiques, souvent dus à une méconnaissance des restrictions morphosyntaxiques, et les erreurs grammaticales, telles que la mauvaise formation du pluriel, dues à des interférences avec la langue première. Les collocations utilisant des prototypes syntaxiques fréquents sont moins sujettes aux déviations que les prototypes moins courants (Nesselhauf, 2005).
- 15 Avant d'aborder la partie méthodologique, il convient de souligner que les erreurs de collocations peuvent être étudiées de différentes manières. Alonso Ramos *et al.* (2010) proposent une typologie des erreurs qui distingue trois dimensions parallèles : la première, dite de *location*, examine si l'erreur concerne un élément de la collocation (la base ou le collocatif) ou la collocation dans son ensemble. La deuxième dimension, centrée sur l'*analyse descriptive* des erreurs, identifie trois types principaux d'erreurs :

lexicales, grammaticales et de registre. Enfin, la troisième dimension, appelée *analyse explicative* des erreurs, se concentre sur la source des erreurs, en distinguant entre les erreurs de transfert (interférence de la L1) et les erreurs d'interlangue, qui résultent d'une connaissance incomplète de la L2 sans interférence de la L1 (Vincze, Alonso Ramos, Mosqueira Suárez & González, 2011 : 282).

- 16 Ainsi, notre analyse des erreurs collocationnelles sera principalement descriptive. Nous adoptons la définition de l'erreur proposée par Hamel et Milicévić (2007), qui englobe tous les écarts morphosyntaxiques, sémantiques et contextuels par rapport à l'usage d'un locuteur expert. De plus, nous suivons Anctil (2011) en ce qui concerne les maladresses lexicales, c'est-à-dire les usages du lexique considérés comme douteux mais qui ne peuvent être catégoriquement jugés comme des erreurs.

### 3. Méthodologie

- 17 Afin d'évaluer les compétences collocationnelles d'un échantillon délimité d'étudiants et d'identifier les difficultés auxquelles sont confrontés les participants lorsqu'ils utilisent des collocations transdisciplinaires, nous avons adopté une approche basée sur l'analyse de productions écrites. Ici, nous présentons d'abord l'échantillon, la méthode de collecte et de traitement des données, ainsi que la construction d'une grille d'erreurs et l'annotation du corpus.

#### 3.1. Profil des participants et procédure de collecte des données

- 18 En analysant les 60 textes constitués de synthèses et de résumés d'articles scientifiques, rédigés par des étudiants en Master 1 (M1) et en Master 2 (M2) entre 2015 et 2020 en sciences du langage et en didactique des langues à l'Université Grenoble Alpes, nous avons réparti les textes comme suit : 30 textes produits par des étudiants francophones, dont 15 en M1 et 15 en M2, ainsi que 30 textes produits par des étudiants allophones, également répartis avec 15 en M1 et 15 en M2. Il est important de noter qu'il ne s'agit pas des mêmes étudiants pour lesquels deux textes ont été recueillis. Les étudiants non natifs proviennent de diverses L1 et ont un niveau avancé de compétence en langue française évalué à C1 selon le *Cadre européen commun de référence pour les langues* (CECRL).
- 19 Les productions, considérées comme des traces écrites visant à répondre aux exigences académiques pour valider leur parcours universitaire, ont été rédigées par des étudiants à la maison. Étant produites à la maison, nous ne pouvons garantir si elles ont été rédigées avec ou sans aide (dictionnaires, traducteurs automatiques, relecture), mais nous pouvons assurer que les textes ont été produits avant l'émergence des intelligences artificielles (IA) et qu'ils n'ont pas été générés avec des outils tels que ChatGPT, par exemple.
- 20 Pour la rédaction, les étudiants suivaient la consigne de « résumer l'article en utilisant vos propres mots tout en préservant les idées principales et les informations essentielles ». La longueur des textes varie de 3 à 6 pages, comprend une partie théorique, une méthodologie, ainsi qu'une analyse et des résultats pour chaque texte produit par les étudiants.

### 3.2. Construction d'une grille d'erreurs

- 21 Parmi les recherches sur l'évaluation des compétences collocationnelles, certaines analysent les erreurs lexicales à travers des grilles d'évaluation. Par exemple, Anctil (2011) propose une grille pour évaluer les productions en L1, tandis que Yan (2017) travaille sur une grille adoptée pour la L2. Ces outils évaluent diverses erreurs lexicales, y compris les associations lexicales en langue générale.
- 22 L'approche d'Anctil (2011) se concentre sur les erreurs lexicales des francophones québécois natifs, structurée en trois catégories : la forme, le sens et la combinatoire restreinte. En revanche, la grille de Yan (2017) se spécialise dans les erreurs verbales dans l'écrit scientifique des apprenants sinophones, classant les erreurs en quatre catégories : 1) morphologiques, 2) sémantiques et cooccurrences lexicales, 3) grammaticales et syntaxiques, et 4) liées au genre et au contexte.
- 23 À notre connaissance, il n'existe pas de grille spécifique en langue de spécialité pour évaluer les compétences des associations lexicales. Pour pallier cette lacune, nous avons développé une grille d'erreurs inspirée des travaux d'Anctil (2011) et de Yan (2017), axée exclusivement sur les erreurs de collocations. Notre grille d'évaluation s'inspire de la sous-partie « Combinatoire restreinte » de la grille d'analyse des erreurs d'Anctil (2011) et de la sous-partie « Cooccurrences lexicales » de la grille de Yan (2017). Elle évalue uniquement les erreurs collocationnelles au niveau grammatical, lexical et contextuel (voir le tableau 1).
- 24 La partie grammaticale de la grille se divise en deux sous-parties et comprend les erreurs au niveau *morphologique* et *syntactique*, telles que les *erreurs de genre*, de *nombre*, de *conjugaison*, de *préposition*, de *déterminant*, de *construction* et d'*ordre des mots*. La partie *lexicale* regroupe des erreurs au niveau du *sens*, comprenant trois types : *collocatif inadapté/sens inadapté*, *base inadaptée/sens inadapté* et *collocatif inadapté/sens adapté*. La dernière partie de la grille concerne les *erreurs contextuelles* liées au registre de langue.
- 25 Nous allons expliquer en détail chaque type d'erreur en donnant des exemples tirés du corpus étudié dans la partie des résultats de cet article.

### 3.3. Repérage et codage des associations

- 26 D'abord, les associations transdisciplinaires de type V+N sont repérées manuellement dans chaque texte. Pour réduire la subjectivité de cette identification, nous avons utilisé des références exemplaires de la norme lexicale, notamment la liste du Lexique scientifique transdisciplinaire (LST) telle que définie dans l'ouvrage de Jacques et Tutin (2018), ainsi que le logiciel Antidote (version 10).
- 27 Ensuite, en nous inspirant de la méthodologie proposée par Nesselhauf (2005), les associations identifiées dans les textes ont été évaluées à l'aide de notre grille d'erreurs par trois enseignants-chercheurs en sciences du langage. Ces derniers sont des locuteurs natifs, avec une maîtrise de la langue spécialisée et un vocabulaire étendu en lexique disciplinaire.
- 28 Une fois que les juges ont identifié et annoté les erreurs, nous avons transféré les données dans Excel pour faciliter l'évaluation des associations, ainsi que la comparaison des annotations des trois juges. Nous avons constaté un certain degré de consensus dans l'évaluation des associations entre les annotateurs. Une analyse

quantitative a été effectuée afin de calculer les fréquences relatives et la distribution exacte des associations dans chaque groupe et de les comparer entre les M1 et M2. Cette analyse visait également à identifier les tendances des types d'erreurs dans un groupe par rapport à l'autre.

Tableau 1. – Grille axée exclusivement sur les erreurs de collocations, inspirée des travaux d'Ancil (2011) et de Yan (2017).

GRAMMATICALE	MORPHOLOGIQUE	Erreur de conjugaison et d'accord (CA)	<b>On obesrvons les résultats</b> mentionnés (observe les résultats)  L'élaboration de notre synthèse sera organisée, selon les <b>étapes suivante</b> (étapes suivantes)
		Erreur d'orthographe (ORT)	Pour étudier les problèmes mentionnés, nous observerons les différences et les solutions utilisées dans les traductions de ces deux langues à partir d'un <b>corpus recueille</b> sur le site internet <i>Opus Corpus</i> . (corpus recueilli)
		Erreur de genre (GEN)	Nous avons construit les <b>tableaux comparatives</b> de système de la flexion nominale du polonais. (tableaux comparatifs)
		Erreur de nombre (NB)	Nous <b>faisons les liens</b> vers les sites d'apprentissage du français (faire le lien)
	SYNTAXIQUE	Erreur de préposition (PRE)	L'ergativité reste donc un sujet complexe comportant des zones d'ombre que les linguistes s'efforcent d'éclairer afin de donner une définition précise et <b>reconnue à ce phénomène</b> linguistique. (reconnue de ce phénomène)
		Erreur d'ordre des mots (ORD)	Dans le cadre de ce travail, premièrement j'ai décrit en examinant l'article de Dirk Geeraerts les <b>courants principaux</b> de la sémantique (principaux courants)
		Erreur de déterminant (DET)	Dans cette étude, nous voulons comparer le système casuel en polonais et ses équivalents français et <b>présenter des difficultés</b> rencontrées (présenter les difficultés)
		Erreur de construction (CON)	Pour l'acquisition guidée, le processus <b>se démarre d'une manière</b> systématique par l'enseignement de langue (démarre d'une manière)
LEXICALE	SEMANTIQUE	Collocatif inadapté sens inadapté (CI)	Les théories évoquées ci-dessus <b>sollicitent l'auteur</b> à faire son travail de recherche abordé dans ce texte. (incitent l'auteur)
		Collocatif inadapté sens adapté (AI)	Pendant l'analyse des exemples en français, nous avons <b>remarqué une présence</b> de la construction aspectuelle GN+Vasp+GV dans laquelle on trouve des verbes aspectuels. (observé la présence)
		Base inadaptée sens inadapté (BI)	Nous parlerons ensuite de la rhétorique contrastive avant de <b>souligner des facteurs</b> socio-discursifs du français et du chinois en rapport à la rédaction de textes argumentés (souligner les aspects)
CONTEXTE	PRAGMATIQUE	Emploi d'une unité lexicale apparentant à la langue générale (REG)	La production écrite, soit en L1 soit en L2, est une <b>activité</b> effectivement <b>compliquée</b> (activité complexe)
		Autres erreurs (AUT)	<b>De manière logique</b> , une langue ergative est une langue qui possède une syntaxe ergative.

## 4. Résultats

### 4.1. La fréquence et l'analyse quantitative des erreurs collocationnelles

- 29 Le tableau ci-dessous présente une vision globale des textes analysés, des associations relevées et des erreurs repérées dans les textes pour les quatre groupes d'étudiants : Locuteur natif Master 1 (LN-M1), Locuteur natif Master 2 (LN-M2), Locuteur non natif Master 1 (LNN-M1) et Locuteur non natif Master 2 (LNN-M2).

Tableau 2. – Récapitulatif des textes analysés, des associations et des erreurs relevées par groupes d'étudiants.

Participants	LN-M1	LN-M2	LNN-M1	LNN-M2	Total
N° Textes	15	15	15	15	60
N° Associations (V+N)	672	664	558	730	2 624

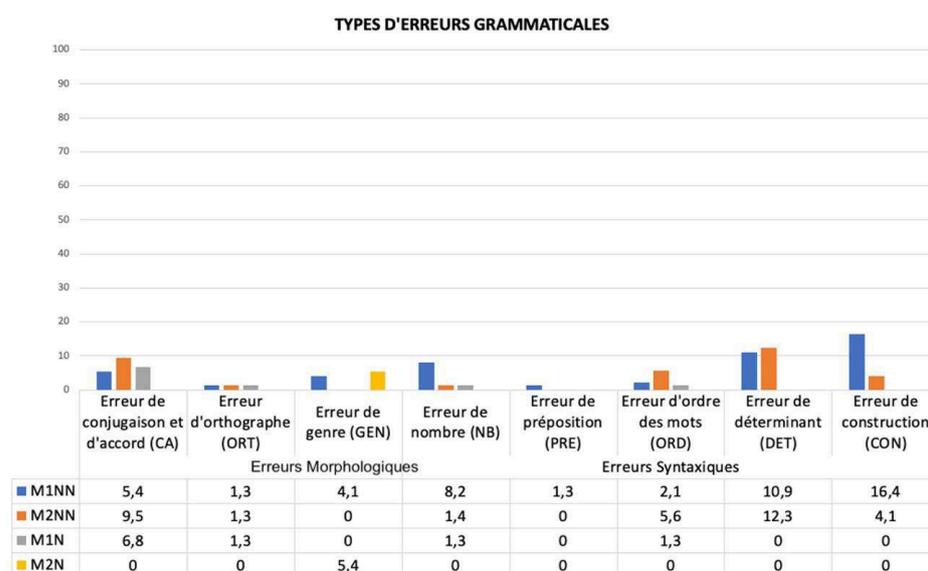
N° Associations erronées	41	27	96	95	259
--------------------------	----	----	----	----	-----

- 30 Parmi les 60 textes examinés, composés de 30 textes d'étudiants en M1 (15 LNN et 15 LN) et de 30 textes d'étudiants en M2 (15 LNN et 15 LN), totalisant ainsi 235 440 *tokens* dans le corpus, nous avons identifié 2 624 associations V+N, qu'elles soient correctes ou inappropriées. En ce qui concerne les associations erronées, nous avons identifié 259 erreurs au total. En observant le tableau 2, on constate que les LNN présentent un nombre d'erreurs plus élevé que les LN tout au long de leur parcours, bien qu'une légère amélioration soit observée en passant de M1 à M2 pour les deux groupes. Les résultats suggèrent que les LNN pourraient bénéficier de soutien linguistique et académique supplémentaire pour améliorer leurs performances, tandis que les LN montrent une amélioration significative en passant de M1 à M2. Ils semblent mieux adaptés aux exigences académiques, mais il y a encore une marge pour réduire davantage les erreurs.
- 31 Par la suite, nous examinerons en détail la performance des LN et LNN en M1 et M2 sur chaque type d'erreur grammaticale, lexicale et contextuelle.

## 4.2. Les erreurs grammaticales

- 32 La partie grammaticale de notre grille se divise en deux sous-parties et évalue les erreurs au niveau *morphologique* et *syntactique*. La figure 1 donne un aperçu de la répartition des erreurs grammaticales dans l'ensemble de notre corpus en fonction de leur appartenance à l'une des huit catégories d'erreurs grammaticales : les *erreurs de conjugaison (CA)*, *d'orthographe (ORT)*, *de genre (GEN)*, *de nombre (NB)*, *de préposition (PRE)*, *d'ordre des mots (ORD)*, *de déterminant (DET)* et *de construction (CON)*.

Figure 1. – Répartition des types d'erreurs grammaticales du corpus.



- 33 L'analyse comparative des erreurs grammaticales entre les étudiants LN et LNN en M1 et M2 révèle des tendances distinctes et des évolutions significatives dans l'utilisation

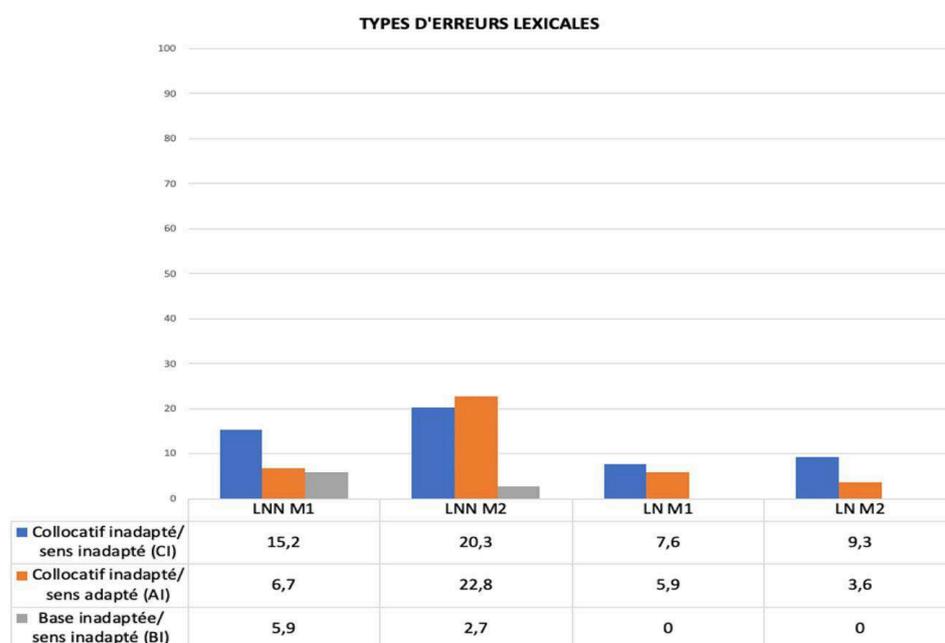
des collocations transdisciplinaires. Globalement, les deux groupes montrent des améliorations dans certaines sous-catégories d'erreurs grammaticales, tandis que d'autres restent des points de défi persistants.

- 34 Au niveau des erreurs syntaxiques, une réduction importante est observée entre les niveaux M1 et M2, témoignant d'une amélioration dans la capacité à manipuler des structures syntaxiques complexes chez les deux groupes. Cependant, des différences apparaissent dans la persistance de certains défauts. Chez les étudiants LN, les erreurs de construction et les erreurs de déterminants indiquent une certaine maîtrise. En revanche, chez les étudiants LNN, bien que la réduction des erreurs soit notable entre M1 et M2, certains défis demeurent, en particulier concernant les erreurs de déterminants, de conjugaison et d'orthographe, où des taux plus élevés sont observés par rapport aux étudiants LN.
- 35 De même, au niveau des erreurs morphologiques, des améliorations sont notables dans la capacité à respecter les règles morphologiques, avec une diminution des erreurs de nombre et de genre entre M1 et M2 pour les deux groupes. Cependant, une nette différence apparaît dans la persistance des erreurs entre les groupes LN et LNN. Alors que les étudiants LN montrent une quasi-disparition des erreurs morphologiques en M2, les étudiants LNN présentent encore des taux non négligeables d'erreurs, bien que réduits par rapport à M1.
- 36 Ces points faibles sont illustrés par des exemples tirés du corpus étudié :
- (1)  
**Nous appuyons** sur l'élément de la langue orale. (nous nous appuyons)
- (2)  
 Dans cette étude, nous voulons comparer le système casuel en polonais et ses équivalents français et **présenter des difficultés** rencontrées (présenter les difficultés)
- 37 Dans l'exemple 1, on observe une erreur de construction. Une erreur de construction se produit lorsque la structure syntaxique ou grammaticale d'une phrase est incorrecte. Cela peut inclure des erreurs telles que des phrases fragmentées, des phrases inversées, des erreurs de concordance de temps ou de voix, des erreurs de subordination ou de coordination, etc. En général, une erreur de construction rend la phrase confuse ou difficile à comprendre pour le lecteur ou l'interlocuteur. L'erreur précise réside dans l'utilisation de la forme réfléchie *nous nous appuyons*.
- 38 Dans l'exemple 2, une erreur dans le choix du déterminant est observée. Une erreur de déterminant peut découler d'une utilisation incorrecte, d'une méconnaissance du genre du nom, ou de l'absence d'un déterminant devant un nom. Dans cet exemple précis, l'utilisation de l'article indéfini *des* devant le mot *difficultés* est erronée. Cependant, pour préciser les difficultés dont il est question dans le cadre de l'étude, il est plus approprié d'utiliser l'article défini *les* au lieu de l'article indéfini *des*. Cela permet de mieux identifier le référent.
- 39 Les sources de ces erreurs peuvent être diverses, telles que la complexité des structures de la langue française et l'influence de la langue maternelle des participants (Nesselhauf, 2005).

### 4.3. Les erreurs lexicales

- 40 Dans notre corpus, nous avons identifié trois types d'erreurs au niveau du sens : *collocatif inadapté/sens inadapté* (CI), *collocatif inadapté/sens adapté* (AI) et *base inadaptée/sens inadapté* (BI). La figure 2 expose une vue d'ensemble de la répartition des erreurs lexicales dans l'ensemble de notre corpus chez les LN et LNN au niveau de M1 et M2. Nous détaillons ensuite chaque type d'erreur.

Figure 2. – Répartition des types d'erreurs lexicales du corpus.



- 41 L'observation des erreurs lexicales chez les M1 et M2 révèle que, pour les erreurs de type CI, le groupe LNN M2 affiche le taux d'erreur le plus élevé (20,3 %), tandis que les locuteurs natifs présentent des taux plus bas (7,6 % pour LN M1 et 9,3 % pour LN M2). En ce qui concerne les erreurs de type AI, LNN M2 montre une augmentation importante des erreurs (22,8 %), soulignant un problème persistant dans le choix des collocations. Les locuteurs natifs, quant à eux, affichent des taux d'erreurs nettement plus bas (3,6 % pour LN M2). Enfin, dans la catégorie des erreurs de type BI, les locuteurs natifs n'ont commis aucune erreur, tandis que les groupes LNN montrent des taux modérés.

#### 4.3.1. Collocatif inadapté/sens inadapté

- 42 Les associations *collocatif inadapté/sens inadapté* (CI) regroupent les erreurs où l'association se produit lorsqu'on utilise un collocatif qui n'est pas couramment utilisé avec la base en question, mais qui a néanmoins un sens qui peut être compris dans le contexte donné. Un locuteur expert dans le même contexte ne produit pas spontanément ce type de combinaisons.

(3)

Les théories évoquées ci-dessus **sollicitent l'auteur** à faire son travail de recherche abordé dans ce texte. (incitent l'auteur)

- 43 Dans l'exemple 3, *sollicitent l'auteur* n'est pas incorrect en soi, mais dans ce contexte, le collocatif *inciter* est plus approprié, car ce verbe exprime mieux l'idée de motiver ou d'encourager quelqu'un à entreprendre une action. *Solliciter* signifie principalement demander quelque chose de manière pressante ou formelle, ce qui ne correspond pas vraiment à l'idée de stimuler ou d'encourager l'auteur à mener des recherches. Par exemple, il serait plus adéquat, dans cette situation, d'opter pour l'association *incitent l'auteur*.

#### 4.3.2. Collocatif inadapté/sens adapté

- 44 L'association *collocatif inadapté/sens adapté* (AI) se produisent lorsque l'on utilise un collocatif qui n'est pas habituellement associé à la base en question, mais qui a cependant un sens qui peut être compris dans le contexte donné.

(4)

Les trois **exemples visibles** au-dessus sont les seuls exemples qu'on trouve dans le sous-corpus OpenSubtitles2016. (exemples cités)

(5)

Pendant l'analyse des exemples en français, nous avons **remarqué une présence** de la construction aspectuelle GN+Vasp+GV dans laquelle on trouve des verbes aspectuels [...]. (observé la présence)

- 45 L'association *exemple visible* dans l'exemple 4 serait classée comme une association inexistante, car le collocatif *visible* n'est pas associé à la base *exemple*. Cependant, malgré cette inadéquation, le sens de l'association demeure compréhensible.
- 46 Nous observons ce type d'erreurs également dans l'exemple 5. Le choix du collocatif *remarqué* est inapproprié avec la base *présence*. En revanche, le sens général de l'association est prévisible, bien que cela constitue une association inexistante. Le collocatif erroné est un synonyme du collocatif approprié. Cette erreur, qui témoigne d'une certaine créativité, résulte d'une méconnaissance des collocations appropriées ou d'une compréhension limitée des nuances sémantiques entre les deux associations. En effet, cela souligne une difficulté potentielle des étudiants à sélectionner avec précision les collocatifs adéquats en fonction du contexte, ce qui peut influencer la clarté et la pertinence des associations verbales créées.
- 47 Le choix inapproprié du collocatif conduit souvent à la création d'associations inexistantes et erronées, reflétant un niveau d'abstraction plus élevé des concepts ainsi que par la maîtrise imparfaite de ces concepts. Il est à noter que nous avons observé un suremploi des collocatifs (le verbe) associés avec la base (le nom). Ce choix limité d'utilisation de certains verbes à haute fréquence tels que *faire*, *avoir*, *voir*, *donner*, *trouver*, etc., peut être expliqué par la possession d'un répertoire lexical restreint et une maîtrise imparfaite de la langue entraînant, comme l'a souligné Kaszubski (2000), une propension des étudiants à recourir à des *safe options* par mesure de sécurité linguistique.

#### 4.3.3. Base inadaptée/sens inadapté

- 48 L'erreur de *base inadaptée/sens inadapté* (BI) se produit lorsqu'une base est associée à un collocatif qui n'est pas approprié pour exprimer une idée ou un concept donné. Dans ce cas, le sens devient inadapté, car la base utilisée ne convient pas à la signification que l'on souhaite transmettre.

(6)

Nous parlerons ensuite de la rhétorique contrastive avant de **souligner des facteurs** socio-discursifs du français et du chinois en rapport à la rédaction de textes argumentés. (souligner les aspects)

- 49 L'exemple 6 présente une erreur de *base inadaptée/sens inadapté* où la base *facteurs* semble inappropriée pour exprimer l'idée souhaitée. Dans ce contexte, *aspects* est préférable à *facteurs*, car il désigne plus précisément des éléments ou dimensions particulières du phénomène discuté, ce qui est plus adapté pour décrire les subtilités socio-discursives liées à la rédaction.
- 50 Le choix de la base dans les associations semble être une compétence relativement maîtrisée par les deux groupes, contrairement au choix du collocatif qui s'avère plus problématique en raison de la variété des collocatifs possibles et des nuances de sens.
- 51 Les travaux de Cavalla (2008, 2015) corroborent ces observations en confirmant la présence d'erreurs au niveau des collocations dans un corpus composé d'écrits universitaires d'étudiants étrangers de niveau master ou doctorat en France. Ces étudiants, d'origines diverses, présentent un niveau avancé en français. Cavalla (2015) met en avant que le type d'erreur le plus fréquent concerne la structure V+N, notant que les apprenants commettent plus d'erreurs dans le choix du collocatif (le verbe) que dans celui de la base (le nom), en raison de la possibilité de déviation du sens du premier. Par exemple, au lieu d'utiliser la collocation correcte *mener une étude*, un étudiant pourrait écrire *faire une étude*, ce qui bien que compréhensible, manque de précision et de naturel linguistique attendu dans les écrits académiques.

#### 4.4. Les erreurs contextuelles

- 52 Cette catégorie englobe les collocations valables mais inappropriées en contexte, dans lesquelles le collocatif est reconnu comme correct, mais ne s'aligne pas sur le registre souhaité. Elle fait référence aux « déviances contextuelles des collocations », selon Bolly (2011), et à « l'emploi d'un collocatif valable, mais qui ne correspond pas au sens visé », selon Anctil (2011).
- 53 L'analyse des erreurs contextuelles dans les écrits des étudiants montre une amélioration notable des compétences collocationnelles avec l'avancement académique, tant chez les LN que chez les LNN. En M1, les LNN affichent un taux d'erreurs contextuelles de 41,5 %, qui diminue de manière significative à 20,1 % en M2. Cette réduction de plus de 50 % indique une amélioration substantielle de leur capacité à utiliser correctement les collocations dans un contexte approprié. Comparativement, les LN commencent avec un taux d'erreurs contextuelles plus bas de 26,1 % en M1, qui baisse à 12,3 % en M2. Bien que l'amélioration soit moins prononcée que chez les LNN, elle reste significative, montrant qu'ils raffinent également leur usage des collocations avec l'avancement académique.

(7)

Les théories évoquées sollicitent l'auteur à **faire son travail de recherche** abordé dans ce texte. (réaliser une recherche)

- 54 Dans l'exemple 7, la maladresse réside dans l'utilisation de la collocation *faire son travail de recherche*, qui n'est pas une erreur en soi, mais plutôt une expression appartenant à un registre de langue général et informel. Pour correspondre au registre académique requis, il serait préférable d'utiliser des expressions telles que *mener une recherche* ou

*réaliser une recherche*, qui reflètent un niveau de formalité et de précision adapté au contexte académique.

## 5. Conclusion

- 55 Nos résultats confirment certaines tendances observées dans la littérature existante sur l'acquisition des collocations, tout en mettant en lumière des aspects spécifiques qui méritent une attention particulière. Ils montrent une distinction claire entre les étudiants natifs et non natifs du français dans l'utilisation des collocations transdisciplinaires.
- 56 Sur le plan quantitatif, les étudiants natifs produisent davantage d'associations transdisciplinaires que les non-natifs et commettent également un nombre d'erreurs grammaticales, lexicales et contextuelles inférieur, tant en Master 1 qu'en Master 2. Il est intéressant de noter que les natifs affichent une progression uniforme et soutenue par rapport aux non-natifs.
- 57 En approfondissant l'analyse qualitative, nous constatons que, bien que les locuteurs non natifs ne rencontrent pas de grandes difficultés avec la syntaxe en raison de la limitation des constructions syntaxiques et morphologiques, qui peuvent être maîtrisées plus facilement, ce sont principalement les aspects sémantiques et contextuels qui posent problème. Ces difficultés sont souvent liées au choix du collocatif plutôt qu'à la base elle-même, qui conserve sa fonction sémantique habituelle. En revanche, le sens du collocatif dépend de la base, ce qui complique son utilisation. Nous avons observé que les difficultés dans le choix des collocatifs peuvent être attribuées à des problèmes de synonymie et de quasi-synonymie, ainsi qu'à un suremploi de verbes fréquents. De plus, le choix inapproprié de collocatifs peut entraîner des erreurs de registre. Ces problèmes peuvent résulter d'un manque de sensibilité aux nuances des synonymes et d'un répertoire lexical limité, conduisant à la création d'unités lexicales qui n'existent pas dans la langue cible ou, bien qu'existantes, ne seraient pas choisies par un locuteur expert dans un contexte académique, ce qui affecte la clarté et la précision de leurs productions écrites.
- 58 Bien que cette étude révèle une amélioration dans certaines dimensions des compétences collocationnelles, ses résultats ne peuvent pas être généralisés à l'ensemble des étudiants universitaires sans tenir compte d'autres facteurs contextuels et individuels. L'étude présente certaines limites, notamment une taille d'échantillon restreinte de 30 étudiants en sciences du langage dans un seul établissement, ce qui réduit la portée des conclusions. De plus, l'analyse d'un corpus restreint diminue la diversité des données. Une étude longitudinale examinant les productions des mêmes étudiants en M1 et M2 aurait permis d'obtenir des résultats plus représentatifs.
- 59 Pour optimiser l'apprentissage et l'enseignement des collocations transdisciplinaires, une attention particulière doit être portée aux nuances lexicales et aux registres. Il est crucial de sensibiliser les étudiants à l'importance de repérer et de comprendre ces structures semi-figées (Cavalla & Legallois, 2020). Il est également essentiel de leur expliquer le rôle clé des collocations dans une communication académique efficace. L'utilisation d'exemples concrets et d'activités ciblées est nécessaire pour aider les étudiants à appréhender les aspects syntaxiques, sémantiques et contextuels des collocations. Une « exposition fréquente et prolongée » aux collocations est

recommandée (Falkert, 2013 : 158). Impliquer les étudiants dans des activités pratiques, telles que des discussions dirigées (présentations, débats) et des travaux écrits (résumés, synthèses), favorise la manipulation active des collocations et renforce la maîtrise de la langue.

- 60 L'utilisation de corpus d'étudiants constitue une approche pertinente, permettant de cibler les besoins linguistiques des étudiants et d'adapter les méthodes d'enseignement en conséquence. Par ailleurs, il est recommandé de développer des critères clairs pour évaluer la maîtrise des collocations, d'utiliser des outils d'évaluation spécifiques et de fournir une rétroaction constructive afin de soutenir la progression continue des apprenants.

---

## BIBLIOGRAPHIE

- ALONSO RAMOS Margarita, WANNER Leo, VINCZE Orsolya, CASAMAYOR DEL BOSQUE Gerard, VAZQUEZ VEIGA Nancy [...] PRIETO GONZÁLEZ Sabela (2010), « Towards a Motivated Annotation Schema of Collocation Errors in Learner Corpora », N. Calzolari *et al.* (dir.), *Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10)*, <[www.lrec-conf.org/proceedings/lrec2010/pdf/751\\_Paper.pdf](http://www.lrec-conf.org/proceedings/lrec2010/pdf/751_Paper.pdf)>.
- ANCTIL Dominic (2011), *L'erreur lexicale au secondaire : analyse d'erreurs lexicales d'élèves de 3<sup>e</sup> secondaire et description du rapport à l'erreur lexicale d'enseignement de français* (thèse de doctorat), Université de Montréal, Montréal, <[https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/5077/Anctil\\_Dominic\\_2011\\_these.pdf?sequence=2&isAllowed=y](https://papyrus.bib.umontreal.ca/xmlui/bitstream/handle/1866/5077/Anctil_Dominic_2011_these.pdf?sequence=2&isAllowed=y)> (26 février 2024).
- ANCTIL Dominic (2012), « Portrait des erreurs lexicales d'élèves de 3<sup>e</sup> secondaire en production écrite et proposition de pistes didactiques », *Pratiques*, 155-156, 7-30, <<https://doi.org/10.4000/pratiques.3447>>.
- BOCH Françoise & RINCK Fanny (2015), « Anaphores démonstratives dans les écrits d'étudiants de Master : comparaison avec les pratiques expertes », *Linx*, 72, 131-150, <<https://doi.org/10.4000/linx.1631>>.
- BOLLY Catherine (2011), *Phraséologie et collocations : approche sur corpus français L1 et L2*, Bruxelles : Peter Lang.
- CAVALLA Cristelle (2008), « Les collocations dans les écrits universitaires : un français spécifique pour les apprenants étrangers », O. Bertrand & I. Schaffner (dir.), *Le français de spécialité : enjeux culturels et linguistiques*, Paris : Éditions de l'École polytechnique, 93-104.
- CAVALLA Cristelle (2015), « Collocations transdisciplinaires dans les écrits de doctorants FLS/FLE », *Linx*, 72, 111-130, <<https://doi.org/10.4000/linx.1624>>.
- CAVALLA Cristelle & LEGALLOIS Dominique (2020), « Caractériser et identifier les unités phraséologiques pour leur enseignement », *Action didactique*, 3(2), 12-30, <<https://asjp.cerist.dz/en/article/199299>>.

- COXHEAD Averil (2002), « The Academic Word List: A Corpus-Based Word List for Academic Purposes », B. Kettemann & G. Marko (dir.), *Teaching and Learning by Doing Corpus Analysis*, Brill, 42, 72-89.
- DU Xiangtao, AFZAAL Muhammad & AL FADDA Hind (2022), « Collocation Use in EFL Learners' Writing Across Multiple Language Proficiencies: A Corpus-Driven Study », *Frontiers in Psychology*, 13, <<https://doi.org/10.3389/fpsyg.2022.752134>>.
- FALKERT Anika (2013), « À propos de l'existence d'une compétence phraséologique chez l'apprenant de FLE. Une étude expérimentale », C. Garcia-Debanç, C. Masseron & C. Ronveaux (dir.), *Enseigner le lexique*, Namur : Presses universitaires de Namur, 153-172.
- FORSBERG Fanny (2006), *Le langage préfabriqué en français parlé L2 : étude acquisitionnelle et comparative* (thèse de doctorat), Stockholm University, Stockholm.
- GONZÁLEZ-REY Isabel (2015), *La phraséologie du français*, Toulouse : Presses universitaires du Midi.
- GRANGER Sylviane & PAQUOT Magali (2008), « Disentangling the Phraseological Web », S. Granger & F. Meunier (dir.), *Phraseology: An Interdisciplinary Perspective*, Amstersam / Philadelphie : John Benjamins, 27-49.
- GYLLSTAD Henrik (2007), *Testing English Collocations: Developing Receptive Tests for Use with Advanced Swedish Learners*, Lund, Suède : Lund University.
- HAMEL Marie-Josée & MILICEVIC Jasmina (2007), « Analyse d'erreurs lexicales d'apprenants du FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage », *Canadian Journal of Applied Linguistics*, 10(1), 25-45, <<https://journals.lib.unb.ca/index.php/CJAL/article/view/19733>> (19 mars 2024).
- HATIER Sylvain (2016), *Identification et analyse linguistique du lexique scientifique transdisciplinaire. Approche outillée sur un corpus d'articles de recherche en SHS* (thèse de doctorat), Université Grenoble Alpes, Grenoble, <<https://theses.hal.science/tel-01690554/document>> (26 février 2024).
- HAUSMANN Franz-Josef (1989), « Le dictionnaire de collocations », F. J. Hausmann *et al.* (dir.), *Wörterbücher: ein internationales Handbuch zur Lexikographie. Dictionaries*, De Gruyter Mouton, 1010-1019.
- HAUSMANN Franz-Josef (1997), « Tout est idiomatique dans les langues », M. Martins-Baltar (dir.), *La locution entre langue et usages*, ENS Éditions, 277-290.
- HIRSH David (2010), *Academic Vocabulary in Context*, Berne : Peter Lang.
- HOWARTH Peter Andrew (1998), « Phraseology and Second Language Proficiency », *Applied Linguistics*, 19(1), 24-44.
- HOWARTH Peter Andrew (2013), *Phraseology in English Academic Writing: Some Implications for Language Learning and Dictionary Making*, Berlin : Walter de Gruyter.
- JACQUES Marie-Paule & TUTIN Agnès (2018), « Le lexique scientifique transdisciplinaire : une introduction », M.-P. Jacques & A. Tutin (dir.), *Lexique transversal et formules discursives des sciences humaines*, Londres : ISTE Éditions, 1-26.
- KASZUBSKI Przemyslaw (2000), *Selected Aspects of Lexicon, Phraseology and Style in the Writing of Polish Advanced Learners of English: A Contrastive, Corpus-Based Approach* (thèse de doctorat), Adam Mickiewicz University, Poznań.
- LAUFER Batia & WALDMAN Tina (2011), « Verb-Noun Collocations in Second Language Writing: A Corpus Analysis of Learners' English », *Language Learning*, 61, 647-672.

- MEL'ČUK Igor (1998), « Collocations and Lexical Functions », A. P. Cowie (dir.), *Phraseology, Theory, Analysis and Applications*, Clarendon Press, 23-53.
- MEN Haiyan (2018), « Data-Driven Learning in Enhancing Learner's Language Idiomaticity », *International Journal of Emerging Technologies in Learning (IJET)*, 15(23), 27-41, <<https://doi.org/10.3991/ijet.v15i23.19023>>.
- NESSELHAUF Nadja (2005), *Collocations in a Learner Corpus*, John Benjamins.
- PECMAN Mojca (2004), *Phraséologie contrastive anglais-français : analyse et traitement en vue de l'aide à la rédaction scientifique* (thèse de doctorat), Université Sophia Antipolis, Nice.
- PETERS Elke (2016), « The Learning Burden of Collocations: The Role of Interlexical and Intralexical Factors », *Language Teaching Research*, 20(1), 113-138, <<https://doi.org/10.1177/1362168814568131>>.
- PHAL André (1971), *Vocabulaire général d'orientation scientifique (V.G.O.S) : part du lexique commun dans l'expression scientifique*, École normale supérieure de Saint-Cloud : Didier.
- POLLET Marie-Christine (2014), *L'écrit scientifique à l'aune des littéracies universitaires : approches théoriques et pratiques*, Namur : Presses universitaires de Namur.
- RINCK Fanny (2010), « L'analyse linguistique des enjeux de connaissance dans le discours scientifique », *Revue d'anthropologie des connaissances*, 4(3), 427-450. <<https://doi.org/10.3917/rac.011.0427>>.
- TUTIN Agnès (2007), « Autour du lexique et de la phraséologie des écrits scientifiques », *Revue française de linguistique appliquée*, 12(2), 5-14, <<https://doi.org/10.3917/rfla.122.0005>>.
- TUTIN Agnès & GROSSMANN Francis (2002), « Collocations régulières et irrégulières : esquisse de typologie du phénomène collocatif », *Revue française de linguistique appliquée*, 7(1), 7-25, <<https://doi.org/10.3917/rfla.071.0007>>.
- VINCZE Orsolya, ALONSO RAMOS Margarita, MOSQUEIRA SUÁREZ Estela & GONZÁLEZ Sabela Prieto (2011), « Exploiting a Learner Corpus for the Development of a CALL Environment for Learning Spanish Collocation », I. Kosem & K. Kosem (dir.), *Proceedings of eLex 2011*, Ljubljana : Trojina, Institute for Applied Slovene Studies, 280-285, <<https://elex2011.trojina.si/Vsebine/proceedings/eLex2011-38.pdf>>.
- WEBB Stuart, NEWTON Jonathan & CHANG Anna (2012), « Incidental Learning of Collocation », *Language Learning*, 63(1), 91-120, <<https://doi.org/10.1111/j.1467-9922.2012.00729.x>>.
- YAN Rui (2017), *Étude des constructions verbales scientifiques dans une perspective didactique : utilisation des corpus dans le diagnostic des besoins langagiers en FLE à l'aide des techniques de TAL* (thèse de doctorat), Université Grenoble Alpes, Grenoble, <<https://tel.archives-ouvertes.fr/tel-01691923>> (26 février 2024).

## NOTES

1. La liste du lexique scientifique transdisciplinaire (LST), élaborée par Tutin et ses collègues dans le cadre du projet TermITH (2013-2016), est disponible à l'adresse suivante : <<http://lst.demarres-shs.fr>>. Voir également *Lexique transversal et formes discursives des sciences humaines* de Jacques et Tutin (2018).

---

## RÉSUMÉS

Cet article examine les collocations transdisciplinaires de type verbe-nom comme *mener une étude*, *effectuer une analyse*, dans les écrits universitaires des étudiants natifs et non natifs du français. En analysant 60 textes constitués de synthèses et de résumés d'articles produits par ces étudiants au cours de leur parcours universitaire, nous avons utilisé une grille d'erreurs pour identifier et évaluer les collocations transdisciplinaires. Cette analyse a mis en lumière les erreurs fréquentes et les tendances observées. Les résultats révèlent un processus d'apprentissage en cours. Les étudiants natifs et non natifs améliorent leur maîtrise des collocations au fil du temps, mais les natifs partent d'un niveau plus élevé et progressent de manière plus uniforme. On observe une diminution des erreurs grammaticales et contextuelles chez les non-natifs en passant de Master 1 à Master 2. Cependant, ils rencontrent des difficultés importantes sur le plan lexical en M2, probablement en raison de la complexité du lexique et des structures qu'ils tentent de maîtriser. Cette analyse souligne l'importance cruciale d'une acculturation aux collocations transdisciplinaires pour améliorer la compétence linguistique des étudiants dans un contexte académique exigeant.

This article examines transdisciplinary verb-noun collocations, such as “*mener une étude*”, “*effectuer une analyse*” in the academic writings of native and non-native French-speaking students. By analyzing 60 texts composed of article summaries and syntheses produced by these students during their university studies, we used an error grid to identify and evaluate transdisciplinary collocations. This analysis highlighted common errors and observed trends. The results reveal an ongoing learning process. Both native and non-native students improve their mastery of collocations over time, but natives start at a higher level and progress more consistently. We observe a decrease in grammatical and contextual errors among non-native speakers from Master 1 to Master 2. However, they face significant lexical difficulties in M2, likely due to the complexity of the lexicon and structures they are attempting to master. This analysis underscores the crucial importance of acculturation to transdisciplinary collocations to enhance students' linguistic competence in a demanding academic context.

## INDEX

**Mots-clés** : collocations transdisciplinaires, erreurs collocationnelles, écrits universitaires, étude comparative, étudiants allophones et francophones

**Keywords** : transdisciplinary collocations, collocational errors, academic writing, comparative study, allophone and francophone students

## AUTEUR

**AZADEH PIROOZ**

Univ. Grenoble Alpes, LIDILEM, 38000 Grenoble, France  
azadeh.pirooz@univ-grenoble-alpes.fr