

**« CACHEZ CETTE ERREUR QUE JE NE SAURAI VOIR ! »
DE L'ORIGINE DES ERREURS LEXICALES DANS LES
PRODUCTIONS ÉCRITES D'ÉTUDIANTS ALGÉRIENS EN
LICENCE DE FRANÇAIS¹**

Résumé : Dans cette étude, nous nous proposons de mettre la lumière sur l'origine des erreurs lexicales dans les productions écrites d'étudiants algériens inscrits en licence de français. À l'aide de typologies des erreurs lexicales issues de la Lexicologie Explicative et Combinatoire, une analyse des erreurs a été effectuée sur un corpus composé de 68 textes rédigés en situation d'examen. Les résultats ont révélé une dominance des erreurs formelles et des erreurs morphosyntaxiques. La méconnaissance de la forme de l'unité lexicale ou de l'une de ces propriétés morphosyntaxiques, l'analogie des constructions et l'interférence constituent les hypothèses les plus avancées afin d'expliquer les formes déviantes relevées.

Mots clés : Analogie- erreur lexicale- explication des erreurs-interférence-Lxicologie Explicative et Combinatoire.

**« HIDE THIS ERROR I CAN'T SEE! »
ON THE ORIGIN OF LEXICAL ERRORS IN THE WRITTEN PRODUCTIONS OF
ALGERIAN UNDERGRADUATE STUDENTS IN FRENCH**

Abstract: In this study, we propose to shed light on the origin of lexical errors in the written productions of Algerian students enrolled in French language license. Using typologies of lexical errors from explanatory and Combinatorial Lexicology, an error analysis was carried out of 68 texts written in an examination situation. The results revealed a dominance of formal and morphosyntactic errors. The ignorance of the form of the lexical unit or of one of its morphosyntactic properties, the analogy of constructions and the interference constitute the most advanced hypotheses in order to explain the deviant forms noted.

Key words: Analogy- Lexical Error- Explanation of errors- interference - Explanatory and Combinatorial Lexicology.

Introduction

L'erreur fut pendant longtemps considérée négativement en classe de langue en raison des représentations socioculturelles et idéologiques à caractère péjoratif qu'on avait d'elle. En effet, même dans une pédagogie récente, les attitudes envers l'erreur se concrétisaient par des pratiques coercitives qui se manifestaient, entre autres, par une sanction négative dans l'évaluation de l'apprentissage, du côté des enseignants, des inhibitions, des stratégies d'évitement et même parfois le silence de peur de commettre une erreur, du côté des apprenants. Au cours des années soixante, on assiste à l'apparition de ce qu'on appelle

¹ Djelloul **Haboul**, Département des Langues Étrangères, Université Ziane Achour de Djelfa, Algérie, djelloul.didac@yahoo.fr

« l'analyse d'erreurs » conçue initialement dans le but de suppléer aux analyses contrastives dans le cas de langues maternelles non décrites ou inconnues des enseignants et des chercheurs. L'analyse des erreurs se propose d'expliquer les erreurs constantes à l'aide des catégories et des techniques de la linguistique moderne. En puisant dans les nouveaux modèles linguistiques ses grilles et ses typologies, l'analyse des erreurs n'a pas encore fait son temps et est toujours appliquée comme approche d'analyse, mais en s'orientant vers la spécification sur des domaines de la langue tels que le lexique (Luste Chaa, 2009, Anctil, 2010) ou l'orthographe (Sautot, 2020).

Nous nous proposons, dans le cadre de cette recherche, d'aborder le lexique sous l'angle de l'erreur lexicale ; nous y défendons le point de vue selon lequel l'analyse des erreurs constitue un outil efficace pour la description de la compétence linguistique chez les apprenants si elle repose sur une base empirique suffisante et des catégories linguistiques bien définies (Granger et Monfort, 1994 : 55).

1. Statut du français en Algérie

Historiquement parlant, et dans les pays du Maghreb, notamment l'Algérie, le français a bénéficié toujours d'un statut particulier. Ainsi, pendant la période de colonisation, le français a été imposé comme langue d'enseignement et d'administration ; l'arabe dialectal jouait le rôle de langue vernaculaire et de langue d'enseignement dans les écoles coraniques. Après l'indépendance, et malgré la proclamation de l'arabe comme langue nationale et officielle, le français a préservé son statut de langue privilégiée dans le paysage sociolinguistique d'alors en tant que langue d'enseignement et d'administration.

Actuellement, et malgré l'arabisation de l'enseignement ainsi que son statut officiel comme première langue étrangère, le français préserve un statut privilégié au sein de la vie sociale. Le paysage médiatique audio-visuel compte une chaîne de télévision et une radio en français ainsi qu'une dizaine de journaux et revue rédigés entièrement dans cette langue. Enfin, dans l'enseignement supérieur, le français est la langue d'enseignement dans la plupart des filières scientifiques et techniques. Qu'en est-il donc de son usage ?

Kebbas et Abbès-Kara constatent que l'Algérie est caractérisée par « une configuration linguistique quadridimensionnelle » composée fondamentalement de

l'arabe algérien, la langue de la majorité, de l'arabe classique ou conventionnel, pour l'usage de l'officialité, de la langue française pour l'enseignement scientifique, le savoir et les relations commerciales [...] et de la langue amazighe plus communément connue sous l'appellation de langue berbère [...] pour l'usage naturel d'une grande partie de la population. (Kebbas et Abbès-Kara, 2012 :35-36)

Les propos de Kebbas et Abbès-Kara montrent clairement que le paysage sociolinguistique algérien, de par sa complexité et sa diversité, pose plus que jamais la question fondamentale de la coexistence des langues traduite entre cohabitation et conflit. Une question s'impose qui consiste à se demander si cette coexistence représente une richesse à préserver ou au contraire un handicap qu'il faut tenter de surmonter.

La réalité sociolinguistique de l'Algérie démontre que le français, de par sa fonction de langue d'accès à l'information et aux connaissances, joue le rôle de *langue seconde*. Un tel constat est partagé par nombreux chercheurs (Benamar, 1997 ; Miled, 2007), pour qui, le FLE n'existerait pas dans les pays du Maghreb du moment où il s'agit de

l'enseignement/apprentissage d'une langue qui ne peut guère être qualifiée d'étrangère pour les populations de cette région. Nous adoptons donc ce point de vue sur le statut du français car il est à même de décrire le français tel qu'il est pratiqué sur notre terrain d'observation et par notre public d'informateurs, à savoir les étudiants du département de français à l'université de Djelfa.

3. Les besoins linguistiques des étudiants du département de français

Actuellement, la prise en compte des profils des apprenants et de leurs besoins linguistiques constituent la clé de voûte du *français sur objectif spécifique* (F.O.S), lui-même, lié historiquement au français fonctionnel. Il s'agit, comme le note Sophie Moirand (1982), de s'interroger sur l'identité de l'apprenant, ses motivations, ses expériences antérieures d'apprentissage, ses connaissances, ses domaines d'expérience et ses habitudes d'apprentissage. Elle souligne dans le même sillage la nécessité de faire adapter le processus d'enseignement/apprentissage au profil de l'institution d'ancrage de l'apprenant, en analysant les demandes et les objectifs de celle-là à travers le contenu de ses programmes.

Dans le cadre de cette recherche, nous nous concentrerons sur une catégorie de public FLS, à savoir des étudiants inscrits en licence de français à l'université de Djelfa. Ces étudiants poursuivent des cours de français et en français et sont censés communiquer dans cette langue en classe ou à l'intérieur du département, du moins, avec les enseignants et le personnel administratif. Le diplôme sollicité est censé leur permettre soit d'entamer une carrière d'enseignant de français au primaire ou au collège, soit de continuer leurs études en master puis en doctorat.

Les étudiants formant notre public d'informateurs possèdent déjà des connaissances de base en français à l'oral et à l'écrit acquises à l'école et pendant les deux premières années de licence. Néanmoins, ces connaissances sont loin d'être suffisantes d'autant plus que nos étudiants éprouvent toujours de nombreuses difficultés notamment en compréhension de l'écrit et en expression écrite. Ce constat devient alarmant lorsqu'on constate que les exigences des études supérieures et même la tâche d'enseignant de français au primaire et au moyen dépassent le niveau réel de nos étudiants. Faire des recherches en rapport avec le cours, préparer des résumés et des synthèses, réussir aux examens, préparer un cours, rédiger des mémoires et les soutenir, demandent de bien comprendre, de bien parler et surtout de bien écrire.

En partant des besoins linguistiques des étudiants inscrits en 3^{ème} année de licence de français, nous avons établi que la compréhension écrite et la production écrite constituent les compétences les plus exigées dans le milieu universitaire. Lire et écrire de manière efficace supposent des « *allers et retours* » assez rapides entre la représentation conceptuelle, *le signifié*, et la forme matérielle de l'unité linguistique, *le signifiant*. Cette opération nécessite la mobilisation du lexique mental au moment opportun de la réception et de la production. Il s'avère donc que le fait de disposer d'un répertoire lexical, riche et varié, influencerait fortement les performances de l'apprenant lors de l'utilisation de la langue.

4. La lexicologie Explicative et Combinatoire comme modèle lexical

Branche lexicale de la Théorie Sens-Texte, la lexicologie explicative et combinatoire (dorénavant LEC) est une théorie lexicologique qui s'oriente vers la description

lexicographique formelle. L'essentiel des descriptions lexicales qu'elle propose s'actualise dans un dictionnaire bien précis : *Le Dictionnaire explicatif et combinatoire du français contemporain. Recherches lexico-sémantiques* (Mel'čuk, et al, 1999).

La LEC est *explicative*, car elle explique le sens de l'unité lexicale (dorénavant UL) par une décomposition de ce sens en des sens plus simples, principe de la définition lexicographique dite *analytique*. La LEC est *combinatoire* en ce sens qu'elle décrit la combinatoire syntaxique et lexicale associée à chaque UL de la langue. On entend par combinatoire syntaxique «*les phénomènes habituellement désignés en linguistique par les termes de sous-catégorisation ou régime*» (Polguère, 1998:25). La combinatoire lexicale désigne, quant à elle, «*les collocations contrôlées par l'unité lexicale en question*.» (*ibid*).

Cette combinatoire est décrite dans le *Dictionnaire explicatif et combinatoire* à l'aide des *fonctions* lexicales ; Celles-ci rendent compte de façon formelle des différents liens lexicaux que l'UL peut entretenir sur l'axe syntagmatique, c'est-à-dire les liens qui connectent l'UL en question à celles avec lesquelles elle est susceptible d'entrer en cooccurrence ou *collocation*, et sur l'axe paradigmatique, c'est-à-dire les liens qui connectent l'UL décrite à celles avec lesquelles elle est liée sémantiquement ou *dérivation sémantique* (synonyme, antonyme, conversif, etc.).

La LEC offre donc une description exhaustive de l'UL dans tous aspects : sémantique, formel et de combinatoire. Les informations qu'elle fournit permettent certainement une meilleure compréhension du fonctionnement des unités lexicales au sein du lexique et une meilleure utilisation de celles-ci par les locuteurs et surtout les apprenants du français.

5. Considérations liées au corpus d'analyse

Les textes ayant constitué notre corpus d'analyse ont été rédigés en classe par des étudiants de 3^{ème} année de licence de français dans le cadre du cours de *Compréhension et Expression Écrite* (CEE) en fin d'année scolaire 2019/2020. Le corpus réuni comprend 68 textes qui comptaient de 500 à 600 mots, ce qui correspond à une longueur moyenne par rapport aux exigences du cours.

Le choix de travailler sur des textes rédigés dans le cadre du module CEE nous a conduit à recueillir des textes de type narratif (rédiger une nouvelle fantastique à partir d'un incipit, d'un proverbe, rédiger un fait divers à partir d'un titre proposé, etc.). Voici les sujets proposés :

Sujet 1 : Choisissez l'un des incipits suivants et continuez le récit pour en faire une nouvelle fantastique :

- a. « *Mon Dieu ! Mon Dieu ! Je vais donc écrire ce qui m'est arrivé ! Mais le pourrai-je ? L'oserai-je ? cela est bizarre, si inexplicable, si incompréhensible, si fou !* »
- b. « *Un homme arrive à Tanger pour passer Noël dans une maison qu'il a abandonnée depuis longtemps...* »
- c. « *Un jeune homme arrive dans une ville anglaise, il cherche une chambre à louer et la trouve chez une vieille dame bizarre qui lui offre le thé...* »

Sujet 2 : Proposez un récit conforme à la morale de l'un des proverbes suivants :

- a. *On est puni par où l'on a péché ;*
- b. *Quand le diable est vieux, il se fait ermite ;*
- c. *A l'œuvre on connaît l'ouvrier.*

Sujet 3 : Rédigez un récit à partir de l'un des titres suivants :

- a. *Le lion en cage*
- b. *Le fidèle compagnon*
- c. *Le merveilleux chat*
- d. *Le mendiant*
- e. *Une rencontre inattendue*
- f. *Ma mère*
- g. *Les cornes, secret du sultan*

Sujet 4 : Rédigez un fait divers pour l'un des titres suivants :

- a. *Millionnaire grâce à la grève des éboueurs !!*
- b. *Il voulait voler un singe ;*
- c. *Le sixième sens des éléphants a sauvé des vies.*

Comme la rédaction des textes constituant notre corpus a été faite en classe et dans des conditions d'examen, le recours au dictionnaire n'était pas permis. De plus, dans ces mêmes conditions, nous pouvions nous attendre à ce que ces textes soient rédigés dans un français standard ou soutenu.

Comme notre objectif est d'analyser des erreurs lexicales dans les textes de nos informateurs, nous nous sommes appuyé sur deux typologies : descriptive et explicative. La première a été développée par Ancil (2010) et est construite autour de trois grandes classes d'erreurs : formelles, sémantiques et de combinatoire restreinte. Cette dernière est subdivisée en trois sous classes : morphosyntaxique, lexicale et pragmatique. Elle couvre l'ensemble des composantes du signe linguistique tel que défini dans la théorie Sens-Texte où il se construit non seulement d'une forme et d'un sens, mais aussi d'une syntactique qui renvoie à sa combinatoire restreinte.

Afin de tenter de comprendre l'origine des erreurs lexicales, nous nous sommes inspiré des travaux de Dominic Ancil pour construire notre typologie qui soit spécifique au contexte de notre recherche (statut du français, type de texte, contrainte de la rédaction, etc.) (Annexe 1).

Mais avant de présenter les résultats de l'analyse de notre corpus, il s'est avéré indispensable pour la correction des textes et l'application des grilles d'analyse de définir ce qu'est *une erreur lexicale*.

6. Notre définition de l'erreur lexicale¹

Nous retenons quatre définitions de l'erreur lexicale qui nous semblent converger dans leur caractérisation de l'erreur en fonction des dimensions de l'UL qu'elle peut mettre en jeu. Mais chacune de ces définitions soulève un problème particulier qui prête à discussion :

[...] erreurs qui sont imputables à une mauvaise connaissance des propriétés inhérentes au mot, que celle-ci soit d'ordre logico-sémantique (le sens dénotatif), stylistique (le registre), collocationnel (la combinaison lexicale) ou lexicogrammatical (les propriétés morpho-syntaxique) du mot.» (Granger et Monfort, 1994 :58-59)

¹Pour une synthèse critique des différentes définitions des erreurs lexicales, voir Haboul, D. (2013), « Réflexion sur la notion d'erreur lexicale : Vers l'élaboration d'une typologie des erreurs lexicales en classe du FLE », *Didactiques*, Vol. 2, N°1, pp.125-140 <https://www.asjp.cerist.dz/en/article/24865>

Tout emploi inapproprié d'une lexie L ayant comme cause des connaissances insuffisantes de ses propriétés sémantiques, formelles et / ou de cooccurrence. Un emploi est jugé inapproprié s'il mène à l'agrammaticalité, mais aussi s'il résulte en une maladresse (Milićević et Hamel, 2007 : 29);

Résultat de la transgression de la norme concernant une propriété d'une unité lexicale (sens, forme, propriété de combinatoire) (Anctil, 2010 :351)

Toute manifestation enfrenant une ou des propriétés d'unité lexicale (sémantiques, formelles ou combinatoires). (Luste- Chaa, 2009 : 166).

Les définitions proposées par Granger et Monfort (1994) et Hamel et Milićević (2007) s'appuient clairement sur une dimension explicative de l'erreur lexicale. Ces auteurs attribuent les erreurs lexicales à des connaissances insuffisantes ou erronées sur l'UL ce qui constitue effectivement la source de ce genre d'erreur, mais elle n'est pas la seule raison générant ce problème ; les confusions entre les UL sont également attribuables à des similarités formelles ou sémantiques entre ces UL. Par ailleurs, nous ne sommes pas d'accord avec le fait qu'on introduise la dimension explicative en termes d'hypothèses sur la source des erreurs dans une définition de l'erreur lexicale, ou du moins, nous souhaiterions ne pas les présenter comme les seules pistes d'explications possibles.

Outre la dimension explicative clairement mentionnée dans la définition de Granger et Monfort (1994), deux problèmes méritent d'être soulevés. Premièrement, l'emploi du terme *mot* ne nous semble pas approprié en raison des confusions qui entourent ce terme. Deuxièmement, le manque de connaissance sur les UL devrait porter, selon ces auteurs, sur les seuls aspects logico-sémantique, stylistique, collocationnel, et lexico-grammatical ; alors que l'aspect formel est pratiquement absent de cette définition. Cet aspect est bel et bien présent dans les définitions proposées par Anctil (2010) et Luste-Chaa (2009). Pour nous, les problèmes formels méritent bien leur place si nous voulons proposer une définition générale de l'erreur lexicale aussi bien pour des natifs que pour des apprenants de L2.

Quant à la définition proposée par Anctil (2005), on lui reproche le fait d'être construite autour de la notion de norme qui n'est pas facile à intégrer dans un travail sur le lexique. La définition de l'erreur lexicale ne doit pas, selon nous, se référer directement à une norme, et surtout à une seule. C'est d'ailleurs ce qu'ont tâché d'éviter Granger et Monfort (1994) dans leur définition.

La définition de Granger et Monfort (1994) nous paraît très intéressante par ce qu'elle fait de la notion d'erreur un élément fondamental, voire même un noyau, et par ce qu'elle intègre la plupart des propriétés inhérentes à l'emploi d'une UL. Il n'en va de même pour la définition d'Olha Luste- Chaa (2009) où l'emploi du terme *manifestation* nous semble flou par rapport au terme *erreur* qui semble le plus approprié. La même remarque pour l'emploi du verbe *enfrenindre* avec le nom *propriété*. Par contre, cette définition a l'avantage de bien résumer les différents aspects pouvant faire l'objet de transgression (sémantiques, formelles ou combinatoires).

Après avoir discuté ces différentes définitions, il est temps de proposer notre définition dans laquelle nous avons essayé de réunir les points de force des définitions précédentes et en même temps d'éviter leurs points problématiques. Une erreur lexicale pourrait être définie, selon nous, comme :

Toute transgression qui concerne une propriété de l'unité lexicale, qu'elle soit d'ordre formel, sémantique, ou de combinatoire restreinte (morphosyntaxique, lexicale ou pragmatique).

Cette définition a l'avantage de pouvoir être appliquée aussi bien dans un contexte de recherche en L1 ou en L2. De plus, les propriétés du lexique mises en question manifestent une séparation entre les aspects du lexique purement formels et ceux qui relèvent de ses propriétés de combinatoire, sans négliger les liens qu'entretient le lexique avec la grammaire (morphosyntaxe) et la pragmatique.

7. Portrait descriptif des erreurs lexicales de notre corpus

Le corpus réuni composé de 68 textes a révélé un total de 731 erreurs. La moyenne est donc de 10 erreurs par rédaction. Par contre, cette moyenne est assez peu représentative puisque le nombre d'erreurs varie énormément d'un étudiant à un autre. La figure ci-dessous montre la répartition des erreurs relevées selon la propriété de l'UL touchée (forme, sens, combinatoire morphosyntaxique, combinatoire lexicale et combinatoire pragmatique) :

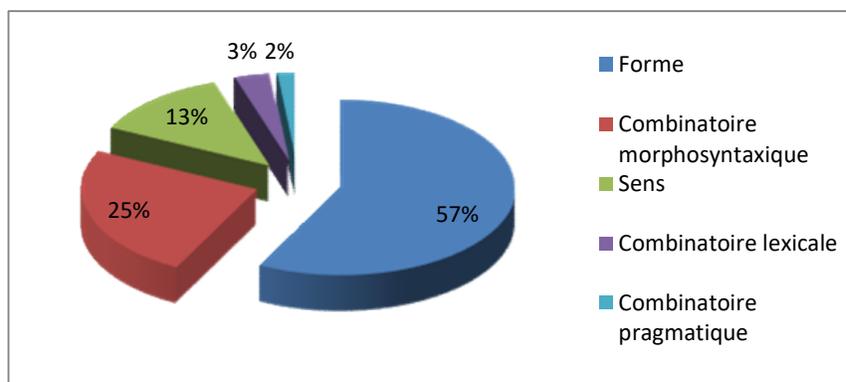


Figure 1 : Répartition des erreurs lexicales de notre corpus

L'observation de cette figure montre le nombre très élevé des erreurs formelles. Celles-ci représentent 57% du total des erreurs traitées, soit près des deux tiers du corpus. Il faut noter que plus de la moitié de ces erreurs formelles (62%) touche l'orthographe d'usage. Les autres erreurs formelles concernent les cas de signifiant déformés, les formes analytiques et les formes dérivées. Les erreurs liées à la combinatoire morphosyntaxique viennent deuxième lieu. Les erreurs affectant ces propriétés sont représentées principalement par l'utilisation de la préposition (choix d'une mauvaise préposition, absence d'une préposition régie ou ajout d'une préposition superflue), le non respect du genre nominal, le non respect du mode verbal, et l'utilisation d'un auxiliaire erroné.

Les erreurs sémantiques, quant à elles, sont dominées par la confusion entre des termes sémantiquement proches et l'utilisation de termes génériques.

Par ailleurs, les erreurs de combinatoire lexicale sont dues principalement au choix d'un collocatif inapproprié, ce qui a donné lieu à des collocations boiteuses ou inexistantes. Les problèmes liés à la combinatoire pragmatique concernent l'emploi d'un terme de registre de

langue inapproprié. Comme il s'agissait de textes rédigés dans un cadre académique, les termes appartenant au registre familier ont été relevés comme des erreurs.

8. Pistes explicatives des erreurs lexicales de notre corpus

Les hypothèses constituant notre typologie explicative ont émergé au fur et à mesure que nous décrivions les erreurs à l'aide de la typologie descriptive. Ceci dit, nous ne disposions pas d'une liste fermée d'hypothèses, mais plutôt de pistes explicatives que nous avons fixées à la fin de notre travail d'analyse. Par contre, nous ne nions pas le recours à certaines pistes explicatives déjà développées par d'autres auteurs, notamment Anctil (2010) dont nous nous inspirons en raison de l'intérêt qu'elles peuvent susciter dans une telle démarche. Voici à présent les principales sources d'erreurs proposées pour expliquer les erreurs relevées.

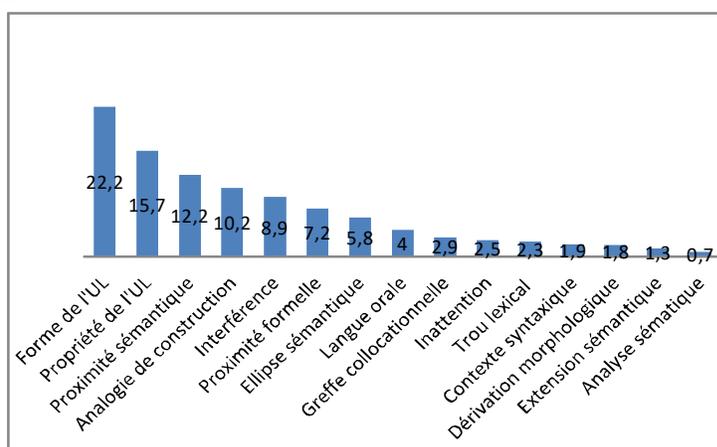


Figure 2 : Proportions des hypothèses explicatives

La figure ci-dessous présente les proportions selon lesquelles apparaissent les hypothèses avancées afin d'expliquer les erreurs de notre corpus. Nous rappelons que certaines erreurs ont fait objet de deux hypothèses considérées selon nous comme susceptibles de rendre compte des sources de ces écarts au même pied d'égalité.

8.1. Problème dans la connaissance de l'UL

Presque la moitié des erreurs analysées (49,5%) semble être due à un problème dans la connaissance de l'UL. La connaissance partielle de la forme de l'UL (PartFor) constitue à elle seule la source d'erreurs la plus fréquente parmi celles que nous avons dégagées. Cette connaissance insuffisante affectant 222 erreurs de notre corpus correspond à la plupart des cas de signifiants déformés rencontrés (222 de 279). Les 57 cas qui restent sont expliqués soit par la proximité formelle (27 cas), la distraction (21 cas), la dérivation morphologique (7 cas) et deux cas dus à l'influence de la langue orale ou d'un registre familier.

Passons maintenant à la connaissance partielle du sens de l'UL. Cette source explique les erreurs qui témoignent du fait que le scripteur ne maîtrise pas tout à fait les nuances de sens

de l'UL employée, ce qui donne lieu à des confusions entre des termes sémantiquement proches ou non. Nous distinguons trois cas de figure : l'ellipse sémantique (5,8%), l'extension sémantique (1,3%) et l'analyse sémantique (0,7%). Commençons par la plus récurrente, soit l'ellipse sémantique.

- (1) *La foudre a provoqué [bloqué] la circulation des trains.*
- (2) *Les anciens romans qui indiquaient [décrivaient, présentaient] la tradition algérienne.*

Dans ces exemples, les scripteurs ne semblent pas maîtriser les sens des UL utilisées. L'étudiant à l'origine de la phrase (1) emploie *PROVOQUER* dans le sens de *BLOQUER*. S'agit-il d'une proximité formelle ou d'une ellipse sémantique ? Nous avons tendance à écarter la première hypothèse vu que deux syllabes distinguent ces deux UL : [ro-vo]. Le verbe *PROVOQUER* est incompatible dans ce contexte et n'exprime pas le sens voulu, soit la foudre à l'origine de l'arrêt pendant une certaine période de la circulation des trains. Dans la phrase (2) l'étudiant ne semble pas conscient qu'*INDIQUER* ne véhicule pas l'idée de description avec détails censée être exprimée par un roman. Ce verbe est plutôt « léger » et ne satisfait pas le sens de la phrase. L'omission ou la non maîtrise de toutes les informations implicites traduisant le sens de l'UL conduisent donc à ces cooccurrences inhabituelles.

Si les erreurs traitées plus haut s'expliquent par une connaissance partielle de la forme ou du sens de l'UL, il n'en est pas de même en ce qui concerne les propriétés de cette dernière. En effet, nous ne pouvons parler que d'une méconnaissance lorsque le problème affecte une des propriétés de combinatoire de l'UL. Cette méconnaissance explique 17% des erreurs de notre corpus et concerne notamment les propriétés de combinatoire morphologique (8,2%) et syntaxique (7,2%) et rarement la combinatoire lexicale (0,6%). Toutefois, cette hypothèse explicative n'est associée à l'erreur que lorsque celle-ci n'est pas due à une influence d'une autre langue déjà connue par l'étudiant.

8.2. L'analogie de construction entre deux UL

L'analogie ne concerne pas seulement le sens ou la forme de l'UL, mais aussi ses propriétés de combinatoire syntaxique. Ainsi, la transposition de la structure syntaxique d'une UL à une autre figure parmi nos sources d'erreurs les plus fréquentes : elle représente 10,2 % de nos hypothèses et explique 11% des erreurs de notre corpus.

- (3) *Il a une grande importance chez [pour] son maître.*
- (4) *Le lion a profité [de] cette occasion et s'est enfui vers la forêt.*
- (5) *Emma a pris son arme et a visé vers la reine [visé la reine].*

Ces phrases présentent toutes des structures déviantes qui, en fait, correspondent aux structures d'UL synonymes ou proches sémantiquement de l'UL erronée : **avoir une grande importance chez qqn* pour *avoir une valeur chez qqn*, **profiter l'occasion* pour *saisir l'occasion*, **viser vers la reine* pour *diriger son arme vers la reine*. Les analogies de construction mises en évidence dans ces phrases semblent donc sémantiquement motivées, puisque l'UL fonctionnant avec la structure syntaxique problématique est un quasi synonyme de l'UL erronée.

Dans le même sillage, l'analogie de construction peut concerner la combinatoire lexicale dans le cas d'une greffe collocationnelle. Il s'agit, en fait, de « *collocations* ou (*expressions*

idiomatiques) déviantes, où un des éléments semble avoir été ‘emprunté’ à une autre collocation valide de la langue » (Polguère, 2007 :01). On appelle *collocation source* la collocation à la base du collocatif emprunté, alors que la *collocation cible* désigne la collocation visée par la greffe collocationnelle. Ce phénomène touche 3,1 % des erreurs de notre corpus et représente 2,9 % des hypothèses avancées. Voici quelques cas rencontrés :

- (6) *Tandis que le troisième jour était plus fertile [fécond], les offres me sont venues¹ [venues] de toute région.*
(7) *Son destin l’a ramené à Tanger, son pays maternel [natal].*

Dans la phrase (6), le collocatif *FERTILE* a été emprunté à la collocation source *la terre est fertile*, la collocation cible devrait être *un jour fécond*. La phrase (7) est similaire : l’étudiant a emprunté l’adjectif *maternel* à la collocation *une langue maternelle* (collocation source) et l’a greffé à l’UL *PAYS*, au lieu de l’UL *NATAL*.

8.3. La proximité sémantique entre l’UL erronée et l’UL visée

La présence de composantes sémantiques communes ou d’un lien conceptuel entre l’UL utilisée et celle visée semble à l’origine de 13,2% des erreurs de notre corpus. Cette hypothèse est avancée pour expliquer les erreurs sémantiques, synonymes ou ayant des sens proches comme dans les phrases suivantes :

- (8) *Et c’est ainsi que le chat merveilleux a su [compris, s’est rendu compte] qu’il ne pourra jamais changer le monde.*
(9) *Cela fait le sultan si perplexe [inquiet].*
(10) *Quand tout le monde a dormi, l’homme a écouté [entendu] un bruit au jardin.*
(11) *Après avoir signé la convention [le contrat] de location.*

Dans la phrase (8), l’étudiant a utilisé le verbe *SAVOIR* pour signifier *COMPRENDRE* ou *SE RENDRE COMPTE*, alors que ces deux verbes ne peuvent pas être des synonymes en raison des nuances de sens qui les distinguent. En effet, le verbe *COMPRENDRE* implique l’idée de connaître *les causes* ou *les motifs*, c’est donc le résultat *d’un enchaînement logique*. Cette nuance de sens est absente dans le verbe *SAVOIR* qui se limite à ‘avoir présent à l’esprit’ ou ‘être conscient de, connaître la valeur, la portée de (tel acte, tel sentiment)’. Ce verbe ne peut donc convenir dans le contexte de la phrase où le chat s’est rendu compte après plusieurs tentatives qu’il lui est impossible de changer la vie des hommes. Toutefois, l’idée de connaître contenue dans les deux UL semble à l’origine de cette confusion sémantique. Maîtriser le sens de l’adjectif *PERPLEXE* aurait pu faire éviter à l’étudiant ayant produit la phrase (9) d’employer cet adjectif pour caractériser l’état d’un sultan qui s’inquiète après avoir perdu son coiffeur personnel, celui à qui il a confié tous ses secrets. L’adjectif *PERPLEXE* s’emploie plutôt pour désigner l’état d’une personne ‘qui hésite, ne sait que penser, que faire dans une situation embarrassante’. Le lien conceptuel existant entre *ÉCOUTER* et *ENTENDRE* pourrait expliquer l’erreur dans la phrase (10). En effet, ces deux UL peuvent être des synonymes dans certains contextes, mais ne peuvent l’être dans la phrase en question ; le bruit dans le jardin a été saisi par l’ouïe de l’homme sans qu’il y applique son attention. Ce lien conceptuel explique plusieurs erreurs

¹ Il s’agit d’une erreur de l’accord qui ne constitue pas l’objet d’analyse dans le cadre de cette recherche.

sémantiques de notre corpus impliquant des verbes de perception tels *OBSERVER* ~ *VOIR*, *ECOUTER* ~ *ENTENDRE*, *REGARDER* ~ *VOIR*. Le lien est encore plus ténu entre *convention* et *contrat*, mais on sait que *CONVENTION* est un 'accord de deux ou plusieurs personnes portant sur un fait précis' alors que *CONTRAT* désigne 'l'acte qui enregistre cette convention'. L'erreur semble donc être due à une confusion entre l'accord et l'acte qui l'enregistre.

8. 4. L'interférence

L'influence de la langue première qui est l'arabe pour les étudiants ayant produit les textes de notre corpus concerne les propriétés lexicomorphologiques et lexicosyntaxiques de l'UL. Dans le premier cas, c'est le genre nominal qui semble être le plus transféré de l'arabe vers le français comme le laissent constater des erreurs telles **UNE VILLAGE* [1] = قرية, **LE RENCONTRE* = لقاء, **LE FLUTTE* = مزمار, **UN MAISON* = منزل, **UN PUBLICITÉ* = إشهار, **UNE AVION* = طائرة. Dans le deuxième cas, c'est l'emploi transitif indirect d'un verbe transitif direct qui est calqué sur l'arabe. Ainsi, certains verbes qui ne nécessitent pas de prépositions pour introduire leurs compléments sont employés avec la préposition *DE*; le verbe *VOULOIR* en constitue le cas le plus fréquent. Nous pensons que ces formes sont calquées sur l'arabe où le verbe correspondant à *VOULOIR* est bien أراد, verbe qui fonctionne avec la préposition *عند* lorsqu'il est suivi d'un deuxième verbe. Les phrases qui suivent illustrent ce cas d'interférence et sont accompagnées de leurs traductions qui laissent constater les similarités ayant conduit à ces erreurs :

(12) *Ils veulent de faire [veulent faire] une publicité.*

(13) أرادوا أن ينجزوا إشهارا

(14) *J'ai voulu de découvrir [voulu découvrir] ce château en minimum de temps.*

(15) أردت أن أستكشف هذه القلعة في أقصر وقت

(16) *Pendant la route, il a voulu de discuter [voulu discuter] avec le chauffeur à propos de la ville.*

(17) خلال الطريق، أراد أن يتحدث مع السائق عن المدينة

En somme, le genre nominal et les constructions verbales semblent constituer les sources d'interférences entre l'arabe et le français les plus fréquentes dans notre corpus. Le recours à l'anglais est rare : seulement deux cas ont été expliqués par cette hypothèse : une erreur de forme fictive (F-Fict) et une forme appartenant à un registre familier :

(18) *Tu étais un brave man [homme].*

(19) *En lui proposant un job [travail].*

La rareté des interférences avec l'anglais s'explique sans doute par le fait cette langue n'est pas d'usage courant d'autant plus que le contact avec celle-ci ne dépasse pas le cadre scolaire, encore moins la séance de l'anglais. En dehors de la classe, l'anglais est quasiment absent de l'environnement de nos étudiants.

8. 5. La parenté formelle entre l'UL erronée et l'UL visée

Un certain nombre de signifiants déformés (14,5%) peuvent être expliqués par une parenté formelle entre l'UL erronée et l'UL visée. Il s'agit de confusions entre des paronymes, c'est-à-dire des UL qui présentent des formes similaires sans aucun lien de sens (phrases 17, 18) ou des homophones qui renvoient à des UL dotées d'un signifiant unique mais d'une multiplicité de signifiés n'ayant pas le moindre élément sémique en commun (phrases 19, 20).

- (20) *C'était la célèbre mission [émission] de «Opra Wenfrey».*
- (21) *Il (le cap central de la ville de Ténès) a été instruit [construit] vers l'an 1700.*
- (22) *A haute voie [voix].*
- (23) *Des repaires [repères] touristiques.*

Cette confusion entre des UL proches sur le plan formel a également affecté des locutions comme dans les phrases suivantes :

- (24) *Mais je voulais comme même [quand même] m'y aventurer.*
- (25) *Autre mot [autrement] dit.*

Dans la phrase (21), l'étudiant a confondu entre les UL *QUAND* et *COMME* dans la locution *quand même*, sans doute parce que c'est l'élément *MÊME* qui fait que phonétiquement parlant on a l'impression que l'on prononce *comme*. Le rapprochement entre les formes *quand même* [kâmëm] et **comme même* [kômëm] est donc à l'origine de cette erreur. Le cas de la phrase (22) n'est pas différent puisqu'il s'agit d'une confusion entre les formes *AUTREMENT* [otRômâ] et *autre mot* [otRømo]. Il s'agit, en fait, d'une opposition *voyelle orale/ voyelle nasale* non maîtrisée qui est à l'origine des confusions survenues dans les locutions évoquées dans ces exemples.

8. 6. L'influence de la langue orale ou du registre familier

Le passage de certaines formes de la langue orale ou du registre familier à la production écrite est à l'origine de 4,3% erreurs analysées. Anctil (2010 :315) parle d'*intégration de l'erreur à l'usage* lorsqu'un élève commet une erreur lexicale parce qu'il a déjà lu ou entendu, cette erreur plusieurs fois. Bien sûr, cette hypothèse explique principalement les erreurs de combinatoire pragmatique, notamment celles liées au registre de langue (*UN TRUC, ÇA, EMBÊTER, SUPER*).

Dans d'autre cas, cette hypothèse est attribuée à certaines erreurs causées par l'absence d'un complément obligatoire de l'UL comme dans les phrases suivantes :

- (26) *Sauf les arbres et la mère en face [du château].*
- (27) *Mais il n'y avait personne autour [de moi], je commençais à paniquer.*

8.7. La distraction

Si certaines erreurs sont attribuables au contexte syntaxique dans lequel figure l'UL erronée, le contexte dans lequel l'étudiant rédige son texte permet d'expliquer 2,5% des erreurs de notre corpus. Il n'est pas donc surprenant de rencontrer des erreurs dues à l'absence d'une relecture ou simplement à la distraction comme **MALHEUREUSEMENT* pour

MALHEUREUSEMENT, **ÉTABIE* pour *ÉTABLI*, ou encore **IMMOBLE* pour *IMMOBILE*. Nous savons tous que lors de l'examen, l'étudiant se trouve dans une situation d'évaluation caractérisée par la peur d'échouer, de commettre des erreurs, de donner des réponses hors sujet. Tout cela s'ajoute au facteur temps : la durée des examens de fin de semestre est d'une heure et demie ; durée qui nous paraît d'ailleurs insuffisante pour produire un texte dont la qualité correspond aux attentes de l'enseignant, de l'université et surtout aux exigences du marché de travail.

Par ailleurs, nous ne nions pas qu'il soit difficile de déterminer, à partir du texte seulement, que l'erreur est due à l'inattention. C'est pourquoi, nous nous sommes assuré que cette hypothèse soit attribuée aux seules erreurs dont la forme correcte apparaît dans le texte, ce qui « innocente » en quelque sorte l'étudiant, sans vouloir bien évidemment l'exempter de sa responsabilité.

8.8. Une lacune dans le vocabulaire du scripteur

Cette explication est avancée dans le cas où l'étudiant n'arrive pas à fournir un terme précis, ce qui correspond donc à un « trou lexical » dans le vocabulaire du scripteur. Les 23 cas ayant été justifiés par cette hypothèse concernent majoritairement les formes analytiques comme dans les phrases suivantes :

- (28) *Ils font comme [imitent] les adultes.*
- (29) *Un échange de poste avec un autre employeur [une permutation].*
- (30) *Des écritures sur les murs [graffitis].*
- (31) *La ville est comme une maison d'abeilles [ruche].*

Dans toutes ces phrases, l'étudiant produit une paraphrase du sens à exprimer, plutôt que de fournir une UL précise. Ces formes analytiques s'avèrent plus ou moins satisfaisantes, mais témoignent soit du fait que l'étudiant n'a pas pu récupérer au moment de la rédaction une UL qu'il connaît déjà, et que dans d'autres circonstances il l'aurait employée comme dans la phrase (25), soit d'une lacune dans le vocabulaire de l'étudiant comme c'est le cas des phrases (26, 27, 28). Il est plutôt surprenant qu'un étudiant en 3^{ème} année de licence ne connaisse pas les UL correspondant à ces formes analytiques qui n'appartiennent pas, selon nous, à un lexique sophistiqué et rarement utilisé, mais plutôt d'un usage courant. Ce constat justifie la nécessité de s'intéresser davantage aux problèmes lexicaux chez nos étudiants.

Dans d'autres cas, l'étudiant emploie une UL qui n'exprime pas le sens visé (29) ou qui véhicule un sens proche (30). Ces cas correspondent à certaines impropriétés et aux erreurs de dérivation sémantique (31) :

- (32) *Le pauvre paysan décida d'accompagner [apprivoiser] le pauvre chien.*
- (33) *Soudain, un bruit [vacarme] règne le bus.*
- (34) *Seul un pauvre loup a dû sortir de sa caverne [trou].*

8.9. Le contexte syntaxique

Le contexte syntaxique dans lequel l'UL apparaît est à l'origine de 1,8% des erreurs de notre corpus. Ces erreurs concernent principalement la combinatoire syntaxique, notamment lors de la pronominalisation d'un complément. Voici quelques cas des plus représentatifs :

- (35) *Il avait tous ce qui lui [le] rend heureux.*
(36) *Son maître a passé plus de deux mois entrain de lui [le] chercher.*
(37) *Enfin, il lui [l'] a trouvé bléssé au bord d'une route entre deux montagnes.*
(38) *Mais cette fois-ci, personne ne va lui [l'] accompagner.*

Les erreurs dans ces phrases sont liées à la combinatoire syntaxique des verbes, mais qui présentent des compléments pronominalisés. Dans toutes ces erreurs, le complément d'objet direct est introduit par le pronom personnel *LUI* au lieu de *LE*. Nous ne pensons pas que le problème réside dans la connaissance de la structure syntaxique régie par le verbe, mais c'est plutôt la pronominalisation qui est à l'origine de ces erreurs alors que dans d'autres contextes syntaxiques, l'étudiant n'aurait pas commis ces erreurs. L'examen de ces phrases nous permet par ailleurs de conclure qu'une règle semble être intériorisée chez nos étudiants : le pronom personnel *lui* s'emploie lorsque le complément, direct ou indirect, désigne une personne. Cette règle erronée fait donc partie du système interlinguistique (ou de l'interlangue) construit par nos étudiants à ce moment de l'apprentissage du français.

8.10. La dérivation morphologique

Dans (1,9%) des erreurs analysées, les étudiants ont appliqué des règles de dérivation morphologique donnant ainsi naissance à des formes « possibles » par ce qu'elles ne sont pas, plus, ou pas encore, attestées dans le lexique conventionnel du français. Voici quelques exemples des formes produites inexistantes en français actuel :

- (39) *Le renard partira hâtement [hâtivement] au lion et lui a dit.*
(40) *J'ai vitement [rapidement] entrée.*
(41) *Comme si j'ai une toxication [intoxication].*
(42) *L'utilisateur de cette route passe par moi.*

Dans les deux premières phrases, l'étudiant a appliqué une règle fondamentale du français visant à former des adverbes à partir d'adjectifs. Cette règle consiste à ajouter le suffixe *-ment* à l'adjectif au féminin. Bien que la règle appliquée soit correcte, son application a produit deux formes adverbiales qui n'existent pas dans le lexique du français actuel : **HÂTEMENT*, **VITEMENT*. Il y a lieu de préciser que seul le TLFi atteste de la forme **VITEMENT*, mais considère son usage comme vieilli ou rare. Sans vouloir dévaloriser le niveau de nos étudiant, mais nous doutons que l'étudiant à l'origine de cette phrase connaisse cette forme et l'ait employée tout en étant conscient de sa présence dans le lexique du français. Il s'agit pour nous d'une création lexicale par l'utilisation d'un suffixe relativement productif en français (Huot, 2001 :172).

Dans les phrases (38) et (39), les formes produites sont respectivement un nom d'action et nom d'agent. Le nom d'action *TOXICATION* a été produit par l'ajout du suffixe *-ation* à l'adjectif *TOXIQUE*, alors que le nom attesté doit résulter d'une formation parasynthétique *INTOXICATION*. Encore une fois, cette forme n'est pas attestée dans le lexique du français. Le cas de la dernière phrase est moins grave puisque l'étudiant a appliqué la règle fondamentale de formation du nom d'agent par l'ajout du suffixe *-eur*, mais non « *les règles mineures postérieures aux RCM* » (Corbin, 1987 :237) qui se traduit par l'emploi des « *segments parasites* ». En français, la règle de construction d'un nom d'agent : [X verbe + eur] = Nom d'agent (ex. *BOXEUR*, *INVENTEUR*, *PENSEUR*, etc.) est souvent complétée par une règle qui ajoute un segment parasite (ex. *ADMINISTRATEUR*, *ASPIRATEUR*,

CONSOMMATEUR). Parfois, on trouve les deux formes attestées avec des significations différentes (par exemple, *serveur* et *serviteur*). La forme produite dans la phrase (39) est donc issue du non ajout de ce segment morphologique (**utilisateur* au lieu d'*utilisateur*).

Conclusion

Mener une recherche pour défricher un champ aussi vaste et complexe que le lexique ne peut prétendre couvrir tous les aspects. Cette recherche n'échappe pas à cette réalité. En effet, aborder le lexique sous l'angle de l'erreur lexicale en adoptant l'analyse des erreurs comme moyen d'investigation nécessite un travail de recherche encore plus approfondi. L'apport de notre travail réside aussi bien dans ses choix méthodologiques que linguistiques et didactiques. L'analyse des erreurs basée sur une modélisation linguistique dont l'intérêt n'a plus besoin d'être démontrée, nous semble un outil efficace pour décrire la compétence lexicale chez nos étudiants à partir d'un corpus de productions écrites.

Cependant, nous pensons qu'un corpus plus large de textes, s'il n'offre pas de résultats différents, permettrait de valider ceux déjà obtenus. De plus, les typologies d'erreurs lexicales appliquées, en raison de la complexité de leur utilisation, ne peuvent se prêter à un usage facile en classe. Soulignons également le caractère hypothétique des scénarios d'explication dégagés ; consulter les étudiants à l'origine de ces rédactions plutôt que d'émettre des hypothèses à partir de la seule trace écrite aurait pu sans doute rendre compte plus clairement des sources des erreurs commises. Le recours aux étudiants ayant produit les textes mérite de faire l'objet d'une recherche à part, par le recours à l'entretien d'explicitation où l'étudiant est appelé à verbaliser à propos d'une tâche d'écriture ou de révision de texte.

Plus concrètement, des pistes didactiques en lien avec les résultats de l'analyse doivent être mises à l'épreuve pour vérifier leur efficacité et l'intérêt qu'ils offrent dans le cadre d'une nouvelle didactique du lexique. Tout cela doit être précédé par une enquête du terrain sur les pratiques réelles d'enseignement/apprentissage du lexique dans les paliers d'enseignement précédant l'université, à savoir, le primaire, le moyen et le secondaire.

Références bibliographiques

- Anctil, D., 2010, *L'erreur lexicale au secondaire*, Thèse de doctorat de l'université de Montréal, Canada.
- Benamar, A., 1997, « Le statut polysématique du FLE dans l'enseignement-apprentissage en Algérie », *Les cahiers de l'Asdifle*, 8, p.199-206.
- Corbin, D., 1987, *Morphologie dérivationnelle et structuration du lexique*, Tübingen, Niemeyer.
- Granger, S. et Monfort, G., 1994, « La description de la compétence lexicale en langue étrangère : perspectives méthodologiques », *AILE*, 3, p.55-75.
- Haboul, D., 2013, « Réflexion sur la notion d'erreur lexicale : Vers l'élaboration d'une typologie des erreurs lexicales en classe du FLE », *Didactiques*, Vol.2, N°1, p.125-140 <https://www.asjp.cerist.dz/en/article/24865>.
- Huot, H., 2001, *Morphologie, forme et sens des mots du français*, Paris, Armand Colin.
- Kebbas, M. et Abbès-Kara, A-Y., 2012, « La variation linguistique en Algérie : facteur de facilitation et/ou de complexification de la communication ? » *Socles*, 1, p.35-44.
- Luste-Chaa, O., 2009, *Les acquisitions lexicales en français langue seconde : conceptions et applications*, Thèse de doctorat de l'université de Metz, France.
- Mel'Čuk, I. Arbatchewski-Jumarie, N., Yordanskaja, L., Mantha, S. et Polguère, A., 1999, *Dictionnaire explicatif et combinatoire du français contemporain : recherches lexico-sémantiques*, Montréal, Presses Universitaires de Montréal.

- Miled, M., 2007, « Le français langue seconde en Tunisie : une évolution sociolinguistique et une didactique spécifique » *Le français aujourd'hui*, 156, p.79-86.
- Milićević, J. et Hamel, M-J., 2007, « Analyse d'erreurs lexicales d'apprenants de FLS : démarche empirique pour l'élaboration d'un dictionnaire d'apprentissage », *Revue canadienne de linguistique appliquée*, 10-1, p.25-45.
- Moirand, S., 1982, *Enseigner à communiquer en langue étrangère*, Paris, Hachette.
- Polguère, A., 1998, « La théorie : Sens-Texte », *Dialangue*, 8-9, p.9-30.
- Polguère, A., 2007, « Soleil insoutenable et chaleur de plomb : le statut linguistique des greffes collocationnelles » <http://olst.ling.umontreal.ca/pdf/Greffescolloc2007.pdf>
- Sautot, J-P., 2020, « Typologie d'erreurs d'orthographe. Pour quoi faire ? » https://scolagram.u-cergy.fr/index.php/content_page/item/318-typologie-d-erreurs-pour-quoifaire_ffhalshs-02913518ff

Djelloul **HABOUL** est docteur en didactique des langues étrangères, maître de conférences au département des langues de l'université de Djelfa (Algérie). Ses recherches portent sur l'analyse des erreurs, et plus récemment elles s'orientent vers l'autonomie de l'apprentissage dans le cadre d'une pédagogie à distance.

Annexe 1

<i>Sources possibles des erreurs</i>		<i>Codes</i>
Connaissance partielle de la forme de l'UL		PartFor
Connaissance partielle du sens de l'UL	Ellipse sémantique	ElpSém
	Extension sémantique	ExtSém
	Analyse sémantique	AnaSém
Méconnaissance d'une propriété de l'UL		MécPro
Trou lexical	Lacune dans le vocabulaire du scripteur ou problème d'accès à un mot en mémoire	TrLex
Analogie	Proximité sémantique	ProxSém
	Proximité formelle	ProxFor
	Analogie de construction	AnaCon
	Choix d'un collocatif appartenant une autre UL sémantiquement proche (greffe collocationnelle)	GrefCo
Interférence	Influence de la langue première ou d'une autre langue déjà connue par l'apprenant	InterF
Influence du contexte syntaxique		ContSyn
Influence de la langue orale ou du registre familier		InfOra
Dérivation morphologique		DérMo
Inattention		Inatt

Tableau 1 : Notre typologie explicative des erreurs lexicales