

République algérienne démocratique et populaire
Ministère de l'enseignement supérieur et de la recherche scientifique
Université Abou-Bakr Belkaïd – Tlemcen
Faculté des lettres et des langues
Département de français



L'Année universitaire 2015 – 2016



Titre de la recherche :

**Analyse des erreurs dans les productions écrites des
élèves du moyen : le cas d'une classe de troisième
année du CEM Mahboub Abdelkader à Ouled
Mimoun**

Mémoire pour l'obtention du diplôme de Master en français.

Spécialité : les sciences du langage

Présenté par :

KHETTAB SID AHMED

BENAINI NASR EDDINE

Sous la direction de :

M. KETTAB DJAFFER

Soutenu publiquement le 05/06/2016, devant le jury composé de :

KHALDI IBTISSEM

Présidente

BENMEDJAHDI MOKHTAR

Examineur

KETTAB DJAFFER

Directeur de recherche

Le titre de la recherche :

Analyse des erreurs dans les productions écrites
des élèves du moyen : le cas d'une classe de
troisième année du CEM Mahboub Abdelkader à
Ouled Mimoun

Remerciements

Nos remerciements à nos familles pour leur soutien tout au long des années d'étude, à tous nos enseignants qui nous ont accompagné durant toutes ces années par leurs connaissances, leurs expériences grâce auxquelles ce travail a pu être réalisé.

Ainsi, nous remercions vivement notre encadrant M. KETTAB de nous avoir aidé à développer en nous le sens critique envers les mots choisis par un ensemble de connaissances et de conseils précieux.

Le sommaire

Introduction	5
Partie théorique	3
Chapitre I : La langue maternelle	3
Chapitre II : La linguistique contrastive	10
Chapitre III : L'analyse des erreurs	17
La partie pratique	27
Chapitre I : Définitions des éléments clés	27
Chapitre II : La construction d'une typologie	45
Chapitre III : Description et interprétation des résultats	48
Conclusion	79
Annexes	73
Bibliographie	95

Introduction

Introduction

La naissance de la linguistique comme étant l'étude scientifique des langues a changé notre conception vis-à-vis des langues. La langue est devenue pour la première fois un objet d'étude. Dans cette perspective, la langue est considérée comme un système ; un système linguistique qui dispose d'une organisation interne.

La linguistique continue son développement à travers le temps et les travaux en linguistique foisonnent donnant naissance à de nouvelles théories et approches.

A travers le temps, les linguistes commencent à s'interroger sur les mécanismes d'acquisition et d'apprentissage des systèmes linguistiques.

Le terme d'*acquisition* évoque le terme d'*apprentissage* d'une langue. Dans ce cadre, on parle souvent d'acquisition des langues maternelles et d'apprentissage des langues étrangères. Dans le deuxième cas l'apprenant met en contact au minimum deux systèmes linguistiques ; le premier est obligatoirement acquis alors que le deuxième est à acquérir par un processus d'apprentissage.

En effet, la confrontation des langues dans l'apprentissage et les influences négatives que peuvent véhiculer les unes sur les autres est l'une des hypothèses fondamentales de la linguistique contrastive, appelée aussi parfois « différentielle ». Cette idée n'est pas arbitraire puisque les linguistes appartenant à ce champ d'étude ont constaté qu'en effet, ce sont les différences entre les différents systèmes linguistiques qui sont à l'origine des difficultés observées durant l'apprentissage des langues étrangères.

La linguistique contrastive n'est pas la seule qui s'intéresse généralement aux erreurs qui peuvent être le résultat d'une interférence entre les langues. L'analyse des erreurs qui vient combler les lacunes de la première dans ce sens, les traite d'une façon plus avancée, voire plus sophistiquée, partant de l'idée que l'apprentissage ne s'effectue qu'à travers les erreurs.

En effet, l'apprentissage de n'importe quelle langue étrangère induit l'apprenant à commettre des fautes et des erreurs. Généralement parlant, on mesure le niveau des compétences par rapport à leur occurrence. Donc on estime qu'il y a une corrélation négative entre les erreurs commises et le niveau des apprenants ; un nombre réduit d'erreurs implique une compétence élevée et vice-versa. Cette relation entre compétences et erreurs a stigmatisé l'erreur en lui attribuant une connotation péjorative. Toutefois, cette vision négative à l'égard des erreurs des apprenants des langues étrangères a été remise en cause puis dépassée. Désormais, on attribue à l'erreur un rôle crucial et indispensable, voir incontournable dans le processus d'apprentissage.

Notre recherche s'inscrit dans ce mouvement qui accorde de l'importance à l'erreur afin de faciliter l'apprentissage des langues étrangères. A cet effet, nous allons essayer d'analyser les erreurs des productions écrites des élèves d'une classe de troisième année de moyen du CEM Mahboub Abdelkader à Ouled Mimoun, Tlemcen, et ce dans le but de comprendre les difficultés qu'éprouve l'apprenant au cours de sa production écrite.

Afin d'atteindre ce but, nous allons essayer tout d'abord de demander aux élèves en question de nous rendre une production écrite en langue étrangère accompagnée d'une traduction ou explication de leur production afin de comprendre le sens dans le cas où ce dernier n'est pas clair, ou bien pour comprendre les lettres et ce dans le cas où elles sont mal écrites. Ensuite, nous identifierons les erreurs suivant une typologie que nous allons construire à travers les définitions que nous trouvons dans le livre

introductoire au traitement automatique des langues naturelles, écrit par Pierrette Bouillon (1998) et ses collaborateurs. Outre la grande importance de la typologie dans l'identification et la classification de l'erreur, elle va nous permettre par la suite d'effectuer des analyses statistiques à travers lesquelles nous essayerons d'aboutir aux résultats qui mettent en évidence les *réurrences* qui ont le sens de *difficulté*.

Les buts de la recherche sont donc, d'une part, essayer de savoir les difficultés qu'éprouve les apprenants au cours de leur productions écrites à travers les réurrences des erreurs. Et, d'autre part, celui de rendre un bilan constituant des « *informations d'importance didactique* » (Desoutter :120). Nous essayerons, si possible, d'expliquer les causes des erreurs en les catégorisant comme des erreurs interlinguales qui résultent d'une interférence entre deux langues, ou comme des erreurs intralinguales qui n'émergent que comme le résultat d'un manque de connaissance des règles de la langue en apprentissage.

Pour la collecte des données, nous chargerons l'enseignante responsable de la classe de la mener en nos lieux et places pour placer les élèves dans un contexte naturel sous lequel leurs productions écrites puissent refléter leurs niveaux plus ou moins naturels.

Afin d'aider les élèves de nous rendre le maximum de contenu, nous leur proposons une question qui ne demande qu'une petite réflexion comme nous leur permettons d'utiliser des dictionnaires et même poser des questions à leur enseignante ou mutuellement entre eux.

D'une façon générale, la population choisie pour cette enquête semble homogène dans la mesure où les langues qu'elles partagent sont les mêmes : généralement, la langue maternelle et de socialisation de la plupart est l'arabe algérien ; la langue de scolarisation à travers laquelle la plupart des matières sont inculquées est l'arabe moderne, alors que le français constitue la première langue étrangère qu'on

commence à apprendre dès la troisième année du primaire. Les élèves de ce niveau ont déjà passé cinq années d'apprentissage de cette langue.

Le type d'enquête de cette recherche est descriptif puisque nous envisageons d'utiliser les méthodes de la statistique descriptive qui synthétisent les informations d'une façon efficace. La recherche est aussi analytique puisque l'un de nos buts est d'essayer, si possible, de trouver les causes des erreurs.

Nous pouvons résumer la problématique de cette recherche dans les questions suivantes :

Quelles sont les erreurs les plus récurrentes ?

Quelles sont les véritables causes de ces erreurs ?

Nous signalons que cette recherche est de type inductif. Au contraire aux recherches hypothético-déductives, celle-ci ne nécessite pas forcément d'hypothèses. Le but donc est de découvrir les récurrences des erreurs en passant par les phases suivantes : l'identification de l'erreur ; la classification de l'erreur en catégories et en fin l'utilisation de la statistique afin de quantifier les récurrences. La phase explicative consiste dans l'identification de l'origine de l'erreur comme nous l'avons déjà mentionné, pourrait être interlinguale ou intralinguale.

En ce qui concerne l'organisation générale du document, nous préférons diviser la recherche en deux parties fondamentales : en partie théorique, dans laquelle nous expliquons les grandes lignes des théories et concepts qui ont une relation directe avec cette recherche ; et en une partie pratique, dans laquelle nous expliquons en détails les éléments clés de la recherche, nos prises de positions et la méthode qu'on a suivie pour créer et appliquer une grille d'analyse avant de passer à la description et l'explication des résultats.

La partie théorique va comporter trois chapitres. Dans le premier, nous allons essayer d'initier le lecteur à la notion de *langue maternelle* qui se répétera tout au long de la recherche dans les autres chapitres et parties. Dans le deuxième, nous essayerons d'initier le lecteur à la linguistique contrastive de multiples perspectives avant d'entamer le dernier chapitre de cette partie dans lequel nous exposerons la relation entre la linguistique contrastive et l'analyse des erreurs qui vient combler le vide que laisse la première dans plusieurs points concernant l'erreur.

De l'autre côté, la partie pratique va comporter trois chapitres : dans le premier, nous allons rendre une description détaillée des éléments clés de la recherche : la population, la période, le milieu et les circonstances dans lesquelles se déroulait l'enquête. Dans le deuxième, nous essayons d'expliquer comment nous avons constitué et appliqué une grille d'analyse. Dans le troisième chapitre, nous allons décrire le résultat des analyses statistiques et en même temps essayer, si possible, de trouver les causes des erreurs.

Partie théorique

Chapitre I :

La langue maternelle

Dans ce chapitre nous essayerons de présenter un ensemble de définitions propre à la notion de *langue maternelle* tout en montrant les limites de chacune étant donné le fait que cette notion « [...] est bien loin d'être aussi claire et aussi stable que le laconisme rassurant des dictionnaires pourrait le laisser croire : » (Urbain, 1982:7)

Le terme de langue maternelle est primordial dans l'identification de la langue antérieure à celle de l'apprentissage. Néanmoins, il s'avère que les réalités linguistiques de différents pays nécessitent l'élaboration, l'adaptation d'une ou de plusieurs définitions propres à chaque réalité.

1. La langue maternelle

1.1. L'impasse définitoire

En effet, les définitions trouvées dans les dictionnaires, concernant le terme langue maternelle, sont souvent concises ne reflétant pas ainsi la réalité linguistique complexe de cette notion qui suscite beaucoup de questions telles que celles posées par Jean Didier Urbain (1982) :

« Qu'est-ce au juste que la langue maternelle ? Existe-t-elle seulement ? Ne s'agit-il pas plutôt d'une fiction théorique ? Et si elle existe vraiment, comment existe-t-elle ? D'où vient-elle ? Quel est son statut social ? Sa fonction ? Quel est son sens intrinsèque ? Et sur la base de quels principes ou dans quel(s) contexte(s) socio-historique(s) a-t-elle été différenciée puis perpétuée ? Cette langue a-t-elle seulement un sens et est-il intrinsèque ? Rien n'est évident en ce qui concerne en propre la langue maternelle. » (Urbain, 1982 :8)

Louis Dabène pose le même problème en s'interrogeant :

« Comment pourrait-on ne pas savoir ce qu'est la langue maternelle dès lors que tout locuteur en possède une et que celle-ci est bien souvent un des traits les plus fondamentaux de son identité ? » (Dabène, 1994 :8)

Il est clair que le concept linguistique de *langue maternelle* est sujet de controverse, loin de faire l'unanimité, ce concept continue à engendrer des polémiques linguistiques.

D'emblée, il apparaît comme un concept univoque, simple à définir et à cerner. Toutefois, *« nous nous trouvons en face d'une entité sémiotique polymorphe et équivoque » (Urbain, 1982 :8)*

Selon Louise Dabène (1994 :15), la conscience de ce problème a poussé des chercheurs à abandonner cette expression au profit d'autres termes jugés moins ambigus. *Dans la même perspective Jean-Didier Urbain, note :*

« En effet, depuis longtemps déjà, des linguistes ont essayé de contourner cette notion, finalement fort embarrassante, en la requalifiant sans cesse, en la spécifiant diversement ou en édifiant tout un paradigme de locutions « équivalentes ». W. D. Whitney suggère ainsi une équivalence entre langue maternelle et langue native. » (Urbain, 7-8)

1.2. Vers un consensus définitoire

Afin d'aboutir à un consensus définitoire, Louise Dabène (1994 :10-15) s'interroge sur les différents critères qui peuvent caractériser une langue et lui attribuer le qualificatif de langue maternelle : s'agit-il de la langue de la mère ? la première langue acquise ? la langue la mieux connue ou celle acquise naturellement ? Ou par contre, une langue comme un système unique à travers lequel s'opèrent des fonctions symboliques primordiales, et le développement de la personnalité ?

1.2.1 La langue de la mère

Le concept, qui vient de l'étymologie, selon lequel la langue maternelle est la langue de la mère, même s'il semble irréfutable n'est pas valide pour toutes les sociétés. Cette définition ne s'applique pas au sein de certaines sociétés ou groupes ethniques à l'instar de la Gambie ou certaines tribus amazoniennes. Dans ces cas-là les enfants sont souvent exposés aux langues des autres familles puisqu'ils sont souvent élevés par d'autres familles.

1.2.2 La première langue acquise

Dans l'ordre d'acquisition, la langue maternelle est forcément la première, Cette position lui donne un privilège, puisqu'elle est acquise dans la période la plus favorable de la vie. Cette antériorité d'appropriation est synonyme d'une compétence supérieure par rapport aux langues étrangères. C'est-à-dire, un des critères qui distingue la langue maternelle est l'antériorité d'appropriation. Néanmoins, Il est clair que la règle ne peut pas être appliquée à toute une société où au lieu de parler d'une langue maternelle, on parlera des langues maternelles.

1.2.3 La langue la mieux connue

Ce critère-là ne fait pas, à l'instar des autres, l'unanimité parmi les spécialistes à cause de la complexité des réalités linguistiques. Ce critère peut être appliqué à certains groupes, communautés mais non pas à d'autres.

1.2.4 La langue acquise « naturellement »

La notion d'acquisition est toujours présente quand on évoque le concept de la langue. Le processus d'acquisition de la langue maternelle est un processus naturel. C'est-à-dire, il n'y a presque pas d'efforts à fournir lors d'acquisition. Comme il n'y a pas d'intermédiaire quelconque pour faciliter cette acquisition. L'enfant acquiert la langue maternelle par le contact et l'interaction avec son entourage, le plus souvent la famille. Cette définition est clairement limitée du fait que la langue maternelle n'est pas la seule qu'on acquiert naturellement.

1.2.5 Langue et langage

Il semble que l'organisation de la fonction langagière en tant que fonction symbolique primordiale, aussi bien que la construction de la personnalité s'opère à travers la langue maternelle comme un système linguistique unique qui exclue les autres. Toutefois, rien ne prouve que ces fonctions et constructions de la personnalité relèvent d'un seul système linguistique. Bien au contraire, pour une portion importante de l'humanité, ces fonctions sont le résultat d'une synergie de différents systèmes linguistiques.

Louise Dabène (1994) explique l'ampleur de l'impasse définitoire de la notion langue maternelle en se basant sur les critères déjà mentionnés et ajoute :

« [...] , aucun des critères définitoires généralement invoqués ne résiste vraiment à l'analyse, et ne suffit à construire solidement le concept. [...] il ne faut pas s'étonner dès lors de la difficulté rencontrée dans la définition précise d'un concept qui ne correspond, en fait, à aucune réalité scientifiquement maîtrisable. » (1994 :15)

En essayant d'adapter ces définitions à la réalité linguistique de la sphère dans laquelle nous menons cette recherche, nous considérons généralement la langue maternelle comme la première langue de socialisation ; la langue de premier contact avec l'entourage parental immédiat et celle performée au sein de la société. Dans cette perspective, la famille est considérée comme une micro société qui incorpore la société. On ne choisit pas sa langue maternelle par ce qu'elle nous est imposée dès notre naissance.

Afin de sortir de cette impasse définitoire et de cette polémique linguistique, nous, auteurs de cette modeste recherche allons parler, simplement, des acquis linguistiques ou langagiers antérieurs par rapport à l'apprentissage d'une langue étrangère, en l'occurrence, la langue française en utilisant les concepts langue source et langue cible, concepts appartenant au jargon scientifique de la linguistique contrastive.

Dans le chapitre suivant, nous allons essayer d'exposer la linguistique contrastive comme champ d'étude ayant relation avec les notions d'*erreur*, de *langue maternelle* et de l'*apprentissage*.

Chapitre II :

La linguistique contrastive

1. La linguistique contrastive

La linguistique contrastive pourrait être définie d'une façon simple comme « [...] une branche de la linguistique appliquée dont l'objectif est la comparaison des systèmes linguistiques de deux ou de plusieurs langues afin de faciliter leur enseignement et leur apprentissage ainsi que la traduction. » (Sóres, 2008 :18)

Les premières études contrastives remontent aux années 1950 aux Etats-Unis. L'ouvrage de Robert Lado (1957) *Linguistics Across Cultures* est considéré comme l'ouvrage fondateur de la linguistique contrastive, dans lequel R. Lado a élaboré les fondements théoriques de cette discipline sur des bases déjà échafaudées par Ch. C. Fries, qui est le pionnier de ces recherches. Et parmi les autres figures emblématiques de cette discipline on trouve les noms de Uriel Weinreich, R. L. Politzer, Ch. A. Ferguson, R. P. Stockwell et J. B. Carrol. Il faut noter ici que U. Weinreich, est considéré aussi comme un des pères fondateurs de ce champ d'étude.

Lado dans la préface de son ouvrage *Linguistics Across Cultures* (1957) met en évidence l'hypothèse de la linguistique contrastive qui consiste dans la prédiction et la description des structures qui peuvent induire aux difficultés d'apprentissage :

“The plan of the book rests on the assumption that we can predict and describe the patterns that will cause difficulty in learning, and those that will not cause difficulty, by comparing systematically the language and culture to be learned with the native language and culture of the student.”¹ (Lado in Wardhaugh 1970:3)

¹ « Le plan du livre se base sur l'hypothèse qui nous permet de prédire et décrire les structures qui vont provoquer des difficultés dans l'apprentissage, et ceux qui ne posent pas de difficultés,

Comme nous l'avons déjà signalé, Lado n'est pas le premier à faire une introduction à la linguistique contrastive, mais le premier travail revient à celui de Fries. Selon Wardhaugh Ronald, (1970) Lado continue à citer Fries afin d'appuyer sa proposition en se référant à une de ses citations *dans son livre Learning English as a Foreign Language (1945)* ; dans laquelle apparaît le terme de *comparaison* entre deux langues qu'on décrit scientifiquement :

“The most efficient materials are those that are based upon a scientific description of the language to be learned, carefully compared with a parallel description of the native language of the learner.”² (ibid.)

Ainsi, la naissance de la linguistique contrastive coïncide avec l'émergence d'une grammaire comparée qui avait les mêmes vocations de faciliter l'apprentissage des langues étrangères, Francis Debyser (1970) note que c'est :

en comparant systématiquement la langue et la culture qu'on envisage à apprendre avec la langue native et la culture de l'apprenant. » [Notre traduction]

² « *Les matériels les plus efficaces sont ceux qui se basent sur une description de la langue à apprendre, soigneusement comparée avec une description parallèle de la langue native de l'apprenant* » [Notre traduction]

« À partir des années 50, que l'on parvint à établir des grammaires comparées d'un nouveau genre, afin de faciliter l'apprentissage de langues étrangères, et, d'une manière générale, le passage d'une langue à une autre. C'est ainsi qu'est née, dans une perspective d'application, la linguistique contrastive dont les ambitions de départ étaient qu'une comparaison « terme à terme, rigoureuse et systématique » de deux langues et surtout de leurs différences structurales ... » (1970 :31)

Parmi les ambitions qui ont contribué à la naissance de la linguistique contrastive est le besoin d'élaborer une pédagogie spécifique et différenciée des langues étrangères, partant de la supposition que l'apprentissage de la langue maternelle (L1), appelée aussi dans ce contexte, la langue source, ne pose pas les mêmes problèmes que l'apprentissage des langues cibles.

La linguistique contrastive ne trouve sa valeur et sa signification qu'en présence de deux ou plusieurs langues. Elle contraste au minimum deux systèmes, deux codes dans une perspective d'apprentissage. En l'occurrence, la langue source et la langue cible. Son rôle majeur est de prévoir, décrire et d'expliquer les difficultés dues à l'influence de la langue maternelle sur la langue cible. Cette linguistique dite aussi différentielle s'intéresse surtout aux différences entre les langues en contact.

Banathy, Trager et Waddle dans Valdman (1966) renforcent l'hypothèse de la linguistique contrastive en affirmant implicitement l'existence des problèmes liés à l'apprentissage à cause des différences dans la culture et la langue native de l'apprenant et celles de la langue étrangère en apprentissage. Afin de dépasser ces problèmes, ils mettent en évidence les tâches de chaque acteur impliqué dans la modulation de la sphère de l'apprentissage des langues étrangères dans les lignes suivantes :

“ . . . the change that has to take place in the language behavior of a foreign language student can be equated with the differences between the structure of the student's native language and culture and that of the target language and culture. The task of the linguist, the cultural anthropologist, and the sociologist is to identify these differences. The task of the writer of a foreign language teaching program is to develop materials which will be based on a statement of these differences; the task of the foreign language teacher is to be aware of these differences and to be prepared to teach them; the task of the student is to learn them.”³ (Banathy, Trager et Waddle dans Wardhaugh 1970 :3-4)

En effet, la linguistique contrastive comme son nom l’indique confronte au minimum deux systèmes linguistiques à différentes échelles, l’un est obligatoirement acquis pendant que l’autre est à acquérir. Le but de cette approche est de faciliter leur enseignement et leur apprentissage.

La linguistique contrastive recoure à la linguistique descriptive dans une perspective d'application, « *La linguistique contrastive a nécessairement pour objet l'application. Bien qu'elle rende parfois hommage, en passant, à Bopp et à W. Von Humboldt, elle n'a pratiquement rien hérité de la grammaire comparée du XIXe siècle,*

³ « ... le changement qui doit prendre place dans le comportement langagier d'une langue étrangère d'un apprenant serait assimilé aux différences dans les structures de la langue et la culture de l'étudiant et celles de la langue et de la culture cible. La tâche du linguiste, l'anthropologue culturel et le sociologue est d'identifier ces différences. La tâche de l'écrivain d'un système d'enseignement d'une langue étrangère est de développer du matériel fondé sur un relevé de ces différences ; la tâche de l'enseignant de la langue étrangère est d'être conscient de ces différences et de se préparer afin de les enseigner ; la tâche d'un étudiant est de les apprendre. » [Notre traduction]

et n'est ni classificatoire ni philologique ni historique. Elle utilise, sans se substituer à elle, la linguistique descriptive... » (Debyser, 1970 :32).

La linguistique contrastive part de l'idée que tous les êtres humains partagent un seul langage. Le langage est universel pendant que leurs langues sont différentes. Et ce malgré la diversité des langues, la linguistique contrastive dispose d'un instrument théorique et métalinguistique assez général et assez unitaire pour rendre comparable des objets différents. C'est ce rapport entre l'unité du langage avec la pensée qui a fondé cette discipline.

De prime abord, la linguistique comparative semble identique à la linguistique contrastive, mais en réalité la première se distingue de la deuxième dans plusieurs points : la linguistique comparative est plus théorique et éminemment diachronique, notamment lorsqu'elle essaie de prouver que deux ou plusieurs langues sont descendantes d'une seule protolangue par la comparaison d'une liste de termes. La protolangue peut être reconstruite à partir des langues descendantes. C'est une technique utilisée pour démontrer la relation génétique entre certaines langues. Cependant, l'aspect synchronique est également à considérer lorsqu'il s'agit de comparer deux langues ou plus à une époque précise. La linguistique comparative a trouvé son élan à travers l'étude des langues indo-européennes.

En ce qui concerne la linguistique contrastive, Lado affirme en se basant sur des données déjà échafaudées par Fries que les éléments similaires à la langue maternelle sont faciles à assimiler alors que les éléments différents ne le sont pas. Il affirme aussi – influencé par l'école béhavioriste – que ce transfert négatif ou interférence est le résultat de transfert des habitudes de la langue source vers la langue cible. Voici une citation prise de son ouvrage *Linguistics across culture* (1957), qui affirme explicitement ce qui a été dit :

“The student who comes in contact with a foreign language will find some features of it quite easy and others extremely difficult. Those elements that are similar to his native language will be simple for him, and those that are different will be difficult.”⁴ (Lado in Desoutter, 117)

2. Le terme d’interférence

Les études sur le bilinguisme ou les "langues en contact", qui ont été menées par U. Weinreich (1953) ont donné naissance au terme d’« interférence » pour faire référence à toutes les différences qui existent entre le discours d’un monolingue et d’un bilingue ; dans ce cadre, les interférences constituent des déviations par rapport à la norme de chaque langue que parle le bilingue et que l’on constate dans son discours. Ainsi, un locuteur bilingue pourra commettre des interférences aussi bien dans sa première langue que dans sa seconde. Il est important de signaler qu’en linguistique contrastive, le terme d’*interférence* vient remplacer des termes comme « transfert négatif » qui est considéré comme un terme véhiculant une connotation péjorative.

Pour mieux cerner le phénomène interférentiel, il est nécessaire de l’approcher de différents angles en l’examinant selon diverses disciplines. Francis Debyser (1970) expose les définitions générales du terme en psychologie, linguistique et pédagogie.

⁴ « L’apprenant qui prend contact avec une langue étrangère va trouver quelques éléments faciles et d’autres extrêmement difficiles, ces éléments qui sont similaires à ceux de sa langue native lui seront simples, mais ceux qui sont différents lui seront difficiles. » [Notre traduction]

2.3. En psychologie

En psychologie, l'interférence linguistique est perçue comme une contamination de comportement, une habitude antérieure qui va exercer une influence négative sur une nouvelle habitude. Le langage peut être défini, dans ce contexte, comme un ensemble de comportements, S'agissant de comportements acquis et appris. Ces comportements sont fixés en habitudes linguistiques. En effet ; il y a une force inconsciente, omniprésente et bien antérieure, à savoir ce désir ou tentation refoulée de s'exprimer dans la langue maternelle. En effet, ces comportements de la langue maternelle entravent l'acquisition d'autres comportements linguistiques. (Ibid. :34-45)

2.4. En linguistique

La linguistique structurale considère l'interférence comme un accident de bilinguisme situé au niveau de la parole. En effet, le lieu de rencontre entre deux langues est toujours un locuteur individuel, bilingue. Et cela quel que soit le contexte sociologique du bilinguisme. (*Ibid.*)

La grammaire générative de N. Chomsky (1957) illustre mieux le phénomène d'interférence. Dans cette perspective, on peut distinguer deux situations de bilinguismes, l'un est parfait tandis que l'autre s'inscrit dans un contexte d'apprentissage par un apprenant d'une langue étrangère. Dans la première, le locuteur bilingue possède deux compétences. Tandis que dans la deuxième, il y a un conflit de deux compétences différentes. Dans cette dernière situation le phénomène interférentiel apparaît au niveau de la performance de l'apprenant. (Debyser, 1970 :34-59)

2.5. En pédagogie

En pédagogie des langues vivantes, on emploie souvent des termes, à l'instar de « déviations », de « glissements », de « transferts », de « parasites », pour qualifier l'interférence qui est considérée comme un type particulier de faute, un obstacle qui entrave le processus d'apprentissage par l'émergence des éléments de la langue maternelle dans la langue cible sous l'effet des habitudes ou de structures déjà acquises. (Ibid. :34-61)

Comme nous l'avons exposé suivant l'explication de Francis Debyser, le concept d'interférence est défini différemment selon chaque discipline. Toutefois, en linguistique contrastive, et c'est ce qui nous intéresse le plus : l'interférence est l'usage, dans la langue cible, des éléments appartenant à une autre langue ; la langue source ou maternelle. Cette définition est équivoque, puisqu'elle renvoie aussi à la notion de code-switching ou alternance codique qui est le passage conscient d'une langue à une autre, alors que l'interférence est involontaire et accidentelle.

Le concept d'interférence linguistique est le pivot de la linguistique contrastive. Ce concept a été échafaudé par les linguistes afin de remplacer la notion de transfert négatif qui avait comme nous l'avons déjà dit une connotation péjorative. On parle d'interférence quand on rencontre dans la langue cible, la langue visée par l'apprentissage, des éléments linguistiques indésirables appartenant à la langue source, ou maternelle. Le passage suivant par Sörös (2008) résume ce que nous venons d'exprimer comme suit :

« La langue maternelle exerce une influence qu'on appelle « le transfert négatif » sur l'apprentissage de la langue étrangère. Afin d'éviter la nuance négative, le phénomène aussi appelé « interférence » : un trait de L1 apparait dans L2, ou un trait de la langue « source » fait son apparition dans la langue « cible ». (2008 :18)

3. Les limites de la linguistique contrastive

L'analyse contrastive a reçu des critiques portant sur la méthode même de la contrastivité, ainsi que la valeur prédictive des analyses contrastives. On a réalisé très tôt qu'il est très difficile, voir, dans certains situations, utopique de recourir à l'analyse contrastive.

Robert Lado (1957) a peut-être négligé la complexité et le dynamisme des processus d'apprentissage. Outre la valeur prédictive de la discipline qui est limitée aux niveaux linguistiques qui ne peut pas être toujours à l'origine de l'erreur puisque l'apprentissage des langues étrangères ne se limite pas seulement à l'appropriation de ses systèmes linguistiques, mais s'étend encore aux conditions d'apprentissage, la cognition de l'apprenant et d'autres facteurs extralinguistiques. L'intégration de la pragmatique dans la discipline vient donner plus d'extension à la pratique, constituant un grand développement. De ce point de vue, les démarches linguistiques ne se limitent plus à la description de tel ou tel élément, mais aussi à expliciter sa valeur dans le contexte.

Maintenant que nous avons exposé les définitions et les conditions dans lesquelles la linguistique contrastive est née et s'est développée, nous allons essayer d'exposer une autre approche qui traite les erreurs d'une façon systématiques, et qui a une relation directe avec notre recherche.

Chapitre III :

L'analyse des erreurs

1. L'analyse des erreurs

Partant d'une vision empirique, les chercheurs ont développé, à partir des années 1960, une nouvelle approche qui est en étroite relation avec l'apprentissage et l'acquisition des langues étrangères. *L'analyse des erreurs* a été, donc, développée afin d'appréhender les différentes contraintes qui entravent l'assimilation des langues cibles.

Le pionnier de cette nouvelle approche est l'illustre P. Corder, qui a échafaudé les bases et les principes de cette approche en se basant plus au moins sur des travaux antérieurs échafaudés par d'autres chercheurs, à l'instar de N. Chomsky, afin d'élaborer certains concepts. Dans cette perspective Corder explique que :

« On a vu dans l'étude contrastive exhaustive des systèmes linguistiques de la langue étrangère (LE) et de la LI de l'apprenant une contribution majeure des linguistes à l'enseignement des langues. Il devait en ressortir un inventaire des zones de difficultés rencontrées par l'apprenant, cet inventaire contribuant à orienter l'attention des enseignants vers les zones justiciables d'un soin et d'une insistance particuliers, afin de surmonter, ou même d'éviter, les difficultés ainsi prévues. Une telle contribution n'a pas toujours convaincu les enseignants, déjà familiarisés, de par leur expérience pratique, avec ces difficultés. Aussi n'ont-ils pas perçu cette contribution des linguistes comme un apport d'informations nouvelles et pertinentes. Ils ont constaté, par exemple, que nombre d'erreurs qui leur étaient familières n'étaient de toute façon pas prévues par les linguistes. Les enseignants se souciaient moins en effet de l'identification de ces difficultés que de la façon de les traiter, d'où ils ont estimé, à juste titre, que les linguistes n'avaient finalement que peu d'informations à leur apporter. » (Corder, 1980 :9)

En effet, la vision essentiellement prédicative, de la linguistique contrastive s'est avérée insuffisante, comportant de nombreuses lacunes.

Toutefois, il est crucial de noter dans cette perspective que l'analyse des erreurs a émergé comme une nouvelle discipline suite aux travaux menés en linguistique contrastive, En d'autres termes, la linguistique contrastive précède l'analyse des erreurs dans le temps.

En outre, il faut recourir impérativement à la linguistique contrastive afin d'expliquer certains mécanismes qui contribuent à l'émergence de certaines erreurs. En d'autres mots, dans les phases explicatives des erreurs, la linguistique contrastive s'avère, parfois, incontournable.

En effet, Cette procédure analytique vient pour compléter et non pas supplanter la linguistique contrastive. La relation donc entre ces deux approches est réciproque. Il y a ici simplement un changement de stratégie :

« Les chercheurs travaillant avec du matériel réel produit par les apprenants n'ont pas pu ne pas remarquer qu'un grand nombre des erreurs prédites par CA n'apparaissaient pas du tout, alors que de nombreuses erreurs apparaissaient sans avoir été prédites par CA. Dans une telle situation, il est naturel de renverser la procédure en démarrant du matériel véritable, d'abord en le décrivant systématiquement, ensuite en analysant les possibles causes d'erreurs. » (Ringbom, 1994)

Selon Corder (1980), partant d'une perspective méthodologique, il y a deux courants de pensées en ce qui concerne les erreurs des apprenants : le premier réside dans l'élaboration des méthodes adéquates afin de prévenir et d'empêcher l'apparition des erreurs, en l'occurrence, l'apparition des erreurs des apprenants est le résultat d'une

technique pédagogique inadéquate. Le deuxième conceptualise que l'ingéniosité humaine doit élaborer des techniques de traitement des erreurs, après leur production.

Ces deux angles de vue selon Corder « *sont compatibles avec une seule et même position théorique sur les langues naturelles et sur leur apprentissage : cette position relève du behaviourisme, pour ce qui est de la psychologie, et du structuralisme, pour ce qui est de la linguistique.* » (1980 :9)

Selon Corder (1980), l'hypothèse behavioriste est supposée s'appliquer dans les deux cas, en l'occurrence, l'acquisition et l'apprentissage des langues.

En effet, l'accent est mis dans les récentes recherches sur l'apprentissage des langues que sur l'enseignement. Ce déplacement de préoccupation n'est pas arbitraire car le domaine de l'appropriation des langues relève plus de la linguistique et de la psychologie que de la pédagogie. Autrement dit, l'accent est mis sur le processus psycholinguistique d'acquisition des langues que sur les méthodes d'enseignements, sans pour autant négliger ou marginaliser la dimension pédagogique.

En effet, F. Debyser (1970) note que l'analyse contrastive focalise sur l'enseignement des langues étrangères que sur leur apprentissage ; nous évoquons ici la dichotomie enseignement, apprentissage pour se référer à l'apprenant d'une langue étrangère :

« En général les indications qu'elles apportent sont destinées aux professeurs et aux auteurs de manuels ou d'exercices plutôt qu'aux élèves ; en conséquence la linguistique contrastive semble concerner davantage l'enseignement (teaching) que l'apprentissage (learning), soit qu'elle ambitionne d'établir une programmation skinnerienne de L2 (algorithme spécifique d'approche de L2 en fonction de Ll), soit qu'elle se limite à armer le professeur d'un arsenal d'exercices correctifs. » (Debyser,1970 :33)

Il ajoute aussi que : *« [...] la linguistique contrastive s'est trouvée liée à une pédagogie de la réussite quelque peu behavioriste, qui considère la faute comme une mauvaise herbe à extirper ; le meilleur parcours pédagogique à proposer à l'élève est en ce cas établi par une progression sans failles, fortement contraignante, où l'on cherche à éviter le contact avec l'erreur. » (Ibid.)*

Dans cette nouvelle perspective, en l'occurrence, l'acquisition et l'apprentissage des langues naturelles. Il est crucial de s'interroger sur ces deux processus d'appropriation des langues naturelles.

En effet, il y a selon les linguistes deux processus distincts ; l'acquisition et l'apprentissage : l'acquisition est un processus d'appropriation naturel, implicite, inconscient, implique une focalisation sur le sens, il ne nécessite pas beaucoup d'efforts. L'apprentissage, par contre, est artificiel, explicite, conscient, et qui implique une focalisation sur la forme et nécessite beaucoup d'efforts. *(Cuq et Gruca, 2005 :109)*

Selon Corder (1980), comme déjà indiqué, l'hypothèse behavioriste est supposée s'appliquer dans les deux cas, en l'occurrence, l'acquisition et l'apprentissage. Cela implique que l'apprenant d'une langue étrangère adopte une stratégie qui ressemble – même si elle n'est pas totalement identique – à celle adoptée par l'enfant durant l'acquisition de sa première langue :

« Je propose donc, comme hypothèse de travail, l'idée qu'au moins quelques-unes des stratégies adoptées par l'apprenant de LE sont en substance les mêmes que celles permettant l'acquisition de la première langue. Ceci n'implique pas que le déroulement, la progression de l'apprentissage soit identique dans les deux cas. » (Corder, 1980 :11)

L'abondance des erreurs de langue chez les enfants est visible pour tout le monde. Le foisonnement de ces erreurs ne suscite pas d'habitudes notre curiosité sur l'origine ou la nature de ces erreurs puisque nous considérons l'occurrence de ces erreurs comme des phénomènes langagiers normaux.

« Si nous postulons l'existence d'un même mécanisme dans les deux cas, nous pouvons alors postuler également que les procédures ou les stratégies adoptées par l'apprenant de LE sont fondamentalement identiques. Le principal trait distinctif entre les deux opérations est la présence ou l'absence de motivation. Si l'acquisition de la première langue n'est autre que la réalisation d'une prédisposition à développer un comportement langagier, l'apprentissage d'une deuxième langue implique alors que quelque autre motivation remplace la prédisposition existant chez l'enfant. » (Ibid.)

En effet, F. Debyser (1970) rejoint Corder dans son raisonnement en s'interrogeant lui aussi sur les similitudes qui peuvent caractériser les deux processus :

« On se demande [...] si l'apprentissage d'une langue étrangère ne procède pas comme pour la langue maternelle, par une série de structurations et de restructurations successives, où des hypothèses provisoires, impliquant des fautes à la fois inévitables et « nécessaires », sont tour à tour formulées et testées par l'élève, et constituent autant de paliers discontinus dans l'organisation progressive de sa compétence. » (Debyser, ibid.)

Le concept pivot, donc, de l'approche de l'« analyse des erreurs » est la notion d'erreur. D'une manière très simple l'erreur est cet écart entre la norme, la forme correcte et la déviation de cette norme dans l'usage et la pratique qui se traduira par la production d'énoncé incorrect dans la langue cible.

Dans l'analyse des erreurs, on se réfère souvent aux erreurs systématiques. *« L'emploi du terme « systématique » dans ce contexte implique bien entendu qu'il puisse exister des erreurs qui ne le soient pas, ou, plus exactement, dont la nature systématique n'est pas facile à discerner. Cette distinction entre erreurs systématique et non-systématique est importante. » (Corder, 1980 :13)*

En d'autres termes, dans l'analyse des erreurs, on distingue deux entités d'erreurs, les erreurs systématiques et les erreurs non-systématiques :

Corder fait appel à la dichotomie compétence vs performance, échafaudée par l'illustre N. Chomsky, afin d'expliquer les nuances qui existent entre ces deux notions. Par définition, les erreurs systématiques sont des erreurs de compétence alors que les

erreurs non-systématique sont des erreurs de performance. « *Les erreurs de performance seront par définition non -systématiques, et les erreurs de compétence systématiques.* » (Ibid.)

Cependant, l'utilisation de ces deux concepts, n'est pas arbitraire, en fait elle est basée sur les mêmes bases échafaudées par Chomsky.

« Aussi sera-t-il commode désormais d'appeler « fautes » les erreurs de performance, en réservant le terme d'« erreur » aux erreurs systématiques des apprenants, celles qui nous permettent de reconstruire leur connaissance temporaire de la langue, c'est-à-dire leur compétence transitoire. » (Ibid.)

En effet, Corder fait une distinction cruciale entre erreur et faute ; en se basant, toujours sur la dichotomie de N. Chomsky compétence vs performance, il réutilise ces deux concepts afin de différencier l'erreur de compétence de l'erreur de performance, c'est-à-dire, l'erreur de la faute. La première est systématique alors que la deuxième est non-systématique.

La faute, dans ce contexte, ne constitue pas, selon Corder, un élément d'analyse systématique, elle est accidentelle ; elle ne relève pas d'un système sous-jacent.

« Nous savons tous qu'un adulte, s'exprimant ordinairement dans sa langue maternelle, commet de nombreuses erreurs de différentes sortes. Elles sont dues, [...], à des défaillances de mémoire, à des états physiologiques tels que la fatigue ou à des états psychologiques, comme par exemple une émotion forte. Ce sont là des aléas accidentels de la performance linguistique, qui ne reflètent pas les lacunes de sa connaissance de la langue. D'ordinaire, nous

nous en rendons compte immédiatement lorsqu'ils se produisent et sommes capables de les corriger de façon plus ou moins assurée. » (Ibid.)

La faute est souvent aléatoire, elle n'obéit pas à un système quelconque, elle est éphémère ; elle émerge selon l'état général, l'humeur de chacun, les conditions de la situation psychologique ou physique de l'apprenant.

« En toute logique, il faut bien s'attendre à de tels lapsus (« lapsus linguae » ou « lapsus calami ») chez l'apprenant de langue étrangère, lui-même contraint par des conditions similaires, externes ou internes, quand il utilise sa L1 ou une LE. Il nous faut alors distinguer les erreurs qui sont dues au hasard des circonstances de celles qui reflètent à un moment donné sa connaissance sous-jacente, ou, comme on pourrait l'appeler, sa « compétence transitoire ». Les erreurs de performance seront par définition non -systématiques, et les erreurs de compétence systématiques. Comme le dit MILLER (1966), « quel intérêt y a-t-il à énoncer des règles conduisant à des fautes ? » Aussi sera-t-il commode désormais d'appeler « fautes » les erreurs de performance, en réservant le terme d'« erreur » aux erreurs systématiques des apprenants, celles qui nous permettent de reconstruire leur connaissance temporaire de la langue, c'est-à-dire leur compétence transitoire. » (ibid.)

2. Les lapsus

En effet, le lapsus est la rencontre de deux intentions ; l'une va remplacer l'autre en l'effaçant ou en la déformant résultant en la substitution d'un énoncé par un autre indésirable, d'une façon spontanée. *Le lapsus linguae* concerne l'oral, et le *lapsus calami* concerne l'écrit.

Dans ce contexte, la faute n'a pas vraiment une valeur significative dans l'analyse.

« Les fautes ne sont pas significatives par rapport au processus d'apprentissage d'une langue. » (ibid.)

Toutefois, diagnostiquer une erreur d'une faute, reste une tâche qui n'est pas facile, il faut être perspicace et avoir des connaissances et un savoir-faire.

« Toutefois, il n'est pas toujours aisé de reconnaître ce qui est une faute de ce qui est une erreur ; il faut, pour les distinguer, procéder à une recherche ou à une analyse d'erreurs plus poussée que celle pratiquée habituellement. » (Corder, :13)

Par contre l'erreur, est hyper significative par rapport au processus d'appropriation d'une langue étrangère.

Dans notre recherche, l'analyse statistique nous permet de distinguer entre les fautes et les erreurs par le nombre de répétition de chacune. Si la faute n'est définie que comme un lapsus, son apparition dans les résultats finaux ne peut être qu'insignifiante par rapport à l'erreur. Dans ce sens la récurrence a le sens *des erreurs commises* par l'ensemble des élèves.

3. La compétence transitoire ou (Dialectes idiosyncrasiques)

Corder (1980) a constaté que la langue de l'apprenant au cours de l'apprentissage d'une langue étrangère (LE) constitue, selon les critères de la linguistique, un système linguistique à part entière. Ce nouveau système linguistique est appelé, donc, compétence transitoire ou dialectes idiosyncrasiques.

*« [...] la langue que parle l'apprenant, ou, éventuellement, certaines « communautés » d'apprenants, constitue un type particulier de dialecte. »
(Corder, 1980 :17)*

Selon Corder (1980), durant l'apprentissage de la langue étrangère, l'apprenant va posséder, ou développer un dialecte propre à lui. Frauenfelder et al. (1980) dit à ce propos que :

« La langue de l'apprenant (LA) et celle du locuteur natif de sa langue cible (LC), entretiennent donc, d'un point de vue linguistique du moins, une relation de « dialecte » [...]. Etant donné ceci, et puisque LC est une langue naturelle, on peut dire à la suite de CORDER qu'« au moins certaines des règles grâce auxquelles on décrit la langue de l'apprenant sont identiques à celles dont on a besoin pour décrire une langue naturelle. » (Frauenfelder et al., 1980 :43)

Cependant, ce dialecte n'est pas sa langue maternelle ni sa langue cible qu'il est en train d'acquérir mais une langue intermédiaire. En d'autres termes, ce processus d'appropriation d'une langue est lui-même un « dialecte » :

« En résumé, pour nous, la connaissance d'une LE se manifeste chez l'apprenant par un système linguistique intériorisé, qui évolue en se complexifiant, et sur lequel il possède, en principe, des intuitions. Ce système est différent de celui de la langue maternelle (LM) (même si l'on en trouve des traces), et aussi de celui de la LC ; il ne peut non plus être considéré comme un simple mélange de l'une et de l'autre, mais comporte des règles qui lui sont propres : chaque apprenant [...] possède à un stade donné de son apprentissage un système spécifique. » (Ibid. :46)

Toutefois, Corder (1980) note que ce système linguistique dispose de traits distinctifs qui le distinguent des dialectes ordinaires, en l'occurrence, les dialectes sociaux.

« Pour le moment, il suffit de distinguer les dialectes qui constituent la langue d'une communauté (que j'appellerai « dialectes sociaux ») et ceux qui ne constituent pas la langue d'une communauté (et j'appellerai ceux-ci « dialectes idiosyncrasiques »). J'avance donc des arguments d'ordre linguistique, et non sociologique, pour appeler ces derniers des dialectes. » (Corder, 1980 :17)

Ce concept a révolutionné le domaine de l'analyse des erreurs, Ce qui a conduit beaucoup de chercheur à l'adopter, en lui attribuant d'autres appellations :

« Ce que nous avons appelé, suivant CORDER, la « langue de l'apprenant », a reçu selon les auteurs différentes appellations : « compétence transitoire » (CORDER 1967), « dialecte idiosyncrasique » (CORDER 1971), « système approximatif » (NEMSER 1971), « interlangue » (SELINKER 1972), « système intermédiaire » (PORQUIER 1974), « système approché » (NOYAU 1976). » (Frauenfelder et al., 1980 :44)

Cependant, parmi toutes ces appellations, la notion d'interlangue, est devenue la plus pertinente dans ce champ d'étude. Néanmoins, cette notion est une reformulation pure et simple des concepts, déjà échafaudés par Corder.

« . L. SELINKER (1969) a proposé d'utiliser le terme interlangue pour cette classe de dialectes idiosyncrasiques, ce qui implique que ce dialecte partage certains traits de deux dialectes sociaux, ou de deux langues, que ceux-ci aient ou non des traits en commun. Cette question reste ouverte, et renvoie au problème des universaux du langage. » (Corder, 1980 :20)

Selon Selinker ; l'interlangue est *« une structure psychologique latente activé par l'apprenant durant son appropriation d'une langue cible (L2). Ces structures sont, en l'occurrence, des « transfert de langue, transfert d'apprentissage, stratégies d'apprentissage de L2, stratégies de communication en L2, surgénéralisation d'éléments linguistiques de L2 ».* (Selinker, in Frauenfelder et al., 1980 :44-45)

L'erreur aujourd'hui a perdu sa connotation péjorative et est devenue une condition *sine qua non* de l'apprentissage, Corder (1980) signale que :

« [...] les erreurs de l'apprenant manifestent le système linguistique qu'il utilise [...] elles sont indispensables à l'apprenant, car on peut considérer l'erreur comme un procédé utilisé par l'apprenant pour apprendre. C'est pour lui une façon de vérifier ses hypothèses sur le fonctionnement de la langue qu'il apprend. Faire des erreurs, c'est alors une stratégie qu'emploient les enfants dans l'acquisition de leur L1, et aussi les apprenants de LE. » (Corder, 1980 :13)

4. Erreurs interlinguales et erreurs intralinguales

Dans le domaine de l'analyse des erreurs, il est vital de distinguer deux catégories d'erreurs de l'apprenant. Les erreurs intralinguales et les erreurs interlinguales :

4.1. Les erreurs intralinguales

L'origine des erreurs intralinguales est la langue ciblée par le processus d'apprentissage. Autrement dit, ces erreurs ne trouvent pas leur origine dans la langue source. En effet, ces erreurs ne sont pas le résultat des phénomènes interférentiels. Parmi les causes qui contribuent à l'émergence de ce type d'erreur ce sont les caractéristiques intrinsèques de la langue cible. Ainsi, elles peuvent être le résultat d'une défaillance ou d'un manque de performance, c'est-à-dire un manque de connaissances de la langue étrangère.

4.2. Les erreurs interlinguales

Les erreurs interlinguales sont des erreurs typiquement interférentielles. En d'autres mots, ces erreurs sont causées, souvent, par l'impact de la langue source sur la langue cible. C'est-à-dire, l'influence par un transfert négatif des éléments du système linguistique déjà acquis sur l'apprentissage de la langue étrangère.

5. Les étapes de l'analyse des erreurs

La linguistique contrastive, l'analyse contrastive, la linguistique différentielle, sont étroitement liées à l'analyse des erreurs.

L'analyse des erreurs, méthodologiquement parlant comporte trois étapes importantes et incontournables :

5.1. Trouver l'erreur

Il s'agit ici de dégager l'erreur, cette étape requiert de la part des chercheurs une maîtrise des normes de la langue cible afin de diagnostiquer ces aberrations linguistiques à l'égard des normes.

5.2. Décrire l'erreur

Dans cette étape les chercheurs procèdent à une description minutieuse de chaque anomalie linguistique dans les textes du corpus.

5.3. Expliquer l'erreur

Arrivant à l'étape explicative, qui consiste à expliquer ces aberrations vis-à-vis de la norme linguistique. Ceci requiert des compétences, non seulement de la langue cible, mais aussi de la langue source. Aussi, nous essayerons de revenir à la linguistique contrastive pour essayer d'appréhender ces aberrations ou anomalies par rapport aux normes.

Dans la partie qui suit, nous essayerons de reprendre les éléments clés de la recherche avant d'expliquer comment nous avons construit une typologie propre à cette recherche, puis à la fin nous allons décrire les résultats et essayer de comprendre les causes des erreurs.

La partie pratique

Chapitre I :

Définitions des éléments clés

1. Comment cette partie va être organisée.

Dans cette partie, nous essayerons, dans un premier temps, de rendre une description des éléments clés de la recherche : la population, la période, le milieu et les circonstances dans lesquelles se déroulait l'enquête. Nous allons ensuite expliquer comment nous avons constitué notre grille d'analyse tout en indiquant, si nécessaire, comment nous avons repéré les erreurs avant d'entamer la description des résultats de l'analyse statistique ; nous préférons à ce stade, partager d'une façon brève la méthode que nous avons suivie pour établir une hiérarchie allant des types les plus généraux des différents niveaux linguistiques jusqu'à l'erreur⁵.

Nous rappelons que notre recherche a pour but de relever les erreurs récurrentes dans les productions des élèves de troisième année moyenne afin de comprendre les difficultés qu'éprouve les apprenants au cours de leurs productions écrites. Comme nous rappelons aussi que le but ultime est donc de remettre un bilan constituant « *des informations d'importance didactique* » (Desoutter :120). Pour pouvoir atteindre ces buts, nous avons fixé un ensemble d'objectifs : la sélection d'une classe, d'une population, la collecte des données à partir de leurs productions écrites, la construction et l'application d'une grille d'analyse, et enfin l'identification et le dénombrement de toutes les erreurs.

Dans les prochaines lignes, nous allons essayer d'apporter plus de détails sur chacun des éléments précédents.

⁵ Nous allons partager un schéma résumant la hiérarchie établie.

2. La collecte des données

Nous avons confié le travail de la collecte des données à une enseignante de français. Nous lui avons demandé de nous apporter des productions écrites des élèves de troisième année du moyen en langue française accompagnée d'une explication ou d'une traduction de son contenu en arabe Algérien. Celle-ci nous a aidé à comprendre non seulement le sens des phrases, mais aussi des mots et ceci dans le cas où d'une part leurs expressions seraient mal construites et d'autre part où les phrases seraient illisibles.

La collecte des données a été réalisée dans la première semaine après les vacances du printemps ; Le Jeudi, 07 Avril 2016. Pour aider les élèves à nous rendre le maximum de contenu, nous leur avons proposé comme thème une question dont la réponse ne nécessite qu'une petite réflexion ; la question était : « Qu'est-ce que vous avez fait pendant les vacances ? ». Ainsi, avec la collaboration de leur enseignante, nous leur avons permis d'utiliser des dictionnaires, de lui poser des questions ou entre eux mutuellement.

3. Population d'enquête

Nous avons sélectionné les élèves de troisième année du moyen du CEM Mahboub Abdelkader à Ouled Mimoun, Wilaya de Tlemcen. Généralement, les élèves de ce niveau ont entre 14 et 15 ans. La population semble homogène dans la mesure où les langues qu'elles partagent sont les mêmes : De façon générale, la langue maternelle et de socialisation de la plupart est l'arabe Algérien ; la langue de scolarisation à travers laquelle la totalité des matières sont inculquées est l'arabe moderne. La langue française, d'autre part, constitue la première langue étrangère à apprendre depuis la troisième année du primaire. Les élèves de troisième année du

moyen ont déjà acquis un ensemble de règles qui régissent le bon fonctionnement de la langue et ce dans les cinq années qui ont précédé leur niveau actuel.

A partir d'une classe de troisième année du moyen, nous avons relevé vingt-six copies dont neuf ont été rejetées puisqu'elles étaient identiques. Ces feuilles pourraient nuire aux résultats finaux qui nécessitent de la rigueur. Les dix-neuf copies sélectionnées pour l'analyse statistique appartiennent à 10 filles et 7 garçons. Nous précisons que nous n'avons pas pris la variable « sexe » dans l'analyse. Ce qui nous intéresse est alors la récurrence qui a le sens de difficulté, prise d'une façon générale.

4. La méthode d'analyse

Afin d'aboutir au but de notre recherche, il était nécessaire de relever et catégoriser les erreurs dans des types généraux qui constituent des niveaux linguistiques adaptés à l'écrit : à partir du livre introductoire au traitement automatique des langues naturelles par Pierrette Bouillon et ses collaborateurs, nous avons pu constituer une typologie qui nous permet de mettre les erreurs dans des catégories et des sous-catégories.

La typologie vis-à-vis de notre problématique est d'une importance capitale. Elle nous permet d'organiser les informations recueillies d'une façon plus ou moins précise, facilitant leurs traitements statistiques. Nous rappelons que le but de cette recherche est de découvrir les récurrences des erreurs d'une façon générale. Les récurrences des erreurs dans ce sens ont la signification de *difficultés – les difficultés* qu'éprouvent le groupe des élèves sélectionnés au cours de leurs productions écrites.

5. Type d'enquête

Le type d'enquête de cette recherche est descriptif suivant la méthode statistique descriptive qui a pour but de :

« [...] résumer l'information contenue dans les données de façon synthétique et efficace. Elle utilise pour cela des représentations de données sous forme de graphiques, de tableaux et d'indicateurs numériques (par exemple des moyennes). Elle permet de dégager les caractéristiques essentielles du phénomène étudié et de suggérer des hypothèses pour une étude ultérieure plus sophistiquée. Les probabilités n'ont ici qu'un rôle mineur. » (Gaudoin, p.8)

Grace à cette méthode et à la typologie construite, nous avons pu représenter les informations en chiffres dans des tableaux et des graphies qui facilitent leur lecture.

Le type d'enquête de cette recherche est aussi analytique puisqu'on essaie de trouver les causes des erreurs, qui pourraient être interlinguales, c'est-à-dire une interférence entre une langue antérieure acquise ou apprise, comme elles pourraient être intralinguales ; un manque de connaissance des règles qui régissent la langue en question.

Après avoir exposé comment nous avons collecté les données, la population d'enquête en question, la méthode d'analyse et enfin le type d'enquête, nous allons par la suite expliquer comment nous avons constitué cette typologie à partir des définitions extraites de ce livre.

Chapitre II :

La construction d'une typologie

1. Les catégories simples

Nous appelons catégories simples celles qui ne comportent qu’un seul niveau linguistique, contrairement aux catégories combinées que nous allons expliquer par la suite. Les niveaux linguistiques constituant cette partie sont : les erreurs morphologiques, syntaxiques, et sémantiques.

Avant de commencer l’explication de chaque catégorie, nous signalons que dans le système de TAL, la langue est présentée comme un ensemble de connaissances (Bouillon, 1998 :12) à partir desquelles nous avons déduit cette typologie.

1.1. Les erreurs morphologiques

Les connaissances morphologiques concernent « *la manière dont les mots sont construits à partir des unités minimales de signification, appelées morphèmes* » (*ibid.* :13). Ils peuvent être des mots entiers comme ceux qu’on trouve dans le dictionnaire, « *pomme, par exemple* » (*ibid.*), comme ils peuvent être des parties de mots ou plutôt des affixes, « *par exemple, anti+ ou +tion* » (*ibid.*). Les affixes s’attachent au mot ; au début, ils sont appelés des préfixes, et à la fin, se sont des suffixes. Enfin, la partie centrale est appelée *racine*. Le tableau (*ibid.*) suivant l’illustre d’une façon claire :

mot	préfixe	racine	suffixe
penser		pens	+er
pensable		pens	+able
impensable	im+	pens	+able

Nous empruntons ce tableau au livre déjà mentionné, *ibid.*

On subdivise les connaissances morphologiques, en connaissances flexionnelles, dérivationnelles et compositionnelles. La morphologie flexionnelle « désigne le processus par lequel les mots reçoivent des affixes qui expriment des informations, sans changer leurs catégories syntaxiques » (*ibid.*) Le « s » du pluriel, à titre d’exemple, ne changera pas la catégorie grammaticale des mots. La morphologie dérivationnelle est « le processus de formation des mots par lequel les mots reçoivent des affixes qui peuvent cette fois changer la catégorie grammaticale du mot. » (*Ibid.*) Par exemple, le verbe changer pourrait être un adjectif par le simple ajout du suffixe +able à la fin. En fin, la morphologie compositionnelle désigne « le processus qui détermine comment les mots eux-mêmes peuvent être combinés pour former de nouveaux mots. » comme grand-père, porte-parole, etc.

A partir de ces définitions, nous avons déduit que les erreurs morphologiques ne peuvent concerner que la forme des mots ou des parties de mots. La catégorie des erreurs morphologiques donc inclue des sous-catégories qui distinguent les différents types d’erreurs : Les *erreurs dérivationnelles* concernant la mauvaise construction du mot à partir de ses affixes, qui changent la catégorie grammaticale du mot. Ainsi, les erreurs de *la forme canonique* qui concerne la mauvaise forme de la racine du mot, ou en d’autres termes, le mot qui ne semble pas identique à celui qui figure dans les dictionnaires.

Outre les deux sous-catégories mentionnées, nous en avons ajouté d’autres suivant notre corpus comme la manipulation de la majuscule ou de la minuscule dans le texte.

En ce qui concerne *Les erreurs flexionnelles* qui concernent la forme des affixes qui ne changent pas la catégorie grammaticale du mot, nous préférons les mettre dans la catégorie qui combine la morphologie et la syntaxe puisqu’ils ont relation avec d’autres éléments de la phrase. A titre d’exemple, un mot ne prend le « s » du pluriel

qu'en fonction du sens et de ce qui précède le mot comme l'article défini ou indéfini du pluriel.

Le schéma suivant résume la hiérarchie propre à ce type :

Catégorie : Morphologique

Sous-catégorie : Majuscule

Précision : Première lettre en majuscule au début du texte

Précision : Première lettre en majuscule pour les noms propres

Sous-catégorie : Minuscule

Première lettre minuscule à l'intérieur du texte

Sous-catégorie Forme canonique

Précision : Mot mal écrit

Précision : Nom propre mal écrit

Précision : Manque : s du pluriel

1.2. Les erreurs syntaxiques

Les connaissances syntaxiques « *décrivent la manière dont les mots se combinent en phrases syntaxiquement correctes et encodent aussi leur régularité structurale ; elles permettent d'opposer la phrase (a), syntaxiquement bien formée (a) à la phrase (b) mal formée (où, par convention, l'astérisque indique une phrase syntaxiquement mal formée) » (ibid. :14)*

- a. « J’étudie l’anglais »
- b. « *L’anglais étudie je »

La syntaxe donc s’intéresse à l’emplacement et aux fonctions des éléments de la phrase. Dans le TAL, on subdivise les informations syntaxiques en informations catégorielles comportant les différentes catégories syntaxiques : nom, verbe, pronom, etc. (ibid. :36) Ainsi en informations sous catégorielles qui :

« Distinguent différentes sous-classes à l’intérieur des catégories syntaxique, en spécifiant le cadre syntaxique dans lequel le mot apparaît. Les verbes, les adjectifs, les prépositions et certains noms diffèrent en effet en fonction de leur structure argumentale, c’est-à-dire du nombre et du type d’arguments qu’ils requièrent (Grishman, 1990). » (Ibid.)

Le verbe pourrait être transitif, intransitif ou bitransitif selon les arguments qu’il peut prendre, Pierrette (ibid) l’explique de la manière suivante :

« Un verbe, par exemple, peut prendre un seul argument (verbe intransitif), deux arguments (un sujet et un complément d’objet direct ou indirect – verbe transitif) ou même trois arguments (un sujet, un complément d’objet direct et un complément d’objet indirect – verbe bitransitif). Il peut aussi avoir plusieurs structures argumentales (comme donner qui est soit transitif soit bitransitif) » (ibid.)

Le tableau (ibid.) suivant, à partir du même livre, l’illustre encore mieux :

Structure argumentale des verbes		
a.	Je pars. *Je pars le chien.	<i>intransitif</i>
b.	Ma sœur me manque. *Ma sœur manque. Je donne un cadeau.	<i>transitif</i>
c.	Je donne un cadeau à ma sœur.	<i>bitransitif</i>

A partir de ces définitions, nous déduisons que les erreurs syntaxiques ne doivent être liés qu’avec le mauvais emplacement des éléments de la phrase ou sa constitution d’une façon générale.

Parmi les erreurs que nous avons relevées il y a : l’ajout des éléments superflues à la phrase, la suppression des éléments obligatoires tel qu’un sujet, verbe ou préposition, des phrases lourdes etc.

Le schéma suivant représente cette catégorie :

Catégorie : Syntaxique

Précision : Eléments de la phrase

Précision : Ajout d’un élément superflu

Précision : Suppression d’un élément obligatoire

Précision : Suppression de plusieurs éléments obligatoires

Précision : Substitution d’un élément par un autre inadéquat

Précision : Virgule devant "et"

Précision : Une phrase lourde

Précision : Éléments de la phrase mal placés

1.3. Les erreurs sémantiques

Les connaissances sémantiques « *concernent les sens des mots et la manière dont les sens se combinent pour former le sens global de la phrase* ». *Ce qui nous permet de détecter l’asémantisme du sens est en premier lieu sa relation avec le monde extérieur. Une phrase syntaxiquement correcte comme « ? Les idées oranges sont populaires. » (Ibid. :14) est sémantiquement erronée puisque dans le monde réel, « les idées ne désignent en effet pas des objets de couleur orange (Cann, 1993 :6-7 in Ibid.)*

Les erreurs sémantiques donc sont des erreurs qui ont une forte relation avec le sens seulement, à condition que la phrase soit syntaxiquement correcte. Si ce n’est pas le cas, nous catégorisons l’erreur dans la catégorie qui combine la sémantique et la syntaxe.

Nous avons observé et relevé de notre corpus trois types d’erreur que nous avons appelés : Inusité, ambiguïté et confusion lexicale. L’inusité désigne la combinaison des mots dont le sens n’est pas compatible avec la langue étrangère en question. L’exemple suivant extrait de notre corpus, et modifié pour une meilleure illustration, montre que l’élève n’avait pas utilisé les mots qu’il fallait pour dire : des vêtements neufs.

Nous avons acheté des vêtements nouveaux.

L’ambigüité dans notre corpus désigne l’absence ou la multitude du sens, l’exemple suivant, extrait et modifié à partir de notre corpus, l’illustre bien :

Et le manège le dernier jour.

En fin, la confusion lexicale concerne la confusion entre un mot et un autre. Par exemple à la place d’utiliser pendant, on utilise pour, etc.

Cette catégorie, conformément aux définitions mentionnées, englobe les erreurs de sens ; ce dernier n’est identifié qu’à partir des mots et leurs référents dans le monde extérieur.

Les catégories simples englobent les niveaux linguistiques les plus adaptés à l’écrit : les erreurs morphologique, syntaxiques et sémantiques. Nous concluons que les erreurs morphologiques ne concernent que la forme des mots, ensuite, les erreurs syntaxiques n’ont de relations qu’avec les éléments de la phrase, leurs fonctions, et leur emplacement. Enfin, les erreurs sémantiques concernent le sens des mots et leurs relations avec le monde extérieur.

Dans la prochaine partie, nous expliquerons comment et dans quelle mesure nous avons constitué des catégories qui combinent les différents niveaux déjà mentionné.

Nous résumons cette catégorie dans le schéma suivant :

Catégorie : Sémantique

Sous-catégorie : Sens

Précision : Inusité

Précision : Ambiguïté

Précision : Confusion lexicale

2. Les catégories combinées

Les catégories combinées comportent des niveaux linguistiques combinés, que nous avons réparties en : erreurs morphosyntaxiques et erreurs sémanticosyntaxiques. Pendant que la première combine la morphologie et de la syntaxe, la deuxième combine la sémantique et de la syntaxe.

La combinaison entre les niveaux linguistiques est nécessaire du fait que la langue fonctionne d’une façon à combiner les connaissances linguistiques afin de dépasser, ce qu’on appelle dans le système de TAL, l’ambiguïté. Par exemple, un mot polysémique ne pourrait être définie qu’en relation avec les autres éléments de la phrase, excédant le champ de la sémantique. Et c’est la même pour la syntaxe et de la morphologie.

L’erreur de même, n’échappe pas à cette règle puisqu’elle peut toucher à plusieurs niveaux.

2.1. Les erreurs morphosyntaxiques

Outre les définitions de chaque niveau constituant la morphosyntaxe, nous ajoutons que leur combinaison est nécessaire afin de représenter l’erreur qui touche la forme du mot qui ne prend forme que grâce aux autres éléments de la phrase : par exemple, l’accord du participe passé ne peut être réalisé que grâce aux autres éléments de la phrase y compris le sujet et le verbe, par conséquent, une erreur de forme est catégorisée comme étant morphosyntaxique.

A cet catégorie morphosyntaxique, nous avons ajouté une sous-catégorie comportant deux sous-catégories correspondant à chaque niveau, respectivement appelés : la forme canonique qui concerne *la forme* du mot ou *flexionnelle* concernant les affixes qui ne changent pas la catégorie grammaticale des mots et *les éléments de la phrase* concernant les différents éléments qui la constituent. Dans chaque sous-catégorie, nous avons ajouté des précisions sur la nature de l’erreur. Le schéma suivant, le montre d’une façon détaillée :

Catégorie : Morphosyntaxique

Sous-catégorie : Eléments de la phrase

Sous-catégorie : Forme canonique

Précision : le genre des noms

Précision : Conjugaison

Précision : confusion entre être et avoir

Précision : confusion entre "et" et "est"

Précision : article indéfini pluriel

Précision : article défini féminin

Précision : article indéfini singulier féminin

Précision : mot mal écrit

Précision : ajout d'un "s" superflue

Sous-catégorie : Eléments de la phrase

Sous-catégorie : Flexionnelle

Précision : manque : s du pluriel

Précision : accord du participe passé

Précision : accord de l'adjectif

Précision : article indéfini devant l'adjectif

Précision : conjugaison

Précision : confusion on, en, a, à

Précision : confusion "a" et "à"

Précision : Infinitif après les prépositions : à, de, pour, sans

Précision : le genre des noms

Précision : conjugaison

2.2. Les erreurs sémanticosyntaxiques

Les erreurs sémanticosyntaxiques combinent la sémantique et la syntaxe. Les erreurs sémanticosyntaxiques sont des erreurs qui touchent au sens et qui ont une incidence sur les éléments de la phrase. Les sous-catégories correspondant à ces deux niveaux combinés, sont respectivement appelés : *sens, éléments de la phrase* et *sens, temps, éléments de la phrase*. La première englobe les erreurs de sens qui ont une relation avec les éléments de la phrase ; la deuxième englobe les erreurs de la conjugaison qui ne respectent pas le temps réel de l’évènement. Le schéma suivant illustre bien la hiérarchie :

Catégorie : Sémanticosyntaxique

Sous-catégorie : sens

Sous-catégorie : Éléments de la phrase

Précision : suppression d’un élément obligatoire

Précision : suppression de plusieurs éléments obligatoires

Précision : mauvaise structure

Précision : incohérence

Précision : confusion dans l’usage des prépositions

Précision : substitution d’un élément par un autre inadéquat

Précision : homonyme

Sous-catégorie : sens, temps

Sous-catégorie : Éléments de la phrase

Précision : durée de l’évènement

Précision : suppression de plusieurs éléments obligatoires pour la compréhension

Jusqu’ici nous avons expliqué comment nous avons constitué la typologie qui nous permettra d’organiser les informations afin de les soumettre aux analyses statistiques souhaitées.

Nous allons essayer, dans le prochain chapitre, de réaliser la description des résultats statistiques afin d’arriver à les interpréter.

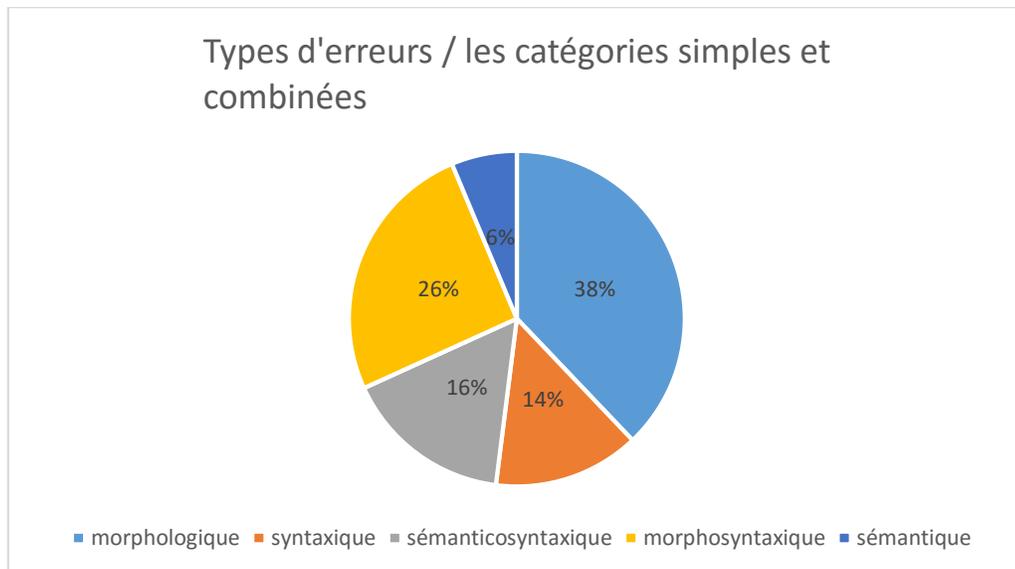
Chapitre III :

Description et interprétation des résultats

Description des résultats

Avant de commencer la description et les tentatives d'interprétations des résultats, nous rappelons que suivant notre problématique, nous allons décrire les erreurs les plus récurrentes et pourrions négliger d'autres qui semblent minimales. Ainsi, dans chaque catégorie, nous partagerons les résultats de l'analyse statistique ; nous les décrirons en donnant des exemples concrets à partir du corpus, puis essayerons de trouver les causes qui ont induit les élèves à les commettre.

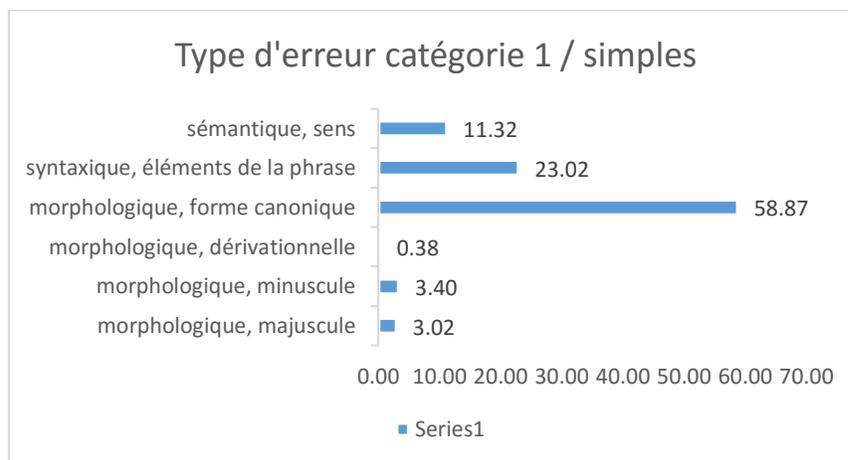
1. Les types d'erreurs (catégories simples et combinées)



Sur l'ensemble de 475 erreurs que nous avons relevées, les élèves ont commis 38% d'erreurs de type morphologique suivies de 26% de type morphosyntaxique. Nous constatons dans les deux, la présence de la morphologie qui, comme nous l'avons identifié grâce aux définitions que nous avons mentionnées dans la partie précédente, ne concerne que la forme des mots ou des parties de mots. Les erreurs sémanticosyntaxiques, d'autre part, constituent 16% de toutes les erreurs commises, suivies de 14% d'erreurs syntaxiques et 6% d'erreurs sémantiques.

A ce stade, nous ne pouvons pas identifier les causes de tous les types d'erreur, cependant, les erreurs morphologiques concernant la forme, sont clairement le résultat du manque de connaissance des règles de la langue française de la part des élèves et de la difficulté de la langue française dans ce sens. Nous les catégorisons donc comme étant des erreurs de type intralinguales qui n'ont pas de relation directe avec leurs acquis précédents, qu'ils soient de leur langue maternelle ou de l'arabe moderne.

2. Les catégories simples



Le nombre total des erreurs de l'ensemble de ces catégories est de 265. Les erreurs les plus récurrentes sont les erreurs *morphologiques, forme canonique* constituant 58.87%, c'est-à-dire celles dont la forme n'est pas identique à celle des dictionnaires de la langue française. Prenons deux exemples de notre corpus :

Erreur 1 : Le vacance

Correction : Les vacances

Explication : Les vacances est un nom féminin pluriel.

Erreur 2 : Otal

Correction : Hotel

Explication : La forme du mot n'est pas identique à celle des dictionnaires.

Pour plus d'exemples, veuillez consulter les annexes B.

Les erreurs *syntaxiques, éléments de la phrase* constituent 23.02% sur l'ensemble des erreurs de ces catégories. Les erreurs syntaxiques comme déjà définies dans la partie précédente concernent l'emplacement des éléments de la phrase. Voici deux exemples extraits de notre corpus :

Erreur 1 : La plage s'est un [...]

Correction : La plage est un / une

Explication : L'élève a ajouté un élément superflu, le « s » devant l'auxiliaire être.

Erreur 2 : J'ai joué avec mes amis au ballon

Correction : J'ai joué au ballon avec mes amis

Explication : une phrase lourde.

Pour plus d'exemples, veuillez consulter les annexes B.

Les erreurs de *sémantique, sens* ne constituent que 11.32% sur l'ensemble des erreurs de ces catégories. Les erreurs de *sémantique, sens* ne concernent que le sens des mots en relation avec le monde extérieur. Prenons deux exemples de notre corpus :

Erreur 1 : un visage très ravi

Correction : un visage ravi

Explication : l'élève a utilisé un sens qui n'est pas utilisé, nous l'avons donc catégorisé comme étant : *inusité*. Que le visage soit ravi ou pas ; la gradation est une anomalie.

Erreur 2 : [...] Avec mes deux frères et le petit

Correction : ambiguïté

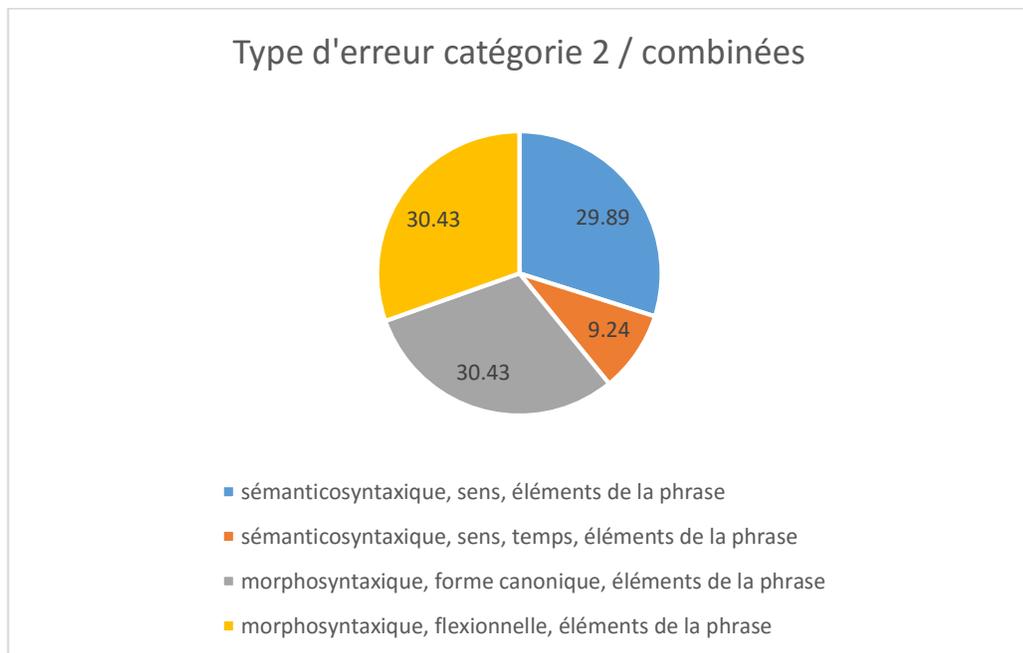
Explication : le petit par rapport à qui ? son enfant, son petit frère ? l'élève n'a rien indiqué de pertinent auparavant.

Pour plus d'exemples, veuillez consulter les annexes B.

Les erreurs de la manipulation de la majuscule, de la minuscule et les *erreurs dérivationnelles* ne sont que des erreurs minimales par rapport à l'ensemble des erreurs de ces catégories.

A ce stade, nous sommes toujours dans les généralités où nous ne pouvons pas constater, voir généraliser les causes des erreurs, à l'exception des erreurs de type morphologique que nous avons identifiées comme relevant des *erreurs intralinguales*.

3. Les catégories combinées



Sur un ensemble de 184 erreurs concernant *les catégories combinées*, que nous avons définies comme des combinaisons de niveaux linguistiques ; les erreurs les plus fréquentes sont des erreurs *morphosyntaxiques, flexionnelles, éléments de la phrase* 30.43% et des erreurs *morphosyntaxiques, forme canonique, éléments de la phrase* 30.43%. *Les erreurs sémanticosyntaxiques, sens, éléments de la phrase* sont aussi nombreuses, constituant 29.89% de toutes les erreurs de cette catégorie.

Les erreurs morphosyntaxiques, flexionnelles, élément de la phrase sont des erreurs qui, comme nous l'avons déjà signalé dans la partie précédente, touchent à la forme des mots qui ne peuvent prendre forme que grâce aux autres éléments de la phrase, elles touchent plus particulièrement aux affixes qui ne change pas la catégorie grammaticale du mot. Prenons deux exemples de notre corpus :

Erreur 1 : Mes idée

Correction : Mes idées

Explication : manque du « s » du pluriel.

Erreur 2 : J'ai demander

Correction : J'ai demandé

Explication : problème de conjugaison.

Pour plus d'exemple, veuillez consulter les annexes B.

Les erreurs morphosyntaxiques, forme canonique, éléments de la phrase constituent un nombre important de 30.43% sur l'ensemble des erreurs de cette catégorie. Cette catégorie se distingue de la précédente par la partie du mot affectée. L'erreur donc touche à la forme intérieure du mot, appelée *la forme canonique*. Voici deux exemples tirés de notre corpus :

Erreur 1 : En à

Correction : On a

Explication : confusion entre on, en et a, à.

Erreur 2 : du vacances

Correction : des vacances

Explication : confusion entre du et des.

Pour plus d'exemples, veuillez consulter les annexes B.

Les erreurs sémanticosyntaxique, sens, éléments de la phrase sont aussi d'une grande importance constituant 29.89% sur l'ensemble des erreurs de ces catégories. Cette catégorie concerne les erreurs de sens qui ont une incidence sur la structure de la phrase. Prenons les deux exemples suivants, extraits de notre corpus :

Erreur 1 : J'suis allée à une plage pour changé mes idées

Correction : Je suis allée à la plage afin de relaxer mon esprit. (Sens dérivé de sa production en arabe dialectal)

Explication : le sens de cette phrase n'est pas clair, nous l'avons catégorisé comme étant : inusité.

Erreur 2 : Les personnage heureux

Correction : Les gens étaient heureux

Explication : l'élève confond entre les personnages et les gens.

Pour plus d'exemples, veuillez consulter les annexes B.

Les erreurs qui apparaissent le moins dans ces catégories sont les erreurs *sémanticosyntaxique, sens, temps, éléments de la phrase*. Nous les avons définis dans le chapitre précédent comme des erreurs qui ont relation avec le temps de l'évènement.

Prenons les deux exemples suivants, tirés de notre corpus :

Erreur 1 : cette vaconce est pleine de joint

Correction : ces vacances étaient pleines de joie.

Explication : le temps choisi n'est pas compatible avec le temps réel des évènements.

Erreurs 2 : cette journée été très belle

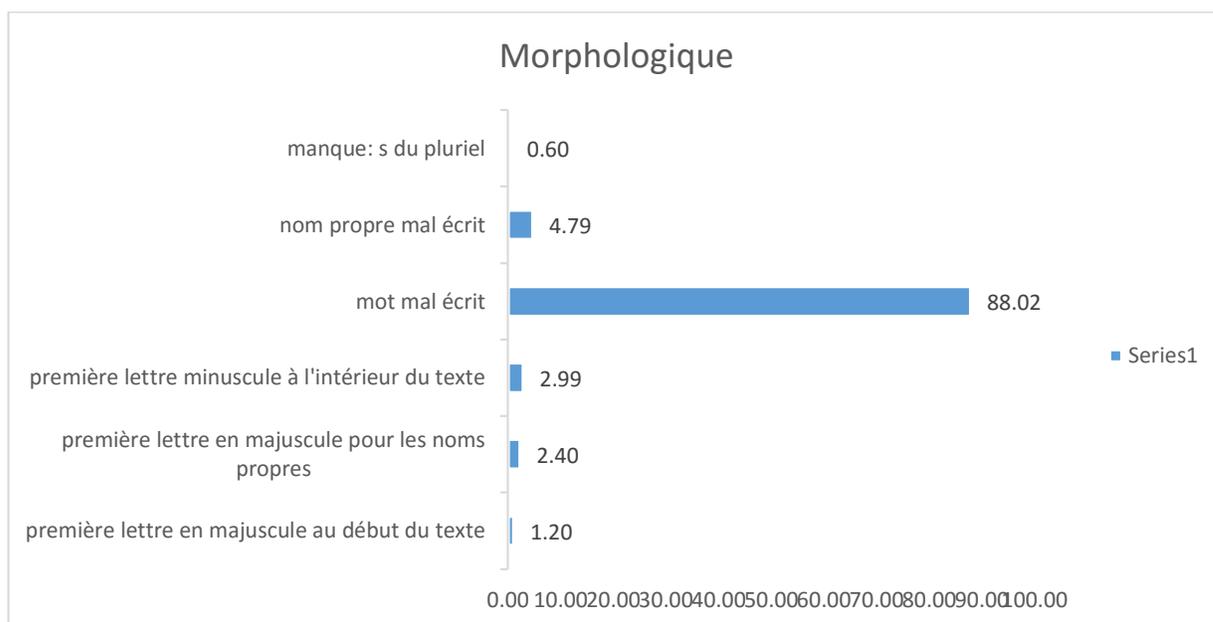
Correction : cette journée était très belle

Explication : le temps choisi n'est pas compatible avec le temps réel des évènements.

Pour plus d'exemples, veuillez consulter les annexes B.

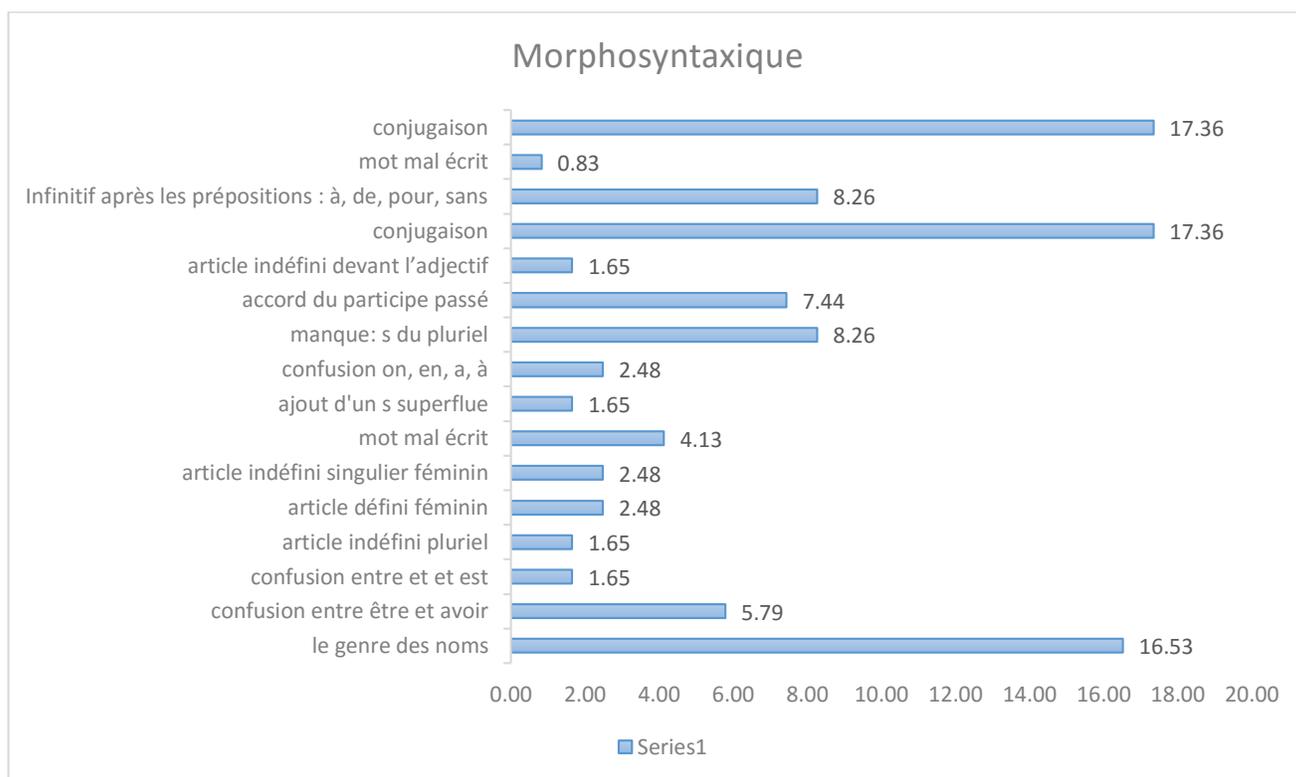
Dans les prochaines lignes, nous allons essayer d'entamer chacune avec plus de précision où nous commentons les erreurs les plus récurrentes et négligeons les autres, conformément à la problématique de notre recherche.

4. Les erreurs morphologiques



La plupart des erreurs morphologiques sont des mots mal construits, les résultats de l'analyse statistique que nous avons obtenus montrent que les mots erronés ont atteint 88.02%. La cause de ces erreurs ne pourrait être que le résultat d'un manque de connaissance de leurs graphies par les élèves en question. Les erreurs morphologiques sont des erreurs typiquement intralinguales.

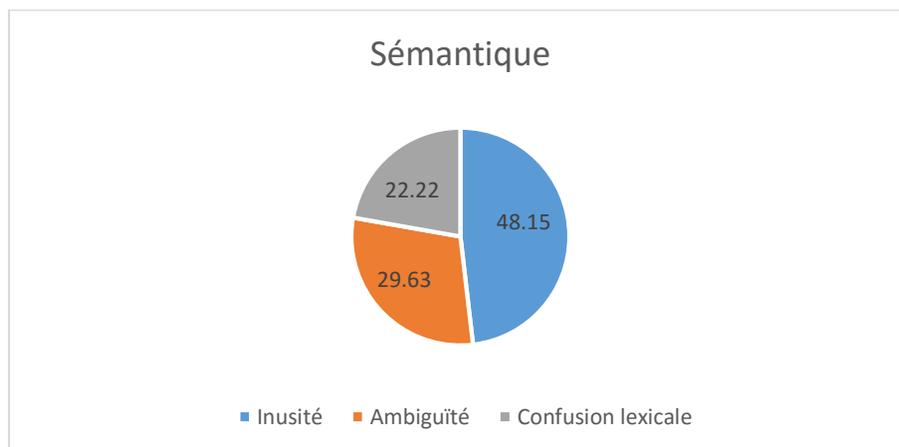
5. Les erreurs Morphosyntaxiques



Les erreurs morphosyntaxiques les plus récurrentes sont : la conjugaison constituant 17.36% des erreurs, le genre des noms (16.53%), l'infinitif après les prépositions : à, de, pour, sans (8.26%), le manque du « s » du pluriel (8.26%) et la confusion entre être et avoir (5.79%)

Il est clair que la conjugaison, l'infinitif après les prépositions : à, de, pour, sans et l'absence du « s » du pluriel, accord du participe passé ne sont que le résultat d'un manque de connaissances des règles de la langue plutôt qu'interférences. Le genre des noms d'autre part, peut être le résultat d'interférence.

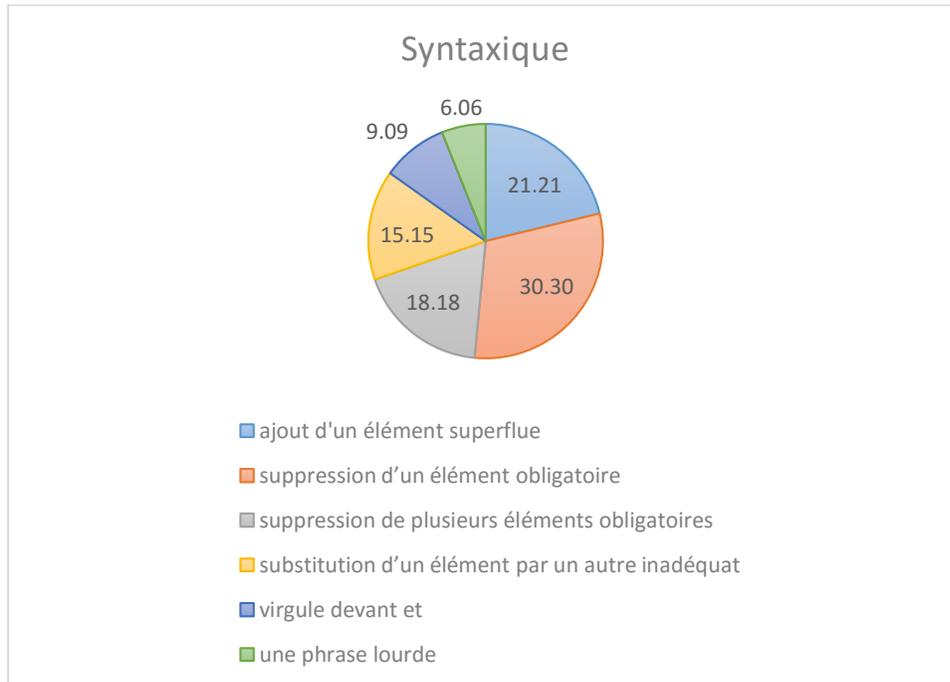
6. Les erreurs sémantiques



Les erreurs sémantiques sont généralement des erreurs de l'inusité (48.15%) et de l'ambiguïté (29.63%)

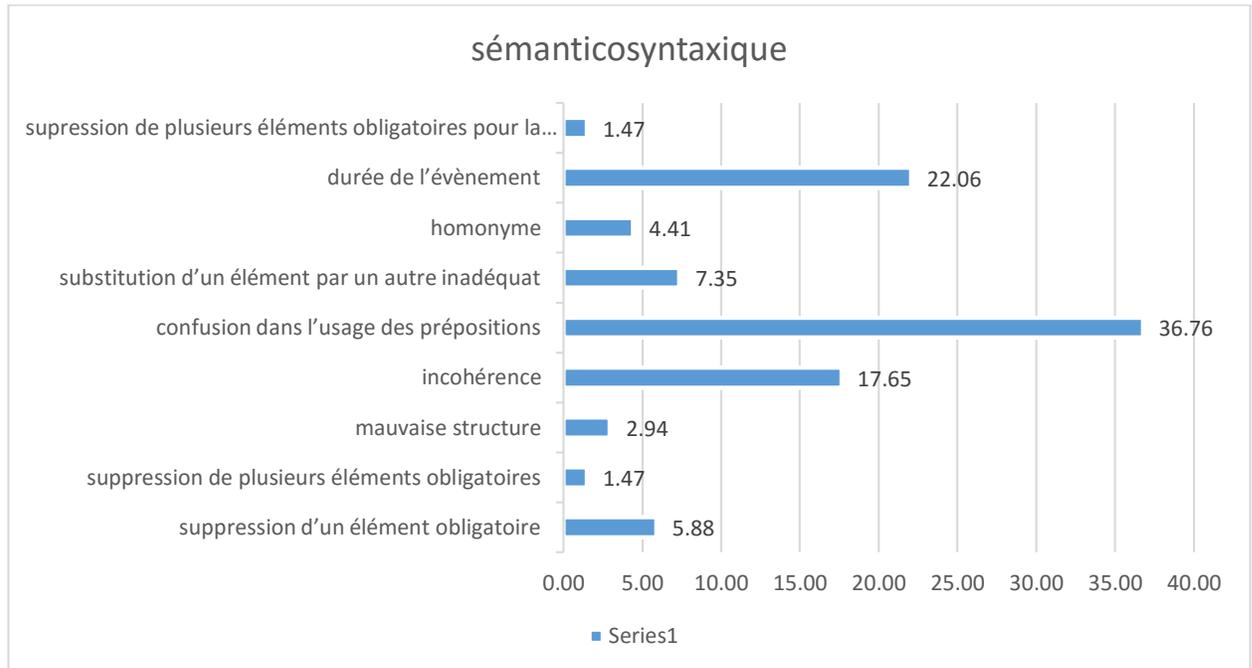
Nous avons constaté que les erreurs de l'inusité émergent quand l'élève ne trouve pas les mots qu'il faut pour s'exprimer. Alors que les erreurs de l'ambiguïté émergent quand l'élève ne complète pas la phrase ou ne précise pas le référent dans le monde extérieur. Les erreurs les plus fréquentes dans cette partie sont des erreurs intralinguales.

7. Les erreurs syntaxiques



La plupart des erreurs syntaxiques sont des erreurs de suppression d'un élément obligatoire dans la phrase, constituant 30.30% suivis des erreurs d'ajout d'un élément superflue (21.21%). A ce stade, nous ne pouvons pas généraliser les causes de ces erreurs, sauf si nous prenons les exemples cas par cas, ce qui nécessite une autre analyse statistique.

8. Les erreurs sémanticosyntaxiques



Dans les erreurs sémanticosyntaxiques, les erreurs qui se répètent sont ceux qui confondent l'usage des préposition (36.76%), la durée de l'évènement (22.06%) et de l'incohérence (17.65%).

Dans cette partie aussi nous ne pouvons pas généraliser les causes de l'erreur puisqu'elles nécessitent une analyse statistique plus précise.

Dans ce troisième chapitre, nous avons décrit les résultats des analyses statistiques que nous avons menés en utilisant une typologie hiérarchique allant des niveaux linguistiques les plus généraux jusqu'aux erreurs.

Conclusion

Conclusion

Dans cette recherche, nous avons effectué une analyse des erreurs dans les productions écrites d'une classe d'élèves de troisième année du moyen du CEM Mahboub Abdelkader à Ouled Mimoun, Wilaya de Tlemcen. Notre but était donc de relever les erreurs récurrentes afin de comprendre les difficultés qu'éprouvent l'apprenant au cours de sa production écrite. Dans ce sens la récurrence a le sens de *difficultés – les difficultés* qu'éprouvent le groupe des élèves sélectionnés au cours de leurs productions écrites. Le but ultime est alors de remettre un bilan constituant « des informations d'importance didactique ». Pour pouvoir atteindre ces buts, nous avons fixé un ensemble d'objectifs : la sélection d'une classe, d'une population, la collecte des données à partir de leurs productions écrites, la construction et l'application d'une grille d'analyse et enfin, l'identification et le dénombrement de toutes les erreurs.

Afin d'aboutir au but de notre recherche, il était nécessaire de relever et catégoriser les erreurs dans des types généraux qui constituent des niveaux linguistiques adaptés à l'écrit : à partir du livre introductoire au traitement automatique des langues naturelles par Pierrette Bouillon (1998) et ses collaborateurs, nous avons pu constituer une typologie qui nous permet de mettre les erreurs dans des catégories et des sous-catégories. Cette typologie était d'une importance capitale. Elle nous a permis d'organiser les informations recueillies d'une façon plus ou moins précise, facilitant leurs traitements statistiques.

Nous avons subdivisé la typologie en catégories simples (erreurs morphologiques, syntaxiques et sémantiques) et en catégories combinées (erreurs morphosyntaxiques et sémanticosyntaxiques) ; dans chacune nous avons incluse des sous-catégories (forme canonique, flexionnelle, dérivationnelle, sens et éléments de la

phrase) et à l'intérieur une autres sous-catégorie, appelé *précision* renvoyant aux erreurs⁶.

Le type d'enquête est descriptif dans la mesure où nous avons utilisé la méthode statistique descriptive qui nous a permis de synthétiser les informations d'une façon claire facilitant leurs lectures. Il est aussi analytique dans la mesure où l'autre but est d'essayer d'expliquer les causes des erreurs qui peuvent être interlinguales ou intralinguales.

De façon générale, nous avons relevé un total de 475 erreurs dont 38% sont morphologiques et 26% morphosyntaxiques. Il est clair que les erreurs les plus récurrentes ont relation directe avec la forme des mots ou des parties de mots. Les erreurs morphologiques sont visiblement des erreurs intralinguales.

Le nombre total des erreurs propres aux catégories simples est de 265. Les erreurs les plus récurrentes sont les erreurs qui touchent à la forme canonique du mot constituant 58.87%. Les erreurs syntaxiques qui touchent à la structure de la phrase est de 23.02%

Le nombre total des erreurs dans *les catégories combinées* atteint 184 erreurs dont les plus récurrentes sont les erreurs morphosyntaxiques, flexionnelles (30.43%) morphosyntaxiques, dérivationnelles (30.43%) et sémanticosyntaxiques, sens (29.89%)

D'une façon plus précise, la plupart des erreurs morphologiques dans la catégorie morphologique, sont les mots mal construits (88.02%). Les erreurs morphosyntaxiques les plus récurrentes sont : la conjugaison (17.36%), le genre des noms (16.53%), l'infinitif après les prépositions : à, de, pour, sans (8.26%) et le manque du « s » du pluriel (8.26%) et enfin la confusion entre être et avoir (5.79%).

⁶ Vous pouvez consultez cette hiérarchie en détail dans la partie des Annexes C.

Pendant que la plupart de erreurs morphosyntaxiques sont des erreurs intralinguales ; d'autres nécessitent plus de catégories afin de connaître leurs causes. Les erreurs syntaxiques les plus récurrentes sont les erreurs qui confondent l'usage des préposition (36.76%), la durée de l'évènement (22.06%) et l'incohérence (17.65%).

Quantitativement parlant, le problème que nous avons rencontré durant cette description est celui de trouver les causes des erreurs d'une façon précise, il nous semble qu'il nécessite toute une recherche focalisant sur une seule catégorie. Qualitativement parlant, il est clair que les causes de la plupart des erreurs sont dues au manque de connaissances des élèves des règles qui régissent le bon fonctionnement de la langue française.

Les informations statistiques que nous avons relevées ont une très grande importance didactique. Elles montrent d'une façon détaillée les difficultés qu'éprouve l'apprenant lors de sa production écrite. Il est possible, à titre d'exemple, grâce à ces informations, d'élaborer des leçons spéciales dans le but d'éliminer ou de réduire ces difficultés.

Annexes

Annexes A :

Les copies des élèves

C1 F

Un jour dans le vacance je suis allée
à Bel Abbese chez Matente en à visitée
la ville pour acheter des nouvelles
vêtement et après sa nous somme allées
au bord de la mer pour amuser et puis
des photos et on à passer un belle soirée.
cette vacence est pleine de joie et je
vous jure par dieu tant puisant cette
journée très belle et je n'oubliera pas.

في يوم واحد في ايام العداوة كنت قد رجت عن طياتي
ليلي انا يا انا يا انا يا انا يا انا يا انا يا انا
وكلنا رجا لنا طرفة البحر يا انا يا انا يا انا
في تلك السهرة متابعي كان جوري في كان صدمت
بالفرحة واطمينة ونفسي لكم يا انا يا انا يا انا
كان شباب براف وصدقتيك نساها -

CSM

j'ai passé les vacances
 à la maison, la première
 semaine en faisant mes
 devoirs de temps en temps
 j'ai regardé la télé
 un peu de temps à l'ordinateur
 j'ai joué avec mes amis
 au ballon.
 j'ai passé la deuxième
 chez ma grand mère et
 mon grand père avec

mes tantes et mes voisins.

أنا فوت لعمل خدائ سيمان اول
 قريت صوالحي. وكنت ننتفح
 فكتي فريوت وكنت تا نيك بعب
 فليجروا وقلعشيا نلعب مع صحابي
 بالون.
 وسيمان ازا وجا كنت عنو جدي
 وجداتي وكذا فوت لعمل.

Annexes B :
Les tableaux d'études

Exemple 1

Enoncé	Lieu d'erreur	Type d'erreur	catégorie 1	catégorie 2	précision	Correction
J'ai passé les vacances à la maison	La maison	morphologique	forme canonique		mot mal écrit	La maison
La première semaine, en faisant mes devoir de temps en temps	Mes devoir	morphosyntaxique	flexionnelle	éléments de la phrase	manque: s du pluriel	Mes devoirs
La première semaine, en faisant mes devoir de temps en temps	La première semaine, en faisant mes devoir de temps en temps	sémanticosyntaxique	sens	éléments de la phrase	mauvaise structure / incompréhensible	Mauvaise structure / incompréhensible / La première semaine, j'ai fait mes devoirs. (Sens dérivé de sa production écrite en langue arabe dialectal)
J'ai regardé la télé un peu de temps à l'ordinateur	J'ai regardé la télé	sémanticosyntaxique	sens, temps	éléments de la phrase	durée de l'évènement	J'ai passé du temps à regarder la télé.
J'ai regardé la télé un peu de temps à l'ordinateur	Un peu de temps à l'ordinateur	sémanticosyntaxique	sens, temps	éléments de la phrase	suppression de plusieurs éléments obligatoires pour la compréhension	J'ai alloué un peu de temps à l'ordinateur.
J'ai joué avec mes amis au ballon	J'ai joué avec mes amis au ballon	syntaxique	éléments de la phrase		une phrase lourde	J'ai joué au ballon avec mes amis
J'ai passé la deuxième chez ma grand mère et mon grand père	Ma grand mère	morphologique	compositionnelle		mot mal écrit	Ma grand-mère

avec mes tantes et
mes voisins.

J'ai passé la deuxième
chez ma grand mère
et mon grand père
avec mes tantes et
mes voisins.

Mon grand père	morphologique	compositionnelle		mot mal écrit	Mon grand-père
----------------	---------------	------------------	--	---------------	----------------

Exemple 2

Enoncé	Lieu d'erreur	Type d'erreur	catégorie 1	catégorie 2	précision	Correction
Pour le vacance, j'ai passé des bonnes mouments	Pour	sémantique	sens		confusion lexicale	Pendant
Pour le vacance, j'ai passé des bonnes mouments	Le vacance	morphologique	forme canonique		mot mal écrit	Les vacances
Pour le vacance, j'ai passé des bonnes mouments	Des bonnes	morphosyntaxique	flexionnelle		accord de l'adjectif	De bons
Pour le vacance, j'ai passé des bonnes mouments	Bonnes mouments	morphosyntaxique	flexionnelle	éléments de la phrase	article indéfini devant l'adjectif	Bons moments
Pour le vacance, j'ai passé des bonnes mouments	Mouments	morphologique	forme canonique		mot mal écrit	Moments

Annexes B : Les tableaux d'études

Pour la semaine prochaine, je suis allé à la plage et puis des photos	Pour	sémantique	sens		confusion lexicale	Pendant
Pour la semaine prochaine, je suis allé à la plage et puis des photos	Pour la semaine prochaine	sémantique	sens		confusion lexicale	Pendant la deuxième semaine
Pour la semaine prochaine, je suis allé à la plage et pris des photos	Je suis allé à la plage et pris des photos	syntaxique	éléments de la phrase		substitution d'un élément par un autre inadéquat	Je suis allé à la plage où j'ai pris des photos
J'ai passé un belle soiré et mangé les glâces	Un belle soiré	morphosyntaxique	forme canonique	éléments de la phrase	article indéfini singulier féminin	Une belle soirée
J'ai passé un belle soiré et mangé les glâces	Soiré	morphologique	forme canonique		mot mal écrit	Soirée
J'ai passé un belle soiré et mangé les glâces	Les glâces	morphologique	forme canonique		mot mal écrit	Les glaces
J'ai passé un belle soiré et mangé les glâces	J'ai passé un belle soiré et mangé les glâces	sémanticosyntaxique	sens	éléments de la phrase	incohérence	Mauvaise construction/ sens
La plage s'est un belle pelase	La plage s'est un	syntaxique	éléments de la phrase		ajout d'un élément superflue	La plage est
La plage s'est un belle pelase	Pelase	morphologique	forme canonique		mot mal écrit	Place
La plage s'est un belle pelase	Un belle	morphosyntaxique	forme canonique	éléments de la phrase	article indéfini singulier féminin	Une belle

Annexes C :
La hiérarchie

Morphologique (37.89%)

Majuscule

Première lettre en majuscule au début du texte (1.20%)

Première lettre en majuscule pour les noms propres (2.40%)

Minuscule

Première lettre minuscule à l'intérieur du texte (2.99%)

Forme canonique

Mot mal écrit (88.02%)

Nom propre mal écrit (4.79%)

Manque : s du pluriel (0.60%)

Morphosyntaxique (25.47%)

Forme canonique

Éléments de la phrase

Le genre des noms (16.53%)

Confusion entre être et avoir (5.79%)

Confusion entre et et est (1.65%)

Article indéfini pluriel (1.65%)

Article défini féminin (2.48%)

Article indéfini singulier féminin (2.48%)

Mot mal écrit (4.13%)

Ajout d'un s superflue (1.65%)

Confusion on, en, a, à (2.48%)

Flexionnelle

Manque : s du pluriel (8.26%)

Accord du participe passé (7.44%)

Article indéfini devant l'adjectif (1.65%)

Conjugaison (17.36%)

Infinitif après les prépositions : à, de, pour, sans (8.26%)

Mot mal écrit (0.83%)

Conjugaison (17.36%)

Sémantique (6.32%)

Sens

Inusité (48.15%)

Ambiguïté (29.63%)

Confusion lexicale (22.22%)

Syntaxique (14.11%)

Éléments de la phrase

Ajout d'un élément superflue (21.21%)

Suppression d'un élément obligatoire (30.30%)

Suppression de plusieurs éléments obligatoires (18.18%)

Substitution d'un élément par un autre inadéquat (15.15%)

Virgule devant et (9.09%)

Une phrase lourde (6.06%)

Sémanticosyntaxique **(16.21%)**

Sens

Éléments de la phrase

Suppression d'un élément obligatoire **(5.88%)**

Suppression de plusieurs éléments obligatoires **(1.47%)**

Mauvaise structure **(2.94%)**

Incohérence **(17.65%)**

Confusion dans l'usage des prépositions **(36.76%)**

Substitution d'un élément par un autre inadéquat **(7.35%)**

Homonyme **(4.41%)**

Éléments de la phrase

Sens, temps

Durée de l'évènement **(22.06%)**

Suppression de plusieurs éléments obligatoires pour la compréhension
(1.47%)

Bibliographie

Bibliographie

- Bouillon, P. & Vandooren, F. (1998). *Traitement automatique des langues naturelles*. Paris: Duculot.
- Corder, S. (1980). Dialectes idiosyncrasiques et analyse d'erreurs. *Langages*, 14(57), 17-28. doi:10.3406/lgge.1980.1834
- Corder, S. (1980). Que signifient les erreurs des apprenants ?. *Langages*, 14(57), 9-15. doi:10.3406/lgge.1980.1833
- Cuq, J. & Gruca, I. (2005). *Cours de didactique du français langue étrangère et seconde*. Grenoble: Presses universitaires de Grenoble.
- Dabène, L. (1994). *Repères sociolinguistiques pour l'enseignement des langues*. Paris: Hachette.
- Debyser, F. (1970). La linguistique contrastive et les interférences. *Lfr*, 8(1), 31-61. doi:10.3406/lfr.1970.5527
- Desoutter, C. Quelle est la place aujourd'hui des études contrastives en didactique des langues étrangères ?. *Gerflint*, 117-126. Consulté à l'adresse <http://gerflint.fr/Base/Italie2/desoutter.pdf>
- Frauenfelder, U., Noyau, C., Perdue, C., & Porquier, R. (1980). Connaissance en langue étrangère. *Langages*, 14(57), 43-59. doi:10.3406/lgge.1980.1837
- Gaudoin, O. *Principes et Méthodes Statistiques* (1st éd., p. 8). Consulté à l'adresse <http://www-ljk.imag.fr/membres/Olivier.Gaudoin/PMS.pdf>
- Sörös, A. (2008). *Typologie et linguistique contrastive*. Bern: Peter Lang.
- The Encyclopedia of language and linguistics. (1994). Consulté à l'adresse <http://phonetiquedufle.canalblog.com/archives/2008/04/30/9006958.html>

Urbain, J. (1982). La langue maternelle, part maudite de la linguistique?. *Lfr*, 54(1), 7-28. doi:10.3406/lfr.1982.5275

Wardhaugh, R. (1970). The Contrastive Analysis Hypothesis. *TESOL Quarterly*, 4(2), 123. doi:10.2307/3586182

Table des matières

Table des matières

Le sommaire.....	3
Introduction.....	3
Introduction.....	5
Partie théorique	3
Chapitre I :	3
La langue maternelle	3
1. La langue maternelle.....	12
1.1. L’impasse définitoire.....	12
1.2. Vers un consensus définitoire	14
1.2.1 La langue de la mère	14
1.2.2 La première langue acquise	14
1.2.3 La langue la mieux connue	15
1.2.4 La langue acquise « naturellement ».....	15
1.2.5 Langue et langage.....	15
Chapitre II :	10
La linguistique contrastive	10
1. La linguistique contrastive.....	19
2. Le terme d’interférence	24
2.3. En psychologie	25
2.4. En linguistique.....	25
2.5. En pédagogie.....	26
3. Les limites de la linguistique contrastive	27
Chapitre III :.....	17
L’analyse des erreurs.....	17
1. L’analyse des erreurs	30
2. Les lapsus.....	37
3. La compétence transitoire ou (Dialectes idiosyncrasiques)	39
4. Erreurs interlinguales et erreurs intralinguales	42

4.1. Les erreurs intralinguales	42
4.2. Les erreurs interlinguales	43
5. Les étapes de l'analyse des erreurs	43
5.1. Trouver l'erreur	43
5.2. Décrire l'erreur	43
5.3. Expliquer l'erreur	44
La partie pratique	27
Chapitre I :	27
Définitions des éléments clés	27
1. Comment cette partie va être organisée.	47
2. La collecte des données	48
3. Population d'enquête	48
4. La méthode d'analyse	49
5. Type d'enquête	50
Chapitre II :	45
La construction d'une typologie	45
1. Les catégories simples	52
1.1. Les erreurs morphologiques	52
1.2. Les erreurs syntaxiques	54
1.3. Les erreurs sémantiques	57
2. Les catégories combinées	59
2.1. Les erreurs morphosyntaxiques	60
2.2. Les erreurs sémanticosyntaxiques.....	62
Chapitre III :	48
Description et interprétation des résultats.....	48
Description des résultats	65
1. Les types d'erreurs (catégories simples et combinées)	65
2. Les catégories simples	66
3. Les catégories combinées	69
4. Les erreurs morphologiques	72
5. Les erreurs Morphosyntaxiques.....	73

6. Les erreurs sémantiques	75
7. Les erreurs syntaxiques	75
8. Les erreurs sémanticosyntaxiques.....	76
Conclusion	79
Annexes.....	73
Bibliographie	95

Résumé

Dans cette recherche, nous avons essayé d'analyser les erreurs des élèves d'une classe de troisième année du moyen afin, d'une part, de découvrir les erreurs les plus récurrentes qui ont le sens de difficultés que les apprenants éprouvent lors de leurs productions écrites, et d'autre part, de rendre un bilan constituant « des informations d'importance didactique » (Desoutter :120). Pour atteindre ces buts nous avons fixé un ensemble d'objectifs : la sélection d'une classe, d'une population, la collecte des données à partir de leurs productions écrites, la construction et l'application d'une grille d'analyse, et enfin l'identification et le dénombrement de toutes les erreurs. Il est possible, grâce aux informations que nous avons relevée, d'élaborer des leçons spéciales dans le but d'éliminer ou de réduire ces difficultés.

Mots-clés : analyse, erreurs, productions, écrites, élèves, moyen, récurrences, difficultés, apprentissage, didactique, typologie, quantification, linguistique, appliquée
