English Lexical Errors Committed by Algerian Learners
The Case of Second Year University Students

Abstract

This paper investigates the lexical errors committed by Algerian learners of English in their written compositions; our main aim is to analyse their lexical deviations following James's (1998) taxonomy, and to find the reasons that lead them to err. The obtained results revealed that errors of distortion are the most frequent error type followed by collocations. The main source of errors is the learners' serious lack of vocabulary knowledge in form and meaning.

Meriem MAZRI

Faculty of Letters and Languages
Department of Foreign Languages
University of Mentouri
Constantine

Introduction

Different scholars agree that despite being the most frequently occurring category of errors in written English, lexical errors have remained under-researched. To justify their little attention towards investigating the area of lexical errors, researchers claim that compared to the less complexity and systematic rules of morphology, syntax and grammar, lexical items "represent the idiosyncratic, non-generalizable features of language" (Warren, 1982:209).

ملخص

يدرس هذا المقال الأخطاء المفرداتية التي يرتكبها الطلبة الجزائريون عند كتابتهم باللغة الإنجليزية. هدفنا الرئيسي يكمن في تحليل أخطائهم المفرداتية وفق تصنيف جيمس (1998) والوقوف على الأسباب التي تؤدي بهم إلى الخطأ. أظهرت النتائج المتحصل عليها أن الأخطاء الأكثر شيوعا هي أخطاء التحريف تليها أخطاء التلازم اللفظي. المصدر الرئيسي للأخطاء هو نقص المتعلمين في معرفة المفردات شكلا و مضمونا.

Similarly, Duskova (1969) maintains that "errors in lexis presented a much less homogeneous material for study than errors in grammar" (p.24). Accordingly, lexical errors are considered as a source of disruption; hence, they deserve a better attention. In this paper, we intent to explore the various types of lexical errors produced by Algerian learners to gain more insights about their occurrences and answer the following questions:

- What are the most frequent types of lexical errors committed in Algerian learners' English compositions?

- What can be the main reasons that lead them to produce lexical errors?

From the above asked questions, we hypothesize that distortions and collocations are the main lexical errors students face.

1. Research in Lexical Errors

Various inquiries related to lexical errors agreed on the significant complexity of this type of deviations. Referred to by Yang and Xu (2001) as "creative errors", they are undoubtedly unsystematic and irregular with reference to the unstable nature of the lexicon; subsequently, they are under-researched. Among the studies that have been undertaken, Duskova (1969) identifies during her analysis of 50 Czech postgraduate students' writing compositions four types of lexical errors: confusion of words due to formal similarity, due to relatedness of meaning, assumed equivalents between Czech and English and spelling errors. Also, Ringborn (1983) focuses his attention on interference errors that derive mainly from the influence of L1 in addition to previously learned languages (L3, L4). His study, actually, is limited in the sense that he centers his investigation only on lexical errors produced by interference with previous linguistic knowledge, disregarding other causes like L2 induced lexical errors. Moreover, Engber (1995) characterizes lexical errors between errors in the lexical choice and those in the lexical form. The former involves wrong choice of individual word and of combinations whereas the latter comprises derivational, verb forms, phonetically related and spelling errors. Furthermore, based on his studies on Yugoslavian students of English, Djokic (1999) develops three main types of lexical errors, namely, substitution, omissions and additions. In the light of her research on advanced Spanish-speaking learners of Italian, Ambroso (2000; cited in Agustin Llach, 2011) establishes a classification of lexical errors due to stylistic errors (pragmatics), syntactic errors, order errors (collocation), semantic errors and idiosyncratic errors. Still, with the ambiguous examples provided in her findings, the use of this taxonomy is questioned when the languages concerned are more lexically different than Italian and Spanish. From these distinct inquiries on lexical errors, we confirm the instability and unsystematic occurrence of such type of deviations. They are difficult to predict; henceforth, they need further considerations.

2. Definition of Lexical Errors

Generally, most studies and researches are concerned with the identification, description and classification of lexical errors without giving a clear definition. Theoretically speaking, two kinds of 'inconsistencies' are related to the term 'lexical error', namely, '*polysemy of meanings*' and '*polysemy of terms*'

Polysemy of meanings alludes to the various explanations researchers associate under the term 'lexical error'. For some, it is considered as the category of errors that is disassociated from grammar (spelling, semantic errors, errors of word choice and pragmatic errors). For others, it is a subordinating term that attends to present a heading for diverse other classes of errors like 'word formation', 'collocation', 'confusion' and 'relatedness' in a form of subcategories (Agustin Llach, 2011, p.73)

Polysemy of terms, on the other hand, refers to the application of various terms that deal with the same phenomenon which is errors related to lexis. Terms like '*wrong lexical choice*', '*errors in the lexical choice*', '*lexical deviances*', '*vocabulary errors*', '*semantic deviations*', '*synforms*', '*lexical deficiencies*' and '*lexical approximation*' are used to substitute the phrase 'lexical error' with regard to the aspects emphasized. For example, Laufer (1991) uses 'lexical confusions' or 'synforms' based on her limited study on wrong word choice. Similarly, Zimmerman (1987) applies 'semantic deviation' with regard to his study on confusion of semantically related words. (ibid.p.74)

In the light of the review of studies presented above, a lexical error is defined, according to Agustin Llach (2011) as a "deviation in form and/ or meaning of target language lexical word" (p.73); she maintains that a shift in form here covers deviations at the orthographic or phonological level within the limits of words only in addition to false collocations. Deviations related to meaning, on the other hand, emerge as a result of misuse of lexical items in context; it substitutes errors caused by incorrect choice of lexical item or inconsiderable use of some semantic features.

3. Lexical Errors Classification

Previous studies on lexical errors set up a great number of different error taxonomies, as illustrated by Kallkvist (1998): "no two previous studies on lexical errors have adopted the same error typology, and categorizing lexical errors is far from a straightforward exercise" (p.82). They have been designed in order to describe and explain learners' errors to find, by then, the most common areas of difficulties learners face in their written and spoken productions.

As far as this research is concerned, our expanded framework for lexical error analysis is mainly extracted from James's (1998) taxonomy. In the light of his studies, James (1998) views lexical errors from two different perspectives. Based on the classic word knowledge framework proposed by Richards (1976) referring to the necessity of knowing a word from its morphology comprising its spelling and pronunciation, syntactic behavior, functional and situational restrictions, semantic values, secondary meaning or connotations, word association and frequency of use, James (1998) differentiates between '*form oriented*' and '*content oriented*' lexical errors by attributing the former to 'formal errors' and the latter to 'semantic errors'. These two main classes, in their turn, are divided into subgroups, as it is illustrated[1] and described below:

3.1. Formal Errors

3.1.1. Formal Misselection

Also called '*synforms*' by Laufer (1991), this sort of errors comprises deviations of similar lexical forms; they include, as James (1998) claims, errors of the malapropism type where learners get confused between pairs of words that look and sound similar. It compiles words which are different in **suffix** (*sick/ sickness)*, **prefix** (*place/ replace)*, **vowel** (*meet/ meat*) or **consonant** (*save/safe*).

3.1.2. Misformation

Moreover, misformation errors result from the invention of a non-existing L2 word. According to James (1998) they are the result of the mother tongue influence on learners' language production. Here, three types of errors are distinguished:
a) **Borrowing:** it is using a word from the L1 without changing it (sel/salt)
b) **Coinage**: it is inventing a word from the L1 (exerced/exercer)
c) **Calque**: it is translating a word or phrase from L1 as in (rich <u>by vitamins</u>/

<div dir="rtl">غني <u>بالفيتامينات</u>)</div>

3.1.3. Distortion

Errors that arise from letter **overinclusion** or **addition**(*untill/ until*), **omission** (*hapiness/ happiness*), **misselection** (*unclode/include*), **misordering** (*specailly/specially*) and **blending** (*bigg/ big + bigger*) are called distortions.

3.2.   Semantic Errors

Semantic errors, on the other hand, are subdivided into two main categories, confusion of sense relations (CSR) and collocations.

3.2.1. Confusion of Sense Relation (CSR)

This category of errors occurs when there is a confusion of words related semantically by using a **general term instead of a specific one** or vice versa (shop/supermarket), applying a **wrong nearsynonym** (empty time/leisure time) or an **inappropriate co-hyponym** (daisy/orchid)

3.2.2. Collocation Errors

It is known that the term collocation refers to the common co-occurrence of words or phrases in a way that sounds natural and correct for native speakers. James (1998) specifies three degrees of misapplication of collocation, namely, **semantic word selection** (to escape diseases/ to prevent) **statistically weighted preferences** (to stop/to quit), and **arbitrary combination** (body's person/ person's body).

4. Research Design
4.1.   Participants
The participants were fifty students (females: 40-males:10, age range: 20-23) in their second year of study in the Department of Englishat the University of Constantine1 in Algeria. They all share a similar educational and linguistic background.

4.2.   Procedure

To analyse the types and frequency of lexical errors in students' written English compositions, the subjects were asked to write an example essay of five paragraphs (introduction, three developmental paragraphs, and conclusion) within one and a half hours on a topic related to the things people do to stay healthy, during the lecture of written expression.

4.3. Results and Discussion

Throughout the analysis of the data, we have noticed a variety in the length of the participants' written compositions. This difference, accordingly, mirrors the frequency and variation of the errors committed. Table1 shows that the fifty essays written by second year students of English at University of Constantine1 yielded 781 lexical errors. With a total number of 501 deviations, formal errors present the most problematic error category (64.15%); they are roughly twice as frequent as semantic errors (35.85%).

Error Types	N°	%
Formal Errors	501	64.15
Semantic Errors	280	35.85
Total	**781**	**100**

Table.1. Distribution of Lexical Errors

4.3.1. Formal Errors

Of the three main formal categories of errors, distortions are the most common ones; they represent (78.85%) of the total formal deviations followed by formal misselection (13.58%), and misformation (7.59%) as it is shown in table 2.

Formal Errors	N°	%	N° of papers
Formal Misselection	**68**	**13.58**	**35**
Suffix type	44	8.78	26
Prefix type	01	0.2	01
Vowel based type	17	3.4	11
Consonant based type	06	1.2	06
Misformation	**38**	**7.59**	**23**
Borrowing	27	5.39	19
Coinage	03	0.6	03
Calque	08	1.6	06
Distortion	**395**	**78.85**	**50**
Omission	101	20.16	47
Overinclusion (addition)	107	21.36	42
Misordering	13	2.6	09
Misselection	174	34.73	45
Blending	00	00	00
Total	**501**	**100**	**50**

Table.2. Distribution of Formal Errors

As far as distortions are concerned, we have noticed that our subjects have serious problems in misspellings. Being committed in fifty papers signifies that second year students of English have difficulties in writing correct English words. Indeed, the

83

occurrence of this type of formal errors is related to the learners' lack of lexical knowledge of the target language (TL), their concentration on '*finding the right words and sentences to express their meaning instead of editing*' (Raimes 1985:247), or simply their anxiety or fear of not having enough time to finish the writing task. As it is shown in table 2, misselection represents 34.73% of the total formal errors. They are manifested in forty five papers in the use of words like "suger", "spesially", "vegitebles", "animia", "obisity", "defficult", "anough", "frech", "whech" and "appal" where students failed in selecting the appropriate letter(s). Besides, some error types like overinclusion (21.36 %) and omission (20.16%) are less common. The former is illustrated in the occurrence of words like: "powerfull", "usefull", "untill", "protean", "plaice", "awhay" and "whay" whereas the latter is noticed in the following: "helthy", "plaing", "realy", "befor", "therfore" and "exale". Corresponding to 2.46% of the whole number of formal deviations, misordering is relatively infrequent, they are appeared in "thier", "dei", "brian", "tow", "contian" and "avioding" in which our participants ignored their correct form. In addition to that, no occurrence of blending is identified.

A propos of formal misselection (FM) deviations, they are found in 68 cases representing 13.58% of formal errors. Committed in 35 papers, they are the result of the wrong use of particular word class (noun, verb, adjective or adverb.), the insufficient knowledge of the word family, and the confusion between words similar in pronunciation. Hemchua and Schmitt (2006) suggest that the use of bilingual dictionaries is the main source of such type of errors since the majority of words are translated as single words with no reference to their use in context. As it is shown in table 2, of all FM deviations, wrong suffixation is the most recurrent subgroup; it constitutes 8.78% of the total number of formal errors. This kind of errors include the wrong distribution of parts of speech such as the replacement of a noun with an adjective or vise versa as in: "health/healthy", "sickness/sick", "safety/safe", "danger/dangerous", "stress/stressed", "obese/obesity", "interested/interest" ; a noun with a verb as in "breathe/breath", "suffer/suffering", "bless/blessings" ; a verb with an adjective as in "attract/attractive", or an adjective with a adverb like in "happy/happily". Moreover, errors of vowel based type are relatively infrequent (3.4%); being committed in 11 papers, they incorporate the confusion between words that sound similar such as "effect" rather than "affect", "then" instead of "than" or "luck" rather than "lack". Illustrated in examples like "safe/save" and "three/tree", errors of consonant based type represent 1.2% of formal errors followed by one case of deviation related to the prefix type (0.2%): "replace" instead of "place"

Following FM errors, misformation is the least frequent subcategory. Identified in 23 papers, they represent 7.59% of the total number of formal errors. The production of misformation deviations indicates the participants' uncertainty or lack of the English lexical knowledge in addition to the influence of other languages on their writing performance. As a result, they find the necessity to borrow, create or to translate from what they have learned and acquired in other languages. As far as misformations are concerned, they are 27 cases in which our participants borrowed from the French language; this can be illustrated in the use of words like: "positif", "negatif", "docteur", "plaisible" and "medicament". Furthermore, errors compiled in the subcategory of calque constitute 1.6% of all formal deviations. Here, some students translated from

French and Arabic languages as in: "may affect <u>on our health</u>" (قد تؤثر على صحتنا),
"<u>for stay</u> healthy" (للبقاء في صحة جيدة), and "rich <u>by vitamins</u>" (غني بالفيتامينات).
Throughout our analysis, we have noticed the invention of three words due to the
influence of French: "exerced" (exercer), "rechutate" (rechuter) and "nourition"
(nourriture).

4.3.2. Semantic Errors

As noted above (table1), semantic errors are less problematic than formal errors;
nevertheless, of its two subtypes, with a total number of 269 errors, collocations are the
second type of lexical errors, after distortions, which are problematic for the subjects.
They account for 96.07% of the total semantic errors, followed by confusion of sense
relation (CSR) errors as it is illustrated in table 3.

Type		N°	%	N° of papers
Confusion of sense relation (CSR)		**11**	**3.93**	**10**
	General term for specific one	00	00	00
	Overly specific term	01	0.36	01
	Inappropriate co-hyponyms	00	00	00
	Near synonyms	10	3.57	09
Collocation		**269**	**96.07**	**50**
	Semantic word selection	177	63.21	43
	Statistically weighted preferences	32	11.43	24
	Arbitrary combinations	60	21.43	29
Total		**280**	**100**	**50**

Table 3. Distribution of Semantic Errors

The analysis of the students' compositions reveals that they have a serious problem
of collocations in their writing. The most numerous collocational errors type is related
to semantic word selection. Having 177 cases, which is equivalent to 63.21% of the
total number of semantic errors, demonstrates the students' lack and insufficient
vocabulary knowledge. Put otherwise, our participants were unable to convey sufficient
meaning in their writing; indeed, along the analysis, we came across some meaningless
and ambiguous sentences, and wrong word choices. For example, to refer to the
preciousness and importance of health, a student wrote "*health is a crow*" instead of
"*crown*". He could not distinguish the difference between the bird and what a king
wears on his head. Similarly, among the errors identified in this subcategory the
production of "being attention", "on the other hand", "advice you must do", and "our
right" instead of "paying attention", "in other words", "advice you should follow", and
"our duty". Following errors of semantic word selection, deviations related to the
arbitrary combination of lexical items totalled 60 cases, which is equivalent to 21.43%.
This finding demonstrates the influence of the students' mother tongue in addition to
their lack of grammatical knowledge. As an illustration, deviations of arbitrary
combination appeared in "should have to do", "diseases of the heart", "vegetable
eating", "must people do" and "a life better" rather than "should do", "heart diseases",
"eating vegetables", "people must do" and a better life". Last but not least, as it is

shown in table 3, deviations related to word choice preferences constitute 11.43% of the total sum of semantic errors. There are 32 cases in which the use of some words was preferable regarding the students' intended meaning like the production of "to stop smoking" rather than "to quit smoking", "different food" instead of "varieties of food" , or "two times" rather than "twice".

Regarding deviations related to CSR, they are the least frequent ones among lexical errors in general and semantic errors in particular. Identified in 10 papers, they represent 3.93% of the semantic deviations. Starting with errors of near synonyms, they are manifested in 9 papers where 10 cases are diagnosed; this equals 3.57%. Among the errors produced, we notice the use of "let an eye", "dangerous diseases", or "empty time" rather than "keep an eye", "serious diseases", and "free/leisure time". Furthermore, there is one case where a student used an overly general term when a specific one is needed (0.36%). Here, he wrote "shop" rather than "supermarket". In addition to that, no occurrence of errors related to the use of a general term for a specific one and those related to inappropriate use of a co-hyponym are distinguished.

Conclusion

In a nutshell, this study focuses on the lexical errors produced by Algerian learners; nevertheless, they are likely to be problematic to a large number of second /foreign language learners. Our findings demonstrate that deviations of distortion followed by collocations represent the overwhelming majority of lexical errors produced by learners in their compositions. Also, this analysis indicates the learners' serious lack of vocabulary knowledge in form and meaning which, subsequently, lead them to borrow or translate the equivalent words from previously learned languages and, hence, facing the problem of interference. Therefore, it is essential to bear in mind the role of lexical errors to indicate the process of vocabulary acquisition since they provide insights about the areas of difficulties faced by FL learners during their learning phases; accordingly, they contribute to the identification of their lexical gaps. Moreover, they facilitate the teacher's task in emphasizing the problematic areas in L2 vocabulary acquisition in addition to providing the researcher with insights into the learners' lexical competence and how it evolves.

References

- Agustin Llach, M.P. (2011) *Lexical Errors and Accuracy in Foreign Language Writing*. UK: Multilingual Matters.
- Ambroso, S. (2000) Descripcio´n de los errores le´xicos de los hispanohablantes.Ana´lisis de la Produccio´n escrita de IT, El Certificado de Competencia General Enitaliano Como L2. In Agustin Llach, M.P.(2011) *Lexical Errors and Accuracy in Foreign Language Writing*. UK: Multilingual Matters.
- Djokic, D. (1999) Lexical Errors in L2 Learning and Communication. *Rassegna Italina di Linguistica Aplicata*Vol. 31, 123_135.
- Duskova, L. (1969). On Sources of Errors in Foreign Language Learning. International *Review of Applied Linguistics*, 7, 11–36.
- Engber, C.A. (1995) The Relationship of Lexical Proficiency to the Quality of ESL Compositions. *Journal of Second Language Writing* 4 (2), 139_155.
- Hemchua, S. and Schmitt, N. (2006) An Analysis of Lexical Errors in the English Compositions of Thai learners. Prospect Vol. 21, 3-25. In Agustin Llach, M.P.(2011) *Lexical Errors and Accuracy in Foreign Language Writing*. UK: Multilingual Matters
- James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. New York: Longman.
- Kallkvist, M. (1998) How Different are the Results of Translation Tasks? A study of lexical errors. In Agustin Llach, M.P.(2011) *Lexical Errors and Accuracy in Foreign Language Writing*. UK: Multilingual Matters
- Laufer, B. (1991) Some Properties of the Foreign Language Learner's Lexicon as Evidenced by Lexical Confusions. IRAL 29 (4), 317_330. In Agustin Llach, M.P.(2011) Lexical Errors and Accuracy in Foreign Language Writing. UK: Multilingual Matters.
- Raimes, A. (1985). What Unskilled ESL Students do as they Write: A Classroom Study of Composing. *TESOL Quarterly*, 19, 229–258
- Richards, J. C. (1976). The Role of Vocabulary Taching. *TESOL Quarterly*, 10, 77–89.
- Ringbom, H. (1983) Borrowing and Lexical Transfer. *Applied Linguistics* 4, 207_212.
- Warren, B. (1982) Common Types of Lexical Errors Among Swedish Learners of English. *Moderna Sprak* 76 (3), 209_228.
- Yang, X. & Xu, H. (2001). *Errors of Creativity.* An Analysis of Lexical Errors Committed by Chinese ESL Students. University Press of America, Inc.
- Zimmermann, R. (1987) Form-oriented and Content-oriented Lexical Errors in L2 Learners. IRAL 25, 55_67.In Agustin Llach, M.P.(2011) *Lexical Errors and Accuracy in Foreign Language Writing*.UK: Multilingual Matters.

Note:

1-All examples are taken from the participants' compositions except for those related to blending, general term for specific one, and inappropriate co-hyponym.