An analysis of lexical errors in the English compositions of EFL Tunisian learners

Sondes Hamdi
ISLT, Tunis
University of Carthage, Tunisia
sondeshamdi@yahoo.fr

Abstract

Research in applied linguistics emphasizes the significance of learners’ errors in SLA and EFL contexts. Despite the continuous interest in error analysis, to the best of our knowledge, no study has provided a systematic analysis of lexical errors committed by Tunisian learners of English at the ISEAH Institute of Kef (Tunisia). This paper tries to fill up this gap by investigating the lexical errors made by 20 EFL learners of Business English at the ISEAH Institute of Kef. The participants were in their third year of study at the ISEAH Institute of Kef (Tunisia). They had learned EFL for eight years. Arabic and French were the languages they spoke at school. Arabic was the only language they spoke at home and with friends. The participants were two males and 18 females, similar in age, ranging from 19 to 21 years old. They had had little previous English-writing experience in their secondary school, since writing was not emphasized at these levels. The data analysis was conducted within the framework of James’s lexical error taxonomy (1998). The results showed that the participants committed more formal errors (94.44 % of the total number of lexical errors) than semantic errors (only 0.05 % of the total number of lexical errors). The implications of the study for English teachers have been highlighted.

Keywords: EFL Acquisition, English, Arabic, Lexis, Error Analysis, transfer.
Introduction

Teachers contend that EFL/ESL learners inevitably produce diverse types of errors. This is, in fact, the process of learning a language. Error Analysis (EA) as a method of analyzing the learners’ errors played a role in L2 acquisition research in the 1970s. EA became the main methodology used for investigating learner language and L2 acquisition, overriding the Contrastive Analysis (CA) method. This was mainly due to the inherent shortcomings of CA and the motivation to improve pedagogy on the basis of the investigation of errors (Corder, 1975). CA looks only into the contrastive features of the two languages, i.e. L1 and the Target Language, while EA is concerned with the learner language and the process of language learning. Corder (1981, p. 36) proposes three main steps of EA research: Identification of errors, description of errors, and explanation of errors.

According to the literature, errors can be interlingual or intralingual. Interlingual errors are attributable to the native language (NL). They occur when the learner’s L1 habits (patterns and systems) interfere and prevent them from acquiring patterns in the TL (Corder, 1971). In other words, interlingual errors are the result of a negative transfer from the mother tongue (L1) to the TL. To put it differently, interlingual errors are “those instances of deviation from the norms of either language which occur in the speech of bilinguals as a result of their familiarity with more than one language” (Weinreich, 1953, p.1). Intralingual errors, on the other hand, are those due to the TL itself.

Lexical errors are worth-studying for several reasons. To start with, empirical studies (Meara, 1984; Lennon, 1991) demonstrate that lexical errors are the most frequently occurring kind of errors in English. They considerably impact the quality of academic writing (Engber, 1995) and native speakers conceive of them as the most “irritating” errors (Santos, 1988). Secondly, “mistakes in lexical selection may be less generously tolerated outside the classroom than mistakes in syntax” (Carter, 1998, p. 185). This is probably true because inappropriate lexical choices could blur communication by making the message misunderstood. Interestingly enough, even though lexical errors are frequent and serious, they have remained under-researched (Hemchua & Schmitt, 2006). In the same vein, to the best of our knowledge, no study has so far conducted a systematic analysis of lexical errors committed by English learners at the ISEAH Institute of Kef (Tunisia). This paper will try to fill up this gap, at least partially, by looking into the lexical errors committed by EFL third-year students at the ISEAH Institute of Kef (Tunisia).

Thus, in an attempt to give insights into the various types of lexical errors committed by Tunisian learners of English at the ISEAH Institute, this study will try to answer the following questions:

1) What lexical errors do the third-year Tunisian learners at the ISEAH Institute make in their English compositions?
2) Which of the errors are most frequent?
This paper comprises four main sections. The first section surveys a few significant studies on lexical errors in SLA and EFL contexts. The second section defines the lexical error taxonomy used for analysis in this paper. The third section presents the methodology used and describes the data collected. The fourth section analyzes and discusses the data at hand. It also gives insights into the pedagogical implications of lexical error analysis.

1. Literature review

Despite the importance and the frequency of lexical errors in SL and FL students’ English compositions, little research has been conducted on lexical errors. However, among the few studies that have been concerned with the analysis of lexical errors in SLA and EFL contexts, we find Duskova (1969), Henning (1973), Laufer (1991), and Hemchua & Schmitt (2006). Duskova (1969) identified four types of lexical errors when she assessed the writings of 50 Czech students: confusion of words with formal similarity, similar meaning, misuse of words generated by one or several equivalents between Czech and English. These findings corroborate those of Henning (1973), whose study suggested that FL learners stored FL vocabulary in connection with form and meaning. This analysis demonstrated that beginners had the tendency to make form-based associations, while more advanced learners tended to make meaning-based associations. Laufer (1991) gave empirical evidence for FL lexical confusion in case of similarity between the lexical form of target words and errors (i.e. “synformic confusions.”) These are words with the same root but different suffixes or prefixes, words identical in all phonemes except one vowel, diphthong or consonant. Hemchua and Schmitt (2006) scrutinized Thai university students’ English compositions for lexical errors. Their analysis suggested that “near synonyms” were the most frequent ones, followed by “preposition partners” and suffixes. The sources of errors identified in this study were from L2 intrinsic difficulty rather than the L1 transfer.

This paper will use a comprehensive error taxonomy based on James’s (1998) error classification in order to analyze Tunisian third-year university students’ English compositions for lexical errors. This lexical error taxonomy will be exposed in some detail below.

2. Lexical error taxonomy

The framework for lexical error classification is mainly based upon James’s (1998) lexical error taxonomy. James categorizes lexical errors into two major types: formal and semantic features, as defined below in some detail:

2.1. Formal errors

Formal errors are divided into three types: Formal misselection, misformations, and distortions (James, 1998). The sub-types and examples for each type are stated below:

2.1.1. Formal misselection

These are “synforms” that share some phonemes/graphemes., i.e. similar lexical forms (visual and sound similarity). The four major types of these synforms are: the suffix type (Eg.
competition/competitiveness), the prefix type (Eg. reserve/preserve), the vowel-based type (Eg. seat/set), the consonant based type (Eg. save/safe).

2.1.2. Misformations

These are words that do not exist in the TL. The source of errors is from the learner’s L1. Thus, they are “intralingual errors.” James (1998) classifies misformation errors into three types: Borrowing (L1 words are used in the TL without change), coinage (i.e. inventing a word from L1), calque (i.e. translation of a word or a phrase from L1 words).

2.1.3. Distortions

These words do not exist in the TL. However, the errors are the result of misapplication of the TL without transfer from L1. James (1998) divides distortions into five types: Omission (intresting instead of interesting), overinclusion (dinning room instead of dining room), misselection (delitous instead of delicious), blending (travei instead of travel).

2.2. Semantic errors

James highlights two types of semantic errors: Confusion of sense relations and collocation errors. Confusion of sense errors encompasses four types of errors: (1) using a superonym for a hyponym, (2) using a hyponym for a superonym, (3) using inappropriate co-hyponyms, (4) using a wrong near synonym. Collocation is a word or phrase that is frequently used with another word or phrase. Figure 1 summarizes the aforementioned error taxonomy:
I. Formal errors

1. Formal misselection
   1.1. Suffix type
   1.2. Prefix type
   1.3. Vowel-based type
   1.4. Consonant-based type

2. Misformations
   2.1. Borrowing (L1 words)
   2.2. Coinage (inventing based on L1)
   2.3. Calque (Translation from L1)

3. Distortions
   3.1. Omission
   3.2. Overinclusion
   3.3. Misselection
   3.4. Misordering
   3.5. Blending

II. Semantic errors

1. Confusion of sense relations
   1.1. General term for specific one
   1.2. Overtly specific term
   1.3. Inappropriate co-hyponyms
   1.4. Near synonyms

2. Collocation errors

Figure 1: James's lexical error taxonomy

3. Method

3.1. Participants

The participants were 20 Tunisian English students in their third year of study at the ISEAH Institute of Kef (Tunisia). They had learned EFL for eight years. Arabic and French were the languages they spoke at school. Arabic was the only language they spoke at home and with friends. The participants were two males and 18 females, similar in age, ranging from 19 to 21 years old. They had had little previous English-writing experience in their secondary school, since writing was not emphasized at these levels. They took a basic writing course in the first year at the university and a paragraph-writing course in the second year. In addition to classroom English input, the participants were sometimes exposed to English language through the internet and English movies. However, they seldom communicated with English native speakers outside the classroom.

3.2. Procedure

To investigate the types of lexical errors in the students’ written English, the participants were asked to write an argumentative composition of about 300-350 words, without consulting
their dictionaries, within an hour, on a topic related to the advantages and disadvantages of life in the country. In order to determine what types of lexical errors were made by the participants, the compositions were corrected by two experienced English teachers.

4. Results and discussions

The 20 compositions written on the same topic by different participants contained 54 lexical errors. That is, on average, each paper yielded 3 lexical errors. Some error types were common (for example, suffix type, borrowing, overinclusion), others were relatively infrequent (for example, vowel-based type, coinage, misselection), and there were no occurrences of others at all (prefix type, misordering, blending).

4.1. Formal errors

As far as formal errors are concerned, the formal distortion of words was the most problematic (40.74 % of the total number of lexical errors), followed by misformations (31.48 % of the total number of lexical errors), misselection of words was less problematic (22.22% of the total number of lexical errors). Table 1 illustrates the frequency of lexical formal errors found in the participants’ writings:

<table>
<thead>
<tr>
<th>Error types</th>
<th>Number of total errors: 54</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Formal misselections</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1.1. Suffix type</td>
<td>9</td>
<td>16.66</td>
</tr>
<tr>
<td>1.2. Prefix type</td>
<td>3</td>
<td>5.55</td>
</tr>
<tr>
<td>1.3. Vowel-based type</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.4. Consonant-based type</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Misformations</td>
<td>17</td>
<td>31.48</td>
</tr>
<tr>
<td>2.1. Borrowing</td>
<td>11</td>
<td>20.37</td>
</tr>
<tr>
<td>2.2. Coinage</td>
<td>1</td>
<td>1.85</td>
</tr>
<tr>
<td>2.3. Calque</td>
<td>5</td>
<td>9.25</td>
</tr>
<tr>
<td>3. Distortions</td>
<td>22</td>
<td>40.74</td>
</tr>
<tr>
<td>3.1. Omissions</td>
<td>8</td>
<td>14.81</td>
</tr>
<tr>
<td>3.2. Overinclusion</td>
<td>9</td>
<td>16.66</td>
</tr>
<tr>
<td>3.3. Misselections</td>
<td>5</td>
<td>9.25</td>
</tr>
<tr>
<td>3.4. Misordering</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3.5. Blending</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

As shown in Table 1, the participants committed various formal lexical errors of borrowing, suffixation, vowel-based errors, calque, omissions, overinclusion and misselections. As far as borrowing is concerned, the students tended to transfer words from French when writing in English, such as “campagne” (country), “problème” (problem), “in definitive” (definitively), “fondamental ressources” (fundamental resources). Tunisian EFL learners have
problems with suffixation, as reflected in “I hate noisy in town”, “you’ll feel very boring”, “I feel relax” wherein they failed to add the appropriate suffixes to the roots. They showed a predilection to translate phrases from French or from their mother tongue, i.e. Arabic. They wrote “I practice sport” instead of ‘I exercise”, which is a translation of the French sentence “je pratique du sport.” Some students wrote “in our days” to mean nowadays, which is a translation from Arabic “fi hedhi al-ayaam” (nowadays).

The results suggest that distortions or misspellings were problematic for the participants. It is important to consider why the participants misspelled words through omission, overinclusion, misselection. In fact, the participants tended to omit or overinclude sounds, as in “different”, “realy” instead of “really,” “fare” instead of “far.” The participants misselected sounds as in “illiteracy” instead of “illiteracy,” “choose” instead of “chose,” “dissatisfied” instead of “dissatisfied.” They showed confusion with vowel choice, writing “leave” instead of “lave,” “lucky” instead of “lack,” “fund” instead of “find.” Raines (1985, p. 247) suggests that such misspellings occur because unskilled L2 writers “concentrate on the challenge of finding the right words and sentences to express their meaning instead of editing.”

4.2. Semantic errors

Lexical semantic errors were not frequent in the participants’ compositions (4 out of 54 errors, that is 7.40 % of the total number of lexical errors). Table 2 below summarizes the participants’ semantic errors:

<table>
<thead>
<tr>
<th>Error</th>
<th>Number of errors</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.Confusion of sense relations</td>
<td>3</td>
<td>5.55</td>
</tr>
<tr>
<td>1.1 General term for specific one</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.2. Overtly specific term</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>1.3. Inappropriate co-hyponyms</td>
<td>1</td>
<td>1.85</td>
</tr>
<tr>
<td>1.4. Near synonyms</td>
<td>2</td>
<td>3.70</td>
</tr>
<tr>
<td>2.Collocation errors</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

There is one error related to inappropriate co-hyponyms wherein a participant misused a term in English as in “we transpire fresh air” instead of “we breathe fresh air.” There are two errors of near synonyms as in “it allows to get a fresh air” and “they will not like travelling to get to their work” wherein the participants used informal words instead of formal ones. “To get to work” is not entirely incorrect, but the use of “get” seems more appropriate in informal writing. The student seems not to distinguish between the register choice in formal and informal writing.
Table 3 summarizes the frequency of formal and semantic errors found in the participants’ writings.

<table>
<thead>
<tr>
<th></th>
<th>Formal errors</th>
<th>Semantic errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>No of errors</td>
<td>51</td>
<td>3</td>
</tr>
<tr>
<td>%</td>
<td>94.44</td>
<td>0.05</td>
</tr>
</tbody>
</table>

5. Discussion and pedagogical implications

The study’s results suggest that Tunisian EFL learners face more difficulties with formal errors (94.44% of the total number of lexical errors) than with semantic errors (only 0.05% of the total number of lexical errors). There are mainly two major sources of errors in EFL learning: Interference from L1 and intralingual source, i.e. the target language itself. Applied linguists contend that the native language of learners plays a significant role in EFL learning. Indeed, negative interference from L1 is acknowledged in the literature as an important factor in EFL learning. In this study, the participants resorted to calque and borrowing when writing in English by transferring words from French and Arabic as in “I practice sport” and “in these days.” Intralingual errors are due to the difficulty of the target language. Students seem to find it difficult, for instance, to write “different” with the sound “e” while they should pronounce it without the sound “e”. They showed confusion with vowel choice writing “leave” instead of “live,” “luck” instead of “lack,” “fund” instead of “find,” which reflects their incomplete learning that English vowels are able to change the whole meaning of a word. Indeed, changing a vowel in a word in English generates a change in meaning.

The study’s results push us to reflect upon the causes of these lexical errors. It can be argued that the lexical errors found in the participants’ essays are attributable to several factors, such as hypercorrection, faulty teaching, fossilization, an inadequate learning. Touchie (1986) states that sometimes “zealous” efforts of teachers in correcting their students’ errors bring about “induced errors.” It happens that learners’ errors are caused by the teacher, teaching materials, or the order of presentation. Interestingly enough, some errors can fossilize and, therefore, it becomes difficult for the learners to get rid of them. Some errors can be the result of inadequate learning. They can be caused by ignorance of rule restrictions and incomplete learning.

These errors should be treated. In this context, the teacher’s role is important. Touchie (1986) confirms that teachers should correct errors affecting intelligibility. Thus, it is important to correct the students’ lexical errors whenever they interfere with the general meaning and understandability of utterances. Also, teachers should correct the most frequent errors. For example, borrowing from L1 turns out to be a pervasive error in the light of this study. It is, therefore, an error of high frequency and generality that teachers should work on.
Conclusion

The objective of this study was to identify the lexical errors made by EFL students at the ISEAH Institute of Kef. The results suggest that Tunisian EFL learners face more difficulties with lexical formal errors (94.44 % of the total number of lexical errors) than with lexical semantic errors (only 0.05 % of the total number of lexical errors). It has been argued that these lexical errors can be attributed to several factors, such as hypercorrection, faulty teaching, fossilization, an inadequate learning. To remedy these learning difficulties, the teacher’s role is significant in order to ensure an affective correction of errors.

It is worth-noting that this study is contributive to EFL research since it has shed light on the lexical errors made by EFL learners in the Tunisian context. It has revealed the difficulties faced by EFL learners in learning English lexis.
References


