

Lexical Choice Difficulties: A Psycholinguistic Study towards a Solution

Mohammed Q. Shormani

Assistant Professor of Linguistics, College of Arts, Ibb University, P.O. Box: 70270, Ibb, Yemen
Email: dshormani@gmail.com

Abstract

Lexes are the “building blocks” of language and no one could ever think that language acquisition could take place without considering its vocabulary the cornerstone a learner commences with, employing several cognitive and non-cognitive learning strategies in acquiring such vocabulary. In spite of the so many studies tackling different kinds of errors an L2 learner commits, there are fewer studies done on lexical errors and so few, if any at all, probing deeply the sources of such errors and their consequences. Thus, this study is intended to provide a psycholinguistic evidence for the possible sources lexical errors could be ascribed to. 50 essays have been selected randomly from 123 ones written by Arabic-speaking Yemeni learners of English given to them as homework assignments. Errors were identified, classified and tabulated. Then, sources were classified into four categories, viz. *L1-transfer*, *L2-influence*, *mutual* and *unrecognized*. The analysis shows that (44%) of the errors were ascribed to *L1-transfer*, (40%) to *L2-influence*, (12.8%) to *mutual* and (3.2%) to *unrecognized*. The Findings could be applied to ESL/EFL vocabulary teaching-learning contexts.

Keywords: Lexical Choice, Arabic-speaking Yemeni Learners, Psycholinguistics, Learning Strategies, Error Sources

1. Introduction

It goes without saying that vocabulary (lexes) are the most essential part in second language acquisition (SLA) process (Chomsky, 1968; Llach, 2005, Shormani, 2012a, 2012b, 2013; Han, 2004; Gass&Selinker, 2008; McCarthy, 1990; Saville-Troike, 2006;Laufer, 1997; Lennon, 1991, 1992; Wang, & Wen, 2002;Bahns, 1993; Ellis, 1997; James, 1998; Khalil, 1995; Mitchell & Myles 1998; Takac, 2008, among many others). It is a fact that words are the “building blocks” of such a process. No one could think of acquiring any language beginning with something other than words simply because they are “the first linguistic items acquired by the learner” (Llach, 2005, p. 46). Thus, no one could ever think that SLA begins with acquiring a grammatical structure, say, for instance, *present progressive* for how this could happen and the learner does not have any lexis in his/her repertoire let alone the terms “present and progressive.” In addition, for those who consider communication the ultimate goal of language acquisition/learning process, lexis are needed for the development of fluency, proficiency and accuracy. For instance, for the communication process (in a particular situation) to be successful, lexis, and lexis alone are what makes a particular learner more prestigious than another as well as being able to use them. However, in spite of all the great effort exerted and the so many studies done on this aspect, L2 learners’ level in mastering L2 vocabulary, English for instance, is still far behind. This appears vividly when one considers the lexical errors committed by advanced learners of English as it is the case in this study where serious lexical errors still persist.

In addition, knowledge of L2 lexis involves different linguistic components, viz. phonology, morphology, syntax, semantic, pragmatics and orthography. Takac (2008, p. 10) adds “knowledge of conceptual foundations that determine the position of the lexical item in our conceptual system” and how such items are used in different contexts. Up-to-date, only relatively few studies have been concerned with lexical errors L2 learners commit, however. What is more is that even those studies are not satisfactory due to the fact that “error categorization frameworks used in ...[such] studies have addressed only a relatively limited number of lexical error categories” (Hemchua, & Schmitt, 2006, p. 3). Fewer, if any at all, are also those studies which are set to investigate lexical error sources. The paucity of studies on lexical errors is attributed to the fact that semantic knowledge is more difficult to assess (Shormani, 2012a). Moreover, investigating errors is a linguistic phenomenon while investigating and consequently identifying what makes a learner commit a particular error is a psycholinguistic phenomenon which requires linguistic and psychological knowledge.

Thus, this study is set to investigate the possible sources lexical errors committed by advanced Yemeni learners majoring in English could be ascribed to. In fact, it attempts to seek answers to questions such as is it L1, i.e. Arabic which causes them or L2, i.e. English?, Are L1 and L2 the only sources? Are there any other sources that such errors could be ascribed to? Thus, fifty essays selected randomly were analyzed; the errors identified were classified into categories and subcategories. The sources were classified into four categories, viz. *L1-transfer*, *L2-influence*, *mutual* and *unrecognized*.⁽ⁱ⁾ *L1-transfer* includes negative transfer of a rule and/or structure into L2, literal translation of Arabic concepts, words and phrases into L2, i.e. English, hypothesized one-to-one correspondence between L1 and L2, false equivalence, Arabic-English dictionaries, etc. However, *L2-influence* includes misconception, insufficient knowledge in English lexis, internalizing L2 lexicon system, lack of

exposure to English, etc.

2. Previous studies

Committing errors in SLA process puts researchers, scholars and teachers vis-à-vis an inevitable phenomenon disturbing them day and night. Moreover, the study of errors of whatever type dates back to the “50s and well into the 60s of the 20th Century” Shormani (2012a, p. 14) and views and attitudes towards such errors vary. In that, there are views stating that errors and their committing are something sinful on the part of SL learners and thus, have to be eradicated (Brooks, 1960). Others maintain that errors in themselves are of great importance to scholars inasmuch as they are to teachers and learners. This view is held by several scholars and SLA researchers (e.g. Corder, 1973, 1981; Selinker 1992, 1993; Richards, 1972, 1974; Dulay & Burt, 1973, 1974; Dulay *et al.* 1982; Tomasello, 2007; James, 1977, 1998; Han, 2000, 2005; Ellis, 1997, Gass & Selinker, 2008). These studies and scholars have investigated the L2 learner errors thoroughly as to what errors are, their types and classifications. They are studied for their importance in understanding the nature of SLA process, the strategies used by the L2 learner, how and why only few learners achieve native-like competence and/or proficiency in English while almost all learners get fossilized among other related issues (Han, 2003, 2004; Shormani, 2012a; Shormani, 2013). Several studies whether by Arab or non-Arab scholars tackle different types of errors including syntactic, phonological, semantic, etc. However, only few of such studies deal with lexical errors and fewer, if any at all, have deeply probed their sources and consequences.

The paucity of studies on lexical errors and hence, their errors is due to the fact that deeply probing their sources and hence, consequences is a psycholinguistic one and hence, requiring a linguistic knowledge in both L1 and L2, psycholinguistic knowledge of what goes wrong with what that makes an L2 learner commit such errors, a pragmatic knowledge which makes the researcher able to be aware and decide whether a particular use of a particular piece of language in a specific situation is erroneous and knowledge of language acquisition, be it L1 or L2, which enables him/her to deeply probe the nature of this process and the SLA strategies cognitively employed by the learner when he/she commits an error among other related factors.

As far as Arab scholars are concerned, and to the best of my knowledge, there is no study whose scope is mainly to investigate the sources of lexical errors committed by Arab and/or Yemeni learners. However, there are some studies done by some Arab researchers which only allude to the sources of lexical errors because their scope is not such errors. For instance, Shormani & Sohmani (2012) have done a study in which they allude to the sources of lexical errors. They have attributed the semantic errors committed by the subjects they investigate to two sources, viz. Arabic and English. Another study has been done by (Mahmoud, 2011) in which the researcher has classified the errors of vocabulary committed by his subjects into *omission, addition and substitution*. In this study, the errors have been deemed to be interlingual and intralingual. Another study on lexical errors has been done by Zughoul (1991) in which he has done a great job in classifying the errors he identified in his corpus but he was not clear about the sources of such errors. He has attributed the errors committed to L1 and L2 but the way he ascribed them to both languages was not satisfactory. The way he ascribed them is according to their categories. For instance, he ascribes a *similar form* error to *similar forms* claiming that it is the *similar form*, as in **efficient money*⁽ⁱⁱ⁾ in which the learner means *sufficient money*, which makes students commit such an error. In his study, the only type of errors which is clearly ascribed to L1, i.e. Arabic is what he classifies as *literal translation*. Khatib (1984) has done study on lexical errors but his purpose was classifying such errors and not their sources.

As far as the non-Arab scholars are concerned, there are some studies that tackle lexical errors even though their scope is not exactly as such. For instance, Mohanty (2006) has done a study in which he tries to examine the source of one error that Arab learners of English usually commit. This error is the misspelling of the English cardinal number, *two* which is misspelled as *tow*. He ascribes such an error to English where the learner gets influenced by words such as *cow, now, know*, etc. In addition, Duskova (1979) investigates the possible sources of errors committed by her subjects having Czech as L1 and classifying them into four broad categories and ascribing errors in each category to the same category claiming that it is L1, i.e. Czech which is the major source of such errors. However, the scope of her study was not limited to lexical errors alone. However, her paper is one of the seminal studies tackling the error source phenomenon. Dulay, *et al.* (1982) find that errors committed by SL learners can be ascribed to L1, i.e. interlingual, L2, i.e. intralingual, both L1 and L2, i.e. *ambiguous* and those having no identifiable source, i.e. *unique*. Hang (2005) has done study on lexical errors committed by Cantonese students ascribing them only to L1 and L2. Llach (2005) has done a study but she was not clear as to what courses such errors because the scope of the paper was different. Hemchua & Schmitt (2006) have done study on lexical errors ascribing them only to L1 and L2. These are the most available studies on lexical errors, at least to me.

3. Defining a Lexical Error

Error in general has been defined as a deviation of the rules of a particular language, be they linguistic, pragmatic, stylistic, etc. To me, as it actually is, the best definition provided for the term “error” is what has been given by Lennon (1991, p. 182), viewing it as “a linguistic form or combination of forms which, in the same

context and under similar conditions of production, would, in all likelihood, not be produced by the speakers native speaker counterparts.” However and for the purpose of this study, a lexical error could be defined as a deviant committed at the level of lexical choice as a result of violating lexical rules particular to English language and hence, excluding all other errors of whatever type they are.

4. The Present Study

This study intends to provide a psycholinguistic evidence for lexis choice difficulty encountered by Arabic-speaking learners of English by deeply probing the lexical choice errors committed by Yemeni learners majoring in English. The findings have pedagogical implications that could be generalized and applied to L2 vocabulary teaching-learning in ESL/EFL contexts.

4.1. Method

4.1.1. Participants

It is widely held (e.g. Shormani, 2012a, 2013) that a learner can easily learn a word or even memorize it but it is difficult to use it properly in an appropriate context unless he/she has practiced and practiced using it. Consequently, the participants of a study concerning lexical errors should be of an advanced level because they are expected to have reached a considerable level of proficiency and been exposed to such courses as phonology, morphology, semantics, syntax, among other knowledge-based courses in addition to novel, drama poetry, etc. Thus, this study involves 50 Yemeni Arabic-speaking learners of English whose 50 essays have been selected randomly from 123 essays given to them as a homework assignment, in their first semester, fourth-year, English department, Ibb University, Yemen, in the academic year 2012-2013. Their ages range from 24-28 years though some of them may be older but not younger than that. They were 25 female and 25 male though sex and age were not considered in this study. They have studied English for about ten years (seven at school and four at university). They have studied several courses of different academic nature including those practical (skills) like spoken, reading, writing, vocabulary, etc., those of theory like syntax, semantics, morphology, etc. and literature courses such as novel, drama, poetry, etc.

4.1.2. Procedure

The essays involved in this study were written about different argumentative topics. In fact, the topics were left to students themselves to choose and thus coming up with about twelve topics including (*Yemeni Unity, my family, my ambition, my first day at college, internet as a source of information, my village, rain in Ibb city, the importance English language has in the world of today*, etc.) The total words were 15007 and the mean length of the 50 essays was 300.18 words (SD= 102.64, min= 133, Max= 512). The number of essays involved in this study(50) allowed for an in-depth and comprehensive analysis of the lexical errors, their sources and hence, the Yemeni learners of English’s difficulties in receiving/perceiving and producing lexically grammatical pieces of language. The researcher (a native speaker of Arabic, Yemeni Arabic) corrected the participants’ manuscripts. When there was some kind of unintelligibility an error imposes, the manuscript was corrected in the presence of the learner committing it to find out what he/she has meant by saying such and such and then the error was identified, classified and counted in the category it belongs to. After correcting the manuscripts, I have consulted an experienced University teacher (an Indian Professor of Applied linguistics) and almost we both agree to such identification. When there was a doubt in classification, we have consulted an experienced English native speaker (an American). To differentiate between an error and a mistake, only those recurrent deviations were considered errors. Moreover, there were errors of other types such as syntactic, spelling, lexico-grammatical among others which were excluded. It is worth admitting here that there was a category overlap but every possible step was considered to make our error corpus as accurate as possible. As far as the sources of the errors identified in this study are concerned, those whose source was L1 were counted in terms of *L1-transfer*. Those whose source was L2 were counted in terms of *L2-influence*. Those whose source lied within both L1 and L2 were counted as *mutual* and those whose sources belong to “something else” other than L1 and L2 were counted as *unrecognized*. This “something else” though interesting and challenging is beyond the scope of this study and hence, left for future studies.

5. Classification of Lexical Choice Errors

For the purpose of investigating the source(s) of an error or a group of errors, a classification of the errors identified in our corpus has to be developed. Thus, a comprehensive taxonomy has been developed based on some kind of amalgam classification taken from Zughoul’s (1991), Shormani &Sohbani’s (2012), Llach’s (2005) adding our own according to the errors identified in our corpus. Thus, our study will be limited only to the types of errors summarized in Figure 1 below.

Figure 1: Summary of Lexical Choice Error Categories

1. **Paraphrase**
2. **Lit. Translation**
3. **Assumed Synonymy**
4. **Derivativeness**

5. Binary Terms
6. Idiomaticity
7. Similar Forms
8. Overused Terms
9. Analogy

As Figure 1 above shows, there are nine categories in which the lexical errors identified in this study were classified. Error frequency, category and source among other things will be presented in Table (1) below.

6. Results and Discussion

Table 1: Lexical Errors: Category, Source, Frequency and Percentage

Category	L1-transfer		L2-influence		Mutual		Unrecognized		Total	
	F	%	F	%	F	%	F	%	F	%
Assed. Synonymy	89	27%	172	52%	58	18%	11	3%	330	16%
Lit. Translation	311	100%	0.00	0%	0.00	0%	0.00	0.00%	311	15%
Paraphrase	96	34%	101	36%	67	24%	19	7%	283	14%
Derivativeness	111	41%	95	1%	55	20%	9	3%	270	13%
Idiomaticity	82	39%	97	46%	26	12%	5	2%	210	10.4%
Binary Terms	48	26%	120	65%	9	5%	7	4%	184	9%
Similar Forms	31	19%	117	73%	12	8%	0.00	0.00%	160	8.2%
Overused Terms	81	51%	52	33%	17	11%	8	5%	158	7.8%
Analogy	41	30%	58	43%	31	23%	6	4%	136	6.6%
Total	890	-----	812	-----	275	-----	65	-----	2042	100%
%	44%	-----	40%	-----	12.8%	-----	3.2%	-----	100%	-----

As presented in Table (1) above, the errors identified in our corpus will be analyzed in terms of the abovementioned four types of sources, source's frequency and percentage for each of the following categories. It should be noted here that every example presented below involves only one single lexical error though in its original form (as written by the participant), it might have more than one error and of different type(s).

6.1. Assumed Synonymy

Synonymy as a lexical phenomenon does exist in any natural language. For instance, the words, *tall* and *long* are said to be synonymous, however, the former is used when we describe length vertically and the latter horizontally (Shormani, 2012a). What concerns us most here is that we have to conceptualize that synonymy in one way or another does exist in human languages and as far as English as SL is concerned, the existence of such a phenomenon creates some kind of confusion to L2 learners leading consequently to what is so-called *assumed synonymy* which is our main concern here. *Assumed synonymy* simply means that L2 learner assumes two or more words to be synonymous and that such synonymous lexes can be used interchangeably and hence, coming up with a lexical error. In our study, the errors committed due to *Assumed Synonymy* can be ascribed to different sources. It can be attributed to *L1-transfer*, *L2-influence*, *mutual* and *unrecognized*. *Assumed Synonymy* scores the highest number of errors with 330, i.e. (16%) distributed as follows: *L2-influence* with 172 error, i.e. (52%), *L1-transfer* with 89 errors, i.e. (27%), *mutual* with 58 errors, i.e. (18%) and *unrecognized* with 11 errors, i.e. (3%). These different sources will be exemplified below supported by examples from our corpus.

(1)* *The teacher asked us to meet him when he is empty.* (free)

(2)* *In my future, I hope to have work.* (job)

(3)* *and our parents make our country very strong.* (fathers)

(4) *I scared with the teacher.* (am afraid of)

In (1) through (4), the learner uses *empty*, *work*, *parents* and *scared with* for *free*, *job*, *parents* and *am afraid of* respectively. These errors are ascribed to different sources. For instance, in (1), the error is ascribed to Arabic caused by hypothesizing a one-to-one correspondence between English and Arabic because *empty* and *free* have the same equivalent in Arabic, i.e. *fadi*. The error in (2) is ascribed to *L2-influence*. One also could think that it can be ascribed to L1 on the basis that both *work* and *job* have the same equivalent in Arabic *?amal* but if one looks closely at the latter and *work*, there is some kind of difference, i.e. the Arabic term *?amal* can be pluralized while the English one cannot. Another reason is that the learner does not use the article *a* and this indicates that the learner knows well such a syntactic rule and hence, English as a source of this error prevails. (3) can be ascribed to both L1 and L2. It can be ascribed to L1 on the ground that the concept of both *fathers* and *parents* in English has one equivalent, i.e. *?aaba* in Arabic and what the learner does is just transfer the only word in Arabic for such a concept into English. It can also be ascribed to L2 on the basis that the learner misconceptualizes that *parents* and *fathers* can be used interchangeably and hence, coming up with such a lexically deviant expression. However, the error in (4) is of *unrecognized*. It cannot be ascribed to L1 because the adjective *scared* is used as a verb which is not possible in Arabic. The exact equivalent in Arabic of *scared* is

xaaiif which has to be used as an adjective and not a verb and always co-occurs with the preposition *min* (from) and not *with*. It cannot also be ascribed to L2 simply because *scare* can be used as a verb and *scared* as an adjective. If, however, it is used as a verb, the prep *with* should not be present on the one hand, and the fact that a student cannot *scarea teacher* on the other hand. If, however, he/she means to use *scared* in the sense of *afraid*, the prep *of* or even *from* should have been used.

6.2. Literal Translation

As shown in Table (1) above, this category scores the second highest rank in error count with 311, i.e. (15%). In fact, literal translation is considered a learning strategy made use of by the learner when he/she fails to express himself/herself. In our study, literal translation contradicts some studies like Shormani's (2012a) but supports some others like Zughoul's (1991). This category includes only L1-transfer errors. (5) exemplifies such a category.

(5)**At university, doctors provide great help to students.* (professors/teachers/Sirs)

(5) exemplifies *L1-transfer* where the learner uses the word *doctor* for *teacher*. This is one of the most recurrent errors noticed to be committed by Yemeni learners of English. They address their professors as *Doctor!*, *Yes, doctor...*, *Excuse me doctor!* What they do actually is substitute the word *doctor* for *Sir*. Arab learners of English including Yemenis are influenced by the fact that the term *doctor* is used to address anyone having a Ph.D degree in whatever discipline and mostly influenced by other students in other departments like history, geography, Islamic studies, etc. where such a term is used extensively to address their professors. A part from this, the use of *doctor* instead of *sir* by Arabs has also a religious connotation. Arab learners of English think that the term *Sir* whose equivalent in Arabic is *saidi* is prohibited to address anybody in Islam except God. They also think that there is a sense of silver when using this word. I myself have told my students that the word *Sir* has nothing to do with such connotations. The only thing it has and must be used for is that in English, a teacher is addressed as *Sir*. However, as far as English is concerned, the term *doctor* is used to address someone who works at hospitals and whose job is to treat people and diagnose their diseases.

6.3. Paraphrase

Kreidler (2002, p. 302) has defined paraphrase as “[a]n alternative way of expressing the content of a sentence” in other words. Shormani (2012a) sees paraphrase as a restatement of a word, phrase or a sentence in different words. However, as far as the context of this study is concerned, paraphrase will be confined only to word and phrase levels. The way the participants of this study paraphrase is considered unacceptable in English because it is of the “decomposition type” which is different from “lexical paraphrase as those of lexical corporation or decomposition” (Shormani & Sohmani, 2012, p. 131). In fact, paraphrase has been considered one of the most fundamental language learning strategies. What happens here is that sometimes the L2 learners including Arabs cannot find the exact word/phrase to be used in a context and so they try to “exploit” paraphrasing strategy but not as what is expected and hence, resulting in an error. This category scores the third highest number of errors after *assumed synonymy* and *literal translation* including 283 errors, i.e. (14%) distributed as follows *L1-transfer* includes 96, i.e. (34%), *L2-influence* includes 101, i.e. (36%), *mutual*, 67, i.e. (24%) and *unrecognized* 19 errors, i.e. (7%). Examples (6-7) illustrate this point respectively.

(6)**In 1990, we make south and north together.* (unite)

(7)**Many facilities are existing in our college.* (available)

(8)**When I saw her, I was able to know her.* (recognize)

(9)**Internet is a good machine for information.* (source)

In (6) through (9), the learner substitutes *make*, *existing*, *able to know* and *machine* for *unite*, *available*, *recognize* and *source* respectively. In (6), the error is ascribed to *L1-transfer* because the learner having Arabic as a knowledge base has used *make...together* which is not acceptable in English. The error in (2) is ascribed to *L2-influence* simply because the exact equivalent of the word *available*, viz. *mawjuud* does exist in the language. Here, the learner hypothesizes that *existing* means *available* can be used interchangeably and hence, coming up with such an error. The error in (8) is a *mutual*, i.e. it can be ascribed to *L1-transfer* because in Arabic *qadir-un ʔalal-ta ʔruf-i* (able to know) means *recognize*. It could also be ascribed to English because the learner hypothesizes that *know* and *recognize* could be used interchangeably which is not lexically acceptable in English. The error in (9) is ascribed neither to L1 nor L2 but its source is *unrecognized*. It cannot be ascribed to the former simply because its equivalent is not acceptable in Arabic and cannot be ascribed to L2 because of being unacceptable in English, too.

6.4. Derivativeness

Derivativeness in almost all Indo-European languages depends on concatenative mechanism in deriving words. For instance, the word *remarkable* is derived by concatenating (combining) the morphemes *re-+mark+able*. Arabic, however, is of a non-concatenative mechanism which is greatly different from that of English (Shormani, 2012a). Shormani & Al-Shorbani (2012) note this difference and state Arabic derivativeness depends on *root* and *pastern* mechanism where a trilateral or quadilateral root is the basic element from which all derivatives are derived. For instance, from the trilateral root *D R S* (to study), the words, *daras-a* (he studied), *daras-at* (she studied), *madras-a(h)* (school), *madaaris* (schools), *mudarris* (teacher), *daaris* (student), etc. can be formed. In addition, Schmitt & Zimmermann (2002, p. 141) argue that it is not true that when an L2 learner knows “one

member of a word family (e.g., stimulate), the other members (e.g., stimulants, simulative) are relatively easy to learn” though such a process makes the learner’s task easier in deciding which member (of the derivative family) to choose in a particular context, it needs a careful way to choose the other members. Thus, in our study *derivativeness* scores 270 errors, i.e. (13%) distributed as follows: *L1-transfer* scores 111 errors, i.e. (41%), followed by *L2-influence* with 95 errors, i.e. (35%), followed by *mutual* with 55 errors, i.e. (20%) and finally *unrecognized* with 9 errors, i.e. (3%). Examples (10)-(13) illustrate the point in question.

(10)*Life learns me many things. (teaches)

(11)*Tomorrow, I will travel to my village. (go)

(12)*I feel happy in syntax lecture. (class)

(13)*After study, I perform all my homework. (do)

where learners substitute *learn*, *travel*, *lecture* and *perform* for *teach*, *go*, *class* and *do* respectively. In (10), is ascribed to Arabic is due to two reasons: first, the fact that the learner’s original sentence was **thelife learns me many things* where *the* has been removed from the sentence to make it include only one error. As mentioned above, every sentence includes only one error. Second, the sentence is considered a very common expression used in Arabic and by all Arabs so as to express the bitter truth of life in teaching humans. (11) is ascribed to L2-influence *per se* due to the insufficient knowledge in L2 lexes to distinguish the use of *travel* and that of *go*. Had it been a transfer from Arabic, the learner should have used *gobecause* Arabic has *ysaafer-u* (travel) and *yaḏhab-u* (go). As far as English is concerned, the difference between *travel* and *go* is that the former is used in moving to another country where more preparation is needed as luggage and so on and it is so in Arabic. Thus, *L2-influence* is the most likely source of the error (11). The error in (12), however, is ascribed to both L1 and L2. It is ascribed to the former because the term *lecture* is commonly used for *class* in Arabic. However, it could be ascribed to the latter because the learner is still internalizing the linguistic system of the target language and he/she simply has not learned the difference between both terms and when to use either. However, the difference between *lecture* and *class* lies in the fact that the latter is confined to a particular place (classroom), a particular time (timetable) and a particular teacher (here the syntax teacher) which is not the case with *lecture*. The error in (13) can neither be ascribed to English nor to Arabic. It cannot be ascribed to Arabic simply because there is an equivalent term *fardo* which is *yaʿmal*. It cannot be ascribed to English simply because the term *perform* could not be assumed to have been acquired prior to *do*.

6.5. Idiomaticity

Idiomaticity being a feature of human languages is best described as a phenomenon where the meaning of an idiomatic expression, consisting of two or more lexemes, is composed contextually and not as the sum of the meaning of each Shoramni&Sohbani (2012, p. 131). This category includes 210 errors, i.e. (10.4%) distributed as follows: *L2-influence* includes 97 errors, i.e. (46%) followed by *L1-transfer* with 82 errors, i.e. (39%), followed by *mutual* with 26 errors, i.e. (12%) and *unrecognized* with 5 errors, i.e. (2%). Examples (14-17) illustrate the issue in question:

(14)*then I changed my clothes when I go to university. (put on)

(15)*I used to get up at six o’clock and stay for some time on bed. (wake up)

(16)*I thought I lost it but I found it in house. (at home)

(17)*I am very proud you, Sir. (proud of)

where the learner uses *changed*, *get up*, *in house* and *proud* for *put on*, *wake up*, *at home* and *proud of* respectively. For instance, the error in (14) is ascribed to *L1-transfer* on the basis that in Arabic, the word used in such cases is *yuyair* which exactly means *change* and what the learner does here is transfer this very word into English. As advanced learners, it cannot be said that they have not yet acquired *put on* and hence, *L2-influence* as a source of such an error is ruled out. (15) is ascribed to L2 on the basis that the learner hypothesizes that *get up* and *wake up* could be used interchangeably. The error in (16) could be ascribed to L1 and L2. It could be described to *L1-transfer* because in Arabic, there are no such constraints as idiomaticity regarding *bait* (house/home). It could also be ascribed to English on the basis that the learner has not yet acquired such an idiomatic complexity. Both English and Arabic have nothing to do with committing the error in (17) simply because the equivalent to *proud of* does exist in Arabic which is *faxuur-un bi* exactly (proud with). Had it been a transfer from Arabic, the learner should have used *proud with*. It cannot also be ascribed to English because such a sentence is unacceptable in the language.

6.6. Binary Terms

Laufer (1997) calls such a category of errors “relational opposites” where oppositeness is implied. However, to me, as the name suggests, *Binary Terms* indicate some kind of exclusiveness, viz. when one is used the other is not, i.e. they are in complimentary distribution. For instance, *buy* and *sell* or *educate* and *teach* stand in different extremes. This category includes 184 errors, i.e. (9%) distributed as follows. *L2-influence* contains 120 errors, i.e. (65%) followed by *L1-transfer* with 48 errors, i.e. (26%) followed by *mutual* with 9 errors, i.e. (5%) and *unrecognized* with 7 errors, i.e. (4%). Each source along with an example is provided below.

(18)*and this educated me how to deal with others. (taught)

(19)*I usually go home at 2 o’clock. (come)

(20)*...then I told him how do you do, Sir?(how are you)

(21)*and how I can say myself in that situation. (express)

In (18) through (21), the learner substitutes *educate*, *go*, *how do you do*, *say* for *teach*, *come*, *how are you* and *express* respectively. The error in (18) is ascribed to *L1-transfer* for the fact that the equivalent of *educate* *ʿalim* in Arabic can be used for *teach* leading us to ascertain that the source of such error is Arabic *per se*. (19) is ascribed to English *per se* because in Arabic the verb *yaʿāhib-u*(go) cannot be used in the sense of *come*. Thus, had it been a transfer even from Yemeni Arabic, the learner should have used *return* meaning exactly *yurawih*. Thus, this error is ascribed to *L2-influence* on the basis that the learner hypothesizes that he/she could use both terms interchangeably and thus, English as a source of such an error prevails. (20) can be ascribed to Arabic and English. It can be ascribed to the former simply because in Arabic, there is only one expression to greet somebody without any distinction either for the first or whatever time(s), and that under the influence of English-Arabic bilingual dictionaries where *how do you do?* and *how are you?* have the same meaning, i.e. *kaiḥfala-k*, the learner hypothesizes that he/she could use both interchangeably and hence, coming up with such an error. (21) cannot be ascribed to Arabic because the concept in which such a sentence is used does not exist at all. It cannot also be ascribed to English because there is no relation between the verb *say* and *myself*.

6.7. Similar Forms

As stated above, this category of lexical errors has been identified by several researchers as being one source of lexical choice errors committed by L2 learners whatever their L1s are. However, they actually fail to account for the source of the errors because *Similar Forms* is a category and not a source of errors. In our corpus, this category includes 160 errors, i.e. (8.2%) of the total errors distributed as follows: *L2-influence* comprises 117 errors, i.e. (73%), followed by *L1-transfer* with 31 errors, i.e. (19%) and *mutual* including 12 errors, i.e. (8%). This category has no errors whose source is *unrecognized*. Below are some examples.

(22)*...but it was not my intension. (intention)

(23)it makes me go throw many steps. (through)

(24)*My father has a shop for changing money. (exchanging)

In (22) through (24), the learner substitutes *intension*, *throw* and *changing* for *intention*, *through* and *exchange* respectively. (22) is ascribed to Arabic simply because Arabic has one exact equivalent, viz. *qaṣd* for both *intension* and *intention*. This error, however, could not be said to have been caused by *L2-influence* simply because while the former is a linguistics term (particularly of semantics), the latter is a lexical term which is used in common situations. (23) is ascribed to L2 alone simply because such similar forms do not exist in Arabic at all. It could be argued that the learner gets confused because of the same pronunciation of both terms. There is no false hypothesizing here, however. The error in (24) can be ascribed to Arabic and English both. It can be ascribed to the former simply because in Arabic, there is only word meaning both *changing* and *exchanging* which is *ṣarf*. It could also be ascribed to English on the basis that both *changing* and *exchanging* have to a great extent the same pronunciation and orthography.

6.8. Overused Terms

In our study, hyponyms of some superordinates like *big* including, *large*, *huge*, *much*, *many*, *a lot of*, *lots*, etc. and hence, resulting in such pieces of language as *big mountain*, *big bridge*, *big number of teachers*, etc. and *small* resulting in *small brother*, *small ambition*, *small car*, *my small sister*, *small area*, *small tree* among the many others. Some items have been used for the opposite sense or meaning like *great* used in the sense of *bad*. Thus, this category includes 158 errors, i.e. (7.8%) distributed as follows *L1-transfer* includes 81 errors, i.e. (51%) followed by *L2-influence* with 52 errors, i.e. (33%) followed by *mutual* with 17 errors, i.e. (11%) and *unrecognized* with 8 errors, i.e. (5%). (25-28) illustrate such a category.

(25)*...and today, I have syntax and transition and poetry and drama and, etc.

(26)*Our family is so big, my big brothers and their families and my big sisters and grandmother and...

(27)*In the university, we have good teachers and good friends and good classes...

(28)*...but Israel is our greatest enemy. (worst)

In (25), the coordinator *and* has been overused. This error is ascribed to *L1-transfer* because the learner is overwhelmed by the Arabic style where the coordinator *wa*(and) is used in such contexts. In (26), the error lies in overusing the word *big* and using *big* for *joint*. This error is ascribed to *L2-influence* because the learner has not yet acquired the difference between *big* and *joint* and where to use each is used and the main reason behind that is the lack of exposure to English which Arab learners of English including Yemeni suffer from. It is worth mentioning here that this error cannot be ascribed to Arabic for the fact that such long, wordier and circumlocuted phrase/sentences are not acceptable in Arabic. (27) can be ascribed to Arabic on the basis that the learner has overused the coordinator *and* which is acceptable in Arabic. It could also be ascribed to English for the fact that the learner has used the adjective *good* over some others like *close/true* and *clean/big*, etc. Both *L1-transfer* and *L2-influence* have nothing to do with committing the error in (28) because it is unacceptable in both languages.

6.9. Analogy

Analogy has been considered a learning strategy which indicates that the learner is not a passive interlocutor in

the learning process. However, in most of the cases it turns with errors. This category includes 136 errors, i.e. (6.7%) distributed as follows: *L2-influence* including 58 errors, i.e. (43%) followed by *L1-transfer* with 41 errors, i.e. (30%) followed by *mutual* with 31 errors, i.e. (23%) and the last one is *unrecognized* with 6 errors, i.e. (4%). Now, consider (29-32) exemplifying the issue in question.

(29)*...there are many people who entered English department...(joined)

(30)*...and he asked to me difficult questions. (asked me)

(31)*my younger brother still in primary school. (is still)

(32)*I have a father and mother.

In (29) through (31), the learner substitutes *entered*, *asked to*, *still*, for *joined*, *asked me* and *is still* respectively. In (29), the error is ascribed to Arabic because the learner analogizes the use of *enter* with the use of *join* being influenced by Arabic semantics where it is possible and acceptable to use *daxal-a* (*entered*) in the sense of *joined*. In (30), the error lies in the co-occurrence of *ask to* being analogized with the use of *say to* vividly showing a semantic-syntax overlap. Ascribing such an error to *L2-influence* becomes due to the impossibility of such structures in Arabic. The learner here wrongly hypothesizes that as in the case of *say to*, he/she could use *ask to*. The error in (31) could be ascribed to *L1-transfer* because in Arabic, the adverb-like modal *mazaal* (*still*), what is termed as *min axawaatkaan* (a sister of *kaan*) can be used as a main verb especially under the influence of English-Arabic bilingual dictionaries. It could also be ascribed to English where the learner hypothesizes that *still* could be used as a verb. (32) presents a surprising error.

It is ascribable neither to Arabic simply because such a sentence is never used. It could not even be ascribed to L2 for the fact that this expression is never used in English, and I think in any language. From a psycholinguistic perspective, I assume here that the learner, influenced by analogy of such expressions as *I have one brother*, *I have two sisters*, etc. may analogize that he/she could use such an expression.

7. Conclusions and Pedagogical Implications

This study makes it clear that lexis acquisition is so important in L2 acquisition process. As presented in Table(1) above, errors frequency, where 2040 errors have been committed, indicates that such errors are considered barriers to communicative competence and hence, the ultimate goal of language acquisition causing unintelligibility, misunderstanding and/or at least difficulty in comprehending the intended message. *L1-transfer* scoring 890 errors, (44.%) indicates that L1 is the first major source of lexical errors. *L2-influence* scoring 812, i.e. (40%) indicates that L2 is the second major source. However, this study stresses that L1 and L2 are not the merely sources of errors committed by L2 learners supporting earlier studies (cf. Shormani, 2012a, 2012b; Dulay et al, 1982). In this study, *mutual* scores 275, i.e. (12.8.%) and *unrecognized* scores 65, i.e. (3.2%). Such findings undoubtedly indicate that Yemeni students encountered a considerable difficulty rising from L2, L1 both languages and “some factors else.”

It is true that this study focuses on Arab learners of English but the results it comes up with indicate that such results and findings could be expected from other learners of English. It actually supports in one way or another studies done on Arab learners (cf. Mahmoud, 2005, 2011; Rababah, 2003; Khali, 1985; Shormani & Sohmani, 2012; Zughoul, 1991, among others). It also supports some studies done on non-Arab learners of English (cf. Laufer, 1997; Llach, 2005; Taylor, 1975; Schmitt & Zimmerman, 2002; Zimmermann, 1986, among many others). Thus, it seems that learners of English as an L2 have or share a great amount of problematic issues whatever their L1s are.

Arabic-speaking learners of English seem to have the most difficulty in *assumed synonymy*, literal translation and paraphrase. This is actually consistent with many studies done on Arab and non-Arab learners of English (e.g. Lennon, 1991; Zughoul, 1991; Llach, 2005; Laufer, 1997; Hemchua & Schmitt, 2006, among others). As far as *assumed synonymy* is concerned and as has been seen, some errors in this category are caused by *L1-transfer*. This is actually due to false conception where students hypothesize that there is one-to-one correspondence between English and Arabic at the level of synonymy unaware of the fact that here is no exact synonymy in a language, let alone between L1 and L2, here Arabic and English respectively (Griffiths, 2006). The confusion takes place between what is to learn and what has been already learned. Hemchua & Schmitt (2006) suggest that learners should be aware of 1) identifying stylistic level, i.e. formal/informal, colloquial/slang, etc. 2) presenting examples in the grammatical categories, like noun, verb, adjective, etc. 3) contextualizing lexis, 4) providing students with the most common collocates of common lexis. Shormani (2012a) argues that L2 learners have the right lexis but due to some factor, they use the incorrect one suggesting that lexis should be taught in context supplied with where a word could be used and where not in association frames.

As reported in our study, literal translation occupies the second rank among the errors committed. L2 learners hypothesize that there is one-to-one correspondence at word level between their L1 and L2 in addition to what is so called polysemy especially divergent where one word in L1 may have two more polysemous words in L2. Here, teachers should draw students' attention to culture differences between Arabic and English. Students should also be made aware that there is some kind of “lack of equivalence in formality and style (for example, respectful and intimate styles) (Hemchua & Schmitt, 2006, p. 21). Therefore, English should be taught and

learned in its own right and using Arabic in the classroom should be eliminated or at least reduced.

For paraphrase, it is actually a skill which needs more practice. Therefore, students should get trained in paraphrasing by giving them constant assignments in two levels, viz. writing and speaking. The former should be done in Writing and Literature Courses in and out of class. The latter, however, can be done only in class and for this purpose, I assume, Spoken Courses should be exploited. For the former, students might be given particular excerpts from a novel, a short story and asked to paraphrase them at home. For the latter, students might be given such a kind of excerpts and asked to paraphrase orally.

For overcoming the difficulty imposed by derivativeness, I propose that students should first learn the difference between Arabic and English derivativeness and this could be achieved in morphology and syntax classes. In the former, students should learn and master the difference between both languages' morphological systems. They should learn English morphology in its own right because it is of concatenative morphology being completely different from that of Arabic which is non-concatenative and how each works (McCarthy, 1981; Holes, 2004). In the latter, however, students should learn the differences between Arabic and English word classes. They should be introduced to the fact that while in Arabic, a word of a grammatical class could be used for another class as in the case of nouns and adjectives, verbs and adverbs, etc., it is not the case in English where such a phenomenon is rarely used. Schmitt, & Zimmermann (2002, p. 164) stress that students should be reinforced when they succeed in deriving an item and when it is not, "teachers should try to capitalize on the partial knowledge that these errors indicate" and I assume this could be achieved in morphology classes.

As far as idiomaticity is concerned, teachers are advised to make it clear to learners that the meaning of an idiomatic expression is composed of the meaning of its elements altogether in an idiomatic way and not of the sum of the meanings of the elements isolatedly. Learners should, for instance, learn that *at home* is different from *in the house* both syntactically and semantically. They also should be introduced to bilingual differences. For instance, while the former exists in English and has no equivalent in Arabic, the latter exists in both languages. Students could also be given a list of the most widely and commonly used phrasal verbs and common preposition partners. This actually can be achieved in grammar classes for intermediate students and syntax classes for advanced students.

For binary terms, students should be introduced to differences existing between such terms. For instance, the difference between *learn* and *teach*. From a syntactic perspective, the former should be learned as a monotransitive verb while the latter as both a mono- and di-transitive verb. From a semantic perspective, while the former is used in the sense that learning is intrinsic while teaching is extrinsic. They should also be introduced to the fact that *learn* is different from its Arabic equivalent. In that, the Arabic one could be used in the sense of *teach*. This actually could be obtained in both syntax and semantics classes.

For similar forms, students should understand that homophones, be they phonological or orthographical, do not have the same meaning and/or use. Phonology classes should be exploited for such a purpose. For overused terms, students should be taught the differences between superordinates and hyponyms and that while the former could, in some cases, be used for the latter, it is not vice versa. For instance, every *rose* is a *flower* but not every flower is a *rose*. They should also be taught some semantic differences between terms as in the case of *big* and *huge*, *young* and *small*, etc. Though it is a learning strategy, students should be taught to be aware that analogy is not always error-free. In fact, analogy indicates that students are not passive interlocutors but active ones, however, if it is ok in a context, it may be not in many others.

What has been stated above is concerned with suggesting some solutions to difficulties encountered by Yemeni Arabic-speaking learners of English where such difficulties could be related to L1, i.e. Arabic, L2, i.e. English or both represented by *L1-transfer*, *L2-influence* and *mutual* sources respectively. Now, the question is what about those errors having *unrecognized* sources. In fact, investigating and providing solutions to such errors is very interesting and challenging but due to the complexity such errors impose, I suggest that such a topic could be the main focus of a study on its own and thus, left for future research.

References

- Bahns, J. (1993). Lexical collocations: a contrastive view. *ELT*, 47(1), 56-63.
- Brooks, N. (1960). *Language and language learning: Theory and practice*. New York: Harcourt, Brace and World.
- Chomsky, N. (1968). *Language and Mind*. New York: Harcourt Brace Jovanovich.
- Corder, S. (1973). *Introducing Applied Linguistics*. Harmondsworth: Penguin Books Ltd.
- Corder, S. (1981). *Error Analysis and Interlanguage*, Oxford: Oxford University Press.
- Dulay, H. & Burt, M. (1973). Should We Teach Children Syntax? *Language Learning*, 23(2), 245-258.
- Dulay, H. & Burt, M. (1974). Errors and Strategies in Child Second Language Acquisition. *TESOL Quarterly*, 8(2), 129-136.
- Dulay, H., Burt, M. and Krashen, S. (1982). *Language Two*. New York: Oxford University Press.
- Duskova, L. (1979). On Sources of Errors in Foreign Language Learning. *IRAL*, 7(1), 11-36.
- Ellis, R. 1997. *Second Language Acquisition*. Oxford: Oxford University press.

- Gass, S. & Selinker, L. (2008). *Second Language Acquisition: An Introductory Course*. (3rd ed.). New York: Routledge.
- Griffiths, P. (2006). *An Introduction to English Semantics and Pragmatics*. Edinburgh: Edinburgh
- Han, Z.-H. (2000). Persistence of the Implicit Influence of NL: The Case of the Pseudo-Passive. *Applied Linguistics*, 21(1), 55–82.
- Han, Z.-H. (2004). *Fossilization in Adult Second Language Acquisition*. Clevedon: Multilingual Matters.
- Hang, J. (2005). A Study of Lexical Errors in Cantonese ESL Students' Writing. Unpublished MA Dissertation, University of Hong Kong.
- Hemchua, S. and Schmitt, N. (2006). An Analysis of Lexical Errors in the English Compositions of Thai Learners. *Prospect*, 21, 3, 3-25.
- Holes, C. (2004). *Modern Arabic: Structures, Functions, and Varieties*. Washington, DC: Georgetown University.
- James, C. (1977). Judgements of Error Gravities. *ELT*, 3(2), 116-124.
- James, C. (1998). *Errors in Language Learning and Use: Exploring Error Analysis*. Hong Kong: Addison Wesley Longman Limited.
- Khalil, A. (1985). Communicative Error Evaluation: Native Speakers' Evaluation and Interpretation of Written Errors of Arab EFL Learners. *TESOL Quarterly*, 19, 2, 335-351.
- Khatib, A. (1984). A Classification of the Lexical Problems of EFL/ESL Students. ERIC Document No. ED246691, 1-51.
- Laufer, B. (1997). Some Properties of the Foreign Language Learner's Lexicon as Evidenced by Lexical Confusions. *IRAL*, 29(4), 317-330.
- Lennon, P. (1991). Error: Some Problems of Definition, Identification, and Distinction. *Applied Linguistics*, 12(2), 180-195.
- Lennon, P. (1992). Error and the Very Advanced Learner. *IRAL*, 29(1), 31-44.
- Llach, M. P. 2005. The Relationship of Lexical Error and their Types to the Quality of ESL Compositions: an Empirical Study. *PortaLinguarum*, 3, 45-57.
- Mohanty, P. (2006). A Unique Error in the Arabic Speakers' Written English: A Search for Its Sources. *Indian Linguistics*, 67(1), 127-128.
- McCarthy, J. (1981). A Prosodic Theory of Nonconcatenative Morphology. *Linguistic Inquiry*, 12, 373-418.
- McCarthy, M. (1990). *Vocabulary*. Oxford: Oxford University Press.
- Mahmoud, A. (2005). Collocation errors made by Arab learners of English. *Asian EFL Journal*. Retrieved February, 13, 2012, from <http://www.asian-efl-journal.com>
- Mahmoud, A. (2011). The Role of Interlingual and Intralingual Transfer in Learner-centered EFL Vocabulary Instruction. *AWEJ*, 2(3), 28-49.
- Mitchell, R. & Myles, F. (1998). *Second Language Learning Theories*. London: Arnold.
- Rababah, G. (2003). Communication Problems Facing Arab Learners of English: A Personal Perspective. *Journal of Language and Learning*, 3(1), 180-197.
- Richards, J. (1972). Social Factors, Interlanguage, and Language Learning. *Language Learning*, 22(2), 159-188.
- Richards, J. (1974). *Error Analysis: Perspectives on Second Language Acquisition*. London: Longman.
- Saville-Troike, M. (2006). *Introducing Second Language Acquisition*. New York: Cambridge University press.
- Schmitt, N. & Zimmerman, C. (2002). 'Derivative Word Forms: What Do Learners Know?' *TESOL Quarterly*, 36(2), 145–171.
- Selinker, L. (1992). *Rediscovering Interlanguage*. London: Longman.
- Selinker, L. (1993). Fossilization as Simplification. In Tickoo, M. (ed.), *Simplification: Theory and Application*. Singapore: SEAMEO Regional language Centre, 14-28.
- Shormani, M. (2012a). *L2 Acquisition in the Arab World: An Empirical Study of Yemeni Learners' English Syntactic and Semantic Error Gravity*. Saarbrücken: LAMBERT Academic Publishing, Germany.
- Shormani, M. (2012b). Sources of Syntactic Errors in Yemeni Learners' English Compositions: A Psycholinguistic Analysis. *AWEJ* 3(4), 275–296.
- Shormani, M. (2013). Fossilization and Plateau Effect in Second Language Acquisition. *Language in India*, 13, 763- 784.
- Shormani, M. & Sohmani, Y. (2012). Semantic Errors Committed by Yemeni University Learners: Classifications and Sources. *International Journal of English Linguistics*, 2(6), 120-139.
- Sonaiya, R. (1991). Vocabulary Acquisition as a Process of Continuous Lexical Discrimination. *IRAL*, 14(4), 269–284.
- Takac, V. (2008). *Vocabulary Learning Strategies and Foreign Language Acquisition*. Multilingual Matters Ltd.
- Taylor, B. (1975). The Use of Overgeneralization and Transfer Learning Strategies by Elementary and Intermediate Students in ESL. *Language Learning*, 25(1), 73-107.
- Tomasello, M. (2007). Cognitive Linguistics and First Language Acquisition. In Geeraerts, D. & Cuyckens, H. (eds.), *Handbook of Cognitive Linguistics*, Oxford: Oxford University Press, 1092-1112.
- Wang, W. & Wen, Q. (2002). L1 Use in the L2 Composing Process: An Exploratory Study of 16 Chinese EFL

Writers. *Journal of Second Language Writing*, 11(3), 225-246.
Zimmermann, R. (1986). Semantics and Lexical Error Analysis. *EnglischAmerikanischeStudien*, 2(86), 294-305.
Zughoul, M. (1991). Lexical Choice: Towards Writing Problematic Word Lists. *IRAL*, 29(1), 45-60.

ⁱThis classification has been adopted and adapted from Shormani (2012b).

ⁱⁱThe Asterisk * stands for the ungrammaticality of the word/phrase/sentence it is put before.