The Study of Iranian EFL Learners' Mental Lexicon through Word Association Tests

Farid Ghaemi (Corresponding author)
Department of English Language
Islamic Azad University, Karaj, Iran

Shahram Halimi Passand
Department of English Language
Islamic Azad University, Gheshm, Iran
(Corresponding author) sh.halimi@rocketmail.com

Abstract

This study intended to peek into Iranian EFL learners' metal lexicon through word association tests (WATs). 31 male and female EFL learners studying at Bonab and Marageh language institutes participated in this study. A WAT comprised of 8 English words adopted from Roux's (2013) word list administered to the participants. The results analyzed and interpreted according to both WA conventional classification and Fitzpatrick's framework. Within conventional classification (syntagmatic, paradigmatic, and clang), the results confirmed the *syntagmatic to paradigmatic change hypothesis* (S→P) only between intermediate and upper-intermediate levels. The results also indicated that low intermediate learners besides other conventional factors, associate words based on phonological and orthographical relations. Within Fitzpatrick's framework, the results indicated that learners generally associate words according to meaning and position across all proficiency levels. However, at low-intermediate level the rate of meaning-based association overwhelms position-based association. Form-based association and erratic association drew the least attention of the participants respectively. Finally, pedagogical implication of this study along with further research idea is discussed.

Keywords: conventional classification, Fitzpatrick's framework, mental lexicon, word association test

Introduction

Nowadays nearly everyone acknowledges that vocabulary is the central to communicating in a foreign language. Without sufficient words to express a wide variety of meaning, communicating in a foreign language cannot happen in a meaningful way. However, this fact was nearly neglected by the majority of researchers (in academic circles at least) in the literature in the past. For instance, as Milton (as cited in Milton & Donzelli, 2013) points out, in structuralist approaches to language learning and teaching it was thought the number of vocabulary items necessary for learning could be limited only to what was strictly necessary to exemplify or use the grammar. However, its importance to the field has recently been acknowledged. For instance, Long and Richards (as cited in Milton & Donzelli, 2013, p. 441) have described the

vocabulary knowledge as the "core component of all the language skills."

This awareness on the importance of vocabulary knowledge among second language researchers necessitate in depth understanding of how the words (no matter how the term is defined; see Milton & Donzelli, 2013) of foreign language are learned, organized, stored, and retrieved by the learners. The mechanism responsible for handling this problem in the mind is traditionally called *mental lexicon*. It is a mental system which contains all the information a person knows about words (Richard & Schmidt, 2002). Nonetheless, little is known about how L2 lexical information is represented in the mental lexicon and how it functions. Perhaps it is because the representation facet of second language acquisition research has not received its due attention in the past. As from psycholinguistic point of view any adequate theory of second language acquisition should fulfill three interrelated aspects: the study of representation, the study of acquisition, and the study of processing (Jiang, 2000). Hence, the study of second language acquisition is incomplete without representation component, since as pointed out by Levelt (as cited in Jiang, 2000) representation and processes cannot be studied independently of each other. This is more conspicuous in the study of vocabulary acquisition in the L2 acquisition. This fact to some extent can explain the lack of adequate conceptual framework by which the findings of numerous L2 vocabulary studies can be discussed.

According to Aitchison (as cited in Khanzaeene-zhad & Alibabaee, 2013) there are roughly four main methods for investigating the mental lexicon: 1) word searches (tip-of-the-tongue or TOT states) and slip of the tongue, 2) linguistics and linguistic corpora, 3) speech disorders and brain scans, and 4) psycholinguistic experiments. Word association test (WAT) is one form of psycholinguistic experiment employed both in first and second language acquisition studies to investigate the lexical connections individuals hold in their developing mental lexicon (Peppard, 2007); since as Aitchison (as cited in Russ, n. d, p. 1) puts it "words are not stored in the mental lexicon as single independent items but form clusters or webs with other related concepts so that words acquire their full meaning in reference to related terms." The WAT is popular because of its simplicity and ease of administration. Word associations are usually obtained through a simple stimulus-response procedure, whereby the researcher provides a prompt word (PW) and the participant utters the first word that comes to the mind. There are different incarnations involving oral-oral, oral-written, and written-written stimulus-response methods. Some WATs ask subjects to reply with the first word they think of, while others require participants to provide as many words as they can within a given time period (Wharton, 2010).

Considering the significance of lexical know-ledge, it is incumbent on us, as language practitioner, to provide the best pedagogical practices in promoting the students' lexical development. Hence, the present study attempts to investigate how the mental lexicon of Iranian EFL learners is organized. If we take into account

the common ways in which they associate words with each other, we will be in a better position to prepare and present lessons that support the natural way the mind acquires and catalogues lexis. Consequently, both teaching and learning will become more efficient.

Theoretical Background

Traditionally word association responses generally fall into three main classes called syntagmatic (collocation, multi-word items, encyclopaedic knowledge), paradigmatic (co-ordination, hyponemy and hypernymy, synonymy) and clang associations (Gass & Selinker, 2008). Syntagmatic associations are identified if the response forms an obvious sequential link with the stimulus word. In other words, stimulus and response words are from different grammatical form classes (e. g. ball→ catch; run → fast; dog → bark). Paradigmatic associations are recognized if the response and stimulus word are from the same word class (e. g. bus→ train; black→ white; dog→ cat, or animal). Clang associations are considered to be without any clear meaningful link, and are based on similarities in phonology or orthography (e. g. phone→ foam; knife→ knight). Some studies also included a *nil* category to handle unclassifiable responses (Wharton, 2010).

Recently, Fitzpatrick (2006, 2007) by noting the shortcoming of traditional WATs response categories, offers a more sophisticated WAT response category. He uses *meaning-based*, *position-based*, *form-based*, (and sub-classifications within these), and *erratic* associations to represent the mental lexicon more clearly. This classification is presented in the following tables (adopted from Roux, 2013):

Table 1
Fitzpatrick's Model: A description

Meaning-based responses(MBR)	Those determined by semantic characteristics.
Position-based responses(PBR)	Determined by syntactic and collocational characteristics
Form-based responses (FBR)	Determined by phonological, orthographical and collocational characteristics.
Erratic responses	No apparent link between cue and response, or no response.

Note. (Fitzpatrick, 2007, p. 330, cited in Roux, 2013, p. 83)

Table 2
Fitzpatrick's classification of association responses

Descriptor	Definition	Specification
Meaning-based Responses (MBR)	Defining synonym	X means the same as y
	Specific synonym	X can mean y in some specific contexts
	Lexical set/context related	X and y same lexical set: coordinates/ meronyms/superordinates provide context

Descriptor	Definition	Specification
Position-based Responses (PBR)	Conceptual association	X and y have some other conceptual link
	Consecutive xy collocation	Y follows x directly (includes compounds)
	Consecutive yx collocation	Y precedes x directly (includes compounds)
	Other collocational	Y follows/precedes x in phrase with word(s) between them
Form-based Responses (FBR)	Change of affix	Y is plus or minus affix
	Similar form not meaning	Y looks similar to x but has not clear meaning link or is an associate of a word with a similar form to x
Erratic Responses (ER)	No link/blank y has no decipherable	Link to x or no response given

Note. (Fitzpatrick, 2007, p.331 cited in Roux, 2013, p. 83-84)

Related Studies

Early studies into native children on WATs (Khanzaeenezhad & Alibabaee, 2013) found that as children aged, they produced more paradigmatic responses, and less syntagmatic and clang associations. This belief was most commonly referred to as the syntagmatic-paradigmatic (S-P) shift (Peppard, 2007). The finding led most SLA researchers to expect that as L2 learners' proficiency increase they would evidence more paradigmatic responses, whereas weaker learners would produce more clang or syntagmatic associations. This inference was unchallenged for decades (see Khanzaeenezhad & Alibabaee, 2013), as Wolter (as cited in ibid) feels the S-P shift would be better described as a "shift from semantically meaningless response to semantically meaningful responses." However, later studies showed that it was rather hasty analogy (see Roux, 2013; Wharton. 2010).

Studies have shown that the word associations produced by second language learners differs systematically from those of native speakers. For instance, in spite of the fact that L2 learners have smaller and limited vocabulary than native speakers, their responses tend to be more varied and less homogeneous. For example, Meara (1983) in a study comparing the behavior of native speakers with L2 learners on WAT found that L2 learners responses tend to be heterogeneous compared with L1 speakers. On the other hand, Soderman (as cited in Rahimi & Haghigi, 2009) in a similar study on native Finnish EFL students found that the shift from syntagmatic to paradigmatic responses by increasing learners' proficiency was not significant. Yoneoka (2001, ibid) evidences the tendency for Japanese participants to respond more frequently with syntagmatic responses. The reason is not completely clear yet; however, Meara (1983) believes one contributory factor seems to be their inclination toward producing clang association as children, and another is frequently misunderstanding the stimulus word.

Some recent studies questioned the clear-cut division between L1 and L2 lexicons, since in the case of not so much high frequency words as prompts, NS and NNS associations become more similar in the proportion of paradigmatic, syntagmatic, and clang responses produced (Wolter, 2001; Fitzpatrick, 2006). This seems to indicate that the actual organizations of mental lexicons of these groups are not so different. However, Wolter (2006) suggests that the real differences exist between syntagmatic associations (e.g. collocations) rather than paradigmatic associations, as "the process of building syntagmatic connections between words in an L2 appear to be considerably harder than the process of building paradigmatic connections.

On the other hand, some researchers argued against the rigid distinction between syntagmatic and paradigmatic associations, as many responses which share the same word class as the prompt word (PW) can be related sequentially as well (e. g. mountain→ bike; school→ graduation). This shortcoming led some researchers to make some modifications and enhance the traditional classification systems (see). To date, Fitzpatrick's (2007) more detailed word association response categories are the most comprehensive. Using this paradigm one can make clear lots of seemingly obscure associations which in the old paradigms were classified as *clang* or even *nil* categories. Although Fitzpatrick acknowledges that it is a slightly time consuming and laborious process, it is the best bay to accurately categorize responses, hence more representative of the actual mental lexicon being examined.

Russ (n. d) in the same vain of studies (using the old paradigm)found that although no definitive conclusion can be made, it appears that L2 learners tend to organize the mental lexicon much like L1 speakers do. He argues that according to his studies word class is an important feature of lexical organization. Moreover, personal experiences and phonological systematizing also appear to play a role in lexical linkage.

Research Questions

As the majority of word association studies utilizing the conventional classification (syntagmatic-paradigmatic shift) reported above led to a number of inconsistent and contradictory results, it seems there is still a need for more exploration to gain a better understanding of how L2 learners' mental lexicon are represented. Therefore, this small-scale research is an attempt to expand our current understanding of behaviors of Iranian EFL learners on word associations. Since Fitzpatrick's paradigm is more sophisticated and studies conducted based on it usually have led to illuminating results, aside from the traditional classification, the same paradigm is also utilized in this study; as no research utilizing Fitzpatrick' classification along with retrospective interview (to the best of my knowledge) has been done on in Iran. Hence this study intended to investigate the Iranian EFL learners behavior on WAT on different proficiency groups, and the specific research question addressed are as follows:

1. Is there any difference in word association behaviors of Iranian EFL learners regarding the traditional paradigm (paradigmatic, syntagmatic, and clang classification)?

2. Do Iranian EFL learners' proficiency levels affect their word association behaviors regarding paradigmatic, syntagmatic, and clang classification?
3. Is there any difference in behaviors of participants regarding Fitzpatrick's four descriptors?
4. Do EFL learners' proficiency levels have any effect on their behavior on four descriptors?

The associated null hypotheses are as follows:

1. There is not any difference in word association behaviors of participants concerning the traditional paradigm (paradigmatic, syntagmatic, and clang classification).
2. EFL learners' proficiency level does not have any effect on their behavior regarding paradigmatic, syntagmatic, and clang classification.
3. There is not any difference in behaviors of participants regarding Fitzpatrick's four descriptors.
4. EFL learners' proficiency does not have any effect on their behavior on Fitzpatrick's four descriptors.

Method

Participants

The participants in this study were 31 male and female EFL learners studying at Bonab and Marageh language institutes. The age range of these participants was between16-27. Some of them were English students at university or English graduate. The majority of learners at this institutes study in conversation classes ranging from beginner to upper levels. The participants were initially informed of the purpose of the study and they eagerly accepted to co-operate.

As one of the goals of the study was to compare the performance of the participants at different levels, namely, low, mid and high groups, sampling was carried out accordingly. It means that on the bases of questionnaire about their English learning background, their current study level at institutes, and also based on the result of vocabulary part of a proficiency test (see the appendix), they were classified into the three groups. It is worth mentioning that the low group was selected from level 4 to make sure that the participants had the required vocabulary knowledge.

Instruments

Along with the vocabulary part of an English proficiency test which consists of 25, and were administered to the participants in the same session, the main instrument in this study for data collection comprised a word association test (WAT). Since choosing the suitable words as prompt is a very delicate task and also some of the similar studies suffer from not choosing the appropriate words as prompt, in this study I adopted Roux' (2013) word list. It consists of 8 frequently occurring and emotionally neutral English words, making them serve as the stimulus words with learners across a wide range of proficiency levels. Hence it has not pitfalls of some of the similar word list. The word list, word classification, and rationale for the choice of words are set out in the following table:

Table 3
Word list, word classification, and rationale for the choice of words

Cue	Word Class	Rational for Choice
Wash	Verb/Noun	A common word in everyday use; polysemic
Computer	Noun	A word in everyday by all the participants
Green	Adjective/Noun/Verb	A polysemic word that could tap socio- cultural and linguistic meaning
Believe	Verb	Perhaps a less frequently used word, slightly more difficult in conceptualization, but nevertheless postulated to be fairly well known amongst both respondent groups
Train	Verb/Noun	A common word in everyday use; polysemic use less common
Exciting	Adjective	A fairly common word, yet postulated to be used less in spoken than in written language
In	Preposition/Adjective/Adverb/Prefix/Noun	A polysemic, function word occurring regularly and with a variety of uses
Drive	Verb/Noun	A common word in everyday use; polysemic

Procedure

Owing to the nature of the study, *Sequential explanatory strategy (QUAL→qual)* (Creswell, 2009), a popular form of mixed method design, was utilized for data collection; i.e. both quantitative (written test) and qualitative (interview) research methods were utilized. However, in this method, as the notation indicates, the qualitative method is embedded within a qualitative design. In other words, collecting and analyzing follow-up qualitative data were used to explain and interpret quantitative results. Hence, to save space, only the end results of our analyses (in quantitative form) are represented in this study.

The data were collected in three phases: first, proficiency test, then WAT, and finally interview were administered. These procedures were followed during the participant's regular class time. In administering WAT, the directions were explained orally by the researcher as supplement to the written instructions in the test sheet. Moreover, several additional stimulus words had been practiced by the researcher and their teachers before the participant responded to the word list in the test. The participants were required to respond to the stimulus word by writing down the first word which comes to their mind as quickly as possible. They were encouraged to respond even if they think that their responses had no association with the stimulus words. It took them about two or three minutes to finish responding to 8 test items. After collecting the papers I interviewed them individually to obtain the reasons behind their responses. Then I jotted down their explanation in the specific part of their papers. Their explanation served

as invaluable information about the matter in question. Member-checking and peer-debriefing were utilized for the credibility of inferences about the comments.

Results

As a first step, using SPSS program, the mean and the standard deviation of the scores based on conventional classification are calculated. The results are presented in Table 4.

Table 4
The mean and the standard deviation of the scores

						95% Confidence Interval for Mean	
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound
Paradigmatic	lower-intermediate	12	2.6667	.65134	.18803	2.2528	3.0805
	intermediate	10	2.8000	.63246	.20000	2.3476	3.2524
	upper-intermediate	9	3.3333	.50000	.16667	2.9490	3.7177
	Total	31	2.9032	.65089	.11690	2.6645	3.1420
Syntagmatic	lower-intermediate	12	4.6667	.65134	.18803	4.2528	5.0805
	intermediate	10	5.1000	.56765	.17951	4.6939	5.5061
	upper-intermediate	9	4.3333	.50000	.16667	3.9490	4.7177
	Total	31	4.7097	.64258	.11541	4.4740	4.9454
Clang	lower-intermediate	12	.6667	.49237	.14213	.3538	.9795
	intermediate	10	.1000	.31623	.10000	-.1262	.3262
	upper-intermediate	9	.3333	.50000	.16667	-.0510	.7177
	Total	31	.3871	.49514	.08893	.2055	.5687

As the figures represents, while the participants at all proficiency levels produced paradigmatic and syntagmatic responses to the cue word, they produced very few clang responses. However, the rate of syntagmatic responses overwhelms the paradigmatic ones across the three proficiency levels. This finding rejects our first hypothesis that 'there is not any difference in word association behaviors of participants concerning the traditional paradigm (paradigmatic, syntagmatic, and clang classification).' Furthermore, one way AVOVA test is used to compare the mean scores of the groups.

The results are presented in the table 5 below

Table 5
The mean, the standard deviation and the results of ANOVA for comparing the mean among the groups

		M±S	F	Sig
Paradigmatic	lower-intermediate	2/67±0/65	3/33	0/052
	intermediate	2/8±0/63		
	upper-intermediate	3/33±0/50		
Syntagmatic	lower-intermediate	4/67±0/65	4/13	0/027
	intermediate	5/1±0/57		
	upper-intermediate	4/33±0/5		
Clang	lower-intermediate	0/67±0/49	4/49	0/020
	intermediate	0/10±0/32		
	upper-intermediate	0/33±0/5		

According to the results of one-way analysis of variance, there is a significant difference among the three groups in terms of syntagmatic (F=4/13, P<0/5), and clang (F=4/49, P<0/05) categories. Again this finding goes against of the second hypothesis that 'EFL learners' proficiency level does not have any effect on their behavior regarding paradigmatic, syntagmatic, and clang classification. However, no significant differences among the three groups observed regarding paradigmatic category.

On the other hand, since the ANOVA test only shows that there is a significant difference in the scores of the groups, but it does not show exactly where this difference is located, a kind of post hoc test is needed to elucidate this issue. Hence Scheffe test (see Tavakoli, 2012) is used to help us pinpoint where those differences are really located. The results are delineated in table 6.

As the table represents, the mean of intermediate group concerning syntagmatic category is significantly higher than upper-intermediate group (P<0/05). This finding confirms S→P hypothesis to some extant which claims that as learners' proficiency increase, they move from syntagmatic to paradigmatic responses. Nevertheless, this case was not confirmed between lower-intermediate and intermediate levels.

Table 6
Scheffe test for locating the exact differences among the proficiency levels

	GROUP(I)	GROUP(J)	Mean Difference (I-J)	Sig
Syntagmatic	lower-intermediate	intermediate	-0/43	0/241
	lower-intermediate	upper-intermediate	0/33	0/444
	intermediate	upper-intermediate	0/77*	0/028
Clang	lower-intermediate	intermediate	0/57*	0/022
	lower-intermediate	upper-intermediate	0/33	0/255
	intermediate	upper-intermediate	-0/23	0/530

*. P<0/05

Finally, the mean of lower intermediate group in terms of clang category is significantly higher than the intermediate group (P<0/05). This finding indicates that at lower proficiency level learners tend to organize words according to their phonological and orthographical relations.

Regarding Fitzpatrick' classification, the mean and standard deviations of scores are also calculated. The results are depicted at Table 7.

Table 7
The mean and standard deviations of scores based on Fitzpatric' classification

						95% Confidence Interval for Mean	
		N	Mean	Std. Deviation	Std. Error	Lower Bound	Upper Bound
Meaning-based	lower-intermediate	12	4.0000	.42640	.12309	3.7291	4.2709
	intermediate	10	3.3000	.48305	.15275	2.9544	3.6456
	upper-intermediate	9	3.7778	.44096	.14699	3.4388	4.1167
	Total	31	3.7097	.52874	.09497	3.5157	3.9036
Position-based	lower-intermediate	12	3.5833	.66856	.19300	3.1586	4.0081
	intermediate	10	3.8000	.63246	.20000	3.3476	4.2524
	upper-intermediate	9	3.1111	.60093	.20031	2.6492	3.5730
	Total	31	3.5161	.67680	.12156	3.2679	3.7644
Form-based	lower-intermediate	12	.0833	.28868	.08333	-.1001	.2667
	intermediate	10	.4000	.51640	.16330	.0306	.7694
	upper-intermediate	9	.2222	.44096	.14699	-.1167	.5612
	Total	31	.2258	.42502	.07634	.0699	.3817
Erratic Response	lower-intermediate	12	.4167	.51493	.14865	.0895	.7438
	intermediate	10	.4000	.51640	.16330	.0306	.7694
	upper-intermediate	9	.8889	.60093	.20031	.4270	1.3508
	Total	31	.5484	.56796	.10201	.3401	.7567

As the figures indicate, the participants produced meaning-based and position-based responses considerably across the three proficiency levels. Interestingly, they produced very few form-based and erratic responses. This finding again strongly rejects our third hypothesis claiming that 'there is not any difference in behaviors of participants regarding Fitzpatrick's four descriptors'.

To follow our previous procedure, the ANOVA test is used to compare the mean scores of the participants in different groups. This information is depicted at Table 8.

Table 8
The mean and standard deviations of the scores and results of ANOVA test

		M±S	F	Sig
Meaning based	lower-intermediate	4 ± 0/43	6/72	0/004
	intermediate	3/3± 0/48		
	upper-intermediate	3/78± 0/44		
Position based	lower-intermediate	3/58 ± 0/67	2/87	0/074
	intermediate	3/8± 0/63		
	upper-intermediate	3/11± 0/60		
Form based	lower-intermediate	0/08± 0/29	1/57	0/225
	intermediate	0/40± 0/52		
	upper-intermediate	0/22± 0/44		
Erratic Response	lower-intermediate	0/42 ± 0/51	2/51	0/099
	intermediate	0/40± 0/52		
	upper-intermediate	0/55 ± 0/57		

As it may seem clear from the figures, the results of one-way ANOVA test indicate that the differences among the three groups is only significant at meaning-based category (F=6/72). Nonetheless, again in order to pinpoint exactly where the differences are located, Scheffe test is utilized. The results are delineated in Table 9.

Table 9
The results of Scheffe test for comparing the mean of three proficiency levels

	GROUP(I)	GROUP(J)	Mean Difference (I-J)	Sig.
Meaning based	lower-intermediate	intermediate	0/70*	0/004
	lower-intermediate	upper-intermediate	0/22	0/541
	intermediate	upper-intermediate	-0/48	0/086

*. P<0/05

The results of the Scheffe test clearly indicate that at meaning-based category the mean score of lower-intermediate participants is significantly (P<0/01) higher than the intermediate group. This finding again does not bear out the fourth hypothesis that 'EFL learners' proficiency does not have any effect on their behavior on Fitzpatrick's four descriptors.

To sum up, the overall results based on conventional classification are not clear-cut, in spite of the fact that they showed that learners at low proficiency level, besides common factors, tend to organize words according to their phonological and orthographical relations. Nonetheless, regarding Fitzpatrick' classification, the findings are much illuminating. The results suggest that the participants predominantly favored meaning-based and position-based responses considerably across the three proficiency levels. This is in contrast to the mainstream belief (in conventional classification; S→P) that only advanced language users more frequently produce paradigmatic responses. Furthermore, the findings also indicate that at lower intermediate level students predominantly associate words according to their meaning (meaning based). As their proficiencies increase, they tend to associate in position-based way. An interesting finding relates to form-base category which drew the least attention from the learners across the three proficiency levels. Finally, as revealed by the participants' explanation on reason behind their responses, encyclopedic knowledge plays an important role across all levels and categories.

Conclusion

One of the reasons people are interested in the field of second language acquisition is to improve pedagogy. Hence, the findings of this study seem to have some obvious implication for teaching vocabulary. The most important message this paper conveys to language teachers and material developers is that words are meaningfully connected in the mental lexicon and should be taught accordingly. In other words, simply telling students the meaning of new words in de-contextualized way is not enough to fully incorporate them into the mental lexicon. Since in this study majority of participants' responses refer to meaning-based and position-based ones,

it requires teachers to highlight those vocabulary learning activities which relate to those domain. Moreover, students' proficiency levels should be considered when developing materials. For example, beside conventional tasks and activities for learning vocabulary, syllabus designers and teachers should include phonological and orthographical relational materials and activities in their syllabuses for low-intermediate students. In addition, due to the idiosyncratic nature of mental lexicon, teachers should pay attention to the student's learning style preferences, and help learners learn vocabularies accordingly.

A word of cautious is in order here. Since the number of cue words in this study was so small and the participants might not be representative of whole population, the findings could not be generalized confidently to all situations and people. Therefore, similar studies with more participants across different situations and of course with carefully chosen cue words need to be carried out in the future. To the best of my knowledge, Fitzpatrick's classification yields more illuminating results comparing with other existing ones. Nonetheless, it seems that if we want to use word association tests as a way to understand the mental lexicons of individuals, a more robust methodology and model is needed to enhance the construct validity of our testing.

References

Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed method approaches* (3rd ed.). Thousand Oaks, CA: Sage.

Cui, Y. (2009). *The development of lexical organization in Chinese EFL learners at tertiary level.* Retrieved June 27, 2014, from http://www.slat.arizona.edu/sites/slat/files/page/awp16cui.pdf

Khanzaeenezhad, B., & Alibabaee, A. (2013). *Investigating the role of L2 language proficiency in word association behavior of L2 learners: A case of Iranian EFL learners*. Retrieved June 28, 2014, from http://www*ojs.academypublisher.com/index.php/tpls/article/.../tpls0301108115*

Gass, S. M. & Selinker, L. (2008). *Second language acquisition: An introductory course*. (3rd ed.). Abingdon: Routledge.

Jiang, N. (2000). *Lexical representation and development in a second language*. Retrieved June 28, 2014, from http://www2.gsu.edu/~eslnxj/jiang2000.pdf

Meara, P. (1983). *Word association in a foreign language*. Retrieved June 25, 2014, from http://www.lognostics.co.uk/vlibrary/meara1983.pdf

Milton, J., & Donzelli, G. (2013). The lexicon. In J. Herschensohn & M. Young-Scholten, (Eds.), *The handbook of second language acquisition.* Cambridge University Press.

Peppard, J. (2007). *Exploring the relationship between word-association and learners' mental development*. Retrieved June 25, 2014, from http://*www.birmingham.ac.uk/documents/college.../peppardmod2.pdf*

Rahimi, A., & Haghighi, H. (2009). *How learners make mental links with words*. Retrieved June 26, 2014, from http://www.english.fullerton.edu/publications/cln/.../arthamzeh-links.pdf

Roux, P. W. (2013). *Words in the mind: Exploring the relationship between word association and lexical development*. Retrieved June 26, 2014, from *http://www.rcube.ritsumei.ac.jp/.../8-*

Russ, R. (n. d). *Word association*. Retrieved June 27, 2014, from *http://*www.kansaiu.ac.jp/fl/publication/pdf_forum/9/3_robin.pdf

Tavakoli, H. (2012). *A dictionary of research methodology and statistics in applied linguistics*. Boston: Rahnama press.

Wolter, B. (2001). Comparing the L1 and L2mental lexicon: A depth of individual word knowledge model. *Studies in Second Language Acquisition, 23*, 41–69.

Wolter, B. (2006). Lexical network structures and l2 vocabulary acquisition: The role of l1 lexical/conceptual knowledge. *Applied Linguistics, 27*(4) 741–747.

Wharton, C. (2010). *The effectiveness of retrospective interview in L2 word association research.* Retrieved, from *http://* jalt-publications.org/archive/proceedingsE048.

Zareva, A. (2007). Structure of the second language mental lexicon: How does it compare to native speakers' lexical organization? *Second Language Research, 23*(2), 123-153.