



Measuring Productive Derivational Knowledge of the Most Frequent Words

ALIREZA BAROUNI EBRAHIMI*

University of Western Ontario (Canada)

Received: 29/01/2020. Accepted: 14/05/2020.

ABSTRACT

Derivational knowledge is associated with writing and speaking skills. These skills are essential for EFL students who express themselves in oral presentations or written assignments. Therefore, diagnostic measurement of productive derivational knowledge is of vital importance, especially in regard to the most frequent 1,000 word families that cover 81% and 85% of written and spoken text. This study measured 46 Iranian university EFL students' productive derivational knowledge of the words at the 1,000 word frequency level. The findings indicate that while participants had the productive form-meaning knowledge of the words at 1,000 level, they did not seem to have extensive derivational knowledge of the same words. This assists in diagnosing area of weakness and placing instructional emphasis on high frequency words.

KEYWORDS: Derivational knowledge; Productive derivational knowledge; Productive derivational assessment; Measuring productive derivational knowledge; Most frequent words.

1. INTRODUCTION

Lexical knowledge is related to success in reading, writing, general language proficiency and academic achievement (Laufer *et al.* 2004; Milton 2013; Milton *et al.* 2010; Schoonen 2010; Stæhr 2008). Meara (1996) and Schmitt (2010) argue that vocabulary knowledge makes a significant contribution to almost all aspects of second language proficiency. In the same vein, Meara and Jones (1988) assert that vocabulary knowledge is heavily implicated in all language skills (listening, speaking, reading, and writing). Among all words, the ones used frequently in

**Address for correspondence:* Western University, Faculty of Education, 1137 Western Rd., London, ON, Canada N6G 1G7; e-mail: abarouni@uwo.ca

a wide range of spoken and written texts are essential in effective comprehension and communication. The most frequent 1,000 word families reoccur so often in spoken and written texts that they have a much greater value for comprehension and use than the second and the third most frequent ones (Webb & Chang, 2012). Nation (2006) asserts that the greatest variation in vocabulary coverage occurs in the most frequent 1,000 word families and in the proper nouns which cover 78% to 81% of written text and approximately 85% of spoken text.

Traditionally, knowledge of a word refers to the meaning and form of that word. Schmitt (2008) asserts that "many teachers and learners consider a word 'learned' if the spoken/written form and meaning are known" (p.333). Form and meaning link is the basic dimension of vocabulary knowledge and most probably the first to be acquired (Laufer & Goldstein, 2004). That is, other dimensions of vocabulary knowledge are learned after form and meaning aspect. Nation's (2001, 2013) conception of vocabulary knowledge includes not just form and meaning but various other aspects including word parts (affixes). Nation (2013) argues that knowing a word can involve knowing that it is made up of affixes and a stem that can occur in other words. These affixes are mainly divided into two types – inflectional and derivational¹.

Nyssonen (2008) found that the performance of her 56 grade 6th participants on the affixes strongly correlated with writing ability at the level of .73. Generally, productive vocabulary knowledge, including derivational knowledge, is associated with productive language skills – writing and speaking (see Evans & Green, 2007; Laufer & Goldstein, 2004). These two language skills are essential for English as a Foreign Language (EFL) students who wish to study at English universities and colleges in which they would struggle with the language demands of university work (see Angelova & Riazantseva, 1999; Gan, 2012; Liu & Jackson, 2008).

Based on the discussed issues, it can be concluded that productive derivational knowledge of the most frequent 1,000 word families has a significant impact on EFL students' speaking and writing for academic success. Therefore, the necessity of knowing the extent to which these students have productive derivational knowledge of high frequency words is a matter of real concern. The current study measures knowledge of derivative forms (word parts) and corresponding form-meaning connection of the most frequent words. For this purpose, one productive test of word parts (derivations) and one corresponding productive test of form-meaning connection were developed.

2. LITERATURE REVIEW

2.1. Word Parts

For the purpose of measuring word parts productively, the first question which needs to be addressed is the matter of measuring inflections, derivations, or both. To test this aspect,

Nation (2013) asks whether the learner can “produce the appropriate inflected and derived forms of the word” (p.538), and for the purpose of testing this aspect specifically, he asks whether the learner can “provide an affixed form of a known word” (p. 551).

Therefore, Nation targets both inflections and derivations for this purpose, although the way he suggests testing this aspect does not seem to be efficient enough because an affixed form does not necessarily show both the inflected and derived forms of a target word. As a result, literature should be investigated to find how word parts have been measured productively by different researchers.

Schmitt and Meara (1997) provided a line space in front of each target word, and the participants were asked to write all allowable suffixes of the target words. They found that the participants did not perform very well on the suffixes. Out of all possible native-like suffixes, 42% of them near the beginning of the school year and 47% of them near the end of the school year were provided productively by the participants.

Schmitt (1998, 1999) measured grammatical knowledge by the following question: “What word class (part-of-speech) is _____?”. Depending on the answer, he went on to ask “Is there a (noun, verb, adjective, adverb) form? If so, what is it?” for the remaining 3 word class forms.

Schmitt (1998) found that his four participants’ mean scores were 2.45, 2.09, 1.82, and 2.55 out of 4 (approximately 56% on average). However, the participants knew some individual word classes better than others—nouns and verbs were the best mastered while adjectives and adverbs appeared to be learned at a later stage. Schmitt believes that this suggests that adjective and adverb forms might be good candidates for explicit instruction.

Schmitt (1999) found that only a very small number of the participants, even though advanced, gave all four word classes for the target words. He believes that this suggests that L2 learners do not seem to master all members of a word family until relatively late in the acquisition process, and even this may not happen at all.

Schmitt and Zimmerman (2002) presented a series of four similar contextualized sentences for each target word. The participants were required to write the appropriate derivative form of the target word in each blank. They were to place an X in the blank if no derivative form existed and also the participants were informed that the target word could be the appropriate form without alteration. Schmitt and Zimmerman found that the participants’ mean number of derivatives was 37.6 (58.8%) out of a possible 64. This shows that the participants produced approximately two of the four possible derivatives. A relatively small percentage of the cases showed full mastery of the word family (all 4 forms) as even the

advanced graduate participants provided fewer than 45% of the possible derivatives for words they believed they knew.

Pacheco (2005) administered two quizzes, before and after her treatment (instruction of affixes), on forming new words using affixes with the instructions that read “make all the new words and word families you can by adding prefixes and suffixes to the words given” (Pacheco, 2005, p. 175). For 10 target words in the first quiz, her students wrote from 13 to 44 words while the number of correct answers ranged from 10 to 31 (with percentages of correct words ranging from 39% to 91%). For 15 target words in the second quiz, her students wrote from 15 to 43 words while the number of correct words were from 5 to 34 (with percentages of correct words ranging from 33% to 100%).

Ishii and Schmitt (2009) prepared the following test format for measuring derivative forms:

| Targeted word | Noun | Verb | Adjective |
|---------------|-------------|-----------|-------------|
| Stimulate | stimulation | stimulate | stimulating |
| Educate | | | |

Figure 1. An example of a word parts test.

Source. Ishii & Schmitt, 2009, p.10.

They piloted the test and found that the regularity in forming adverbs was problematic, and as deleting adverbs did not lower the reliability of their test, they decided to measure just three parts of speech. In the cases where more than one possible word form existed, the participants were required to write only one of the forms. Ishii and Schmitt found that the participants' mean number of derivatives was 17.61 out of a possible 45. Their learners scored 39.13% of derivatives correctly.

In summary, the reviewed literature shows that between derivative and inflective forms, derivations are prioritized. Therefore, as derivations are generally acquired after inflections (Berko, 1958; Ward & Chuenjundaeng, 2009), measuring derivations may reveal deeper word parts knowledge.

2.2. Form and Meaning

It is believed that form and meaning link is the basic dimension of vocabulary knowledge and most probably the first to be acquired (Laufer & Goldstein, 2004). That is, it is unlikely to have the knowledge of word parts of a group of words (target words) without the knowledge

of the corresponding form and meaning of the same words. For that reason, measuring form and meaning knowledge needs to accompany any measurement of other dimensions of vocabulary including word parts.

Nation (2013) states that learners consider a word known as soon as they know its form (what the word sounds like or looks like) and its meaning. However, he asserts that not only do language learners need to know the form and meaning of a word, but also they need to know how to connect these two. To measure form and meaning connection, Nation asks whether the learner can “produce the appropriate word form to express this meaning” (p.538), and for the purpose of testing this aspect specifically, he asks whether the learner can “translate these words into L2” (p. 551).

Laufer and Nation (1999) validated a vocabulary size test of controlled productive ability that is, in fact, a way of measuring form and meaning. For each item in their test, they presented a meaningful sentence context and the first letters of the target item. The first letters prevent the participants filling in another word which could be used semantically in the provided context. The test measures 18 items at 2,000, 3,000, 5,000, University Word List (Xue & Nation, 1984), and 10,000 word levels. The total mean scores of their high school 10th, 11th, 12th, and university English as a foreign language participants were 21.7, 33.4, 40.1, and 55.8 out of 90 (18 items x 5 word levels) respectively. Their participants’ mean scores for the 2,000 word level was 11.8 (65.55%), 15 (83.33%), 16.2 (90%), and 17 (94.44%) out of 18 (18 items).

Read (2000) believes that it is not clear what the blank-filling test (e.g. Productive Levels Test, Laufer & Nation, 1999) is measuring. In fact, it is assumed that the Productive Levels Test (PLT) measures controlled productive knowledge. However, Read (2000) believes that the PLT seems to assess the receptive knowledge rather than productive. In this regard, Webb (2008) argues that because it is a test of cued recall, the presence of the first few letters could be sufficient to recognize the target word receptively.

Following Laufer’s personal advice and taking into account Schmitt et al.’s findings, Llach and Espinosa (2014) combined the PLT Version A and Version C (equivalent versions) from the parallel Version 1 (Laufer & Nation, 1999) to have a 30 item test. Their 197 EFL Spanish participants’ mean scores of the 2,000 most frequent words were 7.26, 8.78, and 12.30 out of 30 in three successive grades (1st, 2nd, and 3rd grades of secondary education).

Zheng (2009) administered the VLT and PLT to 88 Chinese university students. The mean score of her participants for the 2,000 word level was 15.61/18 (86.74%), indicating that her participants seemed to have mastery over the words at 2,000 level.

3. RESEARCH QUESTIONS

To address the focus of the study, and considering the fact that the available participant pool for the current study was Iranian EFL students, the following research questions guided the design of the study:

1. To what extent do Iranian EFL students have productive knowledge of word parts at the 1,000 word frequency level?
2. To what extent do Iranian EFL students have productive knowledge of the corresponding form-meaning connection at the 1,000 word frequency level?

4. METHOD

4.1. Design (Word Parts)

Nation's (2013) framework targets both inflections and derivations for measuring word parts. Therefore, inspired by Schmitt's and Meara's (1997) study, a test with the following instructions and examples was considered the first attempt to measure both inflections and derivations (word parts).

Add any allowable letters to the beginning or end of the following words to change their meanings, uses or forms according to the provided examples. If no letter can be added to the beginning or end of the word, write NONE in front of the word.

For example:

| | ed, ing, s or es |
|---------|-------------------------------|
| BELIEVE | believed, believing, believes |
| MAKE | making, makes |
| ALREADY | NONE |

This method of measuring inflections, however, did not seem to be a truly productive test. In fact, it looked more like a receptive recall test (see Laufer & Goldstein, 2004) because the possible inflectional suffixes were provided, and the participants were asked to recall receptively which one matched the target words. Therefore, the test did not seem suitable to measure word parts productively.

For the purpose of measuring derivational forms, inspired by Pacheco's (2005) test instructions, the provided instructions of the aforementioned first attempt were rephrased and modified to the following:

Make all the new words by adding any letters (prefixes and suffixes) to the words given. Do **NOT** add *ed, ing, s* or *es*. If no letter can be added to the beginning or end of the word, write NONE in front of the word. Please write your answers as clearly as possible.

For example:

| | |
|---------|---|
| BELIEVE | believable; unbelievable; unbelievably; believer; unbeliever; disbelieve; disbeliever; disbelievingly |
| FRESH | freshness; fresher; freshest; freshly; freshen |
| ALREADY | NONE |

The test seemed to have the potential for measuring derivative forms; however, there were a few problems with the test. The test seemed complicated, time-consuming and tiring for the potential participants including the English L1 students. The evaluation of the test was also challenging and subjective because some items had only a few possible answers while other items had a wide range of answers. As a result, the form of the test did not seem to be appropriate for the purpose of this study.

The reviewed literature showed that between derivative and inflective forms, derivations should be prioritized, and it also seemed that measuring both derivations and inflections, as already discussed, was challenging and unfeasible. Therefore, for the purpose of the current study, it was decided to measure derivations only. Furthermore, as derivations are generally acquired after inflections (Berko, 1958; Ward & Chuenjundaeng, 2009), it seemed it was a better option to measure derivations in the hope of having a better picture of the knowledge of the word parts aspect.

An adaptation of Schmitt (1998, 1999) and especially Ishii and Schmitt (2009) could be an appropriate method to serve the purpose of measuring derivative forms productively. Schmitt's and Zimmerman's (2002) method of measuring derivations also seemed promising. However, considering the fact that at least 30 target words would be tested (the reason will be explained in 4.3), the time required to complete the test and consideration of participants' fatigue were the main reasons for not using an adaptation of Schmitt's and Zimmerman's method.

Participants were required to provide four different word classes for each of the target words. If there was more than one possibility, they only needed to provide one, and if there was no form, they needed to write X in the blank box. The following example illustrates this.

| | Noun | Verb | Adjective | Adverb |
|--------|-------------|-------------|------------------|---------------|
| NATION | nation | nationalize | national | nationally |
| PAY | payment | pay | payable | X |
| FRESH | freshness | freshen | fresh | freshly |

4.2. Design (Form and Meaning)

To measure form and meaning aspect productively, Nation's (2013) framework requires the translation of words into L2. Considering the potential participant pool for the current study, Persian to English word translation tests were to be provided. To design the form and meaning test, the Persian translations of the core meaning (usually the first entry in dictionaries) of each target word were provided. For this purpose, two English to Persian dictionaries were consulted. The author compared what he knew as the translation of the core meaning of words with the translations provided by the aforementioned dictionaries. If there

was an inconsistency between what he believed was the core meaning and what dictionaries showed as the core meanings, the COCA was consulted to judge which meaning was the core meaning. For this purpose, the first 100 examples of the corpus were consulted to select the most frequent meaning of the target word. To help guide the task and to prevent the participants from providing possible synonyms for the target words, letters of each target word presented by dashes and one letter in the target word (neither the beginning nor the ending, but a letter representing an English vowel sound in between) was provided.

After preparing the first draft, the test was checked with a Persian as a first language applied linguist. She provided one high frequency Persian synonym for the translation of four items for the test. These extra Persian words were close to the meaning of the provided translations, and seemed to be very helpful because, based on previous background knowledge, language learners may be more used to a specific translation. Finally, the test was piloted with a considerable number of Persian as a first language (Persian L1) graduate students to ensure that the items did not have potential synonyms.

4.3. Target Words

Brezina and Gablasova (2013) offered an objective approach to the development of the New General Service List (new-GSL) by means of examining frequent general words across a variety of language corpora. In addition, the new-GSL's lexical units are lemmas (a headword, its inflected and reduced forms, Nation, 2013). Giving the preference to lemmas enabled them to limit the wordlist to the most frequent words with greater precision in comparison to following the word family principle. Thus, the extraction of the most frequent 1,000 words from the new-GSL seems reasonable.

A To extract the first 1,000 words from the new-GSL (Brezina & Gablasova, 2013), all words with ranks more than 1,000 were removed from the list. Function words needed to be removed as well because they have little lexical meaning or may have even ambiguous meaning. Following removal of function words, the new-GSL came down to 779 words that served as the list of the most frequent words for the purpose of the current research study.

It was inferred, from Nation's (1983) and Schmitt's, Schmitt's and Clapham's (2001) experience with the number of items in the Vocabulary Levels Test, that 30 words could be statistically representative of a 1,000 word level. However, 34 words were selected for the test because at least 30 would remain even if pilot studies resulted in removing a few items.

To select the appropriate target words for the word parts test, two main criteria for each target word were considered:

- a. The target words were to have at least three or four word classes. The participants would get credit for mentioning a missing word class (either noun, verb, adjective, or adverb form) if the target word had just 3 word classes.

- b. The target words were to have no identical word classes. For example, “develop” is a verb and does not have any identical word classes.

Putting the aforementioned criteria into effect, the 779-word list was reduced to 190 words. Finally, these 190 words were randomized within an Excel spreadsheet and the first 34 words were selected as the target words. The selected target words for the word parts test also made up the target words for the form and meaning test.

4.4. Scoring the Word Parts Test

The scoring approach for the word parts test is mainly inspired by three studies: Schmitt (1998), Schmitt and Zimmerman (2002), and Ishii and Schmitt (2009). Based on the aforementioned studies, a norming list for the purpose of scoring the word parts test was created. In order to compile the answer norms in a principled manner, four sources of information were drawn on. First, Nation’s 1K base word list, published on his website, was consulted and all listed derivatives were extracted. Second, three online dictionaries were consulted and all other possible derivatives were added. Third, the presence of these derivatives in the BNC and COCA was double checked. If a form was not available in the corpora, it was removed. The frequency of these derivatives in the corpora was also checked. If a word class (especially adverbs) had low-frequency counts, both the derivative form and the answer “no form exists (X)” were accepted. Fourth, the test was given to 20 English L1 graduate students. If a new form, which was not in the norming list, was encountered, its availability and frequency in the BNC and COCA were checked. If it had a high frequency count, it was added to the norming list. If it had a low-frequency count, both that specific form and the answer “no form exists (X)” were accepted and added to the norming list. If the form was not in the corpus at all, it was not accepted.

Based on such an approach, as an example, the following answers as four word classes for the word MANAGER were accepted:

| | Noun | Verb | Adjective | Adverb |
|------------|---|----------------------|---|-----------------------------------|
| 1. MANAGER | manager; management; mismanagement; manageress | manage; mismanage | manageable; unmanageable; managed; managerial; managing; unmanaged | manageably; X; managerially |

Based on the norming list, the possible scores ranged from 0 (knowledge for no word class) up to 4 (knowledge for all four word classes) for each item. Misspellings were accepted (although they were not frequent) if the intended derivatives could be discerned, because the test measured knowledge of the derivations not spellings. In addition, considering the fact that

the original intent of the study was measuring word parts not grammatical functions, wrongly positioned derivatives were also accepted. For example, “manageable” was mistakenly considered an adverb form by a participant, yet it was accepted as a correct answer.

4.5. Pilot Studies

Ten English L1 graduate students were invited to take part in the pilot study. The instructions of the test were explained to them, and examples were reviewed. Care was taken to make sure they were comfortable with the metalinguistic notions and terms such as *noun*, *verb*, *adjective*, and *adverb*. The average time they spent on the test was 27 minutes.

The results of the pilot study showed that 81.69% of provided word classes were consistent with the derivatives extracted from Nation’s 1K base word list and the consulted dictionaries. There were some derivatives provided such as “solved, imagined, or reflected” for SOLUTION, IMAGINE, and REFLECT. The words were identified as adjectives in the Babylon online dictionary, and they were used in the corpuses (COCA or BNC) after linking verbs as well, which was proof of them being adjectives. Therefore, credit was given to such answers in a second round of fine-tuning the scores. As a result, the consistency of the participants’ performance with the provided norming list increased from 81.69% to 87.35%. In addition, three items with low rate of consistency with the norming list were removed from the test. The removal of the aforementioned items increased the consistency rate of the test from 87.35% to 89.19%.

Three Iranian graduate students in Canada were also invited to take part in the pilot study. The form-meaning test was administered first. They spent approximately 10 minutes on the test. The word parts test was briefly explained and administered afterwards. Care was taken to make sure participants were familiar with metalinguistic terms such as adjective and adverb. The test took approximately 40 minutes. All in all, the results of the pilot with 3 Iranian graduate students showed no glitch in the study. The tests, the scoring, and the procedures seemed appropriate, and the study was ready to be conducted.

4.6. Main Study

4.6.1. Participants

The study was conducted with 40 graduate and 6 undergraduate students from 2 Iranian universities. Their first language was Persian, and English was their foreign language. They were enrolled in a private language center which was specialized in training students for academic purposes to prepare them for TOEFL and IELTS exams. They were intending to get an admission from a North American university.

Their mean score on the Vocabulary Levels Test (VLT, Schmitt, Schmitt, & Clapham, 2001) at the 2nd 1,000 level was 29 (96.82%). Their mean score on a modified version (Llach

& Espinosa, 2014) of the original Productive VLT (PLT, Laufer & Nation, 1999) at the 2nd 1,000 level was 26.43 (88.11%). All participants that remained in the data pool had answered correctly more than 86.66% (the criterion mastery, Schmitt et al. 2001) of the items of the VLT and PLT at the 2,000 word level.

4.6.2 Procedure

The form-meaning test, the PLT, and the VLT were administered in 10, 20, and 15 minutes respectively. Before each test, the test instructions were explained to ensure that the participants were comfortable with the task. The participants were informed that they could spend as long as needed on the tests. The word parts test was administered afterwards and it took approximately half an hour. The suggested time for the test was 40 minutes; however, they could spend as long as required on the test. The instructions of the test were explained before its administration and it was ensured that the participants were familiar with metalinguistic words such as noun, verb, adjective, and adverb.

5. RESULTS

5.1. Form-Meaning Test

Table 1 shows the descriptive statistics of the participants' performance on the form-meaning test. The table shows that participants scored 27.37/31 in the form-meaning test. Based on Schmitt et al. (2001) criterion mastery, the participants had strong form-meaning knowledge of the target words. Schmitt et al. (2001) considered 86.66% (at least 26 items out of 30) to be the criterion for mastery of the word levels of the VLT.

Table 1. Descriptive Statistics of the Performance on the Form-Meaning Test

| | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
|------------|----------|----------------|----------------|-------------|-----------------------|-----------------|
| FMT | 46 | 17.00 | 31.00 | 27.37 | 3.00 | 9.00 |

Note. FMT: Form-Meaning Test. The mean score is from a total number of 31 items.

The mean percentage of the performance may also provide a better understanding of the result. Table 2 shows the descriptive statistics of the percentage of performance on the form-meaning test. The table shows 88.27% performance on the form-meaning test.

Table 2. Descriptive Statistics of the Percentage of Performance on the Form-Meaning Test

| | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
|------------|----------|----------------|----------------|-------------|-----------------------|-----------------|
| FMT | 46 | 54.83 | 100.00 | 88.27 | 9.68 | 93.62 |

Note. FMT: Form-Meaning Test. The number of items is 31; however, the table shows the mean percentage of the performance.

5.2. Word Parts Test

Table 3 shows the descriptive statistics of the participants' performance on the word parts test. The table shows the total mean score on the word parts test and the mean scores for each word class—noun, verb, adjective, and adverb. The descriptive statistics show that verbs had the highest mean score (26.30/31) followed by nouns (24.83/31), adjectives (23.65/31), and finally adverbs (19.80/31). The table also shows a total mean score of 94.59 out of 124, indicating that while approximately three-fourths (94.59/124) of all possible word classes were provided, the total mean score was considerably lower than the maximum possible score (124/124).

Table 3. Descriptive Statistics of the Performance on the Word Parts Test

| | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
|-------------------|----------|----------------|----------------|-------------|-----------------------|-----------------|
| Noun | 46 | 13.00 | 31.00 | 24.83 | 3.55 | 12.59 |
| Verb | 46 | 20.00 | 31.00 | 26.30 | 2.62 | 6.88 |
| Adjective | 46 | 10.00 | 31.00 | 23.65 | 5.18 | 26.81 |
| Adverb | 46 | 3.00 | 30.00 | 19.80 | 6.74 | 45.45 |
| Total Mean | 46 | 47.00 | 118.00 | 94.59 | 14.95 | 223.63 |

Note. The mean scores for noun, verb, adjective, and adverb are from a total of 31 while the total mean score is from a total of 124.

A repeated measures ANOVA was run to examine whether the differences between the means of the word classes (noun, verb, adjective, and adverb) were statistically significant. Table 4 shows the results. The table shows the differences between the means were statistically significant, $F(3,135)=29.81, p < .001$.

Table 4. Repeated Measures ANOVA on the Mean Scores of the Word Classes

| Source | | Type III Sum of Squares | df | Mean Square | F | Sig. |
|------------------|--------------------|--------------------------------|-----------|--------------------|----------|-------------|
| Wordparts | Sphericity | 1068.016 | 3 | 356.005 | 29.810 | .000 |
| | Assumed | 1068.016 | 1.979 | 539.738 | 29.810 | .000 |
| | Greenhouse-Geisser | 1068.016 | 2.069 | 516.128 | 29.810 | .000 |
| | Huynh-Feldt | 1068.016 | 1.000 | 1068.016 | 29.810 | .000 |
| | Lower-bound | | | | | |

| | | | | |
|-------------------------|--------------------|-----------|--------|--------|
| Error(Wordparts) | Sphericity | 1612.234 | 135 | 11.942 |
| | Assumed | 1612.234 | 89.044 | 18.106 |
| | Greenhouse-Geisser | 1612.234 | 93.118 | 17.314 |
| | Huynh-Feldt | 16.12.234 | 45.000 | 35.827 |
| | Lower-bound | | | |

Bonferroni Post-hoc comparisons were run to examine which mean difference was statistically significant. Table 5 shows the results. The table shows that while the performance on verbs was significantly better than the performance on other word classes, the performance on adverbs was significantly worse than the performance on other word classes. There was no significant difference between performance on nouns and adjectives.

Table 5. Results of Bonferroni Post-hoc Comparisons

| (I) Word_parts | | Mean Difference (I-J) | Std. Error | Sig. ^b | 95% Confidence Interval for Difference ^b | |
|----------------|-----------|-----------------------|------------|-------------------|---|-------------|
| | | | | | Lower Bound | Upper Bound |
| Noun | Verb | -1.478* | 0.410 | 0.005 | -2.609 | -0.347 |
| | Adjective | 1.087 | 0.509 | 0.229 | -0.317 | 2.491 |
| | Adverb | 4.978* | 0.908 | 0.000 | 2.472 | 7.485 |
| Verb | Adjective | 2.565* | 0.641 | 0.001 | 0.797 | 4.333 |
| | Adverb | 6.457* | 0.836 | 0.000 | 4.150 | 8.763 |
| Adjective | Adverb | 3.891* | 0.824 | 0.000 | 1.618 | 6.165 |

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

Further analysis was conducted to investigate the number of items provided with 4, 3, 2, 1, and 0 correct word classes. Table 6 shows the descriptive statistics of the number of items provided by the participants with word classes consistent with the norming list. The table shows that on average 12.02 items (from a total of 31) were provided with 4 word classes consistent with the norming list while the numbers of items provided with 3, 2, 1, and 0 word class(es) were 10.83, 6.17, 1.72, and .26 respectively. This indicated that the majority of the items were not provided with all 4 word classes which was not very impressive for advanced students at university level.

Table 6

Descriptive Statistics of the Number of Items Provided with Word Classes Consistent with the Norming list

| | N | Minimum | Maximum | Mean | Std. Deviation | Variance |
|-----------------------|----|---------|---------|-------|----------------|----------|
| 4 word classes | 46 | .00 | 25.00 | 12.02 | 6.66 | 44.38 |
| 3 word classes | 46 | 1.00 | 26.00 | 10.83 | 4.20 | 17.61 |
| 2 word classes | 46 | .00 | 16.00 | 6.17 | 4.20 | 17.66 |
| 1 word class | 46 | .00 | 9.00 | 1.72 | 2.41 | 5.81 |
| 0 word class | 46 | .00 | 5.00 | .26 | .83 | .69 |

Note. The mean scores are from a total number of 31 items.

More analysis was conducted to examine the extent to which the participants were able to provide the word classes of the items. The analysis showed that while the total minimum mean score for provided word classes was 1.83, the maximum total mean score was 3.59. The analysis also showed that on average 3.05 word classes out of 4 were known and provided by the participants, indicating that the participants tended to know 3 word classes. If the results of the ANOVA test are taken into consideration, the participants seem to have more problem with providing adverbs.

6. DISCUSSION

6.1 To what extent do Iranian EFL students have productive knowledge of form and meaning at the 1,000 word frequency level?

The results show that the participants' performance on the form-meaning test reached Schmitt et al.'s (2001) criterion of mastery as the mean percentage on the form-meaning test was 88.27% (27.37/31). Laufer and Nation (1999) assert that 85% to 90% performance is satisfactory for the 2,000 word level of the productive VLT. The same percentage considered by Laufer and Nation would likely be satisfactory for the first 1,000 word level as well. The main reason could be the fact that the same approximate percentage is also considered the criterion mastery of the Vocabulary Levels Test (Schmitt et al., 2001). Schmitt et al. considered 86.66% (26 items out of 30) to be the criterion mastery of the word levels of the VLT.

To the best of my knowledge, no study has specifically measured productive form-meaning knowledge of the most frequent 1,000 words to date; however, the current result of the form-meaning test is consistent with what Laufer and Nation (1999) and Zheng (2009) found for the 2,000 word level. Zheng (2009) used Laufer's and Nation's (1999) Productive Levels Test (PLT) for measuring the word levels. The total mean scores of their university

student participants were 17 out of 18 (94.44%) and 15.61 out of 18 (86.72%) respectively at the 2,000 word level. The findings of the current study, Laufer's and Nation's study, and Zheng's study show that the form-meaning of the most frequent words (either the first or the second 1,000 words or both) may often be known at university level. It should also be noticed that a modified version of the PLT was used in the current study as a diagnostic test to ensure the participants had at least the productive form-meaning knowledge of the words at 2,000 level. The mean percentage of the PLT in the current study was 88.11, which reconfirms the fact that the productive form-meaning of the most frequent words, as discussed before, seems to be known by the university-level participants in this research.

6.2. To what extent do Iranian EFL students have productive knowledge of word parts at the 1,000 word frequency level?

The study shows 76.27% of responses on the word parts test were consistent with the norming list. The participants did not provide approximately one-quarter of the word classes. This finding shows that participants may not necessarily be able to produce derivatives of all the words at the 1,000 level that have 3 or 4 possible derivative forms. Considering the great coverage and range of the first 1,000 words, this is not a very impressive finding for advanced students at a university level.

Twelve out of 31 items (an approximate mean number of items) were provided with all word classes consistent with the norming list. Thus, maximum performance occurred for just over one-third of the target words (12/31). This reveals that advanced participants at the university level do not seem to be able to produce all word classes for the items at 1,000 level with 3 to 4 derivatives. Schmitt (1999) believes that this suggests L2 learners may be unable to master all members of a word family until relatively late in the acquisition process, and even this may not happen at all.

Approximately 11, 6, and 2 items out of 31 (an approximate mean number of items) were provided with 3, 2, and 1 derivative form(s) consistent with the norming list. Yet, the total mean number of derivatives provided by the participants is 3 out of 4, indicating that the participants can produce approximately three of the four possible derivatives. Adverbs have a lower mean score compared to the other 3 word classes, indicating that the participants tend to have more difficulty with providing adverbs.

The results also show that the verb derivatives were the best known with 84.84% produced. Nouns were next with 80% produced. Adjectives and adverbs were less known, 76% and 63.88% respectively. This finding is consistent with Schmitt's (1998), and Schmitt's and Zimmerman's (2002) findings. They also found that nouns and verbs were the best known while adjectives and adverbs appeared to be less known. Schmitt (1998) believes that this suggests that adjective and adverb forms are unlikely to be learned from general exposure, maybe due to their lower frequency of occurrence, and hence might be good candidates for explicit instruction. In other words, nouns and verbs are the most common parts of speech

found in natural text (Kucera & Francis, 1967; Webb, 2005, 2007), and that may cause more exposure to them which may result in learning them better.

Knowledge of word parts was also found to be relatively weak in previous studies discussed in Literature Review. Performance on the test of word parts in the current study was higher than in the earlier studies; however, it should be noted that the aforementioned studies on word parts evaluated target words from lower levels of frequency than 1K (even words from 4K and 6K). In other words, the lower performance on the word parts of the reviewed studies may have been the result of performing on lower frequency target words. In addition, lack of knowledge or limited knowledge of words in lower levels of frequency than 1K may not necessarily make a significant difference for participants because the coverage and range of such word levels are considerably less than the 1K level. Therefore, the findings of the current study, with the consideration of the great coverage and range of 1K, seem to be consistent with the findings of the previous studies.

7. CONCLUSION

The most frequent 1,000 word families comprise up to 81% of written text and 85% of spoken text (Nation, 2006). This can be at least 80 percent of English that an English learner may receive from listening and reading or produce when speaking or writing. For this reason, when it comes to evaluating learners' vocabulary knowledge of the most frequent 1,000 words, every identified weakness merits attention. The most frequent words are so basic and fundamental components of English that they are usually assumed to be very well known by advanced students who have learned English for years both at the public educational system and private language schools. However, it is a real matter of concern that advanced students do not show strong derivational knowledge of such words. While there can be many possible reasons for this limited knowledge, a lack of awareness of the aspects involved in vocabulary knowledge including derivational knowledge and possibly greater focus on receptive input rather than productive use of the language likely play a large role.

It is worthwhile to note that higher scores in knowledge of form and meaning and lower scores in word parts do not seem to develop equally and in tandem. In fact, the form and meaning is baseline knowledge needed in learning vocabulary and other aspects such as word parts may develop incrementally (Schmitt, 2010) with repeated exposures to words (see Webb, 2007) in varied context of use. However, the findings of the current study highlight the fact that learning of the latter aspect may require more focused instruction.

The research suggests that it is still necessary to highlight terms such as word parts, especially for advanced learners. Furthermore, instructional emphasis on productive outcome of language in general and productive derivational knowledge (productive mode) in particular need to be highlighted. It seems it is necessary to highlight that the language received from

listening and readings needs to be produced in speech and writing, and for this purpose, productive aspects of vocabulary knowledge need to be emphasized.

8. IMPLICATIONS OF THE STUDY

Traditionally, knowledge of a word refers to the meaning and form of that word (Schmitt, 2008). However, the current study can raise awareness of an important aspect involved in vocabulary knowledge other than the link between form and meaning.

The study targeted productive (rather than receptive) knowledge of the most frequent words. To the best of my knowledge, no study, to date, has measured the derivational productive knowledge of the most frequent words with the same detailed and purposeful methodology. In contrast to the many studies investigating receptive vocabulary knowledge, relatively few studies that investigate productive vocabulary knowledge can be found. The current study not only enriches the available literature but also sheds light on the significance of measuring productive vocabulary knowledge. Furthermore, there are crucial pedagogical implications of the current study.

Teachers may help learners through the use of different teaching methods or learning activities to overcome their weaknesses in their knowledge of word parts. Among all available activities and procedures that improve students' derivational knowledge, Nation's (2013) suggestions are particularly helpful and practical.

Nation (2013) classified activities for vocabulary learning according to the different aspects of knowing a word. Rich instruction involves using several of activities to enrich students' vocabulary knowledge. The following is an explanation of activities that may be useful to further develop knowledge of word parts.

Filling word part tables. Learners are required to work in pairs to complete a table like the following. All spaces may not necessarily be filled. Learners check their answers with one another before the teacher provides the answer.

| noun | verb | adjective | adverb |
|----------|----------|-----------|----------|
| argument | | | |
| | evaluate | | |
| | | distinct | |
| | | | normally |

Cutting up complex words. Learners are provided with a list of words that they divide into parts. They can also be required to give the meaning of some of the parts. For example, the teacher may ask learners to divide *effortless*, *disappointed*, *misuse*, *disrespectful* into parts. Learners can also be asked to give the meaning of word parts such as *-less*, *dis-*, *mis-*, *-ful*.

Building complex words. Learners are provided with word stems and are required to make words (using e.g., *mis-*, *dis-*, *un-*) from them. For example, words such as *use*, *respect*,

employed are provided and learners are asked to make their negatives such as *misuse*, *disrespect*, *unemployed*.

Choosing the correct form. Learners are provided with sentences containing a blank and a word stem in brackets. They are required to change the stem to the appropriate inflectional and derivational form to fill in the blank. The following is an example.

I went to the doctor for a _____ (consult).

NOTES

¹ The English inflectional affixes are all suffixes including plural *-s*, *-ed*, *-ing*, third person singular *-s*, possessive *-s*, comparative *-er*, and superlative *-est*. Derivational affixes in English include prefixes and suffixes that may change the part of speech of the words, such as *happy* and *happiness*, or *able* and *enable*.

REFERENCES

- Angelova, M. & Riazantseva, A. (1999). If you don't tell me, how can I know?: A case study of four international students learning to write the US way. *Writing Communication*, 16(4), 491-525.
- Berko, J. (1958). The child's learning of English morphology. *Word*, 14, 150-177.
- Brezina, V. & Gablasova, D. (2013). Is there a core general vocabulary? Introducing the New General List. *Applied Linguistics*, 1-23.
- Evans, S. & Green, C. (2007). Why EFL is necessary: A survey of Hong Kong tertiary students. *Journal of English for Academic Purposes*, 6, 3-17.
- Gan, Z. (2012). Understanding L2 Speaking Problems: Implications for ESL Curriculum Development in a Teacher Training Institution in Hong Kong. *Australian Journal of Teacher Education*, 37(1), 43-59.
- Ishii, T. & Schmitt, N. (2009). Developing an integrated diagnostic test of vocabulary size and depth. *RELC*, 40(1), 5-22.
- Kucera, H. & Francis, W. N. (1967). Computational analysis of present-day American English. Providence, RI: Brown University Press.
- Laufer, B. & Nation, P. (1999). A Vocabulary Size Test of Controlled Productive Ability. *Language Testing*, 16(1), 36-55.
- Laufer, B., Elder, C., Hill, K. & Congdon, P. (2004). Size and strength: do we need both to measure vocabulary knowledge? *Language Testing*, 21(2), 202-226.
- Laufer, B. & Goldstein, Z. (2004). Testing vocabulary knowledge: size, strength, and computer adaptiveness. *Language Learning*, 54(3), 399-436.
- Liu, M. & Jackson, J. (2008). An exploration of Chinese EFL learners' unwillingness to communicate and foreign language anxiety. *Modern Language Journal*, 92, 71-86.
- Llach, M. P. A. & Espinosa, S. M. (2014). Knowledge of form and meaning. In J. Milton & T. Fitzpatrick (Eds.), *Dimensions of vocabulary knowledge* (pp. 60-72). UK: Palgrave Macmillan.

- Meara, P. & Jones, G. (1988). Vocabulary size as a placement indicator. In P. Grunwell (Ed.), *Applied linguistics in society* (pp. 80-87). London: CILT.
- Meara, P. (1996). The dimensions of lexical competence. In G. Brown, K. Malmkjaer & J. Williams (Eds.), *Performance and competence in second language acquisition* (pp. 35-53). Cambridge: Cambridge University Press.
- Milton J., Wade, J. & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse & M. Torreblanca-López (Eds.), *Further insights into non-native vocabulary teaching and learning* (pp. 83-98). Bristol: Multilingual Matters.
- Milton, J. (2013). Measuring the contribution of vocabulary knowledge to proficiency in the four skills. In C. Bardel, C. Lindquist & L. Laufer (Eds.), *L2 vocabulary acquisition, knowledge and use new perspectives on assessment and corpus analysis* (pp. 57-78). European Second Language Association.
- Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, 5, 12–25.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. New York: Newbury House.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. UK: Cambridge University Press.
- Nation, I. S. P. (2006). How large a vocabulary is needed for reading and listening? *Canadian Modern Language Review*, 63(1), 59-82.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd edition). UK: Cambridge University Press.
- Nyssonen, L. (2008). *Understanding English word formation: a study among 6th-grade pupils in Finish comprehensive school* (Unpublished MA Thesis). University of Jyväskylä, Finland.
- Pacheco, E. C. R. (2005). *Affixes as a strategy for vocabulary acquisition in a first year ESL college reading course* (Unpublished doctoral dissertation). University of Puerto Rico, Puerto Rico.
- Read, J. (2000). *Assessing vocabulary*. Cambridge: Cambridge University Press.
- Schmitt, N. & Meara, P. (1997). Researching vocabulary through a word knowledge framework: Word associations and verbal suffixes. *SSLA*, 20, 17-36.
- Schmitt, N. (1998). Tracking the incremental acquisition of second language vocabulary: A longitudinal study. *Language Learning*, 48(2), 281-317.
- Schmitt, N. (1999). The relationship between TOEFL vocabulary items and meaning, association, collocation and word class knowledge. *Language Testing*, 16, 189-216.
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the Vocabulary Levels Test. *Language Testing*, 18(1), 55–88.
- Schmitt, N. & Zimmerman, C. B. (2002). Derivative word forms: What do learners know? *TESOL Quarterly*, 36(2), 145-171.
- Schmitt, N. (2008). Review article, Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329-363.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. UK: Palgrave Macmillan.
- Schoonen, R. (2010). *The development of lexical proficiency knowledge and skill*. Paper presented at the Copenhagen Symposium on Approaches to the Lexicon, Copenhagen.
- Stæhr, L. S. (2008). Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36(2), 139-152.

- Ward, J. & Chuenjundaeng, J. (2009). Suffix knowledge: Acquisition and applications. *System*, 37(3), 461-469.
- Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, 27(1), 33-52.
- Webb, S. (2007). The effects of repetition on vocabulary knowledge. *Applied Linguistics*, 28(1), 46-65.
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30, 79-95.
- Webb, S. & Chang, A. C. (2012). Second language vocabulary growth. *RELC Journal*, 43(1), 113 – 126.
- Xue, G. & Nation, I.S.P. (1984). A university word list. *Language Learning and Communication*, 3(2), 215-229.
- Zheng, Y. (2009). Exploring Chinese EFL learners' receptive and productive vocabulary knowledge: Implications for EFL vocabulary teaching. *The Journal of Asia TEFL*, 6(1), 163-188.