Paper:
http://dx.doi.org/10.7575/aiac.ijalel.v.4n.4p.192

Distributed under the terms of a Creative Commons Attribution (CC-BY-4.0)
Abstract

The present study investigates the relationship between the morphological processing of regular and irregular words and second language (L2) vocabulary acquisition. In considering that monolingual Arabic speakers derive a large number of new words from roots by leaning heavily on the regularity of rules in Arabic (Bar & Dershowitz, 2012; Habash, 2010), they are expected to experience difficulty when developing the lexicon of a language with less regular rules, such as English. To examine this assumption empirically, data were collected by administering an English receptive vocabulary knowledge test that included 100 regular and irregular inflected and derived words to 450 Arabic English as a foreign language (EFL) learners from schools in Saudi Arabia. The test also included pseudowords (non-words) to act as ‘gatekeepers’ against the possibility of guessing. The t-test results revealed a non-significant difference in learners’ uptake of L2 words, whether regular or irregular. However, the study indicates that word frequency plays a statistically significant role in learning L2 vocabulary that is irregularly inflected or derived. The frequency effect on irregular word morphology found in this study is in line with the approach of a dual-route mechanism. This approach suggests that irregular words are not rule-based, so are stored in the mental lexicon as full entries, whereas regular words are not. This article suggests that introducing rules for deriving new words from base forms to EFL learners in the early stages of learning would be very useful for L2 vocabulary development.

Keywords: morphological processing, vocabulary acquisition, frequency, Arabic, mental lexicon

1. Introduction

Understanding the way morphological units are stored in second language (L2) learners’ mental lexicon and how they are retrieved through the use of language is an important area of investigation. As Chomsky (1957) has asserted in his ‘generative grammar’, having a knowledge of a finite set of rules leads to being able to produce an indefinite number of new words, and this concept can be translated into knowing a limited set of morphological rules and being able to generate many new lexical items without having to store a large amount of vocabulary in one’s memory. The question that is addressed in this study is how native Arabic speakers deal with learning the processes of regular and irregular L2 morphological structure. Arabic is a morphologically rich language and native Arabic speakers can derive a vast number of words from the same root (Bar & Dershowitz, 2012; Habash, 2010). Using the same root with different patterns, Arabic speakers can generate words with several different meanings. Therefore, because of this capacity for much greater rule-based generation of words, native Arabic speakers can build a large first language (L1) mental lexicon from a relatively small number of roots. When learning an L2, native Arabic learners might utilise this extensive rule generalisation in learning a language that is highly rule-based. Conversely, if the L2 is not very rule-based, the learning task for Arab learners will be somewhat daunting. The English language is the target language investigated in this study. Unlike Arabic, English has fewer derived words and a greater tendency to create new words or borrows them from other languages, e.g., Latin.

As native Arabic speakers have the capacity through rule-based learning to generate large numbers of words from roots in their L1, they may try to implement this skill when learning English vocabulary. This might result in the more effective learning of words which are regular in English than those which are not. The current study thus seeks to address this assumption and tackle one of the key issues pertinent to the reportedly poor English vocabulary knowledge of Saudi school learners (Al-Bogami, 1995; Al-Hazemi, 1993; Alsaif, 2011). However, before demonstrating the study design, a brief background of how regular and irregular words are processed in L1 and L2 will be presented in the following section.
In second language acquisition (SLA) research there is a long-standing debate in the literature on word recognition. The controversy lies in whether native speakers of a language process irregular (e.g., write-wrote) and regular (e.g., talk-talked) verb forms by a single or dual-route mechanism. One school of thought advocates a single processing system, claiming that words which are similar in form tend to be similar in meaning (Basnight-Brown, Chen, Hua, Kostić & Feldman, 2007; Li, 2006; McClelland & Elman, 1986; Rueckl, Mikolinski, Raveh, Miner & Mars, 1997; Rueckl & Raveh, 1999; Rumelhart & McClelland, 1986; Seidenberg & Elman, 1999). The other line of research supports the dual-route mechanism, which argues that recognition of a regular past tense verb utilises a rule-based process where an -ed affix is attached to the present tense to form the past tense verb. English, however, has about 180 verbs that do not follow this rule-based process but are created in an unpredictable fashion. Consequently, those irregular verbs differ in their orthographic and phonological forms. As irregular verbs are not formed by a rule, they should be stored in rote memory (Pinker, 1991).

In more recent studies (e.g., Pinker, 1999; Ullman, 2004), it is suggested that irregular verbs are stored as separate entries in the mental lexicon and retrieved for use from associative memory. In contrast, regular verbs are more likely to use a computational method in the mental lexicon to create past tense forms. This implies that the base word is stored in the lexicon and affixes, such as -ing, -s and -ed, are added to form inflected word forms when needed (Ullman, 2000). A significant number of studies appear to support the dual-route mechanism, whereby regular English verbs are controlled by rule-based processing and irregular verbs occupy separate entries in the mental lexicon and are retrieved from memory.

One of the most influential pieces of evidence to support a dual system originates from brain-damaged aphasic patients. A pioneer study conducted by Marslen-Wilson and Tyler (1997) examined brain-damaged aphasic patients using auditory priming tasks, including regular and irregular past tense forms. Marslen-Wilson and Tyler reported that two of the three aphasic patients were found to have a problem with processing irregular verb forms, while the other patient had problems with only regular forms. This result of double dissociation was interpreted as suggestive of the dual nature of past tense inflection in English (Platisikas & Marinis, 2013).

Another piece of supporting evidence for a dual-route mechanism comes from an investigation of the effect of frequency on processing regular and irregular past tense inflection. Alegre and Gordon (1999) reported a number of studies that show that when irregularly inflected words are more frequent in a language usage they are produced faster than when they are less frequent. This kind of pattern, however, was not found with regularly inflected words. This finding was interpreted as supporting evidence that regularly inflected forms are processed online. The frequency effect seems only applicable to regular verbs which are ranked very high in the frequency scale. In contrast, irregular forms are stored as full-form entries in the mental lexicon and word frequency was found to be central to their recognition (Prado & Ullman, 2009).

The first part of this section sheds some light on the processing of L1 regular and irregular words. In this part, a brief discussion of processing L2 regular and irregular inflection will be provided. Complete proficiency in L2 is always a matter of debate. According to some scholars (e.g., Grosjean, 1998; Johnson & Newport, 1989), there are a number of factors which have been suggested as influencing the acquisition of an L2. Those sorts of factors raise the question of whether L2 learners can achieve the automated processing that underlies rule application in an L1 or fall short of attaining that level of automatization. One of the factors suggested by L2 acquisition research to play an influential role in acquiring an L2 is age. For instance, if the age of a learner when learning an L2 is beyond a critical period, the automatisation of L2 grammar is claimed to be less successful (Bialystok, 1997; Butler & Hakuta, 2004; Platisikas & Marinis, 2013).

Corresponding to this idea, L2 learners who begin learning an L2 at the age of 11 or above are more likely to experience difficulty mastering regular inflection rules. Ullman (2001b) presented a model concerning the acquisition of grammar rules by L2 learners compared to L1 learners. In his model, Ullman used neurocognitive data to argue that if L2 is acquired after puberty, rule implementation becomes increasingly hard and, therefore, adult L2 learners face limitations in its application. In terms of inflection, the model suggests that the past tense rule will be absent in L2 learning. Consequently, L2 learners might be incapable of decomposing regularly inflected forms into their elements, but they should be able to memorise them as separate lexical entries, as is the case with irregular forms (Ullman, 2001b). Ullman, however, later proposed that it is possible for L2 learners to have access to L2 rules as a result of L2 learning experience (Ullman, 2004).

The discussion reported above on L2 learners’ processing of inflected forms appears inconsistent, although support for dual-route mechanism is stronger than single-rout. This study, however, aims to explore one of the potential causes of the poor English vocabulary uptake by Saudi school learners referred to earlier by investigating the role of word regularity in building a larger L2 lexicon. The processes of inflectional forms and the ability to derive new words from base words in English can cause a problem for Arabic learners learning English vocabulary. Studies report low vocabulary knowledge by Saudi school learners but no clear explanation is found as to why those learners learn, on average, 1,000 English words after seven years of classroom input. Only two studies, to the best of the researchers’ knowledge, have suggested that the nature of the vocabulary presented in learners’ textbooks (Alsaif & Milton, 2012) and lack motivation (Al-Akloby, 2001) might explain the humble uptake of English vocabulary. This study, therefore, hopes to bring more evidence to elucidate the cause of poor vocabulary uptake by Saudi school learners.
1.2 Aims and objectives

As the available studies on vocabulary research in Saudi Arabia have not yet provided a clear picture of why Saudi learners experience difficulty learning English vocabulary, the present study aims to contribute to the field of vocabulary acquisition by providing evidence pertinent to this matter. The paper will attempt to find answers to the following questions:

1. Do native Arab learners of English learn regularly inflected and derived forms more easily than irregularly inflected and derived forms?
2. Does word frequency influence learning regular or irregular words, or perhaps both?

2. Method

2.1 Participants

The participants included in the study were 450 male students from two high schools in Saudi Arabia and in three high school levels, years 10, 11 and 12. They were low-level native Arabic learners of EFL. The data were collected in the second semester of the academic year 2012/2013. All the students who were available were tested. Their ages varied between 16 and 18 (M = 17, SD = 1.21), and the only known source of English they had was classroom input. The participants in this study had attended, on average, 576 EFL classroom hours in year 10, 704 in year 11, and 832 in year 12.

2.2 Material

A receptive vocabulary recognition test devised by the researchers, comprising 100 real words and 20 pseudowords (non-words), was administered in this study. Fifty of the words were regularly inflected and had derived forms (21 inflected and 29 derived forms) and 50 were of an irregular form. Pseudowords were included in test to act as gatekeepers to guard against the possibility of the participants guessing. The test was in Yes-No format, where participants had to check the words they knew. This type of testing remains one of the better-known measures of receptive vocabulary knowledge and is considered reliable and valid (Pellicer-Sánchez & Schmitt, 2012).

The lists of regular and irregular forms, however, were carefully selected based on certain criteria: (a) each pair of regular and irregular forms was very similar in terms of frequency, so frequency was not biased; (b) each pair of words was also tested for repetition across learners’ textbooks; (c) common inflectional and derivational endings were also observed for each pair. Correlation analysis was conducted to ensure that bias between the two groups of words was, as far as possible, eliminated. Table 1 shows the correlation statistics for word frequency and repetition across learners’ textbooks for both the regular and irregular lists of words.

Table 1. Correlations between regular and irregular words in terms of frequency (BNC_list) and repetition (textbooks)

<table>
<thead>
<tr>
<th></th>
<th>Word frequency regular</th>
<th>Word frequency irregular</th>
<th>Word repetition regular</th>
<th>Word repetition irregular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word frequency regular</td>
<td>-</td>
<td>0.969**</td>
<td>0.365**</td>
<td>0.419**</td>
</tr>
<tr>
<td>Word frequency irregular</td>
<td>-</td>
<td>-</td>
<td>0.327*</td>
<td>0.433**</td>
</tr>
<tr>
<td>Word repetition regular</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.906**</td>
</tr>
<tr>
<td>Word repetition irregular</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

*Correlation is significant at the 0.01 level (2-tailed); * correlation is significant at the 0.05 level (2-tailed).

Note: Each category contains 50 words.

2.3 Procedures

The receptive vocabulary test used in this study was administered to the participants by the researchers with assistance from volunteer teachers from the two schools where the data were collected. The test was given to the informants on the same morning at both schools. Before the administration of the test, participants were provided with clear oral instructions of the purpose of the task. They were also asked to write their age on the test sheet and details of extra exposure to the English language outside the classroom, if any. As the test was designed to be user-friendly, it only took participants around 20 minutes to complete it.

2.3 Test scoring system

After collecting the data, all responses were marked manually and imported into the SPSS software package (version 20) for analysis. However, as the test also included pseudowords, to act as gatekeepers against guesswork by the participants, it was decided to exclude responses with more than 30% false alarms (items marked as known but which
are not real words). This threshold is not theoretically supported but we believed it to be a plausible figure. By applying this norm, however, the number of processed responses was reduced to 416 (34 were excluded).

3. Results

3.1 Knowledge of regular forms VS irregular forms

The data analysis presented in Table 2 shows the paired sample statistics for the two types of words tested in the current study. As can be seen from the table, the mean scores for the 416 participants are almost equal, although slightly higher for the regular forms. This indicates that learners’ responses showed no preference for either type of word tested. However, the rate of learners’ responses to both regular and irregular forms was generally very low, about one fifth of the words presented to them. The number of words presented to the test-takers in each category was 50 but, on average, the respondents recognised approximately 10. The significance analysis ($t$-test) is presented for clarification in Table 3.

<table>
<thead>
<tr>
<th>Word type</th>
<th>$M$</th>
<th>$N$</th>
<th>$SD$</th>
<th>$SD.E.M$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular words - responses</td>
<td>10.56</td>
<td>416</td>
<td>8.04</td>
<td>0.39</td>
</tr>
<tr>
<td>Irregular words - responses</td>
<td>10.37</td>
<td>416</td>
<td>7.19</td>
<td>0.35</td>
</tr>
</tbody>
</table>

Note: $N$ = number of participants; $M$ = mean; $SD$ = standard deviation; $SD.E.M$ = standard deviation error of mean.

Table 3. Paired sample test ($t$-scores) comparing the mean scores of regular and irregular forms

<table>
<thead>
<tr>
<th>Paired differences</th>
<th>$M$</th>
<th>$SD$</th>
<th>$SD.E.M$</th>
<th>95% CI</th>
<th>$T$</th>
<th>$df$</th>
<th>$P$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regular and irregular</td>
<td>0.19</td>
<td>5.02</td>
<td>0.23</td>
<td>-0.29128</td>
<td>0.67590</td>
<td>0.782</td>
<td>415</td>
</tr>
</tbody>
</table>

Note: † = standard deviation of mean error; CI = confidence interval; LL = lower limit; UL = upper limit.

As revealed in Table 3, it appears, at least for the learners taking part in this study, that their knowledge of regular and irregular English forms was almost the same. The difference between the means is not significant: $t$ (415) = 0.782, $p < 0.435$. This finding does not seem to coincide with our expectation that native Arabic learners might show better receptive knowledge of words which are regularly inflected and derived. Reasons why Arabic learners appear not to have used their skill in generating new forms of base words in L2 as they regularly do in their L1 will be provided in the discussion section.

3.2 Frequency effect on knowledge of regular and irregular word forms

One of the current study questions was whether frequency had a significant impact on word type uptake. The following table (Table 4) thus reports data analysis corresponding to this question.

<table>
<thead>
<tr>
<th></th>
<th>Regular responses</th>
<th>Irregular responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word frequency</td>
<td>0.060</td>
<td>0.334*</td>
</tr>
<tr>
<td>N</td>
<td>50</td>
<td>50</td>
</tr>
</tbody>
</table>

* Correlation is significant at the 0.05 level (2-tailed).

The figures presented in Table 4 show a positive significant correlation ($r = 0.334$) between the frequency of words and knowledge of irregular forms of English words. This moderate correlation suggests that in relation to words which are irregularly inflected or derived, high-frequency words are more likely to be known by L2 learners. On the other hand, frequency of words does not appear to have an effect on the retrieval and recognition of regular words. The analysis in the above table shows a very weak correlation ($r = 0.060$) between frequency and the recognition of regular English forms. This finding seems to support the line of research which suggests that word frequency can affect recognition of irregular word forms but not of regular forms. A discussion of the findings is presented in the following section.

4. Discussion

The present study investigated the effect of processing regular and irregular word morphology on L2 vocabulary learning amongst low-level native Arabic learners of EFL in two public schools in Saudi Arabia. The study had the intention of establishing whether English word regularity could affect vocabulary uptake by native Arabic learners of
English. The main results can be summarised as follows. First, there was no significant uptake of vocabulary that is regularly inflected or derived compared with words which are irregularly inflected or derived. Second, the frequency effect worked in favour of words which are irregularly inflected and derived. The following sections provide a detailed discussion of the findings.

4.1 Processing of regular and irregular L2 words

The results of this study provide insights for further understanding of the construction of an L2 mental lexicon. The results indicate that the irregularly inflected and derived forms are stored in different parts of the mental lexicon. The main results can be summarised as follows. First, there was no significant uptake of vocabulary that is regularly inflected or derived compared with words which are irregularly inflected or derived. This result illustrates in Figure 1. The scattergram suggests that learners are able to uptake more irregularly formed vocabulary than regularly formed words. It can be seen from the graph that some learners have knowledge of irregular words at the higher end around the diagonal. This could also confirm that learners might have difficulty applying English rules to developing knowledge of more regular words.

One possible explanation is that native Arab learners do not master the regular rules in English, so they cannot apply them to generate regular English forms. This lack of rule knowledge could be attributed to unfocused English morphology teaching at the earlier stages of English language learning. According to McBride-Chang et al. (2005), morphological awareness is to play a unique role in vocabulary acquisition and development with young learners. A second possible reason for the poor level of knowledge of regular forms compared with irregular forms by participants in this study is the variation in rule application in English. Learners may already know that applying rules in English is not as straightforward as in their L1 and, as a result, this demotivates them in relation to learning the appropriate rules in English. In language processing research, Ullman (2000) suggests that regular base words are processed into the mental lexicon as separate entries and that rules, such as –ed, –ing, –s, are stored in different parts of the lexicon. In this case, when learners encounter an inflected word, they need to retrieve the base form and the appropriate morphological endings from different parts of the lexicon. This indicates that a learner must know not only the base form of a word, but also the rules to derive new words. For example, if a learner knows the word ‘talk’ but does not know that adding –ed to the word will form a past tense of the same word, he/she will not recognise the word ‘talked’.

A third possible interpretation for the apparent lack of automation of L2 regular rules in the current study is likely to be the age of the participants. As English is introduced to learners in Saudi schools at the age of 12, this might delay the automation of regular English rules in word derivation. According to some scholars, i.e., Bialystok (1997), Butler and Hakuta (2004) and Ullman (2001a), if the age of L2 acquisition is beyond or near the critical period, the automation of L2 grammatical rules is claimed to be less successful. Nonetheless, there are other factors that might contribute to the lack of complete automation of regular L2 rules. These factors are proficiency, practice and exposure to an L2 (Clahsen & Felser, 2006), which are in fact unavailable for EFL learners in Saudi public schools.

Overall, findings from the current study do not show that knowledge of regular forms is significantly greater than knowledge of irregular forms. The findings indicate almost a balanced recognition of both types. These findings do, however, appear to support the body of research that advocates that irregular forms are mapped onto separate entries in the mental lexicon and that representation of base regular forms are processed without the attachment of regular rules in their lexicons (e.g., Marslen-Wilson & Tyler, 1997; Pinker, 1999; Ullman, 2000, 2004). Interpretation of the findings suggests that learners in this study tended to recall the irregular forms as full words and applied rules minimally in regular word recognition. In contrast with the expectation put forward in this study – that learners will score significantly less in response to irregularly tested words than to regular words – learners tended to score almost equally for both types of words. However, it is worth noting that the vocabulary presented to the test-takers in this study are high-frequency words. Therefore, this might be a factor enabling learners to recognize words which are irregular almost to the same degree as they did with regular ones.

4.2 Effect of the frequency of regular and irregular words on L2 vocabulary learning

The second important finding of this study is that frequency has a significant impact on learning L2 irregularly inflected and derived vocabulary compared to regularly inflected and derived words. This finding accords with the results reported by Prado and Ullman (2009), who found that irregular form uptake is influenced by frequency. The finding from this study and that of Prado and Ullman (2009) support the idea that irregularly inflected and derived words are stored as full-formed entries, with frequency being an essential factor in their recognition, whereas regular forms are computed online.

The frequency effect on the learnability of irregular English forms by native Arabic EFL learners is supported by the result illustrated in Figure 1. The scattergram suggests that learners are able to uptake more irregularly formed vocabulary than regularly formed words. It can be seen from the graph that some learners have knowledge of irregular words at the higher end around the diagonal. This could also confirm that learners might have difficulty applying English rules to developing knowledge of more regular words.
The present findings suggest that learning the regular rules of an L2 are crucial for developing a larger L2 lexicon, and that knowledge of morphological rules in a learner’s L1 might not always apply to learning an L2. It should be noted here that the Arabic lexicon is differently structured from the English, which can, to an extent, justify the deficiency of applying knowledge of L1 rules to deriving L2 regular forms. Therefore, the burden of learning L2 rules and words with appropriate morphological attachment should be adopted in order to develop a larger L2 lexicon.

5. Conclusions

This study described an attempt to investigate the effect of processing regular and irregular words by native Arab learners on EFL vocabulary acquisition. The idea was that Arabic learners might utilise the feature of implementing rules to generate a wide range of vocabulary in their L1 from the same root to learn regular English words significantly more effectively than irregular ones. The main interest in the current paper was in finding out:

- Do native Arab learners of English more easily learn regularly inflected and derived forms or irregularly inflected and derived forms?
- Does word frequency influence regular or irregular word learning, or both?

In addition to an interest in finding answers experimentally to the addressed questions per se, another aim was to explore whether native Arab learners of English can transfer their knowledge of the regularity of rules they have already established in their L1 mental lexicon to learn words which have regular morphological endings in English.

The first conclusion is that evaluating participants’ recognition in terms of word regularity has an obvious limitation. Despite the intuitive appeal of assuming that Arab EFL learners could utilise their skills in applying rules from their L1 to learning regular English words significantly more effectively than irregular ones, there seems to be little evidence to validate this assumption. One possible reason for this finding is that those who took part in this study are less proficient learners of English. Consequently, knowledge of a strong system of rules in their L1 appears not to support learning of L2 regular words. Thus, replicating this study with more proficient L2 learners might yield different results.

The second conclusion is in relation to frequency effect on the type of L2 word regularity. It was found that frequency has a more significant effect on the retrieval of irregular words than regular ones. The finding from the current study appears to be in line with findings from previous research, such as that by Marslen-Wilson and Tyler (1997), Pinker (1999) and Ullman (2000, 2004). These previous studies suggest that irregular words are mapped onto separate entries in the mental lexicon and that their retrieval is based on their frequency of occurrence.

Although the findings from this study do not show a clear effect of regular rule knowledge in L1 on developing knowledge of regular words in L2, future research to investigate this type of knowledge is worth conducting. If irregular words are stored separately in the L2 mental lexicon, as supported by many studies, then guessing base words from inflected and derived forms might hold greater difficulty for L2 learners than for regular forms.
References


