Bound Morpheme Frequencies in the Performance of Iranian English Language Undergraduates and English Language Materials Developers in Written Descriptive Tasks

Rahim Najjari
PhD candidate, Urmia University
Rahim.Najjari@gmail.com

Mohammad Mohammadi
Assistant Professor, Urmia University
Mohammdi680@urmia.ac.ir

Abstract

This mini-corpus, cross-linguistic, comparative, and norm-referenced study intends to render the most frequently and oft-used affixes in the written descriptive tasks in the performance of English language materials developers (ELMDs) and Iranian English language undergraduates (IELUs). Samples of writings of both groups were studied and analyzed through affixation principles. The frequency of nearly 63 affixes both inflectional and derivational morphemes was taken into account. The statistical analysis showed a highly significant correlation (.86) with (.75) coefficient determination (shared variance) between the performance of IELUs and ELMDs. The result of the study bears approximately a close resemblance to the findings by Bauer et al. (1982), Fry and Kress (2006), Honig, Diamond, Gutlohn, and Mahler (2000), and White, Sowell, and Yanagihara (1989). Finally, the implications of the study were taken into account.

Keywords: Affixation, Morpheme Studies, Frequency, Morphological Awareness

Received: January 2015; Accepted: April 2016
1. Introduction

In a thoughtful review of research on the adult second-language learning and its application to teaching practice, Pasty Lightbown (1985) identified three broad categories of research, based principally on methodological differences: descriptive studies, experimental studies and hypothesis-testing studies. The second and third categories are out of concern in this paper. Regarding descriptive studies, researchers collect speech samples from second language speakers- either through spontaneous speech or through various elicitation procedures- and then compare these samples to the target language norms. The goal is to account for consistencies or discrepancies between the second language learners’ use of certain linguistic forms and native speakers’ use. Generally, the study falls in the first type of this classification to account for the performance of Iranian language learners and native speakers in the use of bound morpheme. It concerns with Halliday’s (1988) second strategy in the formation of specialized terminology, creating new words from a native word stock which requires different word-formation processes including compounding, conversion, back formation or derivation. The last process (derivation) is yielded by means of affixation, i.e., adding a derivational affix to a word by means of prefixation, suffixation or infixation. Thus, affixation as the most common morphological process known as concatenative morphology (Fasold & Connor-Linton, 2006) is the main concern of this study.

According to Kirby et al. (2012), there are various methods employed to assess morphological awareness (affixation) through oral, written, or combined oral and written form. Additionally, morphological awareness tasks can assess judgment, production, or decomposition abilities. Some of the chronologically-listed of these studies were done by Berko (1958), Hakuta (1976), Larsen-Freeman (1975), Krashen, Sferlazza, Feldman, & Fathman (1976), Makino
Bound Morpheme Frequencies in the Performance…

(1980), Pica (1983), Bybee (1985), and Tarone (1988). The studies have usually served two purposes: (1) some like Bailey, Madden, and Krashen (1974) studied morphemes as a proof for natural order acquisition to show a predictable order for acquiring language and (2) others like McLaughlin (1991) used morpheme studies for the failure of contrastive analysis in dealing with transfer notion from L1 so as to account for the developmental process (creative construct) and interlanguage.

Recently a new conceptualization of morpheme studies or morphological awareness has crept into the literature aiming to study their contribution to reading competence (Carlisle, 2003; Carlisle & Stone, 2005; Deacon & Kirby, 2004). The status of morphemes as units of meaning and the role of inflections in forming syntax suggests that reading comprehension is the most likely beneficiary of morphological awareness. Mahony (1994), for instance, examined the relation between knowledge of derivational morphology and the reading ability of young adults. Her study showed that the more skilled readers make the correct choice on the nonsense derived forms, as well as on the real words. Carlisle (2000) has argued that morphological awareness may contribute to text comprehension by supporting the interpretation of unknown words. Karimi and Gheitury (2009) have extended and incorporated the effect of morphological awareness studies on Iranian pre-university students’ listening transcription. Their study showed that the ability of the students to transcribe listening passages affects the students’ listening comprehension ability since according to Nation and Newton (2009) knowledge of morphological characteristics of a word is believed to be related to building a substantial vocabulary size, identification of their grammatical categories, and effective word recognition skills which are, in turn, reported to be strong predictors of listening transcription performance.
This study is an extension of morpheme study or morphological awareness to the writing performance of language learners, a new domain that morpheme (bound morphemes) of language learners can be dealt with. In order to render a rather comprehensible framework to assess Iranian English language undergraduates’ (IELUs) morphological awareness, English language materials developers’ (ELMDs) performance was taken as a criterion. On the basis of modality and channel, written tasks of IELUs were studied to closely match with the input they have received from ELMDs’ materials. And also, great attention was exercised that written product of both groups to fall in a specific category (genre) of writing i.e., descriptive writing.

2. Significance of the Study

This study can be taken into consideration as:

(1) a rhetorical classification which determines different academic written discourse domains in terms of derivational and inflectional characterization including narratives, instructions, descriptions, reports, etc. In this paper, the focus is on description which is linked to the specific language function that it serves. In fact, it deals with how derivational and inflectional characterization may play a role in rhetoric variation; in other words, how the use and frequency of affixes may contribute to the morphological composition of descriptive writing; consequently, it would reveal which bound (affixes) morphemes play central roles in semantic characterization of a written product to convey its message like studies done by Nagy, Diakidoy, and Anderson (1993);

(2) a common descriptive type of study prevalent in second language writing research (Polio, 2012), dealing with the analysis of written tasks in the
Bound Morpheme Frequencies in the Performance...

terms derivational and inflectional morphemes (affixes) to focus on problematic areas;
(3) a quantitative study intended to elicit highly frequent bound morphemes in terms of the most frequent ones to the least frequent ones like studies done by Fry and Kress (2006); Honig et al. (2000), and White et al. (1989).
(4) a corpus-based study which mostly deals with the frequency of affixes in the performance of both ELMDs and IELUs to take care of consistency in the frequency of affixes used in inner circle of English with its outer circle;
(5) a content analysis of second language writing. According to Polio (2012), there are eight types of methods or techniques to study second language writing: surveys, interviews, meta-analysis, classroom observation, ethnography, content analysis, text analysis, and process research. This study is a type of content analysis to determine how writing is taught in a specific context, the analysis can provide teachers and students with a description of the language needed to be mastered for writing in a specific context. Swales (2004) has extend content analysis to genre analysis to characterize how research articles are organized in English;
(6) morphological awareness: The ability to reflect upon and manipulate morphemes and employ word formation rules in one’s language or other languages. (Kuo & Anderson, 2006); and
(7) finally, as a correlational study which is going to determine how much IELUs’ performance corresponds with ELMDs’ performance. All of the considerations have also been taken into account while the third, the sixth and seventh have been deeply delved with in this paper.
3. Method

3.1. Participants

IELUs are both male and female language learners majoring English at Shahid Madani University, Khatam-e-nabi University, and Maragheh Azad University. They have fulfilled more than half of their undergraduate education and have already passed some writing and grammar courses during their educations. ELMDs are those native language speakers who write on different subjects for pedagogical purposes. This selection works on the assumption that IELUs are mostly flooded with these materials developers’ input, their feedback is of a great value and shows how much they have grasped and which areas warrant a careful consideration.

3.2. Data Collection and Analysis

First, samples of written descriptive products of IELUs and ELMDs were collected and keyed into computer Word processing. The total number of words was 4.130 for each group. The mini-corpus was used in order to be manageable and as Willis (2011, p. 72) states that ‘it is impossible for most language teachers and course designers to assemble their own research corpus for a particular group of learners, unless the learners’ target discourse is a very narrow, well defined area which is readily researchable,’ and’ a pedagogic corpus is inevitably quite small and needs to be selected in a principle manner (Willis, 2011, p. 71). Pedagogically and objectively, it is also a representative snapshot of a real mini-corpus which is the product of learners’ own performance.

Then, principle of morpheme identification was applied on the data. Morpheme identification is treated as the minimal linguistic sign, i.e., a unit of
form and meaning. For example, ‘un’ in unhappy is an entity that consists of the content or meaning on the one hand, and the sounds or letters which express this meaning on the other hand. It is a unit of form and meaning. Linguistic expressions such as unhappy, whose meaning is a function of the meaning of its parts is called compositional (Plag, 2002). Furthermore, addition of meaning by means of vowel alternation evident in English in certain cases of past tense formation and of plural marking on nouns was also taken into account, for instance saw for see. Taking into account this principle, each group product was analyzed with different affixes through “Computer’s Replace Tool” to obtain their frequency. It was not as easy as it seemed, these raw data should have been manually checked whether the computer-elicited frequency matches the principle of identifying affixation. For example, the prefix ‘un’ had been used 10 times, but before all, it should have been checked whether ‘un’ means ‘not’. Moreover, there were some similar affixes making adjectives or verbs like ‘en’ in the words golden and shorten respectively.

4. Results

IELUs and ELMD have used 1099 and 1213, respectively. Table 1 shows the frequency and proportion of inflections and derivations in both groups.

<table>
<thead>
<tr>
<th>Groups</th>
<th>Inflections</th>
<th>derivations</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency for IELUs</td>
<td>374 (34%)</td>
<td>725(66%)</td>
<td>1099(100%)</td>
</tr>
<tr>
<td>Frequency for ELMDs</td>
<td>499(41%)</td>
<td>714(59%)</td>
<td>1213(100%)</td>
</tr>
</tbody>
</table>

After drawing the total data for each group, they were divided into two groups: inflections and derivations with their frequencies, the second step was to provide for these categories and subcategories (nouns, adjectives, adverbs,
verbs and prefixes) their frequencies, their relative frequencies through tables, charts, figures, and graphs (table 2 & figure 1).

**Table 2. Frequency of Bound Morphemes (Affixes) by Six Categories**

<table>
<thead>
<tr>
<th>Groups</th>
<th>verb</th>
<th>adverb</th>
<th>adjective</th>
<th>noun</th>
<th>prefix</th>
<th>inflection</th>
<th>total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency for IELUs</td>
<td>13</td>
<td>28</td>
<td>292</td>
<td>323</td>
<td>69</td>
<td>374</td>
<td>1099</td>
</tr>
<tr>
<td>Frequency for ELMDs</td>
<td>15</td>
<td>82</td>
<td>267</td>
<td>252</td>
<td>99</td>
<td>499</td>
<td>1213</td>
</tr>
</tbody>
</table>

**Figure 1. Relative Frequency of Affixes (verbs, adverbs, prefixes, adjectives, nouns, and inflection)**

The third step was taking comparatively detailed analysis of frequencies of subcategories in both groups’ performance into account.

### 4.1. Inflectional Affixes

The table 3 shows that ELMDs applied verb inflections 125 times more than IELUs and the number of past tense verb inflections in IELUs’ performance was 20 while in ELMDs’ performance it was 115. A noteworthy point about the inflectional suffixes was the frequency of present tense verbs in ELMDs’ performance was 180 (115 infinitive), (65 third person singular) while in its counterpart group, it was 64 (22 infinitive), (44 third person singular). The verb
inflectional suffixes in latter were in past tense while in the former in present tense. It seems that what they have grasped and achieved is in present tense and interestingly enough sometimes it happens in their writing tasks what they have in mind is in past tense while they have expressed it wrongly through present tense forms.

Table 3. Frequency of Inflectional Affixes in the Performance of IELUS and ELMDS

<table>
<thead>
<tr>
<th>Groups</th>
<th>Categories</th>
<th>Number (-s,-es)</th>
<th>Possessive (-’s,-’s)</th>
<th>3rd person singular present (-s,-es)</th>
<th>Past tense (-ed)**</th>
<th>Past perfect (-ed)**</th>
<th>Progressive or continuous aspect(-ing)</th>
<th>Comparative(-er)**</th>
<th>Superlative</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency for IELUs</td>
<td>215</td>
<td>11</td>
<td>65</td>
<td>20</td>
<td>45</td>
<td>8</td>
<td>6</td>
<td>4</td>
<td>374</td>
<td></td>
</tr>
<tr>
<td>Frequency for ELMDS</td>
<td>225</td>
<td>17</td>
<td>42</td>
<td>115</td>
<td>71</td>
<td>8</td>
<td>12</td>
<td>9</td>
<td>499</td>
<td></td>
</tr>
</tbody>
</table>

Note:** both their regular and irregular forms are considered.

4.2. Adjective

The frequency of adjective making suffixes or suffix families for both IELUs and ELMDs was respectively for -able, -ible (19-18), -al,-ial,-ical (96-54), -ant,-ent (38-13), -ate(6-5),-ed (20-33), -er (6-12), -ful (11-7), -ian (13-20), -ic (29-25), -ing (13-20), -ish (5-10), -ive,-ative,-itive (9-5), -less (1-2), -like (2-5), -ly (4-3), -ous,-eous,-ious (8-9), -y (13-31). To compare both group, see figure 2.
4.3. Nouns

IELUs have applied 72 more noun suffixes than ELMDs: 325 vs. 251 (Table 5). Apart from input and proficiency reasons for this matter, a persuading reason is the role of nouns and verbs in descriptive genre. Nouns and verbs play descriptive functions for readers. With nouns, readers see; with verbs, they feel. So, it stands to reason that only nouns are described in IELUs’ performance and the high frequency of adverb making suffixes (-ly) in ELMDs performance gives an overwhelming reason that IELUs failed to apply different uses of adverb with verbs to make readers feel description.

The frequency of noun making suffixes or suffix families for both IELUs and ELMDs was as follows -age (26-8), -al (5-7), -ance (21-15), -ation (28-15), -ee (1-0), -eer (3-1), -er, -or (17-6), -ian, -an (12, 20), -ing (67-75), -ion (8-11), -ist (10-4), -ity (19-16), -ment (10-13), -ness (6-8), -tion, -sion (60-26), -y (30-26). To compare both group, see figure 3.
4.4. Adverb

The number of suffix forming adverbs in the performance of ELMDs and IELUs was respectively 82 and 28. To account for inconsistency and discrepancy, the following stand to reason:

The first one that can be put forth is the high frequency of linking verbs in IELUs' performance since applying more linking verbs in descriptive tasks involves using adjectives. The second factor is perhaps the low frequency of main verbs in IELUs' performance since using main verbs involve describing the action of verb which sometimes needs using adverbs of manner, but as table (Table 4) shows this cannot be true.
### Table 4. Frequency of Linking and Main Verbs in both Groups’ Performance

<table>
<thead>
<tr>
<th>Categories</th>
<th>Frequency of linking verbs</th>
<th>Frequency of main verbs</th>
<th>Total frequency of both linking and main verbs</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>is</td>
<td>am</td>
<td>are</td>
</tr>
<tr>
<td>Frequency</td>
<td>in</td>
<td>127</td>
<td>3</td>
</tr>
<tr>
<td>IELUs' performance</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frequency</td>
<td>in</td>
<td>59</td>
<td>1</td>
</tr>
<tr>
<td>IELMDs' performance</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The third reason is related to the simplicity of adverb structure in IELUs’ performance. Adverbs of manner are placed after object; mid-position, end-position, initial-position; before adjectives; and before another adverb. While these four options were present in ELMDs’ performance, only options 1 and 2 were present in IELUs’ performance.

### 4.5. Verbs

Since the frequency of verb making suffixes was negligible, it was not treated

### 5. Discussion

On the basis of Pearson correlation formula, there exists .86 coefficient correlation between two groups’ performances (Table 5). In order to show the
Bound Morpheme Frequencies in the Performance…

meaningfulness of correlation, the observed correlation should be squared so as to take for their coefficient determination i.e., how much they share overlapping variance. In this study corresponding squared value, common variance, is .75. One point of noteworthy is that even running nonparametric correlation coefficient like Spearman’s rho and Kendall’s tau_b yields a high correlation coefficient i.e., .88 and .72, respectively (Table 6).

Table 5. Pearson’s Product Moment Correlation in the Performance of ELM DS and IELUS

<table>
<thead>
<tr>
<th>Correlation</th>
<th>ELMDs</th>
<th>IELUs</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELMDs Pearson Correlation</td>
<td>1</td>
<td>.867**</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td>.000</td>
<td></td>
</tr>
<tr>
<td>Sum of Squares and Cross-products</td>
<td>7.266E4</td>
<td>5.801E4</td>
</tr>
<tr>
<td>Covariance</td>
<td>1.153E3</td>
<td>920.770</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

**. Correlation is significant at the 0.01 level (1-tailed).
### Table 6. Spearman’s rho and Kendall’s tau_b Correlations the Performance of ELMDs and IELUs

<table>
<thead>
<tr>
<th></th>
<th>IELUs Correlation Coefficient</th>
<th>ELMDs Correlation Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Kendall’s tau_b</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IELUs</td>
<td>1.000</td>
<td>.722**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td><strong>ELMDs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.722**</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td><strong>Spearman’s rho</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IELUs</td>
<td>1.000</td>
<td>.886**</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.</td>
<td>.000</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
</tr>
<tr>
<td><strong>ELMDs</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation Coefficient</td>
<td>.886**</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.000</td>
<td>.</td>
</tr>
<tr>
<td>N</td>
<td>64</td>
<td>64</td>
</tr>
</tbody>
</table>

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

The total affix frequency order in IELUs’ and ELMDs’ performance shows that 63 affixes enjoy frequency and the rest of them are rarely used for this particular writing task. Twenty suffixes were taken out which have the frequent application of more than 20 times (Figure 4).
**Bound Morpheme Frequencies in the Performance…**

![Graph](image)

**Figure 4. The Most Common Affixes in the Performance of IELUs and ELMDs**

This list also bears a close resemblance to the studies done by Bauer et al. (1982), which are ten suffixes (-able/-ible), -ance/-ence), -ate,-ful, -ity, -less, -ly, -ment, -ness, -tion /-ion /-sion); (2) White et al. (1989), which are 20 suffixes (-s/es (plural), -ed, -ing, -ly, cr/or, -ion/-ation/-ition, -able/able, -al/-ial, -y, -ness, -ity, -ment, -ic, -ous, -en, -er (comparative), -ive/-ative/-itive, -ful, -less, -y); (3) Honig et al. (2000), which are (-ible/able, -al/-ial, -ed, -er (comparative), -en/-es, -ful, ic, -ing, -ly, -ion/-ation/-ition, -y, -ive/-ative/-itive, -less, -ment, -ness, -ous, -s/es (plural), and -y); and (4) Fry and Kress (2006) which are (-s/es, -al/ial/ical, -ing, -ed, -ion, -sion, -y, -ic, -ly).

Interestingly enough, the top ten mostly used derivational affixes were common in both groups. (-al/-ial/-ical), (-ing(n)), (con/-com/ co/-cog/-col/-cor-), (-tion/-sion), (-y (n)), (-ic), (-ly (adv)), (-ed (adj)), (-ity), (-able/-ible). There was also little change in ranking positions for the order of affixes with only two or three places which were inconsequential. Both groups have used more derivational affixes than inflectional affixes, i.e., 66 percent for IELUs and 59 percent for ELMDs.

203
Figure 5. The Most Common Prefixes in the Performance of IELUs and ELMDs

As to the prefixes, the common prefixes in both group’s performances (Figure 5) share a close resemblance to the findings by

1. Bauer et al. (1982), [ad-, circum-, com-, de-, dis-, ex-, in-(not), in(into), inter-, mis-, post-, re-, sub-trans-, and –un];
2. White et al. (1989) [un-, re-, in/im/il/ir-, dis-, en/em-, non-, in/im-, over-, mis-, sub-, pre-, inter-, trans-, super-, semi-, anti-, mid-, and under-]; and
3. Honig et al. (2000) [anti-, de-, en/em-, fore-, in/im-, in/im/il/ir-, inter-, mid-, mis-, non-, over-, pre-, re-, semi-, sub-, super-, trans-, un-, and under-].

6. Concluding Remarks

This study showed that there was a strongly high correlation between two groups’ performance. Additionally, the elicited affixes in this study approximately match the studies of Bauer et al. (1982); Fry and Kress (2006); Honig et al. (2000), and White et al. (1989).

The following advantages on the basis of the study will be accrued: (1) English language learners can weigh their performance up and down and notice what they have gained during their education on the basis of ELMDs’ criteria; (2) English language teachers can provide sufficient amount of input, revise their previous syllabus selection, and provide complementary courses in
the case of inconsistency; (3) non-native English syllabus designers can take this norm for developing descriptive materials; (4) morphologists can benefit by the corpus-based result of this study to determine bound morpheme characterization of a particular frame of writing, descriptive writing in this case; (5) and it can trigger a new line of research for affixation in corpus studies.

7. Limitation of the Study

Differences are inconsequential when the frequency of affix application is quantitively taken into consideration between two groups’ performance. However, when their performance is considered as to some specific morphemes, it warrants a careful attention. A noteworthy point about the morpheme acquisition in the paper is that it has only accounted for the implicit acquisition and application of morphological knowledge while according to Kuo and Anderson (2006) a distinction should be made between the acquisition of morphology and the development of morphological awareness. While the former is concerned with development of the ability to comprehend and produce morphologically complex words in natural speech, the latter focuses on the ability to reflect on and manipulate word formation rules in the absence of a communicative context. Morphological awareness can be seen as the more explicit representation and manipulation of implicitly acquired morphological rules. In order to have a full picture of IELUs’ performance, in addition to their implicit acquisition and application of morphemes, their explicit learning especially through’ noticing’ and ‘awareness-raising activities’ should be studied.
8. Suggestions for Further Research

This paper is the first attempt to examine whether morphological knowledge is quantitatively related to writing skills through bound morpheme frequency, though its effect on reading skills has already been done (Carlisle, 2003; Carlisle & Stone, 2005; Deacon & Kirby, 2004). An experimental study is needed to shed light upon the point whether participants’ exposure to morphologically oriented treatment has effect on their use of relevant bound morphemes in a specific genre or rhetoric. One thing interesting is that there are a lot of studies nowadays done on special vocabulary in English for specific purposes; even it is extended to lexical bundles, group of three, four, five, six, and on, (Ädel & Erman, 2012). A lacuna warranting a careful attention is to explore bound morpheme frequencies in different ESP courses and genres in order to provide both bound morpheme related corpora for different ESP courses.

References


**Bound Morpheme Frequencies in the Performance…**


