The Correlation between Language Proficiency and the Use of Lexical Collocations by Algerian EFL Undergraduate Students in Essay Writings: a Case Study of Verb+ Noun Collocations

Abstract:

‘Pseudo-native speaker’ is the term given to a foreign language learner who has reached a high level of proficiency. Such level of proficiency is characterised by the mastery of grammar, vocabulary, and pronunciation. Consequently, a distinction must be drawn between grammatical proficiency and the idiomatic one. Idiomaticity has three basic levels of compositionality: free combinations, collocations, and idioms. The present research work investigates the mastery of Verb+Noun lexical collocations by Algerian EFL learners in Frères Mentouri university. It is hypothesised that Algerian EFL learners have a low level of mastery in the usage and comprehension of Verb+Noun collocations. It is also hypothesised that the essays of Algerian EFL learners lack the use of appropriate Verb+Noun collocations. An online corpus (BNC) and two other online websites (Collocation Checker & O.O.C.D) were used as tools to measure learners’ collocational competence. Examined collocations were extracted from the learners’ examination essays. The analysis was done manually through identification and corpus usage. The findings of the study indicate that Algerian EFL learners have a low competence in dealing with Verb+Noun collocations. Thus, focus on collocational knowledge is required. Teachers must raise their students’ awareness on the importance of collocations in improving the oral and written quality of production. Learners have to, deliberately, be able to identify collocations, and memorise them within their contexts.

Introduction:

Inadequate research has been carried out in the area of collocation. Although some researchers have shown interest in collocations, extensive research still needs to be undertaken to highlight its practical importance in improving EFL learners’ overall linguistic competence. The topic of collocation raises many questions concerning the measurement of learners’ collocational competence, the best methods to teach collocations, or the types of collocations to emphasize on. In order to be able to answer these questions, it is necessary to, first, identify difficulties in mastering collocations (Nesselhauf 2003).
This study, unlike previous studies (Halliday and Hassan 1976; Benson et al. 1986), focuses on Verb + Noun lexical collocations as the mostly used type of collocation. Hatch and Brown (1995) stated that learners acquire collocations in chunks rather than as separate words. The same thing goes to L1 users. In other words, L2 learners learn collocations like infants. The main problem with some studies about collocation is that the definition is vague. No clear and specific definition is given, which makes experimentation difficult (Hussein 1990; Farghal and Obiedat 1995). Firth (1957) defines collocation as the best company a word keeps. Such definition might seem ubiquitous of the notion; however, with closer attention, one would respond by saying that not all words that co-occur are collocations. In Algeria, research on collocation repeatedly tackles its translation (Moussa 2007; Boussalia 2010). The present study attempts to shed some light on Algerian learners’ knowledge of collocation in academic writing. It investigates the correlation between success in writing and the acquisition of collocations.

1. Literature Review

An essential introduction to the study of language is made by tackling the topic of word knowledge. A distinction is drawn between the three types of word knowledge: knowledge of form, knowledge of meaning, and knowledge of use. These three types make up for the whole study of language. In light of this concept, the notion of collocation can be defined and explained. By understanding what collocations are, one will be able to identify them on the language continuum, and distinguish them from free combinations and idioms.

1.1. Definition of Collocation

Collocations have been recognised to distinguish a native speaker from a foreign language learner. Usually, one way a non-native speaker may offer help is by saying ‘Can I help you?’ Meanwhile a native speaker would say ‘Can I give you a hand?’ (Salkauskiene, 2002, as cited in Farrokh, 2012, p. 56). Although a non-native speaker might succeed syntactically, he/she may not sound a native-like in terms of which words combine with which. English, similar to many languages, has its own share of re-occurring word combinations. These word combinations arbitrarily co-occur. A native speaker says ‘ultimate goal’ and not ‘last goal’; one ‘pays a visit’, but does not ‘make a visit’ (ibid, p. 56). Prodromou (2004, as cited in ibid, p. 56) says knowing the meaning of a word not only requires knowing its dictionary definition; one must also know the type of words with which it is often associated. Collocations, either fixed or more flexible, are the result of many years of habitual use by fluent speakers of the English.
Palmer (1933) introduced the notion of collocation. According to him, it is “a string of words that must or should be learned, or is best or most conveniently learnt as an integral whole or independent entity, rather than by the process of piecing together their component parts” (Ganji, 2012, p. 125). Firth (1957) followed later, and was recognized as the one who introduced this concept to the world.

During the last decades, research in linguistics has given much importance to collocation. Scholars, teachers, and corpus designers understood its importance during the language learning process. They mentioned that the increase in linguistic competence is primary in the enhancement of learners’ communicative competence, and the approximation to native-like fluency (Darvishi, 2011, p. 52). On the same note, teachers’ ability to identify collocational errors to EFL leaners can significantly increase their awareness of the subject. Therefore, studying miscollocations is equally important to studying correct ones, for it helps instructors to focus on the difficult ones (ibid, p. 52-53).

Focus on collocation by many scholars led to a variety of definitions. For Firth (1957, p. 179, as quoted in Salman and Mansoor, 2013, p. 04) “you shall know a word by the company it keeps.” According to Lewis (1997, p.44, as quoted in Farrokh, 2012, p. 57) “collocations are those combinations of words which occur naturally with greater than random frequency. Collocations co-occur, but not all words which co-occur, are collocations.” This states clearly that collocations have their specific characteristics which distinguish them from other word combinations like free combinations and idioms. Hardi (2008) defines collocation as “words that occur together in a text. This co-occurrence, he believes, is frequent enough to make a pair of words as a collocational phrase” (ibid, p. 57). Hardi’s words denote the characteristic of frequency which is primary to identifying collocations. However, frequency hits which qualify a certain word combination to move from free combination to collocation differ according to scholars and corpora. Some scholars set the bar to 40 hits per 100 million words; others suggest the hits should reach 50 hits per 100 million words.

1.2. Collocational Competence

As much as grammatical competence is important to reach native-like proficiency, collocational competence makes a foreign language learner a pseudo-native speaker. There are many advantages that collocational competence offers. According to Henriksen (2013, p. 33), collocational competence helps “1) to make idiomatic choices and come across as native-like; 2) to process language fluently under real-time conditions (Columbus,
2010; Ellis et al., 2008); 3) to establish ‘islands of reliability’ (Dechert, 1983; Raupach, 1984) which enable the language user to channel cognitive energy into more creative production; 4) to disambiguate meaning of polysemous words, e.g. the verb commit in the following collocational contexts: commit a crime, commit oneself, commit to memory; and 5) to understand connotational meaning (what Sinclair, 2004 has described as semantic prosody), e.g. the fact that the verb cause is often associated with negative connotations as in cause an accident.” In other words, appropriate understanding of collocations allows the reader or listener to achieve communicative competence.

1.3. Types of Collocations

The various perspectives the notion of collocation was viewed from led to a different categorisation. According to Chia-Chuan, “Cowie and Mackin (1973) classified idioms and collocations into four categories based on idiomaticity from most to least fixed: pure idioms, figurative idioms, restricted collocations, and open collocations (2005, p. 11). On a different note, Wood (as cited in Nattinger & De Carrico, 1992, and paraphrased in Darvishi, 2011, p. 53) categorised collocations into idioms, colligations, and free combinations. His categorisation was based on semantic and syntactic criteria. For Lewis (1997), collocations were classified into strong, weak, frequent, and infrequent (ibid). The categorisation adopted in this research is developed by Bahns. According to Boussalia (2010, p. 15), collocations can be classified into two main categories: grammatical collocations and lexical collocations.

Grammatical collocations contain prepositions, sometimes occurring with verbs, nouns, or adjectives, e.g. (reach down, put forth), (by car, on foot), or (interested in, happy with) (McKeown & Radev, p. 05). Bahns (1993, p. 57, as quoted in Boussalia 2010, p. 15) notes that they “(usually) consist of a noun, an adjective or a verb, plus a preposition or a grammatical structure such as an infinitive or clause.”

On the other hand, lexical collocations are “restricted lexically” (McKeown & Radev, p. 05-06). Benson (1985, p. 62) explains that “lexical collocations contain no subordinate element; they consist of two lexical components” (Boussalia 2010, p. 15-16). This means that there are no grammatical components within the word combination, only lexical ones (content words). According to Gabrielatos (1994, p. 02), “there are three factors determining the categorising of a lexical collocation: the degree of probability that the items will co-occur, the degree of fixity of the combination (i.e. grammatical restrictions), and the degree to which the meaning of the combination can be derived from its constituent parts.” There are six types of lexical collocations:

1. Verb + noun: pay a visit.
2. Adjective + noun: heavy rain.
The Correlation between Language Proficiency and the Use of Lexical Collocations by Algerian EFL Undergraduate Students in Essay Writings: a Case Study of Verb+ Noun Collocations

5. Verb + adverb: announce happily.

1.4 Factors Underlying Collocational Errors

Li (2005, p. 18) proposes that “the major reason why EFL learners generally lack collocational knowledge is that collocation has been neglected in EFL classrooms and thus, learners tend to ignore learning collocations”. Chia-Lin, Liu (1999) identified three factors for collocational error: negative transfer, synonymy and approximation use, and negative interference (Kuo, 2009, p. 149).

Approximation is the incorrect use of words and structures sharing similar semantic features; the learners’ aim is to express the intended meaning. An example is using ‘fell’ to say ‘fail’.

In synonym use, the learner knows a word’s synonym, but fails to associate it with the appropriate collocate, thus, an erroneous collocation is produced. An example is ‘to accept opinions’ instead of ‘to receive opinions.’ Assuming that since ‘receive’ and ‘accept’ are synonyms, the learner uses a collocate which does not fit in the word combination (ibid).

Influenced by their mother tongue, learners tend to use structures which are not acceptable in the target language. Lack of knowledge of syntactic and semantic differences between languages usually leads to such errors. This is called ‘negative transfer.’ For example, ‘miss the lovers’, which is ‘يشتاق للأحباء.

Three main criteria were chosen to identify collocations in this study:
1. The sense of the verb is so specific that it can only combine with a small set of nouns.
2. The verb in this sense cannot be replaced by their syntactically and semantically possible choices.
3. Word combinations which have high frequency hits in the British National Corpus (40 hits) are considered as well-formed collocations (ibid).

2. The Study

The present work investigates the use of Verb-Noun lexical collocations in the essays of the Algerian second year undergraduate EFL learners at the department of English at Université des Frères Mentouri. It is to discuss the following questions:
1. To what extent do second year EFL undergraduate students of English at Université des Frères Mentouri master V+N collocational language while writing essays?
2. Are second year Algerian EFL undergraduate students of English at Université des Frères Mentouri able to produce and understand V+N Collocational language?

It is hypothesised that:

1. Algerian EFL second year undergraduate students of English at Université des Frères Mentouri have a low level of mastery in the usage and comprehension of V+N collocations.

2. The essays of Algerian EFL second year undergraduate students of English at Université des Frères Mentouri lack the use of appropriate collocations.

2.1. Data Gathering Tools

The study used three research tools: Collocation checker, Oxford Online Collocation Dictionary (O.O.C.D), and the British National Corpus (BNC).

Collocation checker is an online software, which permits the researcher to identify correct and erroneous collocations. It shares the same database of BNC. All word combinations retrieved from students’ essays were analysed via this software.

The second tool is O.O.C.D, which is an online software, too. It provides all the possible collocational hits for the searched word, in addition to examples retrieved from the BNC. Its main function - in this study - is to distinguish collocations from free combinations.

The British National Corpus is the third tool. It is an online corpus that comprises 100 million words. BNC contains authentic data, and is considered as a reliable source for examining collocation. A high-frequency hit in BNC is proof of a well-combined word combination. Examples are provided in Table 01. The basic threshold in this study is 40 hits per word combination. All collocations which did not score 40 hits were considered as free combinations, instead.

2.2. Research Method

The present section describes the research method applied in the study. It contains the frequent examples, the main criteria for selecting collocations, and the method of application of the research tools.
### Table 01: Typical V+N Collocation Type Found in this Study

<table>
<thead>
<tr>
<th>Pattern</th>
<th>Correct Collocation</th>
<th>Erroneous Collocation</th>
<th>Suggestion for Improvement</th>
</tr>
</thead>
<tbody>
<tr>
<td>V+N</td>
<td>attend classes</td>
<td>stare star</td>
<td>watch star</td>
</tr>
<tr>
<td></td>
<td>do homework</td>
<td>promote appetite</td>
<td>increase appetite</td>
</tr>
<tr>
<td></td>
<td>answer a question</td>
<td>do preparation</td>
<td>make preparation</td>
</tr>
<tr>
<td></td>
<td></td>
<td>pay time</td>
<td>spend time</td>
</tr>
</tbody>
</table>

The study was based on three criteria to identify acceptable collocations and erroneous ones. Table 02 illustrates these criteria.

### Table 02: Criteria for Choosing Collocation in this Study

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>The sense of the verb is so specific that it can only combine with a small set of nouns.</td>
</tr>
<tr>
<td>02</td>
<td>The verb in this sense cannot be replaced by their syntactically and semantically possible choices.</td>
</tr>
<tr>
<td>03</td>
<td>Word combinations which have high frequency hits in the British National Corpus are considered as well-formed collocations (40 hits).</td>
</tr>
</tbody>
</table>

The initial two criteria are flexible; however, the third one is fixed. In other words, the verb is considered to be ‘restricted’ if one or both of the restrictedness and substitutability criteria were met. Nonetheless, the combination has to score, at least, 40 hits.

A primary task was to identify erroneous collocations. The three research tools were used in the aforementioned order.

Collocation checker was used to spot wrong collocations. After a word combination is inputted, collocation checker indicates whether the combination is correct or erroneous. However, that does not entail that all acceptable word combinations –according to collocation checker—are collocations. After identifying wrong collocations, the send task was to distinguish free combinations from collocations.
O.O.C.D was used to determine whether the verb + noun investigated collocate or not.

BNC came as a final step to ensure that word combinations labelled as ‘collocations’ by O.O.C.D thoroughly fulfill the third criterion, high frequency of co-occurrence. If a combination were labelled as a ‘collocation’ by O.O.C.D, but did not score, at least, 40 hits in BNC, it would be classified as a free combination.

The efficacy of using all of the three online tools is that they share the same database, which ensures congruence in results.

2.3. Population and Sample

The sample of the study consists of second year EFL students at the University of ‘Frères Mentouri’ (Constantine, Algeria). These students have been studying English for nine years. They studied written expression for 3 sessions a week, which is an equivalent of 4.5 hours for the last 2 years. The samples analysed in the study are students’ essays that were written during the second semester examinations. The reason behind the selection of examinations’ essays is to ensure seriousness in production since these essays were rated by written expression teachers. The topics were mainly social ones: gifts received in the past, students’ dream house, and consequences of moving out of one’s native town. Such topics are common to students, and they, presumably, possess the appropriate vocabulary and ideas to elaborate on them. 68 randomly-selected essays were analysed, which is fifth of the total number.

2.4. Results and Discussion

The study resulted in 616 word combinations (free combinations, collocations, and idioms). In table 03, proportionate numbers of each category are represented.

<table>
<thead>
<tr>
<th>Total</th>
<th>Free Combinations</th>
<th>Collocations</th>
<th>Idioms</th>
</tr>
</thead>
<tbody>
<tr>
<td>616 (100%)</td>
<td>379 (61.53%)</td>
<td>237 (38.47%)</td>
<td>0 (0%)</td>
</tr>
</tbody>
</table>

It is predictable that learners use free combinations more than collocations, and that idioms will be of a lesser use. The reason is that collocations and idioms are context-bound and of arbitrary use. What is noticeable is that in 66 essays written by the students, no idiom was identified. This could be due students’ ignorance of idioms as essential in improving the quality of writing.

237 collocations were found in students’ compositions. Only 38.82 % of them were correct. Table 04 illustrates the findings.
The Correlation between Language Proficiency and the Use of Lexical Collocations by Algerian EFL Undergraduate Students in Essay Writings: a Case Study of Verb+ Noun Collocations

Table 04: Frequency of Acceptable/Unacceptable Collocations in Students’ Compositions

<table>
<thead>
<tr>
<th>Composition</th>
<th>Acceptable Collocations</th>
<th>Unacceptable Collocations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>92 (38.82 %)</td>
<td>145 (61.18 %)</td>
</tr>
</tbody>
</table>

The high percentage of erroneous collocations shows clearly that second year EFL learners at Frères Mentouri university lack collocational competence. Although students were given freedom to choose topics to write about, they were unsuccessful in using the appropriate V+N combinations -based on the context of writing. These results open speculation on students’ ability to use the most common type of lexical collocations (V+N) if writing topics were more difficult and more science-based.

According to Kuo (2009, p. 149), there are three main causes for collocational errors. They are approximation, synonymy, and negative transfer (also known as the interference of the mother tongue). Approximation refers to the incorrect use of some words and structures, which have common semantic features that lead to get the same meaning. An example is the of ‘fail’ to say ‘fell’. Synonymy is the use of synonyms which do not allow collocability in the word combination. For example, ‘make’ can be mistakenly substituted by ‘do’ in (do/make a mistake). The influence of the mother tongue or the use of a foreign language is called negative transfer (ibid). The learner literally transfers a sentence from his mother tongue to the target language. For example, ‘قهوة بالحليب’ is not ‘milk coffee’, but, ‘white coffee’.

The ratio of subcategorized errors was calculated, and yielded the results in table 05.

Table 05: Ratio of Subcategorised Errors in Students’ Compositions

<table>
<thead>
<tr>
<th>Source</th>
<th>Approximation</th>
<th>Synonymy</th>
<th>Negative Transfer</th>
</tr>
</thead>
<tbody>
<tr>
<td>Frequency</td>
<td>12 (8.28%)</td>
<td>51 (35.17%)</td>
<td>82 (56.55%)</td>
</tr>
<tr>
<td>Total</td>
<td>145 (100%)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Expectedly, the main cause for learners errors was the influence of their mother tongue. Students wrote combinations such as ‘ride the bus’ (يركب الحافلة) instead of ‘take the bus’. Synonymy was the second cause. This shows that students are not aware of arbitrary nature of collocations, as they based their word choice on free semantic substitutability, rather than restricted one. On the same note, the results indicate that students do not differ between free combinations and collocations in terms of word selection and choice. Approximation scored a low frequency as students showed high semantic competence in differentiating between words.

Delexical verbs were given special attention in the study because their semantic less-significant nature. The known delexical verbs are ‘have/take/make/give/go/do/get’. Table 06 illustrates the frequency of delexical and non-delexical verbs in students’ compositions.

<table>
<thead>
<tr>
<th>Table 06: Frequency of correct Delexical/Non-delexical Correct collocations in Students’ Compositions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Composition</td>
</tr>
<tr>
<td>Frequency</td>
</tr>
</tbody>
</table>

The significantly high number of delexical verbs shows clearly that students have a poor lexical competence. Their reservoir of lexis contains very few vocabularies; the reason that compels them to use delexical verbs to compensate for their ignorance of the appropriate words to be used in the given contexts. The mostly used delexical verb was ‘have’. Apparently, students use it as a direct substitute for the verb they do not know. Accordingly, students are obliged to change the whole structure of their sentence—which might result in a cluttered and unparalleled one- to use the verb ‘have’. For example, instead of saying ‘the beauty of air travel consists in its speed and ease’, the sentence becomes ‘air travel has its beauty in speed and ease’. The stylistic difference between the two obvious, as the first one is concise and paralleled—which corresponds to the nature of English--; the second sentence contains wordplay, which clutters it.

3. Pedagogical Implications

Several suggestions can be made based on the results obtained from the study. Students’ awareness should be raised toward the importance of collocations in writing. Without proper use of collocations in one’s writing, the quality of the production will significantly degenerate. Learners must be aware that language mastery is based on both correctness and appropriateness.
Structuring a correct sentence guarantees safety, but, choosing the right combinations for the appropriate context will make one’s writing sound native-like. Learners, also, have to be aware that collocation runs through the whole of the English language. No piece of natural spoken or written English is totally free of collocation. Basic intelligibility is not sufficient for the learner to reach a native-speaker-like competence. A student who talks about ‘strong rain’ instead of ‘heavy rain’ might be understood; however, he will create disruption in communication. Poor collocation in exams is also likely to lead to lower marks in examinations. The learner has to know that language that is collocationally rich is also more precise.

On a different note, it is undoubtedly important to include collocations in students’ syllabi and textbooks. Teachers have to put emphasis on the teaching of collocations as idiosyncratic word combinations governed by arbitrariness, restrictedness, and high frequency of co-occurrence. They need to shed light on the three levels of compositionality (free combinations, collocations, idioms) with description of each level and its features. Collocations need to be included in every respect of foreign language teaching, as they are divided into: lexical and grammatical. In the grammar lectures, grammatical collocations (colligations) can be taught through description of which prepositions combine with which nouns, for example. Lexical collocations can be taught in written expression and oral expression. Teachers can highlight commonly used miscollocations in writing, and provide the learners with lists of them. In oral expression, constructive feedback –direct or indirect- can enhance learners’ collocational competence. Reading tasks within the classroom can prove useful by making students identify collocations in texts. Reading, in general, is advisable to improve knowledge of collocations.

**Conclusion**

This paper sketched the collocational status quo of the Algerian EFL learners at ‘Frères Mentouri’ university. It confirmed the hypothesis that Algerian EFL learners have a low level of mastery of Verb + Noun lexical collocations. As a whole, it is to be concluded that considering all word combinations as collocations is incorrect. It is now admitted that identifying collocations from other word combinations is difficult due to their vague nature that falls between free combinations and idioms. The study stated that in order to be able to classify a given word combination as collocation, three criteria –or at least two- need to be met: restricted sense, restricted substitutability, and high frequency of co-occurrence. The obtained results show the indispensability of collocational competence in achieving efficiency in essay writing, therefore, language proficiency. The majority of the essays which contained numerous mistakes were graded from average to low. The study
encourages teachers and instructors to give more importance to teaching collocation as an inseparable factor to becoming a pseudo-native speaker. On the other hand, not only do learners need to be exposed to correct collocations, but, to erroneous ones, too. This is because the trial-error mechanism used by L2 learners cannot be applied to ready-made chunks. To sum, it is necessary to carry out further research in this field of study through the investigation of other types of collocations.

References