The Learner' Levels of Collocation Knowledge and Idiomatic Expressions: Exploring the Relationship of Acquisition of Two Items of Phraseology

Thamer Alharthi*

Faculty of Arts and Humanities, King Abdulaziz University, PO box 21589, Jeddah, Saudi Arabia
Corresponding Author: Thamer Alharthi, E-mail: talharthi@kau.edu.sa

ABSTRACT

Whilst there is an ever growing bulk of research focused solely on either single word items or idioms (formulaic sequences as they are usually known) through explicit instruction, this article contributes towards a somewhat severely under-researched field in English as a Foreign Language (EFL) context: the importance of incidental learning of collocation knowledge in comprehending idiomatic expressions within an authentic classroom setting in Saudi Arabia. To test this assertion, the study comprised a sample of 34 Arabic-speaking learners of English taking summer BA English courses. Learners' familiarity of idiom was measured by a means of a background questionnaire which was administered at the beginning of the study. In addition, they were tested by examining the comprehension of the figurative meanings of 15 collocations and the non-compositionality meanings of 15 idiomatic expressions occurring in their EFL coursebook. Learners completed a cloze receptive task and an active recognition task in which each contained short texts composed of 15 collocations and 15 idiomatic expressions respectively. A closer examination of the data reveals that the learning of idiomatic expressions was significantly lower than that of collocations even when target idiomatic expressions were supplemented by contextualized information. It is argued that most learners' attitudes towards the identification of idioms indicated that they are not familiar with multiword items such as idioms. Interestingly, the results demonstrated that the knowledge of collocations was effective in the learning of idioms. The article concludes by a thorough discussion of these results in relation to their pedagogical implications.

Key words: Collocation, Idiom, Incidental Learning, Figurativeness, Non-compositionality

INTRODUCTION

Within classroom focused research, the acceleration of interest on the value of vocabulary knowledge has flourished rapidly in recent years as a prominent and fruitful field in second language acquisition (SLA). Early pedagogically oriented research has been primarily dedicated to produce the effects of both implicit and explicit word learning i.e. through inferencing in context via reading or other input conditions and through direct study of language items (Laufer, 2009; Schmitt, 2008; Nation, 2013). Despite the considerable empirical evidence for incidental and intentional learning of lexical knowledge, the majority of the literature on word learning has concentrated almost entirely on the acquisition of single word items. In addition, extant studies have demonstrated the usefulness of single word items and other classroom learning activities such as reading (Laufer & Aviad-Levitzky, 2017; Pigada & Schmitt, 2006; Webb & Chang, 2015), listening (Alharthi, 2016; van Zeeland & Schmitt, 2013; Staehr, 2009; Matthews, 2018; Matthews & Cheng, 2015; Wang & Daller, 2017), writing (Folse, 2006)

and speaking (Koizumi and In'nami, 2013; Uchihara and Saito, 2016), highlighting the relative dearth of empirical research on the level of the student classroom knowledge of formulaic sequences such as collocations and idioms.

Hill (2000) could be right in reporting the emerging view that whenever vocabulary is mentioned, it will always be particular to individual word items among teachers and students. Apparently the knowledge of vocabulary is more than individual word items which reflects unique multi-word expressions acting in a more distinct meaning in a discourse (Wary, 2002). The study of formulaic sequences opens a vast area of investigations and appears to promise much in return for learning, (Coxhead, 2015), which is not surprising since 'The fact that language is largely formulaic' (Meunier, 2012, p. 111). Lexical knowledge is made up of individual words and formulaic sequences, also known as multi-word items including co-occurrence of two words collocations (Sinclair, 1991) such as *running water, blow your nose*, idiomatic expressions (Howarth 1996) e.g. *as good as gold, lay down the law* that fulfill the same role as individual words. We shall adopt the umbrella term 'formulaic sequences' in the current

paper. That is in many ways; formulaic sequences play an essential part in conveying the comprehension and establishing the unitary meanings as they are commonly cited as being the key of linguistic competence (Conklin & Schmitt, 2008; Siyanova & Martinez, 2015; Wray, 2002). Although the development of formulaic sequences is reasonably expected to chart in native speakers, it tends to be slow and lag behind for L2 learners (Schmitt, 2010). It would appear good news; therefore, that formulaic sequences help learners to improve their language proficiency (Keshavarz & Salimi, 2007; Hsu and Chiu, 2008). Moreover, the knowledge of formulaic sequences fosters reading (Alharthi, 2018; Pellicer-Sánchez, 2015); oral fluency (Wood, 2006). On a negative note, nevertheless L2 learners struggle with formulaic sequences for production and comprehension standpoints.

As we mentioned earlier, formulaic language is a notion that ranges from word combinations to metaphoric idioms, the knowledge of collocation will be measured in the present study as one variable to lead to the acquisition of the meaning of idioms. Collocations are defined by (Nesselhauf, 2005, p.1) as "arbitrary restricted lexemes combinations such as *make a decision or fully aware*". The fact that collocation knowledge is a distinguishable indicator of fluency (Nesselhauf, 2003); they poses problems for L2 learners (e.g. Barfield, 2003; Wolter & Gyllstad, 2012). Empirical research on collocation acquisition from incidental exposure in comprehension revealed low uptake rate of learning with about 33% (Macis & Schmitt, 2016) and 56% by (González-Fernández & Schmitt, 2015). The type of task is also highly likely to influence the rate of learning of collocations in specific contexts. For example, Alharthi (2019) showed that his participants' rate of learning of verb-noun collocations- nearly one third (34.1%) was too low in terms of aural production measurement.

Within the umbrella of formulaic sequences, namely idiomatic language poses the most challenging type of formulaic sequences for language learners to master. Although idioms, the focus of the present study compose of single word items, the meanings of these constituent parts cannot always be understood. Due to the multimedia core features and characterizations of idioms, conceptualizing the construct of idiomatic knowledge is notoriously problematic for applied linguists. Generally, idioms are defined as everyday expressions whose meanings cannot be predicted from the meaning of the sequence of elements that constitute them in contexts. As this is a rather simplified criterion, such phenomena may not be easy to boost the number of idiomatic expressions and are therefore considered barriers for non native English speakers.

Phraseology scholars have defined idioms based on different criteria, the most primary of which is what routinely has been acknowledged as the non-literality or non-compositionality (Cornell, 1999; Cooper, 1999; Fernando & Flavell, 1981). That is an expression is non-literal or non-compositional if it consists of the impossibility for the reader of getting the meaning of the idiomatic expression from its internal parts. A clear example is *kick the bucket* in which *kick* implies the movement of the foot to hit something

suddenly and violently and *the bucket* refers to a round metal container with a handle attached to its sides. However, by combining these words together, we would not be able to identify the meaning as 'passed away'. On this concept, Moon (1998, p. 44) defines idioms as "The degree to which a multiple-word item cannot be interpreted on a word-by-word basis, but has a specialized unitary meaning". This non-literal criterion is also called semantic opacity in which Zyzik (2011, p. 415) noted as "many idioms previously considered to be semantically opaque or arbitrary are actually decomposable and/or conceptually motivated". Overall, it is difficult to draw any far reaching definitions if we consider various criteria and theories that exist among idioms. In the context of the present study, a more infinite definition is operationalised and adopted where the characteristics of non-literality and non-compositionality of idiomatic expressions are included as target items.

The issue of learning vocabulary is an important goal to improve L2 use and to facilitate successful L2 language competence (Nation, 2001; Schmitt, 2010; Kim, 2016; Martinez & Murphy, 2011). While the present educational policy in Saudi Arabia (SA) encourages for promoting awareness and learning foreign languages which by all means includes various aspects of language knowledge, it is no wonder, therefore that a fundamental to this goal is acquiring formulaic sequences such as collocations and idioms. However, it is still an open question whether the learners' knowledge of idiomatic expressions can be predicted from the meaning of the collocations. This is an empirical question that existing studies has not tackled yet. In fact no studies, to the best of our knowledge, have been conducted examining the combined effects of collocation knowledge and idiom knowledge in EFL setting. An underlying rationale for the current study is to fill a gap in research by exploring the relationship between the knowledge of collocations and idiomatic expressions in the classroom. In order to better understand the development of L2 idiomatic expressions and collocations, it is worth revisiting the findings of relevant literature on the learning of these two units of phraseology.

LITERATURE REVIEW

Idiom Learning in Non-native Speakers

There has been wealth of studies that approached L2 idiomatic language from different methodological perspectives, namely psycholinguistics i.e. processing of idioms and pedagogical treatments i.e. teaching idioms (e.g. Conklin & Schmitt, 2008; Cieślicka, 2006; Underwood, Schmitt & Galpin, 2004; Boers, 2000; Chen & Lai, 2013) With respect to learning formulaic sequences, idiomatic expressions have been controversial and challenging that even highly L2 learners avoid using non-compositional idioms (Grant & Bauer, 2004; Siyanova & Schmitt, 2007). Grant and Bauer (2004, p. 50) conclude that "understanding figurative language involves taking a compositional untruth and extracting probable truth from it by an act of pragmatic reinterpretation", suggesting that interpreting the meanings of non-compositional idioms involve both linguistic competence and

The Learner' Levels of Collocation Knowledge and Idiomatic Expressions:
Exploring the Relationship of Acquisition of Two Items of Phraseology

41

pragmatic competence. An emerging number of studies have investigated the effect of context on L2 idiom acquisition. For instance, Katsarou's (2013) study involved the participation of sixty Greek EFL high school students where they have to identify five types of idioms, one of which is related to contextual guessability. The results demonstrate significant correlation between the success of idiom identification and the students' L2 level of proficiency they attained. That is the higher scores they obtained in the university of Cambridge examinations, the more they were able to notice idioms in a text context. In a more recent study, Türker (2016) observed the effect of L1 conceptual and linguistic knowledge and L1 frequency on L2 comprehension of figurative idioms. She presented fifty four Korean metaphorical expressions in one decontextualized and two contextualized tasks to thirty four L2 Korean learners whose native language is English. With the context condition, the findings indicate that L2 learners' performance did not differ between frequent and infrequent metaphors; leading to the conclusion that context plays a significant role in improving the learners' comprehension of L2 metaphoric expressions.

Another area of enquiry within the cognitive linguistics approach is to identify the metaphoric aspects of idioms and their levels of retention in explicit classroom situation. Boers (2000) examined the metaphor awareness in an experimental study involving 118 Dutch intermediate EFL learners. The experimental group received notes where 13 English figurative expressions were organized according to metaphoric themes. The author reported that metaphoric awareness is enhanced to interpret figurative usages of expressions and the results supported pedagogical approach to raise awareness among learners in the identification of metaphoric themes. Some empirical evidence adopting the cognitive approach was spotted by (Chen & Lai, 2013) who present a study with the focus on the role of conceptual metaphor with Taiwanese EFL learners at college level. Due to the lack of massive amount of exposure to the target language, Chen and Lai (2013, p.18) concluded that using conceptual metaphors for a typical EFL context is "dependable clues in comprehending and learning idiomatic expressions".

The aforementioned pedagogically driven studies furnished evidence that help learners to develop the acquisition of L2 idioms but they did not offer features of incidental learning. The focus on the present study is rather on the incidental receptive learning of idiomatic expressions that are influenced by the knowledge of collocations.

Collocation Learning in Non-native Speakers

As we mentioned earlier that several scholars have pointed out the importance of the knowledge of collocation which offers valuable input for developing receptive fluency (Wray, 2002; Nesselhauf, 2003). In the published literature on L2 phraseology, nevertheless one critical issue of the relationship between the EFL learners' knowledge of collocations and idiomatic expressions has surprisingly been left unaddressed. In the absence of intentional treatments, particularly in the context of EFL, it is observed that incidental learning of collocation might be most effective. Focusing on implicit

collocation learning, Alharthi (2018) examined the degree of EFL learners' productive task of verb-noun combinations at three word frequency bands affected their knowledge of collocations while reading. The results showed that incidental learning of collocation suited to promote the development of productive knowledge and that the frequency of occurrence is more effective at fostering such development. In a laboratory setting, Durrant and Schmitt (2010) demonstrated the effect of the repetition on the learning of adjective-noun collocations by advanced ESL learners. The results revealed that multiple exposures led to a development of collocations, in particular when the target low frequency adjective-noun combinations had been re-encountered in the same sentence. Sonbul and Schmitt (2013) argue that two learning conditions of implicit and explicit knowledge would affect the learning of collocations by a group of thirty five native speakers at a British university. They found that all conditions led to remarkable long term gains from recognition and recall learning over a relatively short period of time. Very often the approach taken influences the caveat that most collocation research to date have mainly considered literal meanings e.g. *powerful computer* (Macis & Schmitt, 2017). However, the present study will share the same concept of idiomatic expression i.e. the learners' knowledge of figurative collocations which are not inferred from the component words e.g. *hot ticket* in which the meaning covers something that is currently very popular or well liked. It was argued that not all collocations will always carry the feature of literal meanings but figurative meanings of collocations "will make up substantial percentage of the total occurrences (i.e. perhaps one-fifth to one-quarter" (Macis & Schmitt, 2016, p. 15). Most importantly, a review of the relevant literature here have provided some insight into the learning of L2 learners' knowledge of idiom as well as collocation but lacked a solid account whether their knowledge of collocation may contribute to development of idiom knowledge. A further novelty of the current research design is the use of multiple formats of measurements which in turn depict a concise picture of phraseology development by L2/EFL learners (Webb, 2005). As such, the present study is motivated by asking the following questions:

1. What degrees of L2 learners' knowledge of the collocation and idiomatic expressions gained from reading?
2. How does the knowledge of collocations relate to the knowledge of idiomatic expressions?

METHODOLOGY

Participants

The participants were thirty-four male native speakers of Arabic from a Saudi public institution, King Abdulaziz University (KAU). They were from two intact classes of third year English majors and their ages ranged from 21 to 23 years-old. At the time of study, they were enrolled in one of the two summer courses (Reading 1 and Research Methods) which they typically dedicated a more intensive format (two to three times a week for 12 hours for two modules) than a few hours a week during the regular academic

year to studying English. They agreed to take part in the study since they earned points toward a certain percentage of their course grade. All participants had been learning English in a classroom setting since the first year of intermediate school, when they were 14 years-old. Within the context where the study was carried out, they were estimated to be of a proficiency level of around A2 to B1 on the Common European Framework of Reference (CEFR) for languages or had achieved an equivalent score approximately 75 percent upwards, based on the in-house English Department entrance examination (oral and written).

Materials

Preparation of pre-self reports on learning idiomatic expression

Taking into account the varieties of constructs that shape up the notion of idiomaticity mentioned in the literature, it was thought crucial to prepare a background questionnaire to ascertain the views and beliefs of the learners in relation to the learning of idiomatic expressions. The questionnaire contains the following items: their own definition of idiom, their own examples encountered in classrooms, their own recognition of idiomatic expressions in their reading and listening activities, and their own rating on the importance of learning idiomatic expressions as a type of lexical item. We believed that the information obtained in this preliminary questionnaire would be useful especially in the interpretation for the later results of the current exploratory study.

Target idioms and collocations: sampling procedure

We extracted the target idiomatic expressions and collocations from a range of different sources. To begin with, a preliminary list of 30 idioms was generated on the basis of non-compositionality nature from the *Collins Cobuild Dictionary of Idioms* (2002); The *Free Dictionary* (http://idioms.thefreedictionary.com) which was drawn from the *Cambridge International Dictionary of Idioms and the Cambridge Dictionary of American Idioms.* Moreover, we wanted an indication of learners' knowledge of idiomatic expression; therefore we consulted a textbook *Pathways, Reading, Writing and Critical Thinking* (Blass, Vargo & Folse, 2014) which is used for the BA program. The textbook comes in series of National Geographic that aims to develop learners' reading skill to achieve academic success. The course-book was carefully inspected to include transparent idioms that appeared in blue and bold. For the purpose of the present study, it was important to spot target idioms through the use of reading texts besides presenting the item candidates in semi-contextualized sentences that appeared in the textbook exercises and activities. To increase ecological validity, we manually screened the materials in the learners' textbook to check the activities formats we intend to apply to our study test that are comparatively used in the real classroom situation. Out of the initial 38 candidates determined, a final list of 15 potential idioms were compiled and included in the current study. This controlled selection of target idioms provides the current study with validity, in which all

participants had possibly exposure to these fifteen idioms in their relevant textbook.

In addition, the present study draws on both the Corpus of Contemporary American English (COCA) (Davies, 2008) and the textbook *Pathways, Reading, Writing and Critical Thinking* (Blass, Vargo & Folse, 2014) the learners had been studying in their BA program to extract 28 figurative collocations. Each selected collocation was checked through COCA and the learners' textbook to ensure that collocation items are familiar and already known to the learners. Besides if collocation candidates do not incorporate figurative meaning in both sources, they are eliminated as a result of this inspection. Another key feature of sampling that collocation items appearing in the textbook was identified based on their salience by boldfacing or existence in the textbook exercises. Following the above criteria, 15 collocation candidates were shown to be well known to the present participants and hence were included in the study. It is worth noting that the learners' textbook we surveyed had collocation and idiomatic expressions exercises in almost every unit.

Assessments tests

Two contextualized tests were designed to assess the participants' knowledge of idiomatic expressions and collocation spotted in their study textbook. The format of the first test was best described as the active recognition task (see Laufer & Goldstein, 2004 for further discussion on degrees of vocabulary knowledge) which was created based on the traditional construction of *Vocabulary size Test* by (Nation & Beglar,2007). The test comprised a 15-item multiple-choice format in which for each test item, the test takers were instructed to choose from four-definition options the target idiomatic expression that best fits the context. To make sure that examinees do not make random guesses, the option 'I don't know' was inserted. Items being measured were presented in short contextualized sentences to ensure ecological validity environment for the recognition task. A sample of idiom item from this task is shown below:

Lately, however, this theory has **fallen on hard** scientific **times**. Several recent studies have undermined the link between longevity and caloric restriction.

a. Travel toward
b. Use as a starting point
c. Not make assumption
d. Begun a difficult time
e. I don't know

The format of the second test was a cloze test to measure the participants' receptive knowledge of 15 target collocation presented within written input. For each item, the participants had to complete the gap with the target collocation out of the four options provided. To better exclude the guessing effects, the item 'I don't know' was added. A sample of collocation item from this task is displayed below:

In addition, nearly a third of the countries of mainland Africa are _____, and the only African river navigable from the ocean for long distances inland is the Nile.
a. Landlocked
b. Impressed by

c.   Caused to happen
d.   Lead with
e.   I don't know

   The two tests were then administered to a native speaker of English and a non-native member of teaching staff to assess content validity and to see that items are not too hard to tackle and that they could be comprehended without ambiguity. Judges were able to answer most questions in both tests with no chances of guessing, indicating that the test items were well written. The wording of other vague or difficult tests items were amended accordingly. More importantly, the non-native teaching staff member confirmed the likelihood of the tested items to the participants in their study materials

Procedure and Scoring

The entire procedure of the data collection was held from week 2 to week 3 of the summer semester on the premises of KAU. First, a pre-data collection was carried out, consisting of the background questionnaire which was administered to all participants at the beginning of their summer semester. The time it took for each participant to fill out the questionnaire ranged from 30 to 40 minutes. Testing sessions were then conducted on two separate days. The participants received the cloze receptive test during their normal class which was completed in approximately 45 minutes. To eliminate fatigue effects and properly measure the participants' knowledge of idiomatic expressions, the active recognition test was administered on the following day and the completion of the task took roughly one hour. It will be recalled that the maximum score for each assessment test was 15 as each participant was measured on 15 target collocation and 15 idiom items. As the present study employed multiple-choice formats, each item was scored dichotomously (correct/incorrect) with a correct answer received one point and an incorrect answer scored of zero. Items left unanswered were marked incorrect and hence, received zero.

RESULTS

Table 1 sumps up the descriptive statistics (means and standard deviations) for the cloze receptive test and the multiple-choice recognition test of collocation and idiomatic expression knowledge respectively. On the basis of the findings obtained, the participants learning knowledge for collocations appear to be higher than the learning knowledge for idiomatic expressions. With regards to specific collocation, the percentage scores on the cloze receptive test ranged from 45% to 100% and the means was 14.77. Although the results confirmed that even when idiomatic expressions were measured in multiple-choice format, the participants' performance on such task was lower than the cloze receptive one, indicating that their knowledge of idiomatic expressions was considerably weak and problematic with no individual scored better than 70%. While the maximum scores of the two tests did not differ significantly, the minimum score on the multiple-choice recognition idiom test was dramatically lower than that on the cloze receptive collocation test.
   To further explore any difference between the participants' scores of collocation and idiomatic expressions, a

paired sample *t*-test was conducted. There was a statistical significant difference between the type of word knowledge i.e. collocation vs. idiom (*df*=33, *t*= -2.76, *p* <.001). This demonstrates that the extent of incidental learning is likely depending on the level of lexical knowledge that is measured. It is worth to consider the absolute pattern of learning collocation and idiom (see Figure 1). More noteworthy is that the participants are particularly lacking in knowledge of idiom (*M*=8.19) even with the help of the contextual cues which may be attributed to the participants' comprehension of idiom as individual words rather than as complete units. In the discussion section, we further discuss the results of this study in light of these observations.
   However, a question remains unanswered at the moment, whether the participants' collocation scores can indicate their knowledge of idiomatic expressions. To inspect this, a multiple regression analysis (MR) was conducted with scores of the cloze receptive collocation task as independent variable on the scores of the multiple-choice recognition idiom task as dependent variable. The MR analysis can be shown in Figure 2. It became apparent that the cloze receptive collocation test offered empirical evidence of the relative strength contribution to the idiomatic recognition test scores. Although the collocation test scores made a small 5.3% of the variance when entered into the regression equation, resulted in a statistically significant to the predictive power of regression model (Beta=.483, *p*<.001).

Table 1. Comprehension of collocations and idiomatic expressions

Study tasks	N	Descriptive statistics of scores across tasks			
		Mean	**Min %**	**Max %**	**SD**
Cloze receptive collocation test	15	14.77	45.00	100.00	1.52
Multiple-choice recognition idiom test	15	8.19	18.00	70.00	2.21

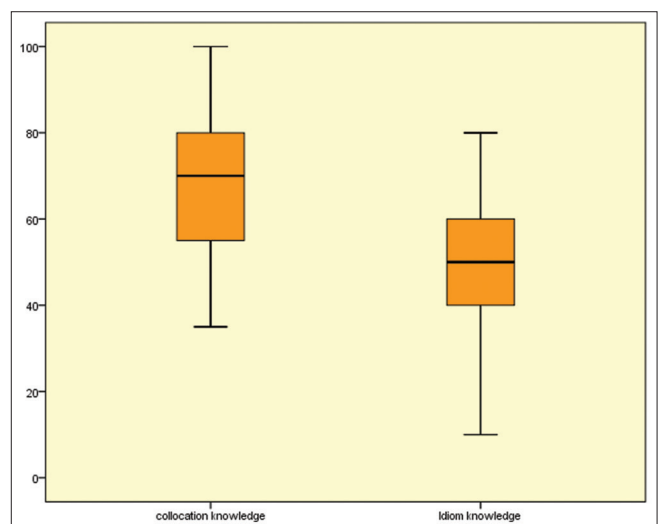*N*=indicates the total number of tested items for each task

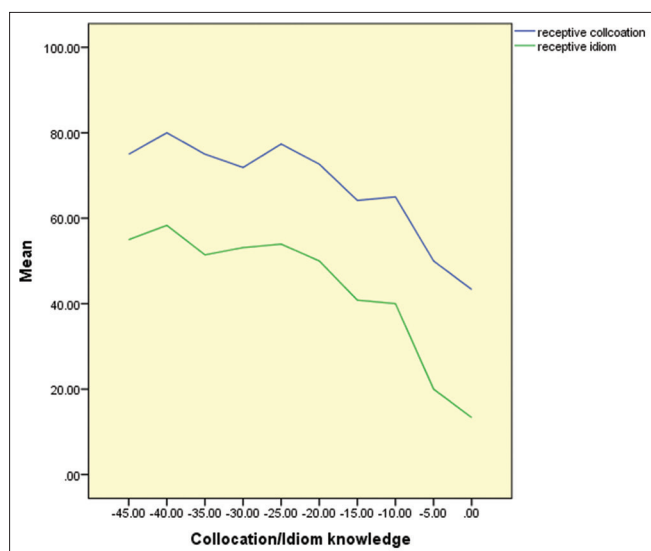**Figure 1.** Collocation vs. idiomatic scores comparison

Figure 2. Multiple regression analysis using collocation knowledge as predictor variable and idiomatic expression learning as criterion variable

The most interesting and encouraging outcome of this analysis is that the direction and the nature of the relationship is positive, indicating that the more collocation knowledge gained by the participants, the higher their scores on the idiom test. In other words, the number of correct figurative collocations supplied by the current participants made a unique contribution 74% to their correct non-literal idiomatic expressions and produced a significant chi-square value ($\chi2$=6.67, df=2, p <.001).

DISCUSSION

Much research in the domain of phraseology in the past has looked at the relationship between the knowledge of single word items and the learning of idiom. The present study differs from previous studies by exploring the contribution of co-occurring figurative collocations that lead outcomes of learning non-literal idiomatic expression. Therefore, in answer to the first research question, the findings indicate that the participants have remarkable figurative collocation knowledge in comparison to that of idiomatic expression with the vast majority scoring between 75% and 80% and this variation is graphically shown in Figure 1. This result suggest that learning is subject to different degrees and variations of knowing the meaning-form link of individual word items to produce the meaning of lexical chunks in a semantically appropriate way in a given sentence such as the knowledge of collocation, an observation that were acknowledged by (Nation, 2001; Schmitt, 2010). That is learners are less prone to produce incorrect figurative collocations which are also a strong indication for demonstrating the meanings of the two component words and how these co-occurrence associations are combined together.

Although the knowledge of figurative collocation is highly useful in reading fluency, their certain meanings are opaque that cannot be easily comprehended in reading input (Boers, 2000; Martinez & Murphy, 2011; Macis & Schmitt,

2017). Clearly in the current study, the presence of context was advantageous and able to strengthen the interpretation of the figurative collocations and led to overly impressive amount of learning, hence showed evidence of incidental collocation learning while reading. The results are in line with recent similar studies on the learning of collocation in classroom settings (Alharthi, 2018; Boers et al, 2014; Durrant & Schmitt, 2010; Sonbul and Schmitt, 2013).

However, the multiple-choice task which measured the present participants' comprehension input using idiomatic expression revealed significant lower scores than that of the collocation test. These results on their own provide evidence that the participants are not able to guess the meaning of idiomaticity, though they were tested at an active recognition task. It should be reminded here that for the comprehension measure of idiomatic items, an option of 'I don't know' was added to the 4 multiple-choice options to reduce guessing. It seems apparent that the feature of non-compositionality of the idiomatic expressions increased the level of difficulty to understand them, though they are supported by contextualized cues. That is the participants tend to be poor at the idiomatic level of knowledge and hence they do not seem to be in a position to benefit from the context surrounding the target idiomatic expressions. It could be possible that the long strings of language such as idiomatic expressions biased the participants to acquire figurative senses of idioms (as opposed to short co-occurrence patterns such as collocations) while reading. Regrettably, the current research does not allow us to be pretty sure with regards to the impact of the length manipulation for multi-word sequences since we did not measure such variable. A plausible explanation is that as participants pay attention to the constituents of the target idiom not as intact whole but rather on individual basis, such the risk of wrong and undesirable associations is increased and caused additional obstacles for them especially that these expressions are non-compositional (Zyzik, 2011). This supports Wray (2002) claim that learners attend to learn formulaic sequences (presumably idioms) in the context sentence word by word.

As mastering various aspects of vocabulary knowledge is important as emphasized by Nation (2001), the participants in the present study evidently appeared to have insufficient knowledge of idiomatic expressions. This explanation is best exemplified by the results of (Kim, 2016; Martinez & Murphy, 2011; Park & Chon, 2018) in which L2 learners' knowledge of idioms was lagging behind that of individual word items. This then raises the question to what extent the learners are intentionally provided with examples of idiomaticity in the classroom. Perhaps the most obvious way of interpreting this result is that in EFL contexts and in other parts of Asian world, learning idioms are ineffectively presented by instructors and our learning context is not the exception. Presumably a social variable such as interaction with native speakers or even with high proficient English learners outside of class are likely to have contributed to the learners' comprehension of idiomatic knowledge. Of course in reality, EFL classrooms such as the present one is rather less conductive in terms of language input and opportunities for practicing the use of the idiomatic expressions. These findings

The Learner' Levels of Collocation Knowledge and Idiomatic Expressions:
Exploring the Relationship of Acquisition of Two Items of Phraseology

45

suggest that EFL learners need to be explicitly trained and instructed in promoting meaningful interactions that lead to full understanding of idioms in meaning-focused input.

Another issue that merits attention is to refer back to the participants' views and attitudes on the importance of learning idiomaticity as an aspect of lexical knowledge. An examination of the study background questionnaire demonstrated that 76% of participants were not able to identify the notion of idiom while 24% said they knew what idiomaticity stands for. This can be shown by few examples of idioms that were reported in their questionnaires and simultaneously appeared in their reading and listening textbooks such as *in essence, not to mention, make assumptions about* and whose meanings can overly be comprehended form the meaning of their internal segments. This particular result although discouraging, serves an important implication for university EFL instructors that the exposure and practice to idiomaticity is rather limited in the classrooms. In other words, learners have insufficient opportunities to encounter idiomatic expressions, if any will explicitly be presented to them. This again helped confirm the above assertion that mean scores on the idiom test were significantly lower than the collocation test. Interestingly, the participants' responses about the importance of learning idiom expressions were measured by means of four rating scales: (1) 'Very much'; (2) 'To a certain extent'; (3) 'Not much'; (4) 'Not at all'. The participants had held the impression that learning idiomatic expressions are very important 83% while the remaining 17% responses indicated that idiomatic expressions are important to a certain extent. The number of responses that indicated little or unimportant were zero. Participants viewed the benefits of learning idiomaticity as they contribute to the increase of native-like features of the language, to help one become immersed in the target culture, and to increase fluency. These attitudes and views towards the usefulness of learning idioms could be the answer to the main concern about allocating time for idiomatic language explicit instruction, which is best, provided by EFL instructors and textbooks writers and would be rather welcomed by EFL learners. On these grounds, it seemed probable that the instruction of idiomatic language should be further employed and expanded in certain circumstances when the learners obtain high levels of language competence. This argument is in line with the principles expressed by Park and Chon (2018, p. 14) as follows: "Also, learning idioms should not be put off until L2 learners reach advanced levels of proficiency".

Despite the low participants' means on the idiomatic expression task, there are promising findings that need to be considered in the context of this study. In an attempt to account whether the learners' higher collocation scores imply more idiomatic knowledge, the MR analyses empirically exhibited statistical significant interaction between the learners' scores for collocation and their scores of idiom. In other words, the cloze receptive collocation test served as a kind of vocabulary assessment that contributes to the learners' knowledge of idiomatic expressions. A possible interpretation of this result, though remains speculative, is that some of the target collocations (e.g. *turning up*, *in the abstract*) and idiomatic expressions (e.g. *in essence, sums up*) composed

similarly of very common individual words which in turn may have helped the learners extended their comprehension to idiomatic expressions. As none of previous empirical research make specific claims about the effectiveness of the collocation knowledge on the comprehension of idiomatic expressions knowledge, it is difficult to compare the current findings to the existing ones. Taken as a whole, this finding is useful for vocabulary instructors as it suggests that developing the knowledge of common single words that make up unusual combinations such as collocations and idiomatic expressions is essential to master these different aspects of lexical knowledge and to lead to successful comprehensions for readers and speakers.

CONCLUSIONS AND CAVEATS

This article explored the relationship between knowledge of figurative collocations and knowledge of non-compositional idioms among learners' development of reading comprehension in the target language. Our results indicate that the participants achieved the most impressive vocabulary gains in their comprehension collocation performance. What we had not predicted as it was not the central focus of the present study was the variable of presence of contextualized clues which had some facilitative impact on figurative meanings of collocations and this effect was shown in the cloze receptive task. There is also growing evidence that textual enhancement using explicit approach can help learners understand collocations knowledge as shown by a number of studies by (Bishop, 2004; Peters, 2012; Choi, 2017). This is a question that calls for further research in which figurative meanings of collocations gains obtained from textual enhancement are compared to unenhanced textual input through implicit learning without formal instruction. In this regards, Teng (2019) pointed out that the inclusion of textual information would be a significant predictor of learning individual words and it might be that this approach works better for other categories of formulaic sequences such as figurative collocations as indicated in the present context. By contrast, the participants made the least progress in the active recognition task that measured their non-compositional idiomatic expressions. This suggests that they have failed to recognize the target idioms even though these items were embedded in informative contexts. Specifically this result indicates that textual information as an aid to L2 acquisition may not always be a powerful indicator of idiom uptake, hence the participants were not very much successful in learning idiomatic expressions to a recognition level of mastery. Proficiency as a linguistic variable would add more credits to explore how different categories of formulaic sequences such as non-literality and figurativeness may interact with learners' levels of proficiency but this was beyond the scope of this study. The most encouraging finding of the study was the contribution of the two-word combinations i.e. collocation knowledge to the learning of idiomatic expressions. However as indicated in MR analyses, although the interaction was positive and significant, the receptive collocation test only brought a slight of 5.3% variance on the idiomatic active recognition test. Given this interesting

result as well as the scarcity of similar relevant studies, it is of paramount importance that further research is required to seek the association between the learners' comprehension of these two items of phraseology before any firm conclusion can be drawn.

PEDAGOGICAL IMPLICATIONS

We feel that the findings of the current investigation have been rich enough to propose some pedagogical implications for EFL teaching and learning. The integration of figurative meanings of collocation into reading texts is clearly desirable and effective in removing potential challenges with incidental learning of these target sequences. This particular phenomenon offers some practical implications for teachers, curriculum designers and learners. That is teachers and coursebooks writers should raise the EFL learners' awareness about the use of text content which will surely help in achieving a higher level of accuracy when recognizing these strings of words incidentally. One of the effective caveats is to encourage the learners to study abroad as argued by (Dörnyei et al. 2004; Macis & Schmitt, 2016) in order to acquire formulaic sequences including collocations. However, engagement with L2 social environment does not always work as simply students cannot afford to do so (Schmitt, 2010). As reading is perhaps the most fruitful source for individual items (Teng, 2019), and for various categories of formulaic sequences (e.g. González Fernández & Schmitt, 2015), the knowledge of collocation can be integrated into the curriculum of a language reading course that would likely affect the learners' comprehension of collocation knowledge. Given the apparently low scores of idiom recognition is explained by the nature of idiom type i.e. the non-literality or non-compositionality. Idiomatic expressions with these types of components are proved to be difficult to recognize in a given reading text. As Cornell (1999, p. 7) suggested that "idioms involve a greater teaching and learning load than normal lexis, hence it is necessary to determine what idioms should receive special attention in a determined language course". It is therefore important that coursebooks writers and teachers draw the learners' attentions to the degree of transparency and salience in which guidelines for target idiom instruction as well as the number of frequency of these expressions should be highlighted in the learners' textbook materials (Thi, Rodgers & Pellicer-Sánchez, 2017). Overall, the present research does not only contribute to our understanding of how figurative meanings of collocations are learned incidentally but also offers new insights into the combined relationship between the learning of collocation and idiomatic expression, hence pinpointing an under researched area which likely deserves more attention from instructors and learners.

REFERENCES

Alharthi, T. (2016). Adult learners' acquisition of vocabulary knowledge from explicit listening: A multidimensional construct. *International Journal of Multidisciplinary and Current Research*, *4*, 626–636.

Alharthi, T. (2018). Minding the gap in vocabulary knowledge: Incidental focus on collocation through reading. *Arab World English Journal*, *9*(2), 3–22.

Alharthi, T. (2019). Formulaic Sequences as Predictors of Listening Comprehension: A Contribution to Research into Incidental Learning of Collocations. *Asian EFL Journal, 23*, 180-190.

Barfield, A. (2003). *Collocation recognition and production: Research insights*. Tokyo: Chuo University.

Bishop, H. (2004). Noticing formulaic sequences: A problem of measuring the subjective. LSO Working Papers in Linguistics, 4, 15–19.

Boers, F. (2000). Metaphor awareness and vocabulary retention. *Applied Linguistics*, *24*, 553–71.

Boers, F., Demecheleer, M., Coxhead, A., & Webb, S. (2014). Gauging the effects of exercise on verb–noun collocations. *Language Teaching Research*, *18*, 54–74.

Blass, L., Vargo, M., & Folse, K. (2014). *Pathways 4 Reading, Writing and Critical Thinking.* Boston, MA: National Geographic Learning.

Chen, Y., & Lai, H. (2013). EFL learners' awareness of metonymy-metaphor continuum in figurative expressions. Language Awareness, 21, 235_248.

Choi, S. (2017). Processing and learning of enhanced English collocations: An eye movement study. *Language Teaching Research, 21,(3),* 403-426.

Cieślicka, A. (2006). Literal salience in on-line processing of idiomatic expressions by second language learners. *Second Language Research*, *22*, 115–44.

Collins Cobuild dictionary of idioms. (2002). Founding editor-in-chief: J. Sinclair. 2nd edition. Glasgow: Harper Collins.

Conklin, K. & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics*, *29*, 72–89.

Cooper, T. (1999). Processing of idioms by L2 learners of English. *TESOL Quarterly*, *33*, 233–62.

Cornell, A. (1999). Idioms: An approach to identify major pitfalls for learners. *IRAL, 37*, 1–21.

Coxhead, A. (2015). Replication research in pedagogical approaches to formulaic sequences: Jones & Haywood (2004) and Alali & Schmitt (2012). *Language Teaching, 51*(1), 1–11. https://doi.org/10.1017/S0261444815000221

Davies, M. (2008). *The corpus of contemporary American English: 520 million words, 1990–present*. Retrieved from: http://corpus.byu.edu/coca (April 2019).

Dörnyei, Z., Durow, V., & Zahran, K. (2004). Individual differences and their effects on formulaic sequence acquisition. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 87–106). Amsterdam: John Benjamins.

Durrant P & Schmitt N (2010). Adult learners' retention of collocations from exposure. *Second Language Research* 26: 163–88.

Fernando, C. & Flavell, R. (1981). *On idiom: Critical views and perspectives*. Exeter Linguistic Studies: Volume 5. Exeter: University of Exeter.

The Learner' Levels of Collocation Knowledge and Idiomatic Expressions:
Exploring the Relationship of Acquisition of Two Items of Phraseology

47

Folse, K. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL Quarterly*, *40*, 273–293.

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *ITL International Journal of Applied Linguistics*, *166*, 94–126.

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? Applied Linguistics, 25, 38–61. doi:10.1093/applin/25.1.38.

Hill, J. (2000). Revising priorities: From grammatical failure to collocational success. In M. Lewis (Ed.), Teaching collocation: Further developments in the Lexical Approach (pp. 47–69). Hove, England: LTP.

Howarth, P. (1996). Phraseology in English academic writing: Some implications for language learning and dictionary making. Tubingen, Germany: Max Niemeyer.

Hsu, J.-Y., & Chiu, C.-Y. (2008). Lexical collocations and their relation to speaking proficiency of college EFL learners in Taiwan. *Asian EFL Journal*, *10*, 181–204.

Katsarou, E.C. (2013). Grasping the nettle of L2 idiomaticity puzzle: the case of idiom identification and comprehension during L2 reading by Greek learners of English. In Major Trends in Theoretical and Applied Linguistics 3: Selected Papers from the 20th ISTAL, ed. W. Lavidas, T. Alexiou and A.M. Sougari, 143–58. London: Walter de Gruyter.

Keshavarz, M.H., & Salimi, H. (2007). Collocational competence and cloze test performance: A study of Iranian EFL learners. *International Journal of Applied Linguistics*, *17*, 81–92.

Kim C. (2016). L2 learners' recognition of unfamiliar idioms composed of familiar words. *Language Awareness* 25(1–2): 89–109.

Koizumi, R., & In'nami, Y. (2013). Vocabulary knowledge and speaking proficiency among second language learners from novice to intermediate levels. *Journal of Language Teaching and Research*, *4*, 900–913.

Laufer, B. (2009). Second language vocabulary acquisition from language input and from form focused activities. *Language Teaching*, *42*, 341–354.

Laufer, B., & Aviad-Levitzky, T. (2017). What type of vocabulary knowledge predicts reading comprehension: Word meaning recall or word meaning recognition? *The Modern Language Journal*, *101*, 729–741.

Laufer, B., & Goldstein, Z. (2004). Testing vocabulary knowledge: Size, strength, and computer adaptiveness. Language Learning, 54, 399_436.

Macis, M., & Schmitt, N. (2016). Not just 'small potatoes': Knowledge of the idiomatic meanings of collocations. *Language Teaching Research,* 1-20. DOI: 10.1177/1362168816645957

Macis, M., & Schmitt, N. (2017). The figurative and polysemous nature of collocations and their place in ELT. *ELT Journal, 71*(1), 50–59. https://doi.org/10.1093/elt/ccw044

Martinez R. & Murphy VA. (2011) Effect of frequency and idiomaticity on second language reading comprehension. *TESOL Quarterly* 45(2): 267–90.

Matthews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, *72*, 23–36.

Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System, 52*, 1–13.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics*, 32, 111–129. DOI: 10.1017/S0267190512000128

Moon, R. (1998). *Fixed expressions and idioms in English*. Oxford: Clarendon Press.

Nation, I.S.P. (2001). *Vocabulary learning in another language*. Cambridge: Cambridge University Press.

Nation, I.S.P. (2013). *Learning vocabulary in another language*. 2nd edition. Cambridge: Cambridge University Press.

Nation, I.S.P., & Beglar, D. (2007). A vocabulary size test. *The Language Teacher*, 31(7), 9–13.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implications for teaching. *Applied Linguistics, 24*(2), 223–242.

Nesselhauf, N. (2005). *Collocations in a learner corpus*. Amsterdam: John Benjamins.

Park, J., & Chon, Y. (2018). EFL Learners' Knowledge of High-frequency Words in the Comprehension of Idioms: A Boost or a Burden? *RELC Journal, 7,* 1-16 DOI: 10.1177/0033688217748024

Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques, *Language Learning Journal, 40,* 65–79. https://doi.org/10.1080/09571736.2012.658224

Pellicer-Sánchez, A. (2015). Learning L2 collocations incidentally from reading. *Language Teaching Research*. doi:10.1177/1362168815618428

Pigada, M., & Schmitt, N. (2006). Vocabulary acquisition from extensive reading: A case study. *Reading in a Foreign Language*, *18*, 1–28.

Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, *12*, 329–363.

Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Basingstoke: Palgrave Macmillan.

Sinclair J (1991) *Corpus, Concordance, Collocation*. Oxford: Oxford University Press.

Siyanova, A., & Schmitt, N. (2007). Native and nonnative use of multi-word vs. one-word verbs. *International Review of Applied Linguistics*, *45*, 119–139.

Siyanova, A. & R. Martinez. 2015. 'The Idiom principle revisited'. *Applied Linguistics* 36/5: 549–69.

Sonbul, S., & Schmitt, N. (2013). Explicit and implicit lexical knowledge: Acquisition of collocations under different input conditions. *Language Learning*, *63*, 121–159.

Staehr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, *31*, 577–607.

Teng, F. (2019). Retention of new words learned incidentally from reading: Word exposure frequency, L1 marginal glosses, and their combination. *Language Teaching Research, 22,* 1-28.

Thi, D. L., Rodgers, M.P.H., & Pellicer-Sánchez, A. (2017). Teaching formulaic sequences in an English-language class: The effects of explicit instruction versus course-book instruction. *TESL Canada Journal/Revue TESL du Canada, 34,(3),* 111-139. http://dx.doi.org/10.18806/tesl.v34i3.1276

Türker, E. (2016). The role of L1 conceptual knowledge and frequency in the acquisition of L2 metaphorical expressions. *Second Language Research* 32, 1: 25–48.

Uchihara, T., & Saito, K. (2016). Exploring the relationship between productive vocabulary knowledge and second language oral ability. *The Language Learning Journal*, 1–12. https://doi.org/10.1080/09571736.2016.1191527

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 153–172). Amsterdam, the Netherlands: John Benjamins.

van Zeeland, H., & Schmitt, N. (2013). Incidental vocabulary acquisition through L2 listening: A dimensions approach. *System*, *41*(3), 609–624.

Wang, Y., & Treffers-Daller, J. (2017). Explaining listening comprehension among L2 learners of English: The contribution of general language proficiency, vocabulary knowledge and metacognitive awareness. *System*, *65*, 139–150.

Webb, S. (2005). Receptive and productive vocabulary learning: The effects of reading and writing on word knowledge. *Studies in Second Language Acquisition*, *27*, 33–52.

Webb, S., & Chang, A. (2015). Second language vocabulary learning through extensive reading: How does frequency and distribution of occurrence affect learning? Language Teaching Research, 18, 1–20. Retrieved from http://ltr.sagepub.com/content/early/2014/11/27/1362168814559800.full.pdf+html

Wood, D. (2006). Uses and functions of formulaic sequences in second language speech: An exploration of the foundations of fluency. *Canadian Modern Language Review 63*(1), 13–33.

Wolter, B., & Gyllstad, H. (2012). Frequency of input and L2 collocational processing. *Studies in Second Language Acquisition*, *35*, 451–482.

Wray, A. (2002). *Formulaic language and the lexicon*. Cambridge: Cambridge University Press.

Zyzik, E. (2011). Second language idiom learning: The effects of lexical knowledge and pedagogical sequencing. *Language Teaching Research*, *15*, 413–433.