Macrothink Institute™

International Journal of Linguistics
ISSN 1948-5425
2018, Vol. 10, No. 6

The Processing of Congruent and Incongruent Prepositional Collocations: Empirical Study of Arab-English Bilinguals

Zenah Dajem

English Department, King Khaled University, Saudi Arabia

E-mail: zdajem@kku.edu.sa

Received: November 11, 2018 Accepted: Nov. 28, 2018 Published: December 19, 2018

doi:10.5296/ijl.v10i6.13886 URL: https://doi.org/10.5296/ijl.v10i6.13886

Abstract

Learning and teaching formulaic sequences such as collocation and idioms, is an increasingly important issue for L2 language learners in the field of language teaching, psycholinguistics and applied linguistics. The aims of the present study are twofold: first, to investigate the effect of congruent and incongruent prepositional collocations on twenty Saudi Arab-English bilinguals in the UK; and secondly to examine the influence of frequency on the processing of these formulaic sequences. Twenty native speakers of English were involved as a baseline for comparison. In the present study, the psycholinguistic software E-Prime was used to measure the reaction time and the accuracy of Saudi Arab-English bilinguals' responses while they are reading 80 sentences on a computer screen to judge their grammatical acceptability, that is, whether they are correct English or not. The results in general indicate that for both groups there was no processing effect in terms of the reaction time of collocation types, but there was a significant difference within the non-native speakers in terms of accuracy. However, regarding the influence of frequency, there was a significant effect of high frequent collocations in both native and non-native speakers of English. This may indicate that similarities and differences between L1 and L2 prepositional combinations are not necessarily due to processing effect, but may be due to high frequency.

Keywords: Formulaic language, Grammatical collocations, Lexical collocations, Bilinguals, Frequency, Congruency, Language processing

1. Introduction

Formulaic language comprises a large part of any language not mention the English language, representing 58.6% of English spoken discourse and 52.3% of English written discourse (Erman and Warren, 2000). These figures give us an idea of its prevalence and its role in the competence of other language users as non-native speakers need to be able to comprehend and produce it effectively, as it is used by native speakers.

The term 'formulaic sequences' has been used to refer to a diverse set of linguistic phenomena. It is "deliberately inclusive, and contains a number of different kinds of patterned language" (Schmitt and Carter, 2004: 9). It can refer to idioms, proverbs, phrasal verbs and fixed binomials/trinomials. Wray (2002: 9) claims that there are over fifty terms that can be referred to as formulaic sequences, such as "chunks, collocation, formulaic speech, formulas, multiword units" or "formulae, formulaic expression, formulaic sequences" (Richards and Schmidt, 2002:210).

It is an undeniable fact that formulaic sequences pose many advantages in L1 and L2 learning and teaching. Generally speaking, their importance pertains to both psycholinguistic and socio-functional aspects (Conklin and Schmitt, 2008). The former refer to the idea of holistic processing where they are stored in the memory as a whole unit and give a chance for other linguistic features to be learnt or acquired (Schmitt et al., 2004; Yamashita and Jiang, 2010); while the latter aspects refer to the ability of formulaic sequences to enhance the effectiveness of communication in daily life, whether with native or non-native speakers (Conklin and Schmitt, 2008; Howarth, 1998; Jackendoff, 1995; Schmitt and Carter, 2004).

The present study aims at providing more insights into the field of psycholinguistics in terms of collocational knowledge. Therefore, the research focus of the current study is on the processing of prepositional combinations among Arab-English bilinguals, particularly Saudi learners of English in the UK. The study looks at several issues of importance to 'verb/noun/adjective + preposition' collocations. In English the collocation is depend on and accustomed to, while in Arabic it is قدم من/ أتى من 'depend on' and إعتاد على--- متعود على 'accustomed on'. There is no clear distinction as to when these are identical across the two languages, which in turn poses a problem for non-native speakers. Furthermore, certain collocations might be very frequent (e.g. go to has a frequency of 58,519 in English as established by the British National Corpus (BNC: online), while others are not (e.g. meddle in occurs 48 times in the BNC). With this in mind, the current research investigates the following research questions:

1. Do similarities and differences in Arabic-English prepositional collocations differentially influence processing?

2. Do different levels of frequency in prepositional combinations, according to BNC, differentially influence processing?

3. Are there any pedagogical implications that we can draw from the results?

1.1 Introduce the Problem

Among Arab students, L2 learning in general is particularly problematic where L1 could have an impact on them in the following areas: the sound system, whether sounds in isolation or sounds in connected speech; vocabulary, whether for word formation or words and meaning including collocation, phrasal verbs, idioms, prepositions and prepositional phrases; types of sentences; sentence expansion; parts of speech and parts of sentences and finally in discourse (Kharma and Hajjaj, 1989).

Additionally, Dajem (2012a: 5) points out that other Arabic empirical studies have investigated the problems of some linguistic aspects of English individually; these include: "Crompton, 2011, in article errors; Habash, 1982, in preposition interference; Ibrahim, 1978, in spelling errors; Kharma, 1981, in definite/indefinite articles; Mahmoud, 2005, in collocation errors; Scott and Tucker, 1974, in error analysis and Tahaineh, 2010, in preposition errors".

The problem of prepositions and prepositional phrases in particular seems salient among L2 learners. Generally, it can be said that prepositions "have earned a reputation for difficulty if not a downright unpredictability" for all L2 learners around the world (Pittman, 1966: 65). In this study, the prepositional combinations or phrases (adj/v/n + preposition) in Arabic language is tested on Saudi students in UK.

2. Literature Review

- Collocations Identification

There is no general consensus among researchers as to what collocations identify with (Evert, 2008); however, there are two main approaches to identifying/defining them: 1) the frequency-based approach and 2) the phraseological approach. Regarding the former, corpus research helps us to understand the field of applied linguistics in general and how to identify formulaic/idiomatic language in particular (Nattinger and DeCarrico, 1992; Schmitt et al., 2004). However, it cannot adequately demonstrate how the language is processed and produced psycho-linguistically (Schmitt et al., 2004). Research can however be used to identify which clusters can be measured, processed and tested, whether spoken or written.

Based on corpus research, one of the methods of identifying collocation is the frequency-based approach (Nesselhauf, 2004) or 'statistically oriented approach' (Herbst, 1996:380), which involves identifying collocation mainly concerned with the co-occurrence of word pairs.

Collocation can be defined in this sense as frequency-based, as illustrated by the following definitions: "an aspect of lexical cohesion which embraces a 'relationship' between lexical items that regularly co-occur" (Carter, 1988: 163) or "the occurrence of two or more words within a short space of each other in a text" (Sinclair, 1991: 170). Such definitions shed light on the nature of frequency-based approach which can also be described as a statistical approach.

The approach is arguably "the simplest method for finding collocations in a text corpus"

(Manning et al., 1999). It can help to identify collocations in any language using a corpus. However, unnecessary words may accumulate which would need to be filtered out according to speech or the strength of association, in order to obtain more reliable results (Justeson and Katz, 1995).

In this regard, Stubbs (1995) argues that identifying collocation cannot rely solely upon co-occurrence frequency but also needs to rely on the association strength between words pairs. Hence, Evert (2008) suggests simple association measures as a criterion of co-occurrence, which is related to the measure the association of the observed frequency (O) and the expected frequency (E) of each word pair in a corpus to make a decision about whether the pair is collocational or not. "The pair is only considered collocational if the observed co-occurrence frequency is substantially greater than the expected frequency" (ibid: 17). Alongside the statistical approach, there is a need for qualitative reasoning in order to provide analysis of word pairs in terms of meaning, suitability and usage.

With regard to the latter approach, i.e. the phraseological approach, it works as an alternative method which clarifies the syntagmatic relation of word pairs (Nesselhauf, 2005). This approach is also known as semantic/substitutional approach or 'significance oriented approach' (Herbst, 1996: 380). It explores the conventionalised underlying meaning and the syntagmatic relation of word pairs. They deal with collocations as "a type of word combination [which] is fixed to some degree but not completely" (Nesselhauf, 2005). Additionally, it is important to distinguish between different types of collocations or word pair combinations. According to Cowie (1994) collocations can be classified as 'free combination' where they are used in a literal sense, 'restricted collocation' where one element of the combination has a literal meaning but not the other part, and 'figurative idioms' where all of its elements have figurative meaning.

Bonk (2001) points out that collocations in this regard can be divided into the broad sense, which is flexible with any fixed formulaic expression, and the narrow sense, which is concerned with lexical rather than grammatical collocations. Under the narrow sense, three classifications may be distinguished (Nesselhauf, 2005: 21); firstly, the syntactic characteristics of collocation (according to the word class); secondly, the semantic characteristics (sense restriction), and, thirdly, the commutability of its elements to substitute or change some elements.

The phraseological approach may depend, when identifying collocation, on the native speakers' intuition (Hasselgren, 1994), wherein they have "extensive knowledge of how words combine in their language, and use this knowledge when they retrieve lexical items and link them appropriately in language production" (Bonk, 2001: 113). Nevertheless, Stubbs (1995) argues that this approach can be considered as a pitfall and not sufficient especially when the statistical approach is required.

In the present study, there is more need for the frequency-based approach where the frequency of prepositional combinations, which are the stimuli, will be calculated according to the British National Corpus. The phraseological approach will not be the focal point in the present research because the stimuli which have been selected in the experiment were

extracted from books on the common grammatical mistakes among L2 learners.

- Language Transfer, similarities and differences

Generally speaking, “differences between the target language and the L1 resulted in learning difficulty and similarities in learning ease” (Ellis, 2008:398). Wolter (2006) also argues that the influence can either facilitate learning or make it difficult, while both lead to language transfer. Nevertheless, Kleinmann (1978) adds that differences can facilitate language learning and acquisition while similar language features may result in language difficulty. For example, Major and Kim (1996) found that Korean-English bilinguals learn the dissimilar sounds in English, e.g. /z/, more easily than some similar sounds such as /dʒ/. However, this cannot be generalised to all language features and contexts, whether in syntax, pragmatics, semantic or phonology, wherein languages are different, although they may have some features in common.

With regard to collocations flexibility, lexical items that can be joined together, e.g. look for or look at, may cause problems for L2 learners when trying to combine words together according to what they think is right based upon their mother tongue, resulting in cross-linguistic relation (Yamashita and Jiang, 2010). In some cases, L2 learners may feel uncomfortable with certain linguistic features due to their difficulty or contrast with their mother tongue, resulting in those learners avoiding using those aspects - a phenomenon called ‘avoidance behaviour’ (Ellis: 2008).

Ellis (2008) cites studies which investigate advanced L2 learners’ avoidance of phrasal verbs in English (including Dagut and Laufer, 1985, advanced Israeli learners; Hulstijn and Marchena, 1989, intermediate and advanced Dutch learners; Laufer and Eliasson, 1993, advanced Swedish learners, Sjoholm, 1995, mixed proficiency Finnish and Swedish learners; and Liao and Fukuya, 2004, intermediate and advanced Chinese learners). These studies show that even advanced L2 learners avoid using certain collocations, i.e. phrasal verbs, which do not match their L1. These findings are in line with the assumption that the similarities between L1 and L2 affect the processing, comprehension and production of L2.

In a more recent study, Wolter and Gyllstad (in press) investigate the processing of congruent and incongruent collocations on high proficient English-Swedish bilinguals and native speakers of English as a baseline for comparison. Response time and error rates of an acceptability judgment task were examined. Both types of lexical items were selected based on their frequency. The frequency of L1 of congruent collocations was taken into consideration to determine whether it made a difference. The findings revealed that “advanced learners are highly sensitive to frequency effects for L2 collocations” (ibid: 2).

In sum, L2 learners, as Siyanova and Schmitt (2008) found, are capable of producing a large number of appropriate collocations, however the underlying intuitions and the fluency with collocations of even advanced learners do not seem to match those of native speakers. Unsurprisingly, even advanced L2 learners encounter occasional difficulty when choosing the appropriate collocations to fit each other (Anwar and Khan, 2012; Dechert and Lennon, 1989; Nesselhauf, 2005). It should be noted, however, that there is another important factor which

may have an impact on L2 language processing, namely the level of frequency of a particular lexical item in a corpus which represents how frequently it is used in spoken and written language. The next section will consider this issue briefly in terms of a recurrence effect on processing.

- Influence of L1 knowledge on L2 processing of collocations

There is not enough research in L1 influence on the development of L2 collocations (Wolter and Gyllstad, 2011); "L1 lexical knowledge can be both a help and a hindrance when forming L2 connections" (Wolter, 2006: 741). Both L1 and L2 play a role in structuring collocational knowledge, combinations of which can be called 'intralexical knowledge'; further investigation is needed on whether L1 knowledge affects L2 processing of collocations (Wolter, 2006). In simple terms, language transfer is "the effect of one language on the learning of another" (Richards and Schmidt, 2002: 294). Such transfer could be a 'positive transfer', as when L1 has similarities with the target language as this can make the learning easier, or a 'negative transfer/interference' as when the target language rules and linguistics patterns overlap with L1 (ibid).

Relevantly, Wolter and Gyllstad (2011) conducted research on Swedish-English bilinguals as well as native speakers of English to explore whether L1 knowledge has an influence on L2 processing using a primed lexical decision task, following a psycholinguistic approach. The focal point of the research is to determine whether the collocations in L2, which have equivalents in L1, can be activated easily rather than those which do not have any equivalents in L1. Results show a considerable influence of L1 intralexical knowledge on L2 collocations. Moreover, the reaction time towards collocations that have equivalents in L1 was shorter than those which do not. Such findings support the assertion that L1 does have an absolute effect on L2 regardless its positivity or negativity.

In a different context but with a similar question, Yamashita and Jiang (2010) investigated the influence of L1 on the acquisition of L2 collocations in Japanese-English bilinguals in EFL and ESL contexts, as well as in native speakers of English as a baseline for comparison. A phrase accessibility judgement task was used based on congruent collocations with an equivalent L1 construction, as well as on incongruent collocations with no equivalents in L1. Error rates and reaction time were examined for accuracy and speed of performance. The results show that L1 similarities to L2 as well as L2 exposure played an important role in the acquisition of L2 collocations. Such findings highlight the fact that incongruent collocations might be problematic for L2 learners.

Another study by Wray (2002) argues that collocations can be formulated mainly based on L1 knowledge for L2 learners. This might indicate that L1 has a positive or a negative effect on L2 learning, comprehension, processing and production. Therefore, L2 learners need to pay close attention to collocations because they are not structured arbitrarily; rather they have some semantic constraints according to the native speakers' use of the language (Wolter and Gyllstad, 2011).

- Problematic Issues of Arabic Context

Mahmoud (2005) investigated errors in the use of collocations and lexical combination among university Arab-Omani learners of English. The context was EFL where it an Arabic speaking country although the learners were also majoring in English. After examining their essays, which were written on a weekly basis, around 420 collocations were found, 64% of which were incorrect. This can be attributed to the negative transfer from L1. The researcher pointed out that relying on L1 to produce L2 collocation can help learners in regard to similarities but could be detrimental when differences occur between L1 and L2. Interestingly, Gas (1979) argues that language transfer cannot occur unless the learner is aware of the similarities and differences between L1 and L2. This further emphasises the fact that L2 learners should in fact be taught collocations as this would enhance their awareness of the similarities and differences.

Mahmoud (2005) drew some pedagogical implications based on these findings to suggest which collocations should be taught directly to Arab students. He argues that "simplified contrastive comparisons between English and Arabic collocations might help students see when to transfer and when not to" (ibid: 7). Additionally, he suggests designing bilingual collocation dictionaries for such language features as well as using a bilingual list of collocation in students' course books. This seems to support Martinez and Schmitt (2012) suggestion of using a phrase expression list rather than a list of single words which may not fit the context.

In a similar study, Anwar and Khan (2012) conducted empirical research on advanced learners of English in ESL Pakistan to investigate the importance of collocation in ESL acquisition. The data was collected through written essay tests from the participants and showed that even advanced learners fell into serious error in collocation (see section 1.5 above for similar studies). However, the main finding of this study revealed the dire need for the "inevitability of direct teaching of collocations" for L2 learners (ibid: 47).

Lakkis and Malak (2000) conducted similar studies to investigate the influence of L1 knowledge on Arab-Lebanese students in acquiring English prepositional usage, particularly in the phrasal verb 'verb+preposition'. The students' essays were analysed based on four categories: 1) when the verb in English has a preposition and does not have one in Arabic; 2) incongruent collocations when the preposition used in L1 is different to L2; 3) congruent collocations when the same preposition is used with the same verb in both languages; and finally 4) when many different prepositions in L1 can be used with the same verb which is not the case in L2.

The findings according to these three categories are as follows. Firstly, students did not notice the importance of adding the required prepositions with those verbs such as wait for, as they do not have a preposition with the verb in Arabic. Secondly, students' answers showed their attempt to prove their knowledge about the differences of preposition usage between L1 and L2, but they still committed serious mistakes. Thirdly, the learners showed a positive performance when they had exactly the same usage. Finally, they still encountered major problems when choosing the appropriate prepositions under the influence of L1 e.g. on 4 o'clock rather than at 4 o'clock.

However, the researcher points out that when the prepositional phrases/preposition were used frequently in the curriculum, the majority of students used them correctly. This is because "the frequency of occurrence of a structure promotes the correct usage of the preposition" (ibid: 6). Based on the findings, Lakkis and Malak's (2000) study has a similar suggestion to the previous one i.e. Mahmoud (2005) who suggested giving students a clear explanation of the differences and similarities between L1 and L2 thereby making them aware of such variation.

Tahaineh (2010) conducted a similar performance analysis study to investigate what kind of errors Jordanian-Arab university students, majoring in English, commit in the use of prepositions by analysing their free compositions. The results indicate that many of the errors are attributed to 'interlingual interference'. In other words, the learners in this study "use the proper prepositions providing equivalents are used in their MT [mother tongue]; select the improper prepositions if equivalents are not used in their MT; omit prepositions if equivalents are not required in their MT and add prepositions if equivalents are required in their MT" (ibid: 77). More interestingly, similar to previous views regarding collocation (see Dechert and Lennon, 1989 and Nesselhauf, 2005), this research argues that even advanced learners could not avoid committing such prepositional errors.

Another important issue regarding the Arabic language is that there are two main varieties of Arabic: modern standard Arabic (MSA) and non-standard Arabic (NSA). Mahmoud (2000) claimed that further research was needed to determine which variety caused more significant problems. This seems to be a very broad issue as one variety would cause problems on certain linguistic features but not on others.

Researchers hinted that Arabic has an influence on English prepositions, although it has not been determined whether the 'culprit' is MSA or NSA. Accordingly, Al-Khresheh (2010) conducted a study on school students in Jordan to investigate which of the two caused the most problems when formulating a correct sentence structure, or 'word order'.

He found that interlingual errors represented 55.3% and 44.7% of the MSA and NSA respectively. This may reveal the significant influence of MSA although this could not be generalised because the participants came from one school and were taught by the same teachers, thus the findings of the study are limited to one context and educational system. It would therefore be useful to conduct the same study in different Arab contexts to determine whether or how the results would vary from one context to another.

- High frequency and low frequency Influence

Frequency of occurrence of formulaic sequences plays an essential role in language learning. Experts in language acquisition and psycholinguists emphasise the importance of lexical patterns in the learner's mental lexicon, whether these lexical patterns are known by individuals or stored as a whole unit in the learner's mental lexicon (Schmitt and Carter, 2004:2).

In this regard, Schmitt et al. (2004) investigated the processing of recurrent strings, derived from a corpus analysis of formulaic sequences, in 34 English native and 45 non-native

speakers, using a combination of corpus and psycholinguistic approaches.

Schmitt et al. tested the psycholinguistic validity of recurrent clusters. They were embedded into a passage and then presented to the participants in the form of a story. Using non-laboratory methodology, this was followed by a dictation task for these items which were chosen based on their frequency from British National Corpus (BNC), CANCODE and MICASE corpora.

The results show that not all recurrent clusters are stored holistically, which might reveal a drawback of relying on the corpus which seems to be insufficient evidence. The researchers suggest that there is no relation between the frequency of occurrence in the corpus and the extent of the holistic processing. In effect, the study shows the importance of combining both corpus and psycholinguistic approaches to obtain more reliable results.

Another study, by Siyanova and Schmitt (2008), investigates L2 processing and adjective+noun collocation. Native and non-native speakers (NNS) were tested using familiarity rating tasks of frequent and infrequent collocations in English. Interstingly, the results show that NNS rated the frequent collocation as less familiar combinations while the NS rated them as familiar, as we would expect. This result occurred regardless of the amount of time that NNS spent on the task; and, on average, NNS were slower compared to the other group. L1 may play a role where NNS use their own intuition to decide what is familiar according to L1 knowledge. Since NNS in may not get as much exposure as NS, it can be inferred that the frequency may not be based only on the recurrence of collocations in a corpus but also on the frequency of input. This suggests the importance of frequency on L2 learners in terms of exposure or input (Monsell, 1991).

2. Methodology

2.1 Participants Selection

In total, there are 40 participants in the current research. Firstly, 20 Saudi learners of English, 15 males and 5 females, who are studying at the University of Nottingham, 3 BA, 7 MA and 10 PhD, with a mean age of 30.10 (SD=4.36, min=22, max=39). None has been to an English-speaking country except for one who has been to United States of America for one year. All started learning English at middle school except for two who started at primary school.

Regarding the English language test, 17 of the 20 participants did IELTS before, 1 did TOEFL and 2 did not take any test. They evaluated their proficiency in English as follows: 30% intermediate-level, 20% advanced for the speaking skill; 20% intermediate-level, 27% advanced, 2% native-like proficiency for the reading skill; 2% beginners, 17% intermediate-level, 27% advanced for the writing skill; 2% beginners, 12% intermediate-level, 35% advanced for the comprehension skill; and finally 17% intermediate-level, 32% advanced for the listening skill. However, their total IELTS grade ranges from 5 to 7. For the four skills grades were: listening (M=5.97, min=5, max=7), reading (M=5.70, min=4.5, max=7), speaking (M=6.20, min=5, max=7) and writing (M=5.85, min=5, max=7).

The remaining 20 participants were monolingual native speakers of 'British' English: 11 males and 9 females, with a mean age of 36.9 (SD=13.3, min=18, max=60). Some of them are students at the University of Nottingham, BA=8, MA=1, PhD=1, and the rest are staff on different teams, e.g. library staff, IT staff, administration staff. They rank their reading ability on a scale from 1 (excellent) to 5 (very poor), and they evaluated their ability as 40% 'excellent', 40% 'good' and 20% 'OK'.

2.2 Material Choice

As mentioned previously, the aim of the present study is to investigate the processing of prepositional combinations in terms of congruent and incongruent collocations based on a selection of the common mistakes that are committed by L2 language learners. I will then categorise them into two groups: the similarities and differences according to Arabic and English. Four language books were reviewed and all the grammatical mistakes were collected in light of those which match Arabic and English in terms of congruent and incongruent collocations.

After collecting these grammatical mistakes, the frequency of occurrence was considered for all collocations based on the British National Corpus. They were divided into three groups in terms of their recurrence in that corpus: high, mid and low. They were then embedded into 80 sentences which were divided into two lists with 40 in each. Each list has 20 sentences which carry those prepositional collocations which are identical in English and Arabic, while the other 20 represent those which have completely different prepositions in the two languages.

Having one list only might have encouraged participants to give a lot of the same answer, therefore two lists were used in order to ensure that they are fairly balanced in terms of correct/incorrect items as well as frequency, thereby making 'Yes' or 'No' responses equally likely for different prepositions for the same verb. Each list was presented to 10 native and 10 non-native speakers. It should be noted, however, that the four collocation types of 'stimuli', which will be mentioned frequently in this paper, are ascribed the following terms:

1. 'Similar correct' means that the English preposition that is used in the collocation/combination is identical to the one which is used in Arabic, and that the sentence has a correct English preposition that should be associated with either a verb/noun/adjective. For example, '*they are going to the hospital*' has exactly the same verb + preposition in Arabic, '*going to*'.

2. 'Similar incorrect' represents the same case as no. 1 but it is intended to comprise an incorrect preposition that associates a verb/noun/adj. For instance, '*they are going at the hospital*' is in contrast to the correct use of '*go to*' in Arabic.

3. 'Different correct' is where the correct English preposition that is used in the sentence is totally different from that which is used in Arabic, e.g. '*it is not good to look at the sun directly*' where '*look at*' is correct English usage while in Arabic the literal translation should be '*look to*'.

4. 'Different incorrect' represents the same situation as the previous point but rather than

using a correct English preposition, a preposition is used which is correct in Arabic but totally wrong in English. For example, 'look *to'* can be considered as a correct English if the Arabic learner translates it literally from Arabic into English which is totally wrong in English.

The remaining 40 sentences (resulting in 80 in total) were included as filler items. The first 20 were simple, acceptable sentences without collocations, while the following 20 were simple sentences that are unacceptable and contain some other sort of grammatical or semantic violation (see Appendix VII and VIII). The same fillers were used for both lists while the prepositional combinations in the two lists were slightly different but the surrounding sentence was the same. For example, 'I am thinking in you' was used in List 1 whereas 'I am thinking of you' was used in the second list. Therefore, each participant will encounter 80 sentences which contain 20 congruent collocations, 20 incongruent collocations, and 40 filler items. Lastly, 5 sentences were presented to the participants at the beginning of the experiment to familiarise them with the experiment. All of these items were presented randomly, edited and adopted based on supervisor feedback and peer review.

2.3 Procedure

A psycholinguistics measure E-Prime was used as a measurement tool to test 1) response time in milliseconds, where 1000 milliseconds equals 1 second, and 2) accuracy – whether they answer correctly or not. As described earlier, participants will read on a computer screen 80 phrases/sentences in English, and judge their acceptability, i.e. whether they are correct or not. Additionally, it was not deemed necessary to offer inducements to the participants due to the small amount of time taken by the experiment, estimated at 10-15 minutes for non-native speakers and 7-10 minutes for natives. It should be noted that a language background questionnaire was administrated as well as the informed consent before the participant starts the experiment.

At the beginning of the experiment, participants were told that they are going to read 80 sentences and they need to decide whether they are correct or not in terms of grammar and meaning. They should choose (✓) on the keyboard if the sentence is correct and (✗) if it is not. After that, they were told that there are a few practical items to which they need to respond, to familiarise themselves with the procedures; the actual study started afterward these preliminary items. It should be noted that the present study followed a similar procedure to that of Schmitt et al. (2004) who used two approaches: 1) corpus-based research where the frequencies were extracted from a corpus and 2) a psycholinguistic approach which can investigate the internal aspect of processing these prepositional combinations.

2.4 Data Analysis Technique

Two types of software were used to collect and analyse the results. Firstly, E-Prime 2.0 Professional was used as it calculates the reaction time in milliseconds and also collects data on accuracy – the correct and wrong responses as well. This software was chosen because the aim of the present study was to examine the processing of a selection of prepositional combinations and determine whether they spend a long or short time on the congruent and in congruent collocations.

Secondly, the statistical software IBM SPSS Statistics 20 was used to measure the differences between all these variables. First, a mixed between-within subjects analysis of variance (ANOVA) was conducted to examine the effect in general of both collocation types and frequency levels. Then, some independent and paired sample t-tests were performed to test the difference between and within the groups.

3. Results and Discussion

3.1 Summary of Major Findings

First, it should be noted that there was a significant difference in general between both native and non-native speakers in terms of reaction time and accuracy. To start off with, the major findings of the effect of collocation types on the processing and production of English formulaic sequences, i.e. prepositional combinations, among Saudi learners of English in the UK, will be summarised as follows:

1. In terms of processing, there was no significant interaction effect between collocation types and participant groups, and no substantial main effect of collocation types in general.
2. In terms of accuracy, there was a significant interaction effect between collocation types and participant groups, and a substantial main effect of collocation types in general.
3. For the reaction time ‘processing’:
    a. For non-native speakers, the data interestingly shows that there was no significant difference in mean reaction time between collocation types.
    b. For native speakers, there was no significant difference between the mean reaction time towards all collocations types.
4. Regarding accuracy:
    a. There was a statistically significant difference in the accuracy of non-native speakers for ‘different correct vs. different incorrect’ and ‘similar incorrect vs. different incorrect’; they achieved better results in ‘different correct’ collocations than in ‘different incorrect’.
    b. Regarding native speakers, there was no significant difference in their accuracy for all collocations types.

Regarding the effect of frequency levels, the major findings are as follows:

1. In terms of processing, there was an interaction effect of frequency types and participant group, and a substantial main effect for frequency types in general.
2. In terms of accuracy, there was no an interaction effect of frequency types and participant group but there was a substantial main effect for frequency types in general.

3. For reaction time ‘processing’:

    a. For non-native speakers, there was a significant difference of frequency levels in which they process high frequent prepositional combinations more quickly than mid and low frequent ones.

    b. For native speakers, the data shows that there was a significant difference of frequency levels in which they process the high frequent ones more quickly than mid frequent ones.

4. Regarding accuracy:

    a. There was no significant difference between non-native responses but they still achieved a higher score in the high frequent ones than in the low frequent ones.

    b. Regarding native speakers, there was a significant difference between ‘high vs. low’ and ‘mid vs. low’ where they achieved better scores in the high and mid ones.

First of all, a mixed between-within subjects analysis of variance ANOVA was conducted before addressing the two main research questions to assess in general the effect of collocation types (similar correct, similar incorrect, different correct and different incorrect – as defined above in Chapter Two); frequency levels (high, mid and low); and participant groups (native speakers of English and native speakers of Arabic, namely Saudi learners of English in the UK).

Firstly, there was no significant interaction effect between collocation types and participant groups, F (3, 36) =.62, p=.60, partial eta squared=.04. Additionally, there was no substantial main effect of collocation types, F (3, 36) =.60, p=.61, partial eta squared=.04.

Regarding the interaction effect of frequency types and participant group, the data reveal that there was a statistically significant interaction between them, F (2, 37) =4.96, P<.05, partial eta squared=.21. Moreover, there was a substantial main effect for frequency types, F (2, 37) =11.64, p<.000, partial eta squared=.38.

Collocation as well as frequency types had a statistically significant interaction effect, F (6, 33) =2.12, p<.01, partial eta squared=.27. This was the opposite for the interaction effect between all three variables, i.e. collocation, frequency types and participant groups, where there was no significant interaction effect between them, F (6, 33) =1.32, p<.27, partial eta squared=.19.

The between subjects effect shows that there was a significant difference between groups in the reaction time, whether for collocation types or with the different levels of frequency, F (1, 38) =54.76, p=.000, partial eta squared=.59.

Table 1 shows the mean and standard deviation for both participant groups across all collocation and frequency types. As we can see, there is a significant difference between the reaction time between native and non-native speakers for all variables. This is expected

because those whose mother tongue is not English are still learning the language and have not yet mastered it, as indicated by the language background questionnaire. The non-natives, in contrast, were taught English grammar at school starting from the middle school but have not been to an English speaking country before. Therefore, it might be expected that they have not fully acquired the language to a certain extent and they might have some grammar rules in mind which they tried to match with what they saw in the grammatical judgement tasks, especially the grammatical judgement task.

Wray (2000 and 2002) argues that advanced L2 learners have an analytical perspective in processing the language which may slow down their processing, whereas this is not the case with native speakers and young L2 learners who process the language holistically. Such holistic processing might occur with L2 learners if L1 could play a role in making a decision about whether a particular linguistic aspect is correct or not. Although beginners are less analytical about language, they might still under the impact of grammar rules to judge whether some linguistic aspects are correct or not.

Table 1. A mixed between-within subjects ANOVA of reaction time for all variables and participants (N=40)

	Native Speaker			Non-native speakers		
Collocation and frequency Types	**N**	**M**	**SD**	**N**	**M**	**SD**
Similar Correct High	20	3425	1387	20	7807	3847
Similar Correct Mid	20	4598	2100	20	8903	6187
Similar Correct Low	20	4143	1654	20	11242	3020
Similar Incorrect High	20	3854	1529	20	7738	3794
Similar Incorrect Mid	20	3713	1641	20	8943	5127
Similar Incorrect Low	20	4159	1638	20	8248	4346
Different Correct High	20	3809	1366	20	7810	3788
Different Correct Mid	20	4999	3287	20	9008	5043
Different Correct Low	20	3859	1644	20	9339	4891
Different Incorrect High	20	3965	1892	20	6466	4863
Different Incorrect Mid	20	4778	2583	20	10088	7886
Different Incorrect Low	20	3643	1915	20	7464	1915

- Accuracy

Regarding the accuracy of their responses, another mixed between-within subjects ANOVA was conducted to assess the accuracy of both group's performances. This will be presented as a percentage of correct answers. The findings indicate that there was a significant interaction effect between collocation types and participant groups, F (3, 36) =6.85, p<001, partial eta squared=.36. Additionally, there was a substantial main effect of collocation types, F (3, 36) =15.71, p<000, partial eta *squared=.56.*

Regarding the interaction effect of frequency types and participant groups, the data reveal that there was no statistically significant interaction between them, F (2, 37) =.45, p=.63, partial eta squared=.02. Moreover, there was a substantial main effect for frequency types in general, F (2, 37) =7.34, p<.002, partial eta squared=.28.

The results also reveal that collocation as well as frequency types had no significant interaction effect, F (6, 33) =1.86, p=.11, partial eta squared=.25. In addition, there was no significant interaction effect between collocation, frequency types and participant groups, F (6, 33) =1.40, p=.24, partial eta squared=.20.

The between subjects effect shows that there was a significant difference between groups in the percentages of correct responses, whether for collocation types or with the different levels of frequency, F (1, 38) =64.02, p=.000, partial eta squared=.62. This is the same case as the reaction time which was expected due to the main difference between those whose mother tongue is English and their peers.

Table 2. A mixed between-within subjects ANOVA of correct responses in percentages for all variables and participants (N=40)

	Native Speaker			Non-native speakers		
Collocation and frequency Types	**N**	**M**	**SD**	**N**	**M**	**SD**
Similar Correct High	20	86.25%	22.17	20	72.50%	32.34
Similar Correct Mid	20	93.33%	13.68	20	83.33%	17.10
Similar Correct Low	20	79.99%	27.36	20	78.32%	19.57
Similar Incorrect High	20	88.75%	20.63	20	66.25%	28.41
Similar Incorrect Mid	20	84.99%	27.52	20	64.99%	36.63
Similar Incorrect Low	20	84.99%	22.87	20	58.33%	32.21
Different Correct High	20	85%	12.56	20	87.08%	3788
Different Correct Mid	20	84.99%	25.30	20	61.66%	36.31
Different Correct Low	20	73.32%	25.59	20	73.33%	27.78
Different Incorrect High	20	82.50%	18.31	20	38.75%	30.85
Different Incorrect Mid	20	79.99%	27.36	20	41.66%	23.87
Different Incorrect Low	20	71.66%	27.09	20	26.66%	20.51

It should be emphasised that there was a significant difference between frequency levels and participant groups in reaction time, while there was a significant difference between collocation types and participant groups in accuracy. This suggests that frequency levels have an impact on both groups.

3.2 Discussion

First question

The present study is interested mainly in ‘similar correct’ and ‘different correct’ collocation types and whether the former can lead to quicker processing while the latter slows processing because it does not match L1 knowledge. However, the results show, in terms of response time, that there was no significant difference between them. This seems to contradict the assumption that has been made in this paper that similarities between L1 and L2 can result in quick processing while the differences can slow down processing. Also, this seems to contradict the notion that differences can facilitate language learning and acquisition while

similar language features may result in language difficulty (Ellis, 2008; Kleinmann, 1978; Wolter, 2006).

Such results could be attributed to the students' level of English. Since all of them have already finished the English language course and started their degree studies, whether BA, MA or PhD, they might not have a problem with processing; however, they might have other problems with the accuracy. This will be discussed in the next section.

Such findings in the present study are not in line with the findings of some studies which investigated the same issue. Wolter and Gyllstad (2011) found that L1 knowledge has a positive influence on L2 collocations processing in a primed lexical decision task especially with those collocations which have similar equivalents in L1. Additionally, in a similar study with similar findings, on Japanese learners of English, Yamashita and Jiang (2010) found that congruent collocations play a very important role in L2 learning and acquisition; however this is not the case with incongruent collocations, which might be problematic for L2 learners.

Unlike the findings of previous research (see Wolter and Gyllstad, 2011; Yamashita and Jiang, 2010), the present study shows that there was no significant difference between collocation types. This might be due to the small size of the sample in the current research; the fact that there were only 20 non-native speakers might have an impact on the results. Moreover, it may be attributed to the educational level of the learners in that most of them are doing their BAs, MAs or PhDs in the UK. They may have reached a level of proficiency such that they can decide whether the presented stimuli are correct or not.

On the other hand, in the case of the latter, i.e. 'similar incorrect vs. different incorrect', it seems that participants committed more mistakes with 'different incorrect' collocations than with 'similar incorrect' which are not only contrasted with L1 but are also completely wrong and unacceptable English. Compared to those 'similar incorrect' collocations, it can be said that L1 knowledge positively helps them to judge the grammar of each sentence because it contrasts with their mother tongue usage. Such findings are similar to Lakkis and Malak's (2000) investigation on Arab-Lebanese learners of English on the use of congruent and incongruent prepositional collocations. They found that the learners showed a positive performance when they have exactly the same preposition in Arabic whereas they encountered a serious problem with collocations which do not, although it seems that they made an effort to prove their knowledge about the differences in preposition usage between L1 and L2.

Nevertheless, the topic of interest for the first question is how non-native speakers performed with those similarities represented in 'similar correct' and those differences represented in those 'different correct'. They performed better in those which are similar to Arabic, namely congruent prepositional collocations, with a difference of 2%. This suggests that, as Mahmoud (2005) argued, relying on L1 to produce and process L2 collocations can help learners in regard to similarities but could be detrimental when differences occur between L1 and L2, with all linguistic aspects.

- Second Question

To test the difference within the non-native speakers regarding the different levels of frequency, and after computing the means as we did in the previous test, a paired-samples t-test was performed to investigate the difference within the group. The findings in Table 3.11 indicate that there is a significant difference in the mean reaction time between the high and low frequent collocations, as well as in the high and mid ones, in which they seem to process the high frequent collocations more quickly than the others.

Arguably, this supports the concept that frequent formulaic sequences are processed more quickly than those with low frequency (Schmitt and Carter, 2004). Furthermore, as Lakkis and Malak (2000:6) claim, "the frequency of occurrence of a structure promotes the correct usage of the preposition". This claim is based on their study of Lebanese Arab learners of English which found that the frequent prepositional phrases/prepositions in the curriculum were used more correctly by the participants in their study than the less frequent ones. However, frequency cannot be based only on their recurrence in a corpus; frequency of input needs to be considered as well because the importance of exposure to the target language (Siyanova and Schmitt, 2008).

This study has demonstrated the positive influence that frequency has on the quick processing or the correct production of collocations. However, it is still not clear whether collocations are processed holistically in the memory or not. Knowing how they are processed can contribute to our understanding of how the mind deals with such formulaic sequences.

4. Experimental Implications

While the present study has contributed some important perspectives on the processing and production of prepositional combinations in English by Saudi learners of English in the UK, it has not ascertained whether these recurrent clusters are stored holistically, as a whole unit in the memory, or not. This study relied on grammaticality judgement task only; using a different psycholinguistic measure, such as eye tracking, may help to investigate the issue further, while using self-based reading could complement the use of grammatical judgement tasks. Schmitt et al. (2004) argue that the frequency of lexical items in a corpus is not enough evidence to judge whether those items are processed holistically or not.

5. Suggestions for Further Research

The current research has contributed to the field of psycholinguistics in terms of the processing and production of prepositional combinations. However, there is a need for further research to investigate both of these aspects further. The eye tracking technique mentioned above could provide more information about how those students process formulaic sequences in their reading, perhaps providing different insights to those provided by grammatical judgement tasks. Additionally, the issue of holistic processing could be investigated using the eye tracker technique to determine whether similar collocations are processed differently from different collocations, and whether formulaic sequences are acquired holistically, as a whole unit, or incrementally, that is, gradually.

Having identified the limitations of the current study, a more comprehensive study that involves a significantly larger number of participants with various analysis techniques would be recommended to endorse the findings of this or any other similarly designed study. Such a proposed study would remove doubts about how formulaic language functions among L2 users and would also provide a rich database for future research.

References

Adolphs, S., & Durow, V. (2004). Social-cultural integration and the development of formulaic sequences. *In:* Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins Publishing Company. p.107-126.

Anwar, R. M. B., & Khan, L. A. (2012). Collocations and Second Language Use: Errors Made By Advanced Learners in Pakistan. *Arts, Social Sciences and Scientific Studies, 5,* 47-73. Arabicorpus. Available: http://arabicorpus.byu.edu/ [Accessed 01-06- 2012].

Benson, M., Benson, E., & Ilson, R. (1986). *The BBI combinatory dictionary of English: a guide to word combinations.* Amsterdam/Philadelphia: John Benjamins.

Bnc. Available: http://corpus.byu.edu/bnc/x.asp?w=1219&h=686 [Accessed 01-07- 2012].

Bolander, M. (1989). Prefabs, patterns and rules in interaction? Formulaic speech in adult learners' L2 Swedish. *In:* Hyltenstam, K. & Obler, L. K. (eds.) *Bilingualism across the lifespan: Aspects of acquisition, maturity, and loss*. Cambridge: Cambridge Univ Press.

Bonk, W. J. (2001). Testing ESL learners' knowledge of collocations. *In:* Hudson, T. & Brown, J. D. (Eds.), *A focus on language test development: Expanding the language proficiency construct across a variety of tests*. Honolulu: University of Hawaii Press. p.113-142.

Carter, R. (1988). Vocabulary, cloze and discourse: An applied linguistic view. *In:* Carter, R. & Mccarthy, M. (eds.) *Vocabulary and language teaching*. Harlow: Longman. p.161-180.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied linguistics, 29*(1)**,** p.72-89.

Cowie, A. P. (1994). Phraseology. *In:* Asher, R. E., & Simpson, J. M. Y. (eds.) *The encyclopedia of language and linguistics*. Oxford: Pergamon Press. p.3168-3171.

Crompton, P. (2011). Article Errors in the English Writing of Advanced L1 Arabic Learners: The Role of Transfer. *Asian EFL Journal,* 50.

Dagut, M., & Laufer, B. (1985). Avoidance of phrasal verbs: A case for contrastive analysis. *Studies in second language acquisition, 7*(1), 3-79.

Dajem, Z. (2012a). The Role of Proficiency among Arabic-English Bilinguals: Empirical Experiment on the Use of Prepositions. Nottingham: University of Nottingham.

Dajem, Z. (2012b). *Teachers' Attitude towards Teaching English Grammar: An Empirical Investigation on Saudi Arabian Middle and Secondary Schools.* Germany: LAP Lambert Academic Publishing.

Dechert, H. (1983). How a story is done in a second language. *In:* Faerch, C., & Kasper, G. (eds.) *Strategies in interlanguage communication*. London: Longman Pub Group. p.175–95.

Dörnyei, Z., Durow, V., & Zahran, K. (2004). Individual differences and their effects on formulaic sequence acquisition. *In:* Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins Publishing Company. p.127-151.

Ellis, R. (2008). *The study of second language acquisition.* Oxford: Oxford University Press.

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text, 20*(1), 29-62.

Evans, D. (2012). Distribution of some English tenses in the spoken corpus and textbooks. *In:* Classroom, Q. G. I. T. (ed.). Nottingham: University of Nottingham.

Evert, S. (2008). Corpora and collocations. *In:* Lüdeling, A. & , M. K. (eds.) *Corpus Linguistics. An International Handbook*. Berlin: Mouton de Gruyter. p.1-53.

Firth, J. R. (1957). *Modes of meaning. In Papers in linguistics 1934-1951*. Oxford: Oxford University Press.

Firth, J. R. (1968). A synopsis of linguistic theory, 1930-55. *In:* Palmer, F. R. (ed.) *Selected papers of J.R. Firth 1952-1959*. Harlow: Longman. p.168-205.

Fitikides, T. J. (1969). *Common mistakes in English.* Orient Blackswan.

Gas, S. (1979). Language transfer and universal grammatical relations. *Language Learning, 29*(2), 327-344.

Gratian, V. (1995). *Prepositions.* Sterling Publishers Pvt. Ltd.

Habash, Z. (1982). Common Errors in the Use of English Prepositions in the Written Work of UNRWA Students at the End of the Preparatory Cycle in the Jerusalem Area. *Retrieved January*, 2 2003.

Halliday, M. (1966). Lexis as a linguistic level. *In:* Ce, B., Jc, C., Mak, H., & Rh, R. (eds.) *In Memory of J.R. Firth*. London: Longman.

Hancock, P. (2001). *Common Errors in English.* Essex: Bluestone Press.

Hasselgren, A. (1994). Lexical teddy bears and advanced learners: A study into the ways Norwegian students cope with English vocabulary. *International Journal of Applied Linguistics, 4*(2), 237-258.

Herbst, T. (1996). What are collocations: sandy beaches or false teeth? *English Studies, 77*(4), 379-393.

Howarth, P. (1998). The phraseology of learners' academic writing. *In:* Cowie, A. P. (ed.) *Phraseology: Theory, analysis, and applications*. Oxford: Oxford University Press. p.161-86.

Hulstijn, J., & Marchena, E. (1989). Avoidance: grammatical or semantic causes. *Studies in second language acquisition, 11*, 242-300.

Ibrahim, M. H. (1978). Patterns in spelling errors. *ELT journal, 32*(3), 207-212.

Jiang, N., & Nekrasova, T. M. (2007). The processing of formulaic sequences by second language speakers. *The Modern Language Journal, 91*(3), 433-445.

Kharma, N. (1981). Analysis of the errors committed by Arab university students in the use of the English definite/indefinite articles. *IRAL-International Review of Applied Linguistics in Language Teaching, 19*(1-4), 333-345.

Kharma, N., & Hajjaj, A. (1989). *Errors in English among Arabic speakers: Analysis and remedy*. Longman.

Kleinmann, H. H. (1978). The strategy of avoidance in adult second language acquisition. *In:* Ritchie, W. (ed.) *Second language acquisition research*. New York: Academic Press. p.157-174.

Krashen, S., & Scarcella, R. (1978). On routines and patterns in language acquisition and performance. *Language Learning, 28*(2), 283-300.

Krashen, S. D. (1982). *Principles and practice in Second Language Acquisition.* Oxford: Pergamon.

Lakkis, K., & Malak, M. A. (2000). Understanding the Transfer of Prepositions: Arabic to English. Lebanon: American University of Beirut.

Laufer, B., & Eliasson, S. (1993). What causes avoidance in L2 learning. *Studies in second language acquisition, 15*(01), 35-48.

Liao, Y., & Fukuya, Y. J. (2004). Avoidance of phrasal verbs: The case of Chinese learners of English. *Language Learning, 54*(2), 193-226.

Mahmoud, A. (2000). Modern standard Arabic vs. non-standard Arabic: Where do Arab students of EFL transfer from? *Language Culture and Curriculum, 13*(2), 126-136.

Mahmoud, A. (2005). Collocation errors made by Arab learners of English. *Asian EFL Journal*, (2).

Major, R. C., & Kim, E. (1996). The similarity differential rate hypothesis. *Language Learning, 49*, 151-183.

Manning, C. D., Schütze, H., & Mitcognet (1999). *Foundations of statistical natural language processing.* MIT Press.

Martinez, R., & Schmitt, N. (2012). A Phrasal Expressions List. *Applied linguistics*.

Monsell, S. (1991). The nature and locus of word frequency effects in reading. *In:* Besner, D. & Humphreys, G. W. (eds.) *Basic processes in reading: Visual word recognition*. Hillsdale: Lawrence Erlbaum Associates. p.148-197.

Nattinger, J. R., & Decarrico, J. S. (1992). *Lexical phrases and language teaching.* Oxford: Oxford University Press.

Nesselhauf, N. (2004). What are collocations. *In:* Nesselhauf, N. & Skandera, P. (eds.) *Phraseological units: Basic concepts and their application*. Basel: Schwabe. p.1-21.

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

O'keeffe, A., Mccarthy, M., & Carter, R. (2007). *From corpus to classroom: Language use and language teaching.* Cambridge: Cambridge Univ Press.

Oppenheim, N. (2000). The importance of recurrent sequences for nonnative speaker fluency and cognition. *In:* Segalowitz, N., & Riggenbach, H. (eds.) *Perspectives on fluency*. Ann Arbor: University of Michigan Press. p.220-40.

Ortega, L. (2009). *Understanding Second Language Acquisition.* London: Hodder Education.

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. *Language and communication, 191*, 225.

Peters, A. M. (1983). *The units of language acquisition.* Cambridge: Cambridge University Press.

Pittman, G. A. (1966). *Activating the use of prepositions.* London: Longmans.

Richards, J. C., & Schmidt, R. (2002). *Dictionary of language teaching and applied linguistics.* Essex: Longman.

Schmitt, N. (2004). *Formulaic sequences: Acquisition, processing, and use.* Amsterdam: John Benjamins Publishing Company.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action. *In:* Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins Publishing Company. p.1-22.

Schmitt, N., Grandage, S., & Adolphs, S. (2004). Are corpus-derived recurrent clusters psycholinguistically valid? *In:* Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins Publishing Company. p.127-151.

Schmitt, N., & Underwood, G. (2004). Exploring the proccessing of formulaic sequences through a self-paced reading task. *In:* Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins Publishing Company. p.173-189.

Scott, M. S., & Tucker, G. R. (1974). Error analysis and English language strategies of Arab students1. *Language Learning, 24*(1), 69-97.

Sinclair, J. (1966). Beginning the Study of Lexis. *In:* Ce, B., Jc, C., Mak, H. & Rh, R. (eds.) *In Memory of J.R. Firth*. London: Longman.

Sinclair, J. (1991). *Corpus, concordance, collocation.* Oxford: Oxford University Press.

Sjoholm, K. (1995). *The Influence of Crosslinguistic, Semantic, and Input Factors on the Acquisition of English Phrasal Verbs: A Comparison between Finnish and Swedish Learners at an Intermediate and Advanced Level.* Abo, Finland: Abo Akademi University Press.

Stubbs, M. (1995). Collocations and semantic profiles: on the cause of the trouble with quantitative studies. *Functions of language, 2*(1), 23-55.

Swain, M. (2000). The output hypothesis and beyond: Mediating acquisition through collaborative dialogue. *In:* Lantolf, J. P. (ed.) *Sociocultural theory and second language learning*. Oxford: Oxford University Press. p.97-114.

Tahaineh, Y. (2010). Arab EFL University Students' Errors in the Use of Prepositions. MJAL2: 1JANUARY2010. ISSN.

Underwood, G., Schmitt, N., & Galpin, A. (2004). The eyes have it: An eye-movement study into the processing of formulaic sequences. *In:* Schmitt, N. (ed.) *Formulaic sequences: Acquisition, processing, and use*. Amsterdam: John Benjamins Publishing Company. p.153-172.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review. *Applied linguistics, 16*(2), 180-205.

Wolter, B. (2006). Lexical network structures and L2 vocabulary acquisition: The role of L1 lexical/conceptual knowledge. *Applied linguistics, 27*(4), 741-747.

Wolter, B., & Gyllstad, H. (in press). Frequency of input and L2 collocational processing: A comparison of congruent and incongruent collocations. *Studies in Second Language Acquisition.*

Wolter, B., & Gyllstad, H. (2011). Collocational Links in the L2 Mental Lexicon and the Influence of L1 Intralexical Knowledge. *Applied linguistics, 32*(4), 430-449.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied linguistics, 21*(4), 463-489.

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

Yamashita, J., & Jiang, N. (2010). L1 influence on the acquisition of L2 collocations: Japanese ESL users and EFL learners acquiring English collocations. *TESOL quarterly, 44*(4), 647-668.

Appendix

Appendix 1. Language Background Questionnaire for nonnative speakers

Please provide the following information by ticking in the box or writing your response.

- Age ……..
- Gender
  - o Male
  - o Female

- Nationality (.....)
- Level of study:
  - o studying an English language course (advanced level)
  - o BA
  - o MA
  - o PhD
- When did you start learning English?
  - o Elementary school
  - o Intermediate school
  - o Other…….……..
- How many years have you been studying English WITHOUT counting school years, i.e. primary, middle and secondary school?
  ……
  ……
- Have you ever lived in an English-speaking country? If yes, please specify for how many years?
  - o No/Yes……
    ……
- Last IELTS (or TOEFL) score:……..
- Mark with an "x" the box that best reflects your level of proficiency in English for each skill:

	Elementary	Intermediate	Advanced	Native-like proficiency
8. Reading				
9. Listening				
10. Speaking				
11. Writing				

Thank you very much for your cooperation.

Appendix 2. Language Background Questionnaire for native speakers

Please provide the following information by ticking (√) in the box or writing your response in the space provided.

- Age ……..
- Gender
  - o Male
  - o Female
- Level of study:
  - o BA
  - o MA
  - o PhD
  - o Staff (please specify ……..……)
- Is English your native language (mother tongue)?
  - o Yes
  - o No
- Which variety of English do you speak?
  - o British
  - o American
  - o Other (…….
- Rate your abilities in the reading skill

□ excellent □ good □ ok □ weak □ very poor

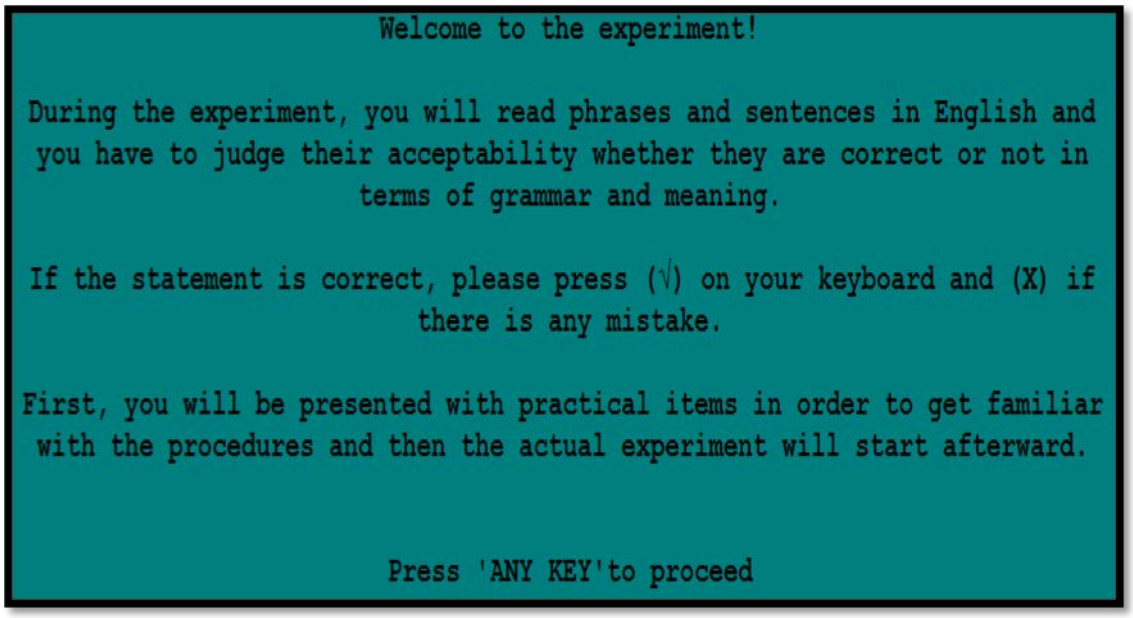
★ How many hours a week do you read in English:

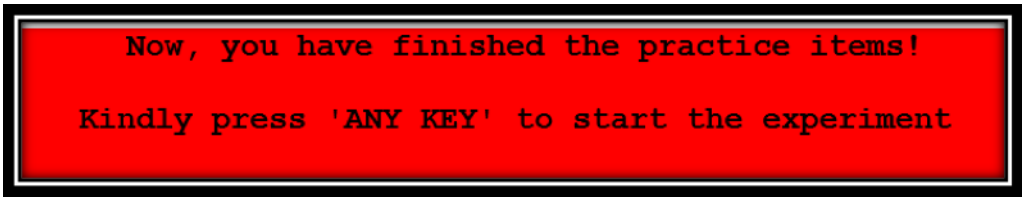
- For Academic Purposes: …… hours.

- For Pleasure: …….. hours.

Thank you very much for your cooperation

Appendix 3. Screenshot of the experiment stages





Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/4.0/)