| **RESEARCH ARTICLE**

A Review of the Effects of Frequency and Congruency on the Processing of Multiword Expressions

Yan Feifei

Postgraduate student, School of Foreign Languages, Ocean University of China, Qingdao, China
Corresponding Author: Feifei Yan, **E-mail**: fayeryan97@163.com

| **ABSTRACT**

More and more attention has been paid to the processing of multiword expressions in recent years. This paper reviews empirical studies that have examined the effects of frequency and congruency and their interactive role on the processing of multiword expressions. The results indicated that although frequency and congruency influence the processing of all kinds of multiword expressions, the studies mostly concentrate on collocations; their interactive role with proficiency has not been specified; research exploring the effect of congruency is limited to translational congruency. Future studies can compare the difference in the processing of different kinds of multiword expressions.

1. Introduction

Research in bilingual language processing extends beyond single words to lexical units known as multiword expressions (MWEs) (Du et al., 2021). Multiword expressions are heterogeneous, consisting of many expression types, such as idioms (*kick the bucket*), lexical bundles (*in the middle of*), binomials (*bride and groom*), collocations (*strong tea*), and other phrasal elements (Siyanova-Chanturia & van Lancker Sidtis, 2019).

These four types of multiword expressions are different from each other. First of all, they are distinctive from each other in terms of their definition. Collocations can be very broadly defined as any frequently co-occurring words or, more accurately, words that co-occur more frequently than we might expect by chance (Biber et al., 1999). They can often be at least partially figurative, and various classifications have been proposed from a phraseological point of view (Howarth, 1998). From a frequency-based perspective, collocations are defined according to certain corpus-derived metrics, such as t-scores (a test of the null hypothesis that there is no connection between two words) or mutual information (MI), which measures the strength of co-occurrence between two words that form a collocation. Typically, an MI score of 3 is taken as the threshold above which a word pair can be considered to be of linguistic interest (Hunston, 2002). Binomials are three-word phrases that are realized in English as an A and B form, where specific word order is preferred (*knife and fork* vs. *fork and knife*) (Benor and Levy, 2006; Carrol and Conklin, 2020). They are highly fixed; that is, the reversed form is rarely used (Carrol and Conklin, 2020). Both form and word order are stressed in the definition of binomials. That is to say, "*A and B*" is a binomial rather than its reversed form "*B and A*", in that "*A and B*" is always more frequent than "*B and A*". For example, "*bride and groom*" is a binomial rather than "*groom and bride*", because "*bride and groom*" is used more frequently than its reversed form. An idiom is a phrase, saying, or a group of words with a metaphorical (not literal) meaning, which has become accepted in common usage. An idiom's symbolic sense is quite different from the literal meaning or definition of the words of which it is made. There are a large number of idioms, and they are used very commonly in all languages. Lexical bundles are word combinations that commonly co-occur in a register and are identified statistically, as opposed to intuitively, based on a cut-off frequency determined arbitrarily by a researcher (Cortes, 2004).

In addition, they have their own characteristics. Collocations are very broadly defined as co-occurring word pairs, and here we define them as combinations of words that are entirely compositional and semantically "free" but which co-occur in conventional and recurrent patterns. Idioms are broadly, but variably, non-decomposable, and in all cases, the meaning of the whole phrase must be retrieved directly from the lexicon to some degree. In contrast, binomials, sequences of *A and B* where specific word order is highly preferred, such as *salt and pepper,* can be literal or figurative and often constitute semantic associates but are highly fixed in the sense that the reversed form is rarely, if ever used. Lexical bundles are the most frequent, recurrent, multiword sequences in a register, which is strictly based on frequency rather than intuitive criteria (Biber, 2004).

The factors affecting second language processing can be divided into three categories according to their sources that is textual factors, learner factors, and lexical ontological factors. Textual factors include text types, the repetition frequency of target words, and the modes of input, which affect learners' language processing. Learner factors include both verbal and non-verbal aspects. Verbal factors mainly refer to learners' proficiency levels. Non-verbal factors include the learner's internal drives, such as metacognitive strategies, learning motivation, learning strategies, and learners' attention to target words in the process of reading. Existing research mainly focuses on the first two categories, while less attention is paid to the third category. Lexical ontological factors refer to lexical ontological features such as part of speech, spelling, morphology, word length, and semantic characteristics of words. These intra-lexical factors also affect the acquisition and processing of target words to a certain extent. There has been growing interest in the effect of frequency and congruency on the processing of English multiword expressions among all these influencing factors.

The focus on the frequency effect is derived from the usage-based view of language, which holds that language learning is experience-driven. The object of second language acquisition is language structures, so the frequency of structures in language input is crucial to acquisition. The different frequency of structures and the degree of use mirror each other. The higher the frequency a word or phrase has, the greater the probability of acquisition and processing. But rule-based opponents insist that phrases are frequency-independent and can be synthesized online directly from mental grammar libraries.

Congruency is always used to measure the influence of the first language (L1) on the second language (L2). It refers to the degree of semantic and structural correspondence between expressions in the first language and the second language. Unlike native speakers in language acquisition, second language learners are inevitably influenced by their mother tongue. Lado(1957), in his masterpiece, indicated that if the language structures of the target language are similar to that of the mother tongue, it will be easy for learners to learn the target language because of the positive transfer from L1; on the contrary, when the language structures of those two languages have many differences in many ways, it will bring about many difficulties for the second language learners, and the second language learners also may make errors about language expressions in the process of learning because of the negative transfer from L1. Before learning a second language, learners have the L1 knowledge and existing concepts in their mind, coupled with the brain's intense tendency to use previous learning experiences; language learners learn the second language consciously or unconsciously compared with previous experiences stored in the first language, and gradually establish a second language structure and category based on that of their first language. Just as the connectionist point of view says, second language learning is the result of analogy with the mother tongue. Language transfer occurs because of this kind of analogy. Existing studies also point out that Chinese EFL learners, no matter at the elementary level or a higher level, tend to rely on the concept categories or patterns of their mother tongue to form their knowledge of English expressions, which leads to transfer bias when there are differences or inconsistencies between Chinese and English concept systems (Lu, 2012).

2. Reviewing the Related Literature
2.1 Studies of Frequency and the Processing of English Multiword Expressions
Frequency is a determining factor in language learning. According to its role in the formation of structure, the frequency can be divided into token frequency and type frequency. The former refers to the frequency of occurrence of a particular word or phrase in specific language input, and the latter refers to the types of expression allowed by a schema construction. Token frequency helps learners to maintain irregular forms and plays an essential role in the construction and representation of language expression. Type frequency helps learners to establish form abstract schemata, which plays an important role in the productivity of language and ensures the creative use of language. Since the former has a greater effect on the overall mental representation and entrenchment, most research focuses on token frequency. The connectionism views of language (MacWhinney, 1997) and the usage-based model (Bybee, 1998) both emphasize the importance of frequency in language acquisition and processing and believe that language users can register the frequency of language input, and at the same time, the representation and processing of language information at all levels can be influenced by experience or frequency of use (Ellis, 2002). Based on this idea, many studies take frequency as an entry, aiming to observe its influence on the processing of multiword expressions.

The research exploring the effect of frequency involves the four kinds of multiword expressions. Most studies found the frequency effect in the processing of multiword expressions, although there were still a few inconsistent results.

2.1.1 Collocations

Since research into frequency intuition had focused primarily on native and, to a lesser degree, nonnative speaker intuitions about single word frequency. Siyanova-Chanturia and Spina (2015) explored the L1 and L2 intuitions about collocation frequency. Native speakers, as well as advanced and intermediate L2 learners, were asked to judge 80 Italian noun-adjective pairings as high frequency, medium frequency, low frequency, or very low frequency. It was found that both L1 and L2 intuitions of high-frequency collocations correlated strongly with corpus frequency. Neither of the two groups of participants exhibited accurate intuitions of medium and low-frequency collocations. About very-low-frequency pairings, L1 but not L2 intuitions were found to correlate with corpora for the majority of the items.

Many studies focused on the psychological reality of multiword expressions in L2. Since many researchers have paid attention to the psychological reality of advanced and intermediate-level L2 learners' processing of collocations; Wang (2015) considered the psychological reality of high-frequency L2 English collocations in elementary and lower intermediate English as foreign language learners to examine whether less competent L2 learners can acquire collocation knowledge. Results showed that learners of both proficiency levels could form psychological representation for the L2 collocations involved, but L2 proficiency cannot affect this representation, thus confirming the learnability of L2 collocations by less proficient L2 learners.

Zhang & Zhang (2020) tested the effect of collocational frequency and directionality on the processing of English verb-noun collocations. Both intermediate and advanced English as L2 speakers and English native speakers did a primed lexical decision task. The results showed that priming effects in NSs were predicted by the scores of association strength and collocational frequency, while those in advanced L2 learners were only predicted by directionality and collocational frequency. What's more, the intermediate L2 learners spent comparable time processing the two types of collocations and the control items.

Jiang et al. (2020) recorded the eye movements of third and fourth graders, as well as adults (first-language Mandarin), as they read phrases varying in frequency embedded in sentence context. Jiang et al. (2020) explored how phrase frequency, operationalized as phrase type (collocation vs. control) or (continuous) phrase frequency, and age might influence participants' reading. Adults read collocations and higher frequency phrases consistently faster than control and lower frequency phrases, respectively. Critically, fourth, but not third, graders read collocations and higher frequency phrases faster than control and lower frequency sequences, respectively, although this effect was largely confined to a late measure. The results reaffirmed phrase frequency effects in adults and pointed to emerging phrase frequency effects in primary school children.

Collocational frequency, individual word frequency, and their association were also explored. Öksüz et al. (2021) investigated the effects of individual word frequency, collocational frequency, and association on L1 and L2 collocational processing. An acceptability judgment task was administered to L1 and L2 speakers of English. Response times were analyzed for three types of adjective-noun pairs: high-frequency, low-frequency, and baseline items. Öksüz et al. (2021) extended previous research by examining whether the effects of individual word and collocation frequency counts differ for L1 and L2 speakers' processing of collocations and also compared the extent to which L1 and L2 speakers' response times are affected by mutual information and log Dice scores, which are corpus-derived association measures. Both groups of participants demonstrated sensitivity to individual word and collocation frequency counts. However, there was a reduced effect of individual word frequency counts for processing high-frequency collocations compared to low-frequency ones. Both groups of participants were similarly sensitive to the association measures used.

The studies concentrating on the processing of collocations have found that even lower intermediate and elementary second language learners can form psychological representations for the processing of second language collocations. However, some studies also indicated that second language learners spent comparable time processing collocations and the controls. Therefore, more attention should be paid to the psychological reality of second language learners of different proficiencies.

2.1.2 Lexical Bundles

Native learners seemed to learn and store frequency information about multiword expressions. There is mounting evidence that native learners are sensitive to distributional information at many grain sizes. Arnon and Snider (2010) focused on the distributional properties of words, the units they consist of (morphemes, phonemes), and the syntactic structures they appear in (verb-categorization frames, syntactic constructions). The results of a series of studies showed that native speakers of English were sensitive to the frequencies of lexical bundles (e. g. *don't have to worry*): more frequent lexical bundles were processed faster. The effect was not reducible to the frequency of the individual words or substrings and was observed across the entire frequency range (for low, mid-and high-frequency phrases).

The above study of the frequency effect focused on the processing of native learners, and the frequency effect has also been found in the processing of second language learners.

Zhong and Cadierno (2015) employed an online grammaticality judgment task on the E-Prime program to investigate the cognitive processing of second language high-frequency lexical bundles on 30 Chinese and 30 Danish English as Foreign Language (EFL) learners. It was found that Second language learners in the Chinese and Danish EFL contexts also had the processing advantages of English lexical bundles as natives. From high frequency to low frequency of the stimuli, the reaction time forms a continuum from fast to slow, providing evidence for the frequency effect.

Xu and Wang (2015) investigated the effects of frequency and English proficiency on Chinese learners' processing of L2 lexical bundles by comparing the performance of a grammaticality judgment task among two groups of non-English majors at two proficiency levels. The results showed that both higher and lower proficiency learners benefited from the processing advantage of high-frequency sequences, and this advantage can be observed not only for phrasal sequences but for non-phrasal ones. The effect of lexical bundles on lower proficiency learners is higher than that on higher proficiency learners, and higher proficiency learners generally perform better than lower proficiency learners on different types of L2 formulaic sequences (with lower error rate and faster speed), but the effect of L2 proficiency is more pronounced on the identification of non-phrasal sequences than on phrasal ones.

The studies concentrating on the processing of lexical bundles have found the frequency effect not only in native speakers but also in second language learners. What's more, it was proved that higher proficiency learners generally perform better than lower proficiency learners.

2.1.3 Idioms

Siyanova-Chanturia et al. (2011) investigated the online processing of idioms in a biasing story context by native and non-native speakers of English using eye-tracking. The stimuli are idioms used figuratively (*at the end of the day* – 'eventually'), literally (*at the end of the day* – 'in the evening'), and novel phrases (*at the end of the war*). Native speaker results indicated a processing advantage for idioms over novel phrases, as evidenced by fewer and shorter fixations. Further, no processing advantage was found for figurative idiom uses over literal ones in a full idiom analysis or a recognition point analysis. Contrary to native speaker results, non-native findings suggested that L2 speakers process idioms at a similar speed to novel phrases. Further, figurative uses were processed more slowly than literal ones.

The processing advantage for idioms was found in native speakers but not in non-native speakers, which is different from the research results in collocations and lexical bundles. More research about the processing advantage for idioms is needed.

2.1.4 Binomials

Siyanova-Chanturia et al. (2011) explored whether speakers are sensitive to the frequency with which phrases occur in language or not. An eye-tracking study was reported by examining the processing of English binomials that differ in phrasal frequency by native and proficient nonnative English speakers. Participants read sentences containing three-word binomial phrases (*bride and groom*) and their reversed forms (*groom and bride*), which are identical in syntax and meaning but differ in phrasal frequency. Mixed-effects modeling revealed that native speakers and nonnative speakers, across a range of proficiencies, are sensitive to the frequency with which phrases occur in English. Results also indicated that native speakers and higher proficiency nonnatives are sensitive to whether a phrase occurs in a particular configuration (binomial vs. reversed) in English.

Using online grammaticality judgment tasks and choosing Chinese English learners of different English proficiency as subjects, Wang et al. (2021) explored the processing mechanism of English binomials with different frequencies. The study showed that binomials are psychologically realistic to some extent. In addition, the results also showed that there is a frequency effect on the processing of binomials of English learners.

Besides exploring frequency using a corpus-driven method, the frequency of occurrence was also viewed as a variable. Conklin and Carrol (2021) monitored native English speakers' eye movements when reading short stories containing existing (conventional) patterns (*time and money*), seen once, and novel patterns (*wires and pipes*), seen one to five times. Subsequently, readers saw both existing and novel phrases in the reversed order (*money and time*; *pipes and wires*). In four to five exposures, much like existing lexical patterns, novel ones demonstrate a processing advantage. Sensitivity to lexical patterns, including the co-occurrence of lexical items and the order in which they occur, arises rapidly and automatically during natural reading.

The frequency effect of binomials has been found in native and higher-proficiency non-native learners. However, the lower-proficiency English as foreign language learners is rarely explored. In addition, more technological methods such as eye-tracking and events-related potentials (ERP) can be used in future studies.

2.2 Studies of Congruency and the Processing of Multiword Expressions

The influence of the first language on second language acquisition and processing is an important topic in the field of second language acquisition and processing. Under the background of interdisciplinary research of applied linguistics, studies on the influence of the first language are no longer confined to the comparison and analysis of the similarities and differences between first and second language. Rather, it becomes comprehensive research involving more such as language, cognition, psychology, and other multilevel variables, so that we can explore the influence of L1 on the processing of a second language more scientifically. Congruency is always used to measure the influence of the L1 on L2. It refers to the degree of semantic and structural correspondence between expressions in the first language and second language. An L2 expression is congruent if it has a word-for-word translation equivalent form in the L1 (Conklin and Carrol, 2018).

Most studies exploring the effect of congruency on the processing of multiword expressions concentrate on collocations because of their characteristics. And recently, some research has gradually focused on binomials.

2.2.1 Collocations

Yamashita and Jiang (2010) investigated first language influence on the acquisition of second language collocations by comparing the performance on a phrase-acceptability judgment task among native speakers of English, Japanese English as a second language (ESL) users, and Japanese English as foreign language learners. The test materials included both congruent and incongruent collocations. EFL learners made more errors and reacted more slowly to incongruent collocations than congruent collocations. ESL users generally performed better than EFL learners (lower error rate and faster speed), but they still made more errors on incongruent collocations than on congruent collocations. However, the L1 effect was not apparent in the ESL users' reaction time. The results suggested that both L1 congruency and L2 exposure affected the acquisition of L2 collocations, with the availability of both maximizing this acquisition. What's more, it was difficult to acquire incongruent collocations even with a considerable amount of exposure to L2, and once stored in memory, L2 collocations were processed independently of L1.

From the perspective of the L2 mental lexicon, Wei and Zhang (2017) explored how Chinese EFL learners' thinking influenced their English collocational links. The results showed that there would be no significant difference between the reaction time used in recognizing congruent collocations and that used in recognizing non-congruent collocations if the degree of collocation familiarity is high. Likewise, the reaction time used in recognizing congruent collocations is significantly shorter than that used in recognizing non-congruent ones if the degree of collocation familiarity is low. These results suggested that Chinese thinking exerts a positive influence on collocational links in the English mental lexicon, which may become diminished when the collocations are familiar to learners. What's more, Chinese thinking can be triggered no matter whether collocations are familiar to learners or not. In addition, these results accounted for the comparable reaction time of ESL and EFL learners in Yamashita and Jiang's study.

Zhang (2017) investigated processing patterns of verb-noun, noun-noun, and adjective-noun collocations by native English speakers and Chinese learners of English via online lexical decision tasks. Congruent, English-only, and Chinese-only collocations were carefully chosen from the British National Corpus, Corpus of Contemporary America, and Chinese Learner English Corpus corpora. Results showed that native speakers spent the shortest time judging legitimate noun-noun collocations while Chinese learners were the most sensitive to verb-noun collocations. The three kinds of collocation have different processing effects in the collocation types with different degrees of semantic correlation.

Sonbul and El-Dakhs (2020) looked at the interaction between congruency and level of proficiency and directly compared untimed collocation recognition assessed through traditional tests to timed recognition evident in psycholinguistic tasks. Sonbul and El-Dakhs administered both types of form recognition measures to 228 female Saudi English as foreign language learners in two separate experiments: a traditional multiple-choice test and a timed acceptability judgment task. The timed acceptability judgment also tested 37 native speakers of English as a baseline for comparison. Congruency, estimated proficiency measured by a vocabulary test, and the interaction between the two was evaluated as predictors of untimed and timed recognition through mixed-effects modeling. Results showed that congruency and estimated proficiency had a clear effect on untimed and timed recognition. In addition, the effect of proficiency was clearer on timed recognition with a gradual decrease in the first language effect as proficiency increased, getting closer to native-like collocation processing.

It was found that congruency had a clear effect on the processing of collocations. Its interactive role with proficiency, familiarity, second language as well as the category of collocation has been explored. More studies can focus on its interactive role related to learners' internal drives, such as metacognitive strategies, learning motivation, and learning strategies.

2.2.2 Binomials
Studies also concentrated on binomials besides collocations in recent years. Du et al. (2021) investigated cross-language influences in the processing of binomial expressions (*knife and fork*), from a first language to a second language and from L2 to L1. Two groups of unbalanced bilinguals (Chinese/L1-English/L2 and English/L1-Chinese/L2) and a control group of English monolinguals performed a visual lexical decision task that incorporated unmasked priming. To assess cross-language influences, three types of expressions were used: congruent binomials, English-only binomials, and Chinese-only binomials translated into English. Lexical decision latencies to the last word (*fork*) in a binomial (*knife and fork*) were compared with response latencies to the same word in a matched control phrase (*spoon and fork*). It was found that Chinese-English bilinguals showed a significant priming effect for congruent binomials but no facilitation for English-only binomials; English-Chinese bilinguals showed a trend toward priming for congruent binomials, which did not reach statistical significance, and no priming for English-only binomials; English monolinguals showed comparable priming for congruent and English-only binomials. Concerning the Chinese-only binomials, none of the three participant groups showed priming for translated Chinese-only binomials over controls.

Cross-language influences were found in the processing of binomials. The interactive role of congruency and other influencing factors can be explored in future studies.

2.3 Studies of the Interactive role of Frequency and Congruency on the Processing of English Multiword Expressions
It was proved that frequency, congruency, as well as proficiency level have an interactive role in the processing of multiword expressions. However, the interactive effect research concentrates on collocations and lexical bundles. Future research can take idioms and binomials into consideration.

2.3.1 Collocations
Wolter and Gyllstad (2013) investigated the influence of frequency effects on the processing of congruent and incongruent collocations in a second language. An acceptability judgment task was administered to native and Swedish advanced nonnative English speakers to assess response times to and error rates for these collocations along with a matched set of unrelated items. The results suggested that advanced learners are highly sensitive to frequency effects for L2 collocations, and there was a continued influence of L1.

Besides the interactive role of collocational frequency, L1 congruency, and L2 proficiency, many research also focused on the effects of word frequency on L2 collocational processing. Choosing congruent collocations, English-only collocations, Japanese-only collocations, and baseline items as experiment materials, Wolter and Yamashita (2017) investigated the effects of collocational frequency, L1 congruency, and L2 proficiency as well as word frequency on L2 collocational processing. Two groups of L1 Japanese intermediate and advanced speakers of English and one group of English native speakers performed an online acceptability judgment task on these four types of adjective-noun constructions. Response times were analyzed using mixed-effects modeling and correlations. In contrast to NSs, NNSs processed congruent collocations significantly faster than English-only collocations. As for frequency, all three groups demonstrated sensitivity to both word-level and collocation-level frequency. However, the distributions differed across the three groups. Wolter and Yamashita concluded that age/order of acquisition effects provided the best explanation for the congruency results. Regarding the frequency results, the findings conflict with claims that NNSs may process formulaic sequences differently compared with NSs.

Some researchers chose congruent and incongruent collocations as their experiment materials. Fang and Zhang (2019) selected low and advanced proficiency groups of Chinese EFL learners as participants. Online acceptability judgment tasks on same-node congruent and incongruent collocations were adopted, and learners' analysis revealed that the two proficiency groups differed significantly in error rates but insignificantly in reaction time of responses to the congruent collocations. Word frequency and collocational frequency interactively affected the reaction time of the low-proficiency group, while only collocational frequency had an effect on the reaction time of the high-proficiency group, showing proficiency influence on lexical frequency, and translational congruency had a significant mediating effect merely on the reaction time of the low proficiency group and this effect varied inversely with L2 lexical frequency. Zhang and Fang (2020) adopted both native speakers of English and low and advanced proficiency Chinese EFL learners as participants. Results showed that both NSs and NNSs processed more accurately same-translated collocations but not faster than judging incongruent collocations. NNSs' language proficiency modulated the effects of constituent word frequency and collocational frequency on the processing output, and lexical frequency played a modulating role in the processing of all types of collocations for both NSs and NNSs.

The studies above concentrated on the low and advanced proficiency groups of Chinese EFL learners and more proficiency levels worth exploring in second language processing. Fang and Zhang (2021) adopted twenty English native speakers and sixty adult Chinese-English intermediate and advanced bilinguals as participants and investigated the effects of first language congruency, word frequency, collocational frequency, second language proficiency, and their combined effects on L2 collocational processing.

Twenty English native speakers and 60 adult Chinese-English intermediate and advanced bilinguals completed an online phrasal decision task. A judgment was made about whether congruent collocations, incongruent collocations, and non-collocational baseline items were commonly used in English. Response time and accuracy data were analyzed with mixed-effects models, with a focus on the interactions of the variables of interest. Results demonstrated similar effects to the former studies. Both L2 groups were significantly more accurate but not significantly faster in judging the congruent collocations than the incongruent collocations, and unexpectedly similar patterns were observed with the L1 group. All participants were sensitive to word frequency and collocational frequency, but they differed in the manner and degree of such sensitivity, and the processing of incongruent collocations showed greater combined effects of word frequency and collocational frequency than in the processing of congruent collocations.

2.3.2 Lexical Bundles

Besides the investigation of collocations, lexical bundles are also a focus. Xu and Wang (2015) investigated the effects of sequence frequency, congruency, and English proficiency on Chinese learners' processing of L2 formulaic sequences, by comparing their performance on a grammaticality judgment task among two groups of non-English majors at two proficiency levels. The results showed that learners' sequence processing is influenced by all three factors. More specifically, both groups of learners benefited from the processing advantage of formulaic sequences, and this advantage can be observed not only in congruent sequences but also in incongruent sequences. The effect of congruency is higher on the less proficient learners than on the more proficient ones, and higher proficient learners generally perform better than their lower proficient counterparts on different types of L2 sequences with a lower error rate and faster speed.

3. Summary and Conclusion

This study reviews studies that explored the effects of frequency, congruency, and their interactive role on the processing of multiword expressions. It was found that frequency and congruency interactively influence the processing of multiword expressions. However, there are also some limitations to this study. This study only focuses on the effect of frequency and congruency. Besides frequency and congruency, there are also many other influencing factors.

Frequency is a textual factor, congruency is a lexical ontological factor, and learners' proficiency level is the verbal aspect of the learner factor. Existing studies have mostly explored the interaction of these three factors while ignoring the non-verbal aspects of the learner factor.

Future research that examines the frequency, congruency, and proficiency should further elucidate the role of non-verbal aspects of the learner factor and other factors not involved, such as metacognitive strategies, learning motivation, text types, and so on. Moreover, lower proficiency second language learners should be given more attention since there were inconsistent results for the lower proficiency second language learners. In addition, it's worth exploring whether the result of one type of multiword expression can apply to all kinds of multiword expressions, so different types of multiword expressions also need to be compared in future studies.

Conflicts of Interest: The authors declare no conflict of interest.
Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations or those of the publisher, the editors, and the reviewers.

References:
[1]   Arnon, I., McCauley, S. M. & Christiansen, M. H. (2017). Digging up the building blocks of language: Age-of-acquisition effects for multiword phrases. *Journal of Memory and Language*, 92, *12*(1).
[2]   Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*(1), 1–48.
[3]   Bod, R. (2006). Exemplar-based syntax: How to get productivity from exemplars. *The Linguistic Review*, 23, 291–320. doi:10.1515/TLR.2006.012
[4]   Bybee, J. L. (2006). From usage to grammar: The mind's response to repetition. *Language, 82*(4), 711–733.
[5]   Bybee, J., & Beckner, C. (2009). *Usage-based theory. In Heine, B, & Narrog, H. (Eds.),* The Oxford Handbook of Linguistic Analysis, Oxford Handbooks online. 1–30.
[6]   Carrol, G., & Conklin, K. (2020). Is All Formulaic Language Created Equal? Unpacking the Processing Advantage for Different Types of Formulaic Sequences. *Language And Speech*, *63*(1), 95-122. http://doi.org/10.1177/0023830918823230
[7]   Conklin, K., & Carrol, G. (2021). Words Go Together Like 'Bread and Butter': The Rapid, Automatic Acquisition of Lexical Patterns. *Applied Linguistics*, *42*(3), 492-513. http://doi.org/10.1093/applin/amaa034
[8]   Davies, M. (2004). BYU-BNC: The British National Corpus. BYE, Brigham Young University. http://corpus. byu.edu/bnc
[9]   Davies, M. (2008). *The corpus of contemporary American English: 520 million words*, 1990-present. BYE, Brigham Young University.

https://corpus.byu.edu/coca/

[10] Ding, C., & Reynolds, B. L. (2019). The effects of L1 congruency, L2 proficiency, and the collocate-node relationship on the processing of L2 English collocations by L1-Chinese EFL learners. *Review of Cognitive Linguistics*, *17*(2), 331–357

[11] Du, L., Elgort, I., & Siyanova-Chanturia, A. (2021). Cross-Language Influences in the Processing of Multiword Expressions: From a First Language to Second and Back. *Frontiers in Psychology*, *12,* http://doi.org/10.3389/fpsyg.2021.666520

[12] Fang, N., & Zhang, P. (2021). L1 Congruency, Word Frequency, Collocational Frequency, L2 Proficiency, and Their Combined Effects on Chinese–English Bilinguals' L2 Collocational Processing. *International Journal of Bilingualism*, *25*(5),1-17. http://doi.org/10.1177/13670069211024747

[13] Gablasova, D., Brezina, V., & McEnery, T. (2017). Collocations in Corpus-Based Language Learning Research: Identifying, Comparing, and Interpreting the Evidence. *Language Learning*, *67*(S1), 155-179. http://doi.org/10.1111/lang.12225

[14] Goldberg, A. (2006). Constructions at work: *The nature of generalization in language*. Oxford, England: Oxford University Press.

[15] Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning*, *66*(2), 296–323.

[16] Hallin, A. E., & Van Lancker Sidtis, D. (2017). A closer look at the formulaic language: Prosodic characteristics of Swedish proverbs. *Applied Linguistics*, *38*(1), 68–89. http:// dx.doi.org/10.1093/applin/amu078.

[17] Howarth, P. A. (1998). Phraseology and second language proficiency. *Applied Linguistics*, *19*(1), 24–44.

[18] Jacobs, C. L., Dell, G. S., Benjamin, A. S., & Bannard, C. (2016). Part and whole linguistic experience affect recognition memory for multiword sequences. *Journal of Memory and Language*, *87*, 38–58.

[19] Janssen, N., & Barber, H. (2012). Phrase frequency effects in language production. *PLoS ONE*, 7(3), e33202. http://dx.doi.org/10.1371/journal.pone.0033202.

[20] Jiang, N. (2012). *Conducting reaction time research in second language studies*. Routledge.

[21] Jiang, S., Jiang, X., & Siyanova-Chanturia, A. (2020). The Processing of Multiword Expressions in Children and Adults: An Eye-tracking Study of Chinese. *Applied Psycholinguistics*, *41*(4), 901-931. http://doi.org/10.1017/S0142716420000296

[22] Jolsvai, H., McCauley, S. M., & Christiansen, M. H. (2020). Meaningfulness Beats Frequency in Multiword Chunk Processing. *Cognitive Science*, *44*(10) http://doi.org/10.1111/cogs.12885

[23] Lau, E., Namyst, A., Fogel, A., & Delgado, T. (2016). A direct comparison of N400 effects of predictability and incongruity in adjective-noun combination. *Collabra*, *2*(1), 1-19

[24] Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, *29*(4), 694–716.

[25] Lee, S. (2021). The Effect of L2 Proficiency, L1 Congruency, and Collocation Frequency on L2 Collocation Processing: an Experimental Study of Korean EFL Learners. *Korean Journal of English Language and Linguistics*, *21*(10), 60-84. http://doi.org/10.15738/kjell.21..202110.1060

[26] Mondini, S., Jarema, G., Luzzatti, C., Burani, C., & Semenza, C. (2002). Why is "red cross" different from "yellow cross"?: A neurophysiological study of non-adjective agreement within Italian compounds. *Brain and Language*, *81*, 621– 634. doi:10.1006/brln.2001.2552

[27] Morgan, E., & Levy, R. (2016). Abstract Knowledge versus Direct Experience in Processing of Binomial Expressions. *Cognition*, *157*, 384-402. http://doi.org/10.1016/j.cognition.2016.09.011

[28] Öksüz, D., Brezina, V., & Rebuschat, P. (2021). Collocational Processing in L1 and L2: The Effects of Word Frequency, Collocational Frequency, and Association. *Language Learning*, *71*(1), 55-98. http://doi.org/10.1111/lang.12427

[29] Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing*, *18*(1), 55–88.

[30] Siyanova-Chanturia, A. (2017). Researching the Teaching and Learning of Multi-word Expressions. *Language Teaching Research*, *21*(3), 289-297. http://doi.org/10.1177/1362168817706842

[31] Siyanova-Chanturia, A., Conklin, K., Caffarra, S., Kaan, E., & van Heuven, W. J. B. (2017). Representation and Processing of Multi-word Expressions in the Brain. *Brain and Language*, *17*(5), 111-122. http://doi.org/10.1016/j.bandl.2017.10.004

[32] Siyanova-Chanturia, A., Conklin, K., & Schmitt, N. (2011). Adding More Fuel to the Fire: An Eye-tracking Study of Idiom Processing by Native and Non-native Speakers. *Second Language Research*, *27*(2), 251-272. http://doi.org/10.1177/0267658310382068

[33] Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. J. B. (2011). Seeing a Phrase "Time and Again" Matters: The Role of Phrasal Frequency in the Processing of Multiword Sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *37*(3), 776-784. http://doi.org/10.1037/a0022531

[34] Siyanova-Chanturia, A., & Spina, S. (2015). Investigation of a native speaker and second language learner's intuition of collocation frequency. *Language Learning*, *65*(3), 533-562. http://doi.org/10.1111/lang.12125

[35] Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism: Language and Cognition*, *18*(3), 419–437.

[36] Sonbul, S., & El-Dakhs, D. (2020). Timed versus untimed recognition of L2 collocations: Does estimated proficiency modulate congruency effects? *Applied Psycholinguistics*, *41*(5), 1197-1222. http://doi.org/10.1017/S014271642000051X

[37] Sonbul, S., & Siyanova-Chanturia, A. (2021). Research on the on-line processing of collocation: Replication of Wolter and Gyllstad (2011) and Millar (2011). *Language Teaching*, *54*(2), 236-244. http://doi.org/10.1017/S0261444819000132

[38] Sosa, A. V., & MacFarlane, J. (2002). Evidence for frequency-based constituents in the mental lexicon: Collocations involving the word of. *Brain and Language*, *83*(2), 227–236.

[39] Tomasello, M. (2003). Constructing a language: A usage-based theory of language acquisition. *Harvard University Press*.

[40] Toomer, M., & Elgort, I. (2019). The development of implicit and explicit knowledge of collocations: A conceptual replication and extension of Sonbul and Schmitt (2013). *Language Learning*, *69*(3), 783–783.

[41] Van Lancker, D., Canter, G., & Terbeek, D. (1981). Disambiguation of ditropic sentences: Acoustic and phonetic cues. *Journal of Speech and Hearing Research*, *24*, 330 –335.

[42] Vilkaitė, L., & Schmitt, N. (2019). Reading collocations in an L2: Do collocation processing benefits extend to non-adjacent collocations?

Applied Linguistics, *40*(2), 329-354. http://doi.org/10.1093/applin/amx030

[43] Wang, Q. (2015). The psychological reality of L2 general collocations by low-proficiency Chinese EFL learners. *Modern Foreign Languages, 38*(2), 206-216.

[44] Wang, Y. (2021). Effects of frequency and lexical-semantic relations of component words on Chinese English learners' processing of English binomials. *Foreign Language Teaching and Research*, *53*(4), 532-544.

[45] Warren, P., Boers, F., Grimshaw, G., & Siyanova-Chanturia, A. (2018). The Effect of Glossing Type on Learners' Intake of New Words During Reading. *Studies in Second Language Acquisition*, *40*(4), 883-906. http://doi.org/10.1017/S0272263118000177

[46] Wolter, B., & Gyllstad, H. (2013). Frequency of Input and L2 Collocational Processing. *Studies in Second Language Acquisition*, *35*(3), 451-482. http://doi.org/10.1017/S0272263113000107

[47] Wolter, B., & Yamashita, J. (2015). Processing collocations in a second language: A case of first language activation? *Applied Psycholinguistics*, *36*(5), 1193-1221. http://doi.org/10.1017/S0142716414000113

[48] Wolter, B., & Yamashita, J. (2018). Word Frequency, Collocational Frequency, L1 Congruency, and Proficiency in L2 Collocational Processing. *Studies in Second Language Acquisition*, *40*(2), 395-416. http://doi.org/10.1017/S0272263117000237

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

[50] 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. http://doi.org/10.5054/tq.2010.235998

[51] Yeldham, M. (2020). Does the presence of formulaic language help or hinder second language listeners' lower-level processing? *Language Teaching Research*, *24*(3), 338-363. http://doi.org/10.1177/1362168818787828

[52] Zhang P. (2017). L1 semantic priming and word type effect: Processing of Chinese EFL learners' English collocations. *Foreign Language Teaching and Research, 49*(6), 818-832.