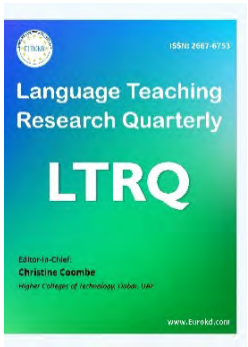

Language Teaching Research Quarterly

2024, Vol. 45, 106–122

Exploring the Relationship between English Proficiency and Influential Factors on Productive Knowledge of Multi-Word Units to Create Effective Learning Materials

Maryam Barghamadi[1]*, Amanda Müller[1], James Rogers[2], Joanne Arciuli[1]

[1]College of Nursing and Health Sciences, Flinders University, Adelaide, Australia
[2]Faculty of Foreign Studies, Meijo University, Japan

Received *11 June 2024* ***Accepted*** *14 November 2024*

Abstract

The study explored the impact of proficiency, frequency, L1 – L2 congruency, and semantic transparency on the knowledge of multi-word units (MWUs) among Iranian L2 learners (N = 256). A gap-filling test was used to assess learners' productive knowledge, employing a high-frequency MWU list created with the lemmatised concgramming method. The list, which included 11,212 MWUs across four frequency levels, was ranked by L1 – L2 congruency and semantic transparency. As revealed by regression models with bootstrapping, the results showed a significant positive correlation between MWU knowledge and proficiency (IELTS scores). It was observed that knowledge of MWUs decreased with lower frequency, and participants scored higher on congruent and transparent items. The study also found that frequency and semantic transparency influenced IELTS scores, with a significant interaction between congruency and incongruent items. These findings underscore the crucial role of L1 – L2 congruency in learning MWUs and provide hope for educators, as they offer insights to enhance learning by prioritising L1 – L2 congruency in English MWU resources. This study's findings can potentially lead to improved language teaching methodologies and curriculum design, offering a pathway to optimise learning outcomes.

Keywords: *IELTS Score, L1 – L2 Congruency, Lemmatised Concgramming Method, Multi-word Units (MWUs), Semantic Transparency*

How to cite this article (APA 7th Edition):

Barghamadi, M., Müller, A., Rogers, J., & Arciuli, J. (2024). Exploring the relationship between English proficiency and influential factors on productive knowledge of multi-word units to create effective learning materials. *Language Teaching Research Quarterly*, *45*, 106-122. https://doi.org/10.32038/ltrq.2024.45.06

* Corresponding author.
E-mail address: barg0008@flinders.edu.au https://doi.org/10.32038/ltrq.2024.45.06

Introduction[1]

Understanding a word involves grasping the relationship between its form and meaning, recognising its typical collocations, and being aware of the contexts in which it is used (Frankenberg-Garcia, 2018). This knowledge is essential for appropriately using multi-word units (MWUs), critical components of fluent language production (Barghamadi, 2024). Recent studies in second-language (L2) research and education have emphasised the importance of MWU knowledge, particularly concerning their processing (e.g., Yamagata et al., 2023). Various terms refer to MWUs, such as *formulas, formulaic sequences, lexical phrases, ready-made utterances, and prefabricated chunks* (e.g., Wray, 2002). While no universally accepted definition of MWUs exists, words that co-occur predictably are traditionally known as *collocations* (Rogers et al., 2021). In this research, the terms' collocation' and 'MWU' are used interchangeably to refer to a single entity, identified by their frequent co-occurrence (Hoey, 2005), as identified through a lemmatised concgramming methodology (see Cheng et al., 2006).

MWUs play a vital role in spoken and written discourse (e.g., Erman & Warren, 2000). While correlation does not definitively establish causation, research suggests a positive relationship between knowledge of MWUs and L2 proficiency metrics (Vu & Peters, 2022b). Thus, MWUs are considered significant components of language acquisition. However, even advanced L2 learners face challenges in producing MWUs (e.g., Boers et al., 2014). Evidence from corpus-based studies indicates that L2 learners' productive knowledge of MWUs is often limited, with studies showing that L2 learners' essays contain about half the number of collocations compared to native speakers (Laufer & Waldman, 2011). Furthermore, L2 learners may either refrain from using certain collocations or overuse those they have mastered. Additionally, gap-filling tests of collocation knowledge have concluded that L2 learners have limited knowledge of MWUs (e.g., Sonbul et al., 2023).

Assessing learners' knowledge of MWUs is crucial for improving language learning programs, as teachers need to focus on specific MWUs while considering various influences that affect acquisition and instructional choices. Several studies have examined factors influencing MWU learnability (e.g., Boone et al., 2022; Ding & Reynolds, 2019; Fang & Zhang, 2021; Gyllstad & Wolter, 2016; Sonbul, 2015; Sonbul et al., 2024a; Wolter & Yamashita, 2018). Key factors include semantic transparency, L1 – L2 congruency, and learners' L2 proficiency. While significant attention has been given to identifying these factors, their application in developing collocational resources has received limited emphasis based on influential factors (Barghamadi et al., 2023; Rogers, 2017). Within this context, Barghamadi et al. conducted a comprehensive analysis of around 11,000 MWUs, focusing on the congruency between learners' first and target languages. Their findings underscored the importance of congruency in selecting collocations suitable for direct instruction, mainly when many items are incongruent.

The present research builds upon the foundational work of Barghamadi et al. (2023), emphasising the critical role of L1 – L2 congruency in developing effective collocational resources. This follow-up investigation aims to delve deeper into the key factors influencing

[1] This article is based on a dissertation submitted to Flinders University, South Australia, in partial fulfilment of the requirements for the degree of Doctor of Philosophy. The original dissertation can be accessed at: (https://theses.flinders.edu.au/view/1697d081-df3e-49a6-8b91-23a36ab666d1/1).

collocation acquisition and further refine these resources. While the previous study highlighted the importance of L1 – L2 congruency, the current research expands its scope to examine how various factors—specifically, L1 – L2 congruency, frequency levels, semantic transparency, and proficiency level—affect the acquisition of MWUs. The objective is to reaffirm the significance of congruency and uncover how these factors interact to influence MWU acquisition. Moreover, evaluating a specific learner group's collocation knowledge can identify areas requiring focus, thus facilitating the creation of adequate learning resources.

Accordingly, this study explores the relationship between the aforementioned influential factors by assessing Persian EFL (English as a Foreign Language) learners' productive knowledge of collocations. To achieve this, we address the following research questions (RQs):
RQ$_1$: Is there a relationship between knowledge of MWUs and Persian-speaking learners' proficiency measured by IELTS scores?
RQ$_2$: To what extent does Persian-speaking learners' productive collocation knowledge change over MWU frequency levels?
RQ$_3$: To what extent do the frequency level, L1 – L2 congruency, and semantic transparency of MWUs predict Persian-speaking learners' proficiency?

This research explores the relationship between English proficiency and the influential factors affecting productive knowledge of MWUs. The study seeks to provide practical guidance for educators and researchers in language teaching and learning by investigating these aspects. Ultimately, the goal is to enhance the effectiveness of collocational resources and instructional strategies, creating more effective learning materials.

Literature Review

Definition and Identification of Collocations

There are two major approaches to defining collocations in L2 research: the phraseological approach (e.g., Cowie, 1994) and the frequency-based approach (e.g., Howarth, 1996). By analysing co-occurrence restrictions among words and semantically transparent and restricted meanings, the phraseological approach distinguishes idioms (*pay their last respects*), collocations (*pay attention*), and free word combinations (*pay a bill*). Conversely, the frequency-based approach defines collocations as any words that co-occur more frequently than one would expect by chance by using statistical measures.

Rogers (2017) combined the frequency-based approach with the lemmatised concgramming method to identify useful MWUs from Corpus of Contemporary American English (COCA) (Davies, 2008). The concgramming method (Cheng et al., 2006) identifies constituency variation (AB, ACB) and positional variation (AB, BA). For example, in the lemma, pair *come/terms*, MWUs such as *come to terms, come to terms with, to come to terms with, coming to terms,* and so forth, are counted together.

Expanding upon the concgramming method, Rogers (2017) developed a novel technique to find exemplary MWUs from lemmatised concgram data. This technique identified an exemplar of each lemmatised concgram. For instance, for the lemma pair *come/terms,* a variety of MWUs are counted together, but which of these are the most common? His technique took the most frequent core (in this case, *come to terms*) and expanded it if any other MWUs contained the core and their frequencies comprised 50% or more of that of the core. For

instance, if *come to terms* occurred 50 times, and *come to terms with* occurred 25 times, then the exemplary MWU was extended to *come to terms with.*

Thus, in the current research, collocations/MWUs are operationally defined as lemmatised concgrams. It departs from the traditional approach based upon the frequency of co-occurrence because the frequency and mutual information corpus data are used to search for collocations using high-frequency pivot words and collocates (only nouns, verbs, adjectives, and adverbs). Whether or not MWUs are phrasal verbs or idioms, et cetera, is not a factor that is considered; thus, collocations and MWUs are treated the same way in this study (see Barghamadi, 2024 for more details).

Factors Affecting Collocational Knowledge

Several variables influence L2 collocation acquisition and processing, with congruency between L1 and L2 being a crucial factor. Congruency has been extensively studied, and research indicates that congruent collocations, which have word-for-word translations in the learner's L1, impose a lower learning burden than incongruent ones (Boone et al., 2022; Ding & Reynolds, 2019). To give an example, *take a shower* (/dɒʃ gerftæn/*دوش گرفتن)* and *take a photo* (*/æks gerftæn/ عکس گرفتن)* are congruent for Persian EFL learners because the word-for-word translation is the same. An example of an incongruent collocation is *take medicine (/dɒro mæsræf kon/ دارو مصرف کن)*, which would literally be *use medicine* in Persian.

Research consistently demonstrates that L1 – L2 congruency significantly influences the learning and processing of L2 collocations. Studies show that congruent collocations are processed more quickly and accurately than incongruent ones (Yamashita & Jiang, 2010). Even advanced L2 learners may produce unacceptable L2 collocations due to overreliance on word-for-word translations from their L1 (Davoudi & Behshad, 2015). For example, Davoudi and Behshad noted that *look for money* (*earn money*), *learn knowledge* (*gain knowledge*)**,** and *bring some reasons* (*state some reasons*) were some errors made by Iranian university students were due to their L1 interference.

However, recent studies show that the effect of congruency varies (e.g., Boone & Eyckmans, 2023; Fang & Zhang, 2021). In this regard, Fang and Zhang reported a significant increase in accuracy but no improvement in speed among the L2 groups for judging congruent collocations as opposed to incongruent collocations with intermediate and advanced Chinese learners of English. These findings underscore the importance of designing teaching materials that promote accuracy and consider the efficiency of language processing. Exploring the effects of congruency between learners' L1 and second languages is essential for developing effective instructional resources. Consequently, educators should consider integrating congruency-focused activities into their materials to better support language learners in comprehending and using linguistic structures.

Frequency also plays a role in the faster processing of MWUs; Siyanova and Schmitt's (2008) study revealed that participants responded more quickly to high-frequency collocations in judgment tasks. Similarly, González Fernández and Schmitt (2015) found that corpus frequency correlated moderately with collocational knowledge among L1 Spanish-speaking English learners. This suggests that both corpus frequency and input frequency are crucial considerations, as evidence suggests a strong link between a collocation's learnability and its frequency. Such insights align with usage-based approaches, such as those advocated by

Christiansen and Chater (2016), which underscore the importance of input, frequency, and experience in language acquisition. Meanwhile, Vu and Peters (2022a) found that Vietnamese learners recalled congruent collocations better than incongruent ones when reading. However, corpus frequency, MI score, and collocation type did not significantly impact the results. This suggests that while frequency plays a crucial role, other cognitive mechanisms may also contribute to the processing and retention of collocations.

Several studies claim that L2 learners are affected by experience factors, as noted above. Although semantic transparency has mainly been investigated in idiom studies (e.g., Howarth, 1996), this criterion might be a factor in the phraseological approach. Studies have produced mixed findings regarding the degree of transparency. Swedish learners of English demonstrated slower processing of collocations in Gyllstad and Wolter's (2016) research due to the semi-transparent nature of their target collocations. However, Macis and Schmitt (2017) investigated Chilean learners and found no correlation between frequency, semantic transparency, and understanding of figurative meanings. They reported that the years spent in college, time spent abroad, and reading time positively correlated with this knowledge, suggesting that completing additional university studies provides a greater understanding of figurative collocations and their combinations.

Determining the relationship between the abovementioned factors is a highly intricate task. While Yamashita and Jiang (2010) observed that higher levels of L2 proficiency among Japanese learners of English and L2 exposure reduced the L1 – L2 congruency effect, Wolter and Yamashita (2018) found that the impact of congruency did not decrease due to L2 proficiency. Consequently, more research is necessary since the findings are inconsistent. Thus, the current study aims to bridge this gap by exploring the factors influencing the acquisition of MWUs. We aim to identify essential factors for direct teaching purposes, contributing to more effective language learning strategies.

Measuring Collocational Knowledge

While standardised resources like the *New General Service List* (Browne, 2014) and the *Vocabulary Level Test* (Nation, 1990) exist for single words, there is a notable absence of standardised tests for MWUs (Gyllstad & Schmitt, 2018). The lack of standardisation stems from variations in defining and categorising MWUs, posing challenges due to the absence of a universal definition. Consequently, researchers face difficulty selecting test items, leading to experiment variations based on different theoretical perspectives.

Assessing productive knowledge of MWUs remains a challenge compared to single-word assessment methods. Corpus-based studies and gap-filling tests are commonly used to measure productive knowledge. For example, Men (2018) conducted a corpus-based study revealing that proficiency may affect the accuracy and frequency of certain collocations, such as verb + noun (VN), noun + noun (NN), and adjective + noun (AN) collocations. However, corpus-based studies may not fully control determinant factors, and learners may exhibit avoidance of certain collocations or overuse mastered ones (Laufer & Waldman, 2011).

On the other hand, gap-filling tests allow researchers to target specific elements like congruency (Sonbul et al., 2023). However, they may fail to comprehensively understand collocational usage in real-time production contexts (Frankenberg-Garcia, 2018). Frankenberg-Garcia developed an alternative gap-filling test format to address this limitation, where

participants must fill gaps in sentence contexts using all possible collocates for academic purposes. With this format, the scoring is not easy, and it is impossible to examine other affecting factors, especially congruency.

The current study used Laufer and Nation's (1999) format to assess productive collocational knowledge and investigate influential factors. Laufer and Nation's (1999) *Productive Vocabulary-Levels Test* for single words was designed based on six frequency bands with 1,000 items. They selected 18 items from each frequency band, providing a contextualised sentence for each item. The test structure is similar to a C-test but uses sentences rather than paragraphs, and the cues are not always half-words. Scoring can be automated since there is just one proper response for each item, and each answer is identified as correct or incorrect. Moreover, previous research on measuring collocation knowledge has generally focused on selected test items from the first three 1,000 frequency levels of English (e.g., Nguyen & Webb, 2017; Sonbul et al., 2023). This research took a different approach by utilising a list of innovative high-frequency MWUs derived from lemmatised concgrams (see Rogers, 2017).

Methods

Participants

Ethical approval for the research was obtained from the Social and Behavioural Research Ethics Committee of Flinders University. Iranian university students (n=256) volunteered to participate in the study. In terms of the study's context, Iran's education system emphasises English as a foreign language, and the students in this study were selected based on their enrollment in English language programs at Iranian universities. Their motivation for learning English was primarily for academic purposes, aligning with their studies in English-related fields.

The participants were undergraduate and graduate students majoring in English. The average proficiency level of participants was B2 to C1 based on the Common European Framework of Reference (CEFR) according to their IELTS scores with a mean of 6.8, *BCa* 95% CI [6.7, 7], SD = .99, and SE = .036. The MWUs test scores had a high level of internal consistency, as determined by a Cronbach's alpha reliability coefficient of 0.88 and a MWUs test score with a mean of 20.75 (62.5%), *BCa* 95% CI [20.11, 21.24], *SD* = 5.41, and *SE* = .20.

Research Instruments

Collocations list and test items

Using a well-constructed corpus and method to identify useful collocations is imperative in such a study. Scholars widely acknowledge the profound impact of corpora on the quality of derived word lists (Dang, 2020), highlighting the significance of meticulous selection. Rogers (2017) utilised data from the Corpus of Contemporary American English (COCA), which boasted a corpus of 450 million tokens at the time of their study and developed a list of high-frequency MWUs derived from lemmatised concgrams (see Rogers, 2017).

Subsequently, Barghamadi et al. (2023) conducted a contrastive analysis on this corpus-derived list comprising 11,212 MWUs. In their research, each MWU was translated into Persian and evaluated for L1– L2 congruency via a precise and objective 12-point system. This system assigned points based on the degree of equivalence between the translated word combinations and their counterparts in Persian. For instance, a perfect translation received a

total score of 12, while discrepancies in translation resulted in proportional point deductions. For example, *do your homework* is word to word equal to Persian, and the point is 12/12. Conversely, *make* in *make an impression is* equivalent to *put* in Persian (تاثیربگذارد /tæʔsiːr bogzɑːrəd/), resulting in a point score of 6. Additionally, the list was classified according to Grant and Bauer's (2004) taxonomy to investigate semantic transparency, dividing MWUs into four subcategories: literals (compositional), one non-compositional element (ONCE), figurative, and core idioms (for a more detailed explanation, see Author 1 et al., 2023). Consequently, each MWU had one L1 – L2 rating point and one transparency classification.

Productive collocational test design and scoring

A gap-filling test was employed to measure the participants' productive knowledge of the MWUs and explored the role of frequency, L1 – L2 congruency, and semantic transparency on proficiency. The MWU list, comprising 11,212 items, was divided into four separate frequency bands (F1, F2, F3, and F4), each containing 2,803 MWUs. Semantic transparency was divided into two subcategories: literal and opaque (figuratives, ONCE, and Core idiom). L1 - L2 congruency was subdivided into ratings of 0-4, 6-8, and 9-12 (see Table 1).

Table 1

The Bank of Items to Choose Test Questions Based on L1 – L2 Rating and Semantic Transparency

L1 – L2 Rating	Literal	Opaque
0-4	1100	689
6-8	3888	649
9-12	4646	240

Next, items with the lowest frequency, L1 – L2 rating and semantic transparency were selected from each frequency band. Accordingly, one item with the lowest frequency was categorised as literal, and one as opaque, receiving an L1 – L2 rating of 0-4, was selected. Since the total number of literal items in the 6 to 12 congruency rating set was three to four times higher than in the 0-4 rating set, two literal items and one opaque item were chosen from these sets (refer to Table 2). Consequently, eight items from each frequency band were selected, totalling 32 items, maintaining an equal ratio of such items in the test.

Table 2

Number of Items from each Frequency Band

Frequency Band	Literal with 0-4 L1 – L2 Rating	Literal with 6-8 L1 – L2 Rating	Literal with 9-12 L1 – L2 Rating	Opaque with 0-4 L1 – L2 Rating	Opaque with 6-8 L1 – L2 Rating	Opaque with 9-12 L1 – L2 Rating	Total Items
F1	1	2	2	1	1	1	8
F2	1	2	2	1	1	1	8
F3	1	2	2	1	1	1	8
F4	1	2	2	1	1	1	8
Total	4	8	8	4	4	4	32

Regarding congruency, an L1 – L2 rating was assigned between congruent and incongruent items as follows:

- 0-3 Very incongruent
- 4-6 Somewhat incongruent
- 8 Mostly congruent
- 9-12 Nearly or totally congruent

An equal number of incongruent and congruent settings were chosen: 16 items with a rating of 6 points and under and 16 items with a rating of 8 to 12 points to control for congruency. These 32 items were also selected based on transparency classification. As a higher ratio of the MWU items was classified as literals than opaque formulations, 20 literals and 12 opaque MWUs were determined. Consequently, the test consisted of 32 items, structured as follows based on Table 3 and Table 4.

Table 3

Structure of MWU Test Items by Congruency and Frequency Bands (F1-F4)

Category	Congruent Items	Incongruent Items	Total Items
F1	4	4	8
F2	4	4	8
F3	4	4	8
F4	4	4	8
Total	16	16	32

Table 4

Structure of MWU Test Items by Sematic Transparency and Frequency Bands (F1-F4)

Category	Literal Items	Opaque Items	Total Items
F1	5	3	8
F2	5	3	8
F3	5	3	8
F4	5	3	8
Total	20	12	32

To design the target test in this study, MWUs are derived from a pivot word (typically refers to the primary or central word around which other words (collocates) are grouped or associated) and its collocate. For test questions, the target answer was either the pivot or the collocate. For test questions, the target answer was either the pivot or the collocate. Cobb's (2013) Vocabprofiler, utilising integrated COCA/BNC data, was employed to confirm the frequency of pivots and collocates in each sentence. For example, the MWU ‘purchase price of’ was identified as the most frequent representation of the lemmatised congram purchase/price. Consequently, the word ‘purchase*’* was selected as the production target answer since it occurs less frequently than ‘price.’ This decision was based on the rationale that the less frequent item is more likely to be predicted by the more frequent pivot word.

Tests with one- and two-letter prompts were trialled with five native speakers, revealing that these types were challenging, and word length negatively affected test takers' ability to answer. Therefore, it was essential to consider other factors to control for possible alternative answers, such as:

- The length of words
- Sharing the same letter(s) prompt (e.g., ‘root’, ‘road’)

For consistency, the C test format was utilised, where half of each item was deleted, accompanied by a second hint indicating the number of letters by dashes. Items with more than three letters were selected for this format. To control for alternative responses, if a target item shared the same onset and number of letters with more than 20 other words, it would be replaced with another item. For instance, 'take root' should have been included, but 38 words exist as alternative responses (e.g., 'road,' 'role,' 'rose,' 'room,' etc.), leading to their exclusion and replacement with an alternative MWU. The following is an example of a test item for the MWU 'housing project':

- *The housing pro _ _ _ _ was initiated to help provide affordable housing to low-income families.*

Procedure

The test was administered online, with a prompt for participants to sign a consent form. To manage online access, participants were granted 40 minutes to complete the test via a Google Docs form, with access restricted after this time. The objective of the test was to assess knowledge of collocations rather than word forms. Minor spelling and grammar errors, such as 'device' instead of 'devise,' were overlooked to ensure consistency in the scoring procedure. Conversely, if part of the speech was disregarded, the item was marked as incorrect (e.g., 'violated' instead of 'violation'). Incorrect or unanswered responses were coded as false. 20 % of the test results were graded independently by one of the authors and an English teacher with native-level proficiency. The interrater reliability coefficient (ICC) value of .97 indicated a high level of agreement between the two raters' scores. Consequently, the author's results were used for statistical analysis.

Statistical Analyses

The data were analysed using version 27 of SPSS. All relevant assumptions were tested before conducting the analysis. Bootstrapping, a resampling technique, was employed to obtain reliable estimates of standard errors and confidence intervals (CIs). Bootstrapping is particularly suitable as a substitute for parametric estimation when the assumptions underlying those methods are questionable (Donaldson, 2019). Additionally, bias-corrected and accelerated (*BCa*) intervals were utilised to provide more accurate and reliable CIs. This approach enhances the precision of the results and mitigates potential biases, thereby ensuring the robustness and validity of the findings.

Results

A Pearson correlation analysis, supplemented with bootstrapping (1000 samples) and bias-corrected and accelerated (*BCa*) intervals, was conducted to examine the relationship between Persian-speaking learners' proficiency as measured by IELTS scores and their knowledge of MWUs (*RQ1*). The results revealed that there was a strong and positive correlation with a large effect size (Frost, 2019) between IELTS scores and MWU test scores (r (256) = 0.754, *BCa* 95% CI [0.703, 0.80], p <.001, R^2= .57), indicating that higher IELTS scores are associated with greater knowledge of MWUs. This suggests that proficiency in English, as measured by IELTS, may be linked to a better understanding and use of MWUs among Persian EFL learners.

The descriptive statistics of participants' scores based on four frequency levels (F1, F2, F3, F4) are displayed in Table 5. The table shows that a total score of 8 is possible for each frequency level; the participants' mean scores decrease across the frequency level from a mean score of 5.90 ± .11, *BCa* 95 % CI [5.66, 6.15] in F1 to a mean score of 4.57±.10, *BCa* 95% CI [4.37, 4.79] in F4. This decline suggests that participants performed better on higher-frequency items than on lower-frequency items, indicating that the MWU test may be considered "valid to a certain extent" (Schmitt et al., 2001, p. 67).

Table 5

Bootstrapping Analysis of Mean Scores for the MWUs Test

Variables	Mean	SE	*BCa* 95% CI	
			LL	UL
F1	5.90	.11	5.66	6.15
F2	5.20	.10	4.98	5.39
F3	5.07	.11	4.86	5.30
F4	4.57	.10	4.37	4.79
Congruent	11.27	.17	10.94	11.60
Non-congruent	9.48	.20	9.10	9.86
Literal	13.7	.24	13.3	14.27
Opaque	6.9	.13	6.6	7.1

Note. Bootstrap results are based on 1,000 Bootstrap samples.

Repeated measures ANOVA determined whether there were statistically significant differences between the frequency levels of MWUs in the productive knowledge test (*RQ2*). The amount of MWU knowledge over four frequency categories differed significantly, F (3, 765) = 54.56, p = < .001, η^2 = .176). Post hoc analysis with Bonferroni adjustment revealed that the difference decreased significantly from F1 to F4. However, the participants exhibited similar knowledge of MWUs at the F2 and F3 frequency levels, suggesting that these two levels were not significantly different.

A hierarchical multiple regression was run using the four frequency levels as independent variables to determine if MWU knowledge affected subsequent IELTS scores. The results revealed that each additional frequency level added to the regression model explained additional variance in IELTS scores. The best predictor of IELTS scores was F4 (the least frequent MWU items). The increase in one score of F4 led to a .193 *BCa* 95% CI, [.134, .257] increase in IELTS score. However, the final model with all four independent variables explained 60.7% of the variance. (R^2= .607, Adjusted R^2=.601), but F2 was not a significant predictor of IELTS scores in this model. This finding suggests that proficiency may be mainly influenced by knowledge of less frequent MWUs, indicating a deeper understanding of the language and its nuances among learners.

To answer *RQ3*, multiple regression analyses examined the relationship between frequency levels, congruency, and transparency. The MWUs test with 32 items consists of 16 congruent and non-congruent items. A total score of 16 was possible for each category. Table 5 illustrates the descriptive statistics of participants' scores based on congruency. The results revealed that the participants' mean score for congruent items (Mean= 11.27±.17, *BCa* 95%CI [10.94, 11.60]) was higher than non-congruent items (Mean= 9.48±.20, *BCa* 95%CI [9.10, 9.86]). The congruency model (Table 6) was found to be statistically significant, F (2, 253) = 167.915, p

< .001, indicating that the independent variables (i.e., congruent and non-congruent MWUs) collectively significantly predicted the dependent variable (i.e., IELTS scores). A significant interaction was found between independent and dependent variables as follows: F (3, 252) = 115.95, p < 0.001; B= 0.11, t = 2.39, p = .017, 95% CI [.002, .021]), corresponding to an increase in the R-squared of the model without interaction from 57% to 58% for the model with interaction. This model suggests that these two independent variables affect the dependent variable by .011. Further examination of the interaction model revealed a rising trend in the knowledge of non-congruent MWUs for IELTS scores. This pattern suggests that participants' proficiency in non-congruent MWUs may significantly impact their overall IELTS scores, indicating a nuanced relationship between MWU congruency and language proficiency.

Table 6

Summary of Regression Analyses for Variables Predicting IELTS Score

Variables	*B*	*SE B*	β	*t*	*p*
Frequency (Constant)	4.069	.159		25.607	.000
F1	.178	.028	.320	6.	.000
F2	-.011	.032	-.018	-.352	.725
F3	.183	.029	.315	6.363	.000
F4	.193	.032	.326	6.065	.000
Congruency (Constant)	3.94	.178		22.16	.000
Congruent	.155	.022	.416	7.004	.000
Incongruent	.125	.019	.398	6.693	.000
Transparency (Constant)	4.006	.159		25.182	.000
Literal	.10	.014	.381	7.35	.000
Opaque	.213	.024	.467	9.013	.000

Note. Dependent Variable: IELTS_Score

The MWUs test with 32 items consists of 20 literal and 12 opaque items. Since the number of items was unequal across transparent and opaque items, we used percentages. The results revealed that the participants' mean score for opaque items (Mean=58% ±1.1, *Bca* 95% CI [55.4, 60.4] was lower than that for literal (Mean = 69% ± 1.1, *Bca* 95% CI [66.4, 71.4]). The model with semantically transparent and opaque items significantly predicted IELTS scores statistically (F (2,253) = 179.86, p < .001, adj. R^2= .58). Examination of the regression coefficients for semantically transparent MWUs (B = .020, 95%CI [.015, 0.26]) and opaque MWUs (B = .026, 95%CI [0.20, .031]) and their respective p-values (p < .001) indicated that both independent variables significantly contributed to the model's predictive power (Table 6). However, opaque MWUs had a more significant effect with the β weight of .467. The interaction model was statistically significant, F (3,253) = 119.63, p < 0.001, but the interaction between literal items and opaque items was not significant (t = .388, p = .69).

Discussion

This study investigated the productive knowledge of MWUs among Persian-speaking English learners, focusing on critical factors such as frequency, L1 – L2 congruency, semantic transparency, and proficiency. The findings provide important insights into the role of these factors in L2 collocation knowledge and its relationship to proficiency, highlighting implications for teaching, curriculum design, and further research. The results revealed a strong

positive correlation between learners' L2 proficiency, measured by IELTS scores, and their knowledge of MWUs (RQ1). This confirms previous research (e.g., Sonbul et al., 2024a; Vu & Peters, 2022b), consistently showing that proficiency and collocational knowledge are closely linked. Learners with higher proficiency are more familiar with a broader range of collocations, suggesting that as learners' L2 proficiency increases, their ability to use MWUs. However, one limitation of this study is the reliance on self-reported IELTS scores, which may introduce bias. Future studies could use official test scores to ensure more accurate proficiency measures. Additionally, future research should consider other variables, such as exposure to L2, prior vocabulary knowledge, and the context in which collocations are used, to gain a more comprehensive understanding of this relationship.

Regarding the influence of frequency on productive knowledge (RQ2), the study demonstrated that learners performed better on high-frequency MWUs than on lower-frequency ones, confirming earlier studies (e.g., Nguyen & Webb, 2017; Sonbul et al., 2023). This finding reinforces the importance of frequency in L2 learning, supporting usage-based approaches that emphasise the role of repeated exposure and practice (Christiansen & Chater, 2016). A notable implication for curriculum design is prioritising high-frequency MWUs in teaching materials, gradually introducing items that are less frequent as learners' proficiency increases. However, the limited number of test items per frequency group may limit the generalisability of these findings, suggesting the need for future studies to include larger item pools to cover a broader range of MWUs.

An unexpected finding related to frequency was that MWUs in the second frequency band (F2) did not significantly predict IELTS scores despite the predictive solid power of less frequent MWUs (F4). This suggests that proficiency development may not follow a linear progression based on frequency alone and could be influenced by other factors, such as the semantic properties of specific MWUs or contextual variables. This aligns with previous research (e.g., González Fernández & Schmitt, 2015; Vu & Peters, 2022a), which shows that frequency effects can be inconsistent. Future research should investigate the specific characteristics of mid-frequency MWUs and their relationship to language proficiency to clarify these patterns.

Building on the frequency-related findings, the last research question (RQ3) examined the role of not only frequency but also L1 – L2 congruency and semantic transparency of MWUs. This progression allows us to understand better the various factors influencing productive collocational knowledge. The role of L1 – L2 congruency in collocational knowledge was examined, with results indicating that learners performed better on congruent MWUs than on non-congruent ones. This is consistent with previous findings (Boone et al., 2022; Vu & Peters, 2022a), who reported that congruent items are more straightforward for learners to process and recall. The substantial predictive value of congruent and non-congruent MWUs on IELTS scores suggests that L1 interference continues to be challenging for advanced learners, mainly when dealing with non-congruent items (Davoudi & Behshad, 2015).

Moreover, previous research (Rogers, 2017; Barghamadi et al., 2023) found that a high ratio of items in a large-scale MWU list was incongruent with their Japanese and Persian equivalents. Therefore, advanced L2 learners may form unacceptable L2 collocations because they rely too heavily on word-for-word translations from their L1. As such, placing more attention on incongruent MWUs in English classrooms and materials is desirable. Thus, L1 –

L2 congruency should be a key focus in language instruction. Explicit teaching strategies, such as contrastive analysis and activities targeting incongruent MWUs, can help learners overcome negative transfer from their L1 and develop more accurate collocational use in English (Boone & Eyckmans, 2023; Laufer & Girsai, 2008). Furthermore, future studies should examine how learners acquire incongruent collocations at different levels of incongruency to gain deeper insights into the impact of varying degrees of incongruency on language learning outcomes.

Finally, the study found that learners performed better on MWUs with transparent meanings compared to those with opaque meanings. However, this finding aligns with previous studies (e.g., Gyllstad & Wolter, 2016); the regression analysis showed that knowledge of opaque MWUs had a more substantial predictive value for proficiency. This suggests that while transparent MWUs are easier to learn, mastering opaque items may be more indicative of advanced language skills, possibly because they require deeper linguistic understanding. These findings highlight the importance of incorporating transparent and opaque collocations in language instruction, particularly for advanced learners, to help them build more nuanced and flexible language skills.

The study also highlighted the interaction between frequency and congruency in shaping learners' ability to acquire opaque items. While transparency was considered, it emerged as a relatively minor factor in determining MWU knowledge. For instance, learners showed tremendous success with opaque items such as *pouring in* (a more frequent, somewhat incongruent MWU) compared to *drift off* (a less frequent, totally incongruent MWU), underscoring the complex interplay between these factors. This result suggests that frequency and congruency, rather than transparency alone, are key drivers of MWU acquisition and proficiency development.

Consequently, General English Phrases[2] developed in this study addresses the challenge of L1 interference by focusing on non-congruent items. Learners practising these complex collocations can reduce negative transfer and improve their English fluency. This approach aligns with findings from Sonbul et al. (2024b), who demonstrated the effectiveness of deliberate learning in collocation retention using equal and expanding spacing schedules. In this study, deliberate practice of collocations—through repeated exposure—proved essential for long-term retention, with both spacing conditions significantly outperforming a control group. This supports the notion that explicit activities, such as using flashcards for spaced repetition, effectively reinforce collocational relationships (Barghamadi et al., 2022; Nakata, 2020). Moreover, moving from explicit learning to more contextualised activities, such as reading and listening in real-world scenarios, allows learners to encounter collocations in authentic contexts, solidifying their understanding and usage. Together, these strategies underscore the value of deliberate learning in mastering complex MWUs in L2 settings.

In conclusion, the findings from this study emphasise the multifaceted nature of L2 collocation knowledge, where frequency, congruency, and transparency each play a significant role. Educators can foster a more robust understanding of collocations among learners by embracing explicit teaching methods, considering contrastive analysis, and tailoring materials to address L1 – L2 congruency and frequency. These pedagogical strategies equip language learners with the tools to navigate the intricacies of collocational usage in their second

[2] General English Phrases List (http://secureapp.au/4600/)

language. Future research should explore more extensive, more diverse datasets to investigate these relationships further, offering a more comprehensive view of the processes underlying L2 collocation acquisition.

Conclusion

This research developed an MWU test based on Rogers's (2017) list to investigate the relationship between productive collocational knowledge and key variables: frequency, semantic transparency, and L1 – L2 congruency. The findings underscore the importance of congruency in learners' proficiency, the predictive power of less frequent MWUs in language proficiency, and the role of semantic transparency in aiding understanding. These results highlight critical considerations for curriculum design, suggesting that educators should emphasise frequent MWUs while integrating contrastive analysis of L1 and L2 collocations. This approach could better support learners in developing a more robust understanding of collocational patterns, enhancing their overall language proficiency.

This study's limitations include a relatively small sample size and reliance on self-reported proficiency measures, which may introduce bias. Although a positive correlation was found between the number of items tested and the accuracy of subsequent estimates, including more than 30 items would negatively affect practicality due to the length of the test (Gyllstad, 2020). While receptive knowledge studies use larger sample sizes, as seen in Nguyen and Webb (2017) with 180 items, productive tests typically involve fewer items. For example, Frankenberg-Garcia's (2018) study used only ten, and Sonbul et al. (2023) included 60. In this study, eight items per frequency level were selected. Future research should consider including more items from the developed list to enhance the robustness of the findings. Additionally, the online test format, potentially influenced by participants' digital literacy, may have impacted their performance. While our study utilised multiple regression and hierarchical regression analysis, the potential of mixed-effect models to provide a more detailed understanding of variable interactions is evident. Future research should also adopt longitudinal designs and examine additional factors, such as individual learner differences and exposure to authentic language. These factors could significantly impact MWU acquisition, underscoring their importance in our field.

ORCID

https://orcid.org/0000-0003-4982-3537
https://orcid.org/0000-0002-8726-8231
https://orcid.org/0000-0002-8184-4859
https://orcid.org/0000-0002-7467-9939

Acknowledgements

The authors wish to thank the editors and anonymous reviewers for their insightful comments and constructive feedback, which have greatly improved the quality and clarity of this manuscript. Their thoughtful suggestions and attention to detail were invaluable in refining the work.

Funding

This research was supported by Flinders University, which provided funding for the creation of a collocation list.

Ethics Declarations

This study was approved by the Flinders University Human Research Ethics Committee (Project number 4365).

Competing Interests

No, there are no conflicting interests.

References

Barghamadi, M. (2024). *Identifying and teaching English collocations for Persian students.* [Unpublished doctoral dissertation], Flinders University, Australia.

Barghamadi, M., Rogers, J., Müller, A., & Arciuli, J. (2023). The use of semantic transparency and L1-L2 congruency as multi-word units selection criteria. *Studies in English Language and Education*, *10*(2), 723-740. https://doi.org/10.24815/siele.v10i2.28644

Boers, F., Lindstromberg, S., & Eyckmans, J. (2014). Some explanations for the slow acquisition of L2 collocations. *Vial-Vigo International Journal of Applied Linguistics, 11*, 41–62.

Boone, G., & Eyckmans, J. (2023). Productive collocation knowledge in L2 German: Study abroad and L1 congruency. In B. L. Reynolds (Ed), *Vocabulary learning in the wild* (pp. 67–86). Springer Nature Singapore. https://doi.org/10.1007/978-981-99-1490-6

Boone, G., De Wilde, V., & Eyckmans, J. (2023). A longitudinal study into learners' productive collocation knowledge in L2 German and factors affecting the learning. *Studies in Second Language Acquisition*, *45*(2), 503–525. https://doi.org/10.1017/S0272263122000377

Browne, C. (2014). A new general service list: The better mousetrap we have been looking for. *Vocabulary Learning and Instruction*, *3*(2), 1–10. http://dx.doi.org/10.7820/vli.v03.2.browne

Cheng, W., Greaves, C., & Warren, M. (2006). From n-gram to skipgram to concgram. *International Journal of Corpus Linguistics, 11*(4), 411–433. https://doi.org/10.1075/ijcl.11.4.04che

Christiansen, M. H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral & Brain Sciences*, *39*, e62. https://doi.org/10.1017/S0140525×1500031X

Cobb, T. (2013). *Vocabprofile*. Retrieved from http://www.lextutor.ca/vp/

Cowie, A. P. (1994). Phraseology. In Asher, R.E. (Ed.). *The Encyclopedia of language and linguistics* (pp. 3168–3171). Oxford University Press.

Dang, T. N. Y. (2020). Corpus-based word lists in second language vocabulary research, learning, and teaching. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (1st ed., pp. 299–303). Routledge.

Davies, M. (2008). *The corpus of contemporary American English: 425 million words, 1990-present.* Bringham Young University. http://corpus.byu.edu/coca/

Davoudi, M., & Behshad, A. (2015). Collocational use: A contrastive analysis of strategies used by Iranian EFL learners. *Theory and Practice in Language Studies*, *5*(12), 2646-2652. https://doi.org/10.17507/tpls.0512.29

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. https://doi.org/10.1075/rcl.00038.din

Donaldson, P. (2019). *Bootstrapping in SPSS v. 22* (Version 1). Figshare. https://doi.org/10.6084/m9.figshare.9275909.v1

Du, X., Afzaal, M., & AlFadda, H. (2022). Collocation uses in EFL learners' writing across multiple language proficiencies: A corpus-driven study. *Frontiers in Psychology*, *13*, 752134. https://doi.org/10.3389/fpsyg.2022.752134

Erman, B., & Warren, B. (2000). The idiom principle and the open choice principle. *Text: Interdisciplinary Journal for the Study of Discourse, 20*, 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Fang, N., & Zhang, P. (2021). L1 congruency, word frequency, collocational frequency, L2 proficiency, and their combined effects on Chinese–English bilinguals' L2 collocational processing. *The International Journal of Bilingualism: Cross-Disciplinary, Cross-Linguistic Studies of Language Behavior*, *25*(5), 1429–1445. https://doi.org/10.1177/13670069211024747

Frankenberg-Garcia, A. (2018). Investigating the collocations available to EAP writers. *Journal of English for Academic Purposes, 35*, 93–104. https://doi.org/10.1016/j.jeap.2018.07.003

Frost, J. (2019). *Regression analysis: An intuitive guide for using and interpreting linear models*. Statistics by Jim Publishing.

González Fernández, B., & Schmitt, N. (2015). How much collocation knowledge do L2 learners have? The effects of frequency and amount of exposure. *ITL—International Journal of Applied Linguistics, 166*, 94–126. https://doi.org/10.1075/itl.166.1.03fer

Grant, L., & Bauer, L. (2004). Criteria for re-defining idioms: Are we barking up the wrong tree? *Applied Linguistics, 25*(1), 38-61. https://doi.org/10.1093/applin/25.1.38

Gyllstad, H. (2020). Measuring knowledge of multi-word items. In S. Webb (Ed), *The Routledge handbook of vocabulary studies* (1st ed., pp. 387–405). Routledge. https://doi.org/10.4324/9780429291586-25

Gyllstad, H., & Schmitt, N. (2018). Testing formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sánchez (Eds.), *Understanding formulaic language*: *A second language acquisition perspective* (pp. 174–191). Routledge.

Gyllstad, H., & Wolter, B. (2016). Collocational processing in light of the phraseological continuum model: Does semantic transparency matter? *Language Learning*, *66*(2), 296-323. https://doi.org/10.1111/lang.12143

Hoey, M. (2005). *Lexical priming: A new theory of words and language* (1st ed.). Routledge. https://doi.org/10.4324/9780203327630

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

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. https://doi.org/10.1093/applin/amn018

Laufer, B., & Nation, I. S. P. (1999). A vocabulary-size test of controlled productive ability. *Language Testing*, *16*(1), 33–55. https://doi.org/10.1191/026553299672614616

Laufer, B., & Waldman, T. (2011). Verb–noun collocations in second language writing: A corpus analysis of learners' English. *Language Learning, 61*, 647–672. https://doi.org/10.1111/j.1467-9922.2010.00621.x

Macis, M., & Schmitt, N. (2017). Not just "small potatoes": Knowledge of the idiomatic meanings of collocations. *Language Teaching Research, 21*, 321–340. https://doi.org/10.1177/1362168816645957

Men, H. (2018). *Vocabulary increase and collocation learning: A corpus-based cross-sectional study of Chinese learners of English*. Springer Singapore Pte. Limited. https://link.springer.com/book/10.1007/978-981-10-5822-6

Nakata, T. (2020). Learning words with flash cards and word cards. In S. Webb (Ed.), *The Routledge handbook of vocabulary studies* (pp. 304–319). Routledge.

Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Newbury House.

Nguyen, T., & Webb, S. (2017). Examining second language receptive knowledge of collocation and factors that affect learning. *Language Teaching Research, 21*, 298–320. https://doi.org/10.1177/1362168816639619

Rogers, J. (2017). *What are the collocational exemplars of high-frequency English vocabulary? On identifying MWUs most representative of high-frequency lemmatized concgrams* [Doctoral dissertation, University of Southern Queensland]. University of Southern Queensland Repository. https://doi.org/10.26192/5bf5ff14ed350

Rogers, J., Müller, A., Daulton, F. E., Dickinson, P., Florescu, C., Reid, G., & Stoeckel, T. (2021). The creation and application of a large-scale corpus-based academic multi-word unit list. *English for Specific Purposes, 62*, 142–157. https://doi.org/10.1016/j.esp.2021.01.001

Schmitt, N., Schmitt, D., & Clapham, C. (2001). Developing and exploring the behaviour of two new versions of the vocabulary levels test. *Language Testing, 18*, 55–88. https://doi.org/10.1191/026553201668475857

Siyanova, A., & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*, 429–458. https://doi.org/10.3138/cmlr.64.3.429

Sonbul, S. (2015). Fatal mistake, awful mistake, or extreme mistake? Frequency effects on off-line/on-line collocational processing. *Bilingualism*, *18*(3), 419–437. https://doi.org/10.1017/S1366728914000674

Sonbul, S., El-Dakhs, D. A. S., & Alharbi, R. (2024a). Rendering natural collocations in a translation task: The effect of direction, congruency, semantic transparency, and proficiency. *International Journal of Applied Linguistics, 34*(1), 117–133. https://doi.org/10.1111/ijal.12482

Sonbul, S., El-Dakhs, D. A. S., & Masrai, A. (2023). Second language productive knowledge of collocations: Does knowledge of individual words matter? *Studies in Second Language Acquisition*, *45*(2), 480–502. https://doi.org/10.1017/S0272263122000341

Sonbul, S., Macis, M., & Gyllstad, H. (2024b). The effect of equal versus expanding spacing practice on the deliberate learning of L2 collocations. *TESOL Quarterly*. Advanced online publication. https://doi.org/10.1002/tesq.3364

Vu, D. V., & Peters, E. (2022a). Incidental learning of collocations from meaningful input: A longitudinal study into three reading modes and factors that affect learning. *Studies in Second Language Acquisition, 44*(3), 685–707. https://doi.org/10.1017/S0272263121000462

Vu, D. V., & Peters, E. (2022b). The role of formulaic sequences in L2 speaking. In T. Derwing, M. Munro, & R. Thomson (Eds.), *The Routledge handbook of second language acquisition and speaking* (pp. 285–298). Routledge.

Webb, S., Newton, J., & Chang, A. C. S. (2013). Incidental learning of collocation. *Language Learning, 63*(1), 91–120. https://doi.org/10.1111/j.1467-9922.2012.00729.x

Wolter, B., & Yamashita, J. (2018). Word frequency, collocational frequency, L1 congruency, and proficiency in L2 collocational processing: What accounts for L2 performance? *Studies in Second Language Acquisition, 40*(2), 395–416. https://doi.org/10.1017/S0272263117000237

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

Yamagata, S., Nakata, T., & Rogers, J. (2023). Effects of distributed practice on the acquisition of verb-noun collocations. *Studies in Second Language Acquisition, 45*(2), 291–317. https://doi.org/10.1017/S0272263122000225

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. https://doi.org/10.1002/(ISSN)1545-7249