overSEAS 2024

This thesis was submitted by its author to the School of English and American Studies, Eötvös Loránd University, in partial fulfilment of the requirements for the degree of Bachelor of Arts. It was found to be among the best theses submitted in 2024, therefore it was decorated with the School's Outstanding Thesis Award. As such it is published in the form it was submitted in **overSEAS 2024** (http://seas.elte.hu/overseas/2024.html)

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

Bölcsészettudományi Kar

BA THESIS

The Relevance of Using Formulaic Language for English Language Learners

A Formulaszerű Nyelvhasználat Jelentősége az Angol Nyelvtanulásban

Written by:

Ulnyrova Sofia

English and American Studies

English track

2024

CERTIFICATE OF RESEARCH

By my signature below, I certify that my ELTE B.A. thesis, entitled *The Relevance of Using Formulaic Language for English Language Learners*, is entirely the result of my own work, and that no degree has previously been conferred upon me for this work. In my thesis I have cited all the sources (printed, electronic or oral) I have used faithfully and have always indicated their origin. The electronic version of my thesis (in PDF format) is a true representation (identical copy) of the printed version. If this pledge is found to be false, I realize that I will be subject to penalties up to and including the forfeiture of the degree earned by my thesis.

Date: 2024.04.14 Signed: Ulnyrova Sofia

EÖTVÖS LORÁND TUDOMÁNYEGYETEM

Bölcsészettudományi Kar

BA THESIS

The Relevance of Using Formulaic Language for English Language Learners

A Formulaszerű Nyelvhasználat Jelentősége az Angol Nyelvtanulásban

Supervisor:

Dóczi Brigitta

Assistant Professor

Written by:

Ulnyrova Sofia

English and American Studies

English track

2024

Table of contents

Abstract

Vocabulary is a significant constituent of the language learning process: lexical phrases, in particular, are essential when developing learners' language skills. Since formulaic is a complex topic, it is necessary to shed light on its features. In this study, I analyse formulaic language use as a whole unit, covering its definition, functions, importance for learners, and teaching implications. The findings highlight the need to emphasise formulaicity in educational institutions. Additionally, the article lists how formulaic language teaching will change thanks to the new technologies.

Keywords: formulaic language use, multiword elements, EFL

1 Introduction

A broad vocabulary is crucial for learning English as a Foreign Language (EFL) because "it underpins all other language skills" (Egamova & Sharofova, 2022, p. 190). A wide range of lexical expressions can help students master English more quickly. Additionally, Sullivan and Alba (2019, as cited in Dakhi & Fitria, 2019, p. 16) stated: "Without grammar very little can be conveyed; without vocabulary nothing can be conveyed". Hence, one should focus on vocabulary development to understand English and express themselves confidently in this language.

Considering the need for a broad lexicon, English poses a challenge for its learners – acquiring formulaic language (FL), one of the main vocabulary elements. The research on this topic has started recently, and the results have already shown its shape in the language: the frequency of FL is approximately 80% in general (Altenberg, 1998) and 50% in writing (Erman & Warren, 2000). Thus, a robust command of FL is significant for English learners, especially at advanced levels. However, at the same time, it also appears to be extremely hard to master (Lundell, 2021) primarily due to a misconception: EFL teachers and students usually think about vocabulary in terms of singular words instead of phrases (Schmitt, 2010) and it leads to a lack of emphasis on FL.

This literature review aims to showcase the importance and applicability of FL in teaching. It analyses the misconception above of how vocabulary is taught in educational organisations in contrast to how people use English in reality – depending on formulaic expressions. Understanding the principle of FL and its acquisition processes enables the development of more efficient methods for teaching these essential communication elements. In the case of structure, the paper will cover the definition and classification of FL, as well as its significance and advantages in language teaching. The general conclusion on FL will be drawn at the end of the paper.

2 Definition and Classification of Formulaic Language

Applied linguistics has provided a new perspective on languages: its significant part accounts for specific patterns or *formulas*. Currently, there are multiple definitions of FL. For Sinclair (1991, p. 110), expressions such *sooner or later* or *to make a decision* refer to "single choices, even though they might appear to be analysable into segments". Wood (2020, p. 30) broadens the term. The researcher claimed that FL is a "multiword language phenomena which holistically represent a single meaning or function, and are likely mentally stored and used as unanalysed wholes, as are single words". Erman and Warrens (2000, p. 31) account for the above definitions but also highlight that native speakers primarily use these multiword constructions: "A combination of at least two words favoured by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalisation". As can be seen, the concept of FL is vague because there are overlapping terms.

Based on the existing literature, the mentioned definitions include three key elements: unity, mentality and nativity. Firstly, it is evident that formulaic sequences are inextricably linked to the agreement of different words, both linguistic (nouns, verbs, adjectives, and adverbs) and functional (conjunctions, prepositions, articles, pronouns), as seen in the examples. Secondly, a mental concept appears because the term is also related to brain function, namely memorising and speech production. Thirdly, FL expressions are primarily used by native speakers of a language.

The features of FL are compositionality and flexibility (Wray, 1999). Compositionality refers to how the meaning of a sentence can be understood by interpreting singular words. Formulaic phrases often violate this principle by conveying a meaning unrelated to their components, such as "*kick the bucket*." Hence, formulaic expressions are non-compositional.

Regarding flexibility, Wray (1999, p. 215) categorises three types: immutable sequences, partially flexible sentences, and formulaic frames. Immutable sequences are fully fixed and do not compromise any changes. Examples are proverbs and quotations *("You can't have your cake and eat it too"*). Partially flexible sentences can form based on grammar and pronouns *("You've really*

done it this time" vs. *"I'd really done it this time"*), but their core remains the same (e.g. *"really done it this time"*). Formulaic frames are the most flexible type of FL element. They offer spaces for words that can be utilised to produce different versions of this phrase (for example, *"X caught Y red-handed"*).

Hence, the flexibility of FL can be visualised as a spectrum:

Figure 1

Spectrum of FL's flexibility (Wray, 1999, author's figure)

There are several ways to determine FL phrases, such as frequency statistics in corpora, results in Internet search engines, and word compatibility. However, it is essential to note that even if sequences appear often, they might not be considered formulaic. For example, a relatively rare phrase, *"In spite of"*, will be viewed as the one since it "shows formulaicity, because it contains words which very commonly occur in this order, and it has a particular unitary meaning or function" (Wood, 2020, p. 38). Therefore, another approach is more reliable for estimating if a sequence is formulaic - mutual information. The tool measures the strength of the association between components of a phrase. The sequence is formulaic if these words are likely to combine and the mutual information index is high.

Another way to check if a phrase is formulaic is the native-speaker checklist, consisting of five criteria (Wood, 2010, as cited in Wood, 2020, p. 41–42). Moreover, non-native phrases were accepted in it, for example, *thanks god* instead of *thank god*. However, more research needs to be done to define FL more explicitly because the FL classification, as stated before, is relatively obscure (Wood, 2020).

The checklist:

1. Phonological coherence and reduction.
2. The taxonomy used by Nattinger and DeCarrico (1992).

   Syntactic strings include NP + Aux + VP and pragmatic phrases such as *"How do you do?"* (Arnaud, 1993b, p. 36).
3. Greater length/complexity than other outputs.

   An EFL speaker might say *I would like* . . . or I *do not understand*, utilising the structures only in these sequences and never successfully using *would* or *do not* in other cases.
4. Semantic irregularity, as in idioms and metaphors.
5. Syntactic irregularity.

(Wood, 2020, pp. 41–42)

40 FL terms were identified 20 years ago (Wray & Perkins, 2000, as cited in Wood, 2020); the main categories are collocations, idioms, metaphors, proverbs, phrasal verbs, and compounds. However, the elements may overlap (Wood, 2020).

Figure 2

The main categories of FL (Wood, 2020, pp. 31–35)

Category	Examples
Collocations	- Collocated pairs of words: *senior manager, single parent* - Syntactic functions: *red herring* - Pragmatic functions: *How are you?*
Idioms	*Spill the beans, burn one's candle at both ends.*

Metaphors	- Vehicle (unusual manner) + topic (the referent of the vehicle) scheme: *Time is a healer* - Simile: *Life is like a box of chocolate*
Proverbs	*A stitch in time saves nine, like death and taxes*
Compounds	- Closed forms: *hardcore, laptop* - Hyphenated forms: *open-handed* - Open forms: *real estate*
Phrasal verbs	- Verb + preposition: *give up, let down* - Verb + particle: *show off, cave in* - Verb + particle + preposition: *run out of*
Lexical phrases	- Poly words: *so far so good* - Institutionalized expressions: *nice meeting you* - Phrasal constraints: the ___er the ___er - Sentence builders: I think that X

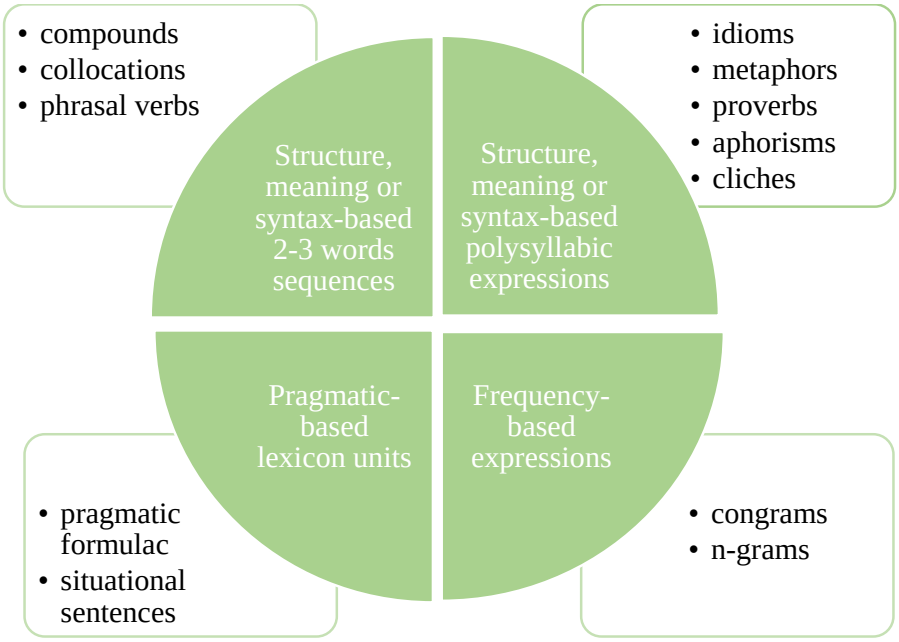
Another crucial category is lexical bundles, which frequently require three or more words occurring in texts **(**Biber et al., 1999, as cited in Wood, 2020, p. 36). There are three categories of lexical bundles – referential, stance, and discourse (Biber, 2006, as cited in Wood, 2020, p. 36).

1. Referential bundles show actual or abstract objects or characteristics, such as *"the size of the . . ."* or *"one of the things"*.
2. Stance bundles refer to levels of certainty or attitudes: *". . . are likely to be . . ."* or *"what do you think . . ."*
3. Discourse bundles explain the connections between topics: *"on the other hand"* and *"as well as . . ."*

Additionally, formulaic phrases can be divided into four categories: structure-based, meaning-based, syntax-based, pragmatic-based, and frequency-based.

Figure 3

Main categories of FL based on their nature (Wood, translated from Dóczi, 2024, chapter 4)



The following part of the section details the three most popular FL categories: collocations, idioms, and phrasal verbs. "Lexical collocations consist of nouns, verbs, adjectives, and adverbs. They usually do not contain clauses, infinitives or prepositions", according to Rogošić (2014, p. 172). Hence, the most crucial collocation combinations are adjective + noun, noun + verb, verb + noun, adverb + adjective, and verb + adverb.

"A collocation is made up of two parts – a pivot word (also called a 'node') which is the focal word in the collocation and its collocate(s), the word or words accompanying the pivot word" (Shin, 2007, p. 202). As the researcher claims, there are four criteria for collocations:

1. The pivot word must be a content one;

2. The pivot word and its collocate(s) must all be in the most frequent words of English.

   According to the author, learning collocations should strengthen and enrich students' vocabulary and not add burden by covering unknown, lower-frequency words.

3. The collocation should occur frequently.

   It should be frequent enough to get into the high-frequency words of the language.

4. The collocation should be grammatically correct to be complete and fit in a sentence.

   Grammatical well-formedness means that a collocation can be a meaningful unit for teaching and learning.

(Shin, 2007, p. 203)

According to Talmy (1985, 2000, as cited in Spring, 2018, p. 122), there is a theory of event conflation regarding phrasal verbs. It suggests that conflation occurs in five events: motion, change of state, the realisation of goals, aspect, and correlation of actions.

1. Motion: Jack skipped across the park;
2. Change of state: Jack tied together the boxes.
3. Realization of goals: Jack chased down the criminal.
4. Aspect: Jack ran on, even though he was tired and wanted to quit.
5. Correlation of actions: Jack sang along with the radio.

(Talmy, 1985, 2000, p. 122)

Phrasal verbs have multiple meanings due to the variety of particles. For example, “look up” can be “to look upwards” and “to investigate”. The first implies an action – looking – and the direction – up. The second meaning is completing an action, where “up” shows that the “looking” process was finished. Additionally, the following articles are in modern English: aback, aboard,

about, across, after, ahead, along, apart, around, ashore, aside, astray, away, back, behind, by, down, forth, etc. (Thim, 2012, p. 11).

Idioms is another FL category; finding a unified definition is challenging. One of these definitions is: “Conventionalized multi-word expressions often, but not always non-literal” (Fernando, 1996, as cited in Lui, 2003, p. 673). These language elements must always be clear, specific, and systematic. There are three categories of idioms identified:

Figure 4

Three categories of idioms ([a] from Fernando 1996, p. 32, as cited in Lui, 2003, p. 673)

Category	Examples
Pure	*kick the bucket, pull someone’s leg, make off with* [a]
Semiliteral	*fat chance* [a]*, a use something as a step stone, go through*
Literal	*according to, in sum* [a]*, throw away*

In conclusion, the ubiquity and significance of multiword constructions are evident; however, there was a lack of emphasis. Moreover, formulaic sequences are considered one of the most challenging components of English vocabulary, making learning even more complicated. Nevertheless, the use of formulaic phrases by EFL students could change after new research in this field and offering cutting-edge educational methods.

3 Significance of Formulaic Language in Language Teaching

The complexity of the term “formulaic language” indicates its significance. Although scholars classify FL elements differently, they still provide a large percentage of the frequency of those phrases. Referring to the number mentioned in the introduction, according to Altenberg

(1998), formulaic expressions account for over 80% of the vocabulary of an adult native speaker and thus make up a considerable part of English language. These statistics imply the authentic nature of the FL. The emphasis on native proficiency highlights the need for students to familiarise themselves with FL phrases because it will ultimately reduce the difference between in-class vocabulary and real-life communication.

There is a clear correlation between higher proficiency in English and a more robust command of FL, according to Rafieyan (2018). As the author claims, the main reason for this phenomenon is the ubiquity of formulaic phrases. When EFL learners are more exposed to English, they naturally encounter multiword sequences. Constant repetition stores these phrases in long-term memory, making language acquisition more effective. However, the study also finds that advanced FL skills might not lead to native-like fluency because of the difference in how native and non-native speakers process formulaic phrases – as whole units and piece by piece. Nevertheless, Rafieyan (2018) still proves the positive impact of FL on language proficiency.

Even though FL is ubiquitous among native speakers, EFL speakers do not utilise it this often. Kecskés (2007) analysed how people communicate using English as a Lingua Franca. The author states it is "a language model" (Kecskés, 2007, p. 213), a language phenomenon rather than proper English. Within this language model, people avoid shortcuts and slang and focus on being literal and understood by another person. The author provides an example of a conversation where the formulaic phrase was misinterpreted:

> "Chinese student: – I think Peter drank a bit too much at the party yesterday.
>
> Turkish student: – Eh, **tell me about it.** He always drinks much.
>
> Chinese student: – When we arrived he drank beer. Then Mary brought him some vodka. Later he drank some wine. Oh, too much.
>
> Turkish student: – Why are you telling me this? I was there.
>
> Chinese student: – Yes, but you told me to tell you about it".
>
> (Kecskés, 2007, p. 191)

Even though English as a Lingua Franca is a somewhat simplified form of English, it is still necessary to be aware of FL, as one might be in various social situations with native and non-native speakers where knowledge of multiword expressions will be constructive.

3.1 The Impact of Formulaic Language on Writing and Speaking Skills

Since the FL concept is paramount, multiword sequences are essential for EFL learners. Martinez and Schmidt (2012) listed some of them in the article. The first reason is that FL "phrases realise many referential, communicative, and textual functions in discourse". In writing, these are expressing concept (*take into account* = considering something), signposting discourse organisation (*on the other hand* = signalling transition), and efficiently conveying ideas by including technical phraseology in the text (using a *figure of speech* in a non-literal environment) (Schmidt & Carter, 2004, p. 16). Such phrases may help EFL learners express concepts comprehensibly and explicitly. Moreover, FL sequences can implement people's needs in various contexts, including academic and scientific ones (Biber et al., 2004). Therefore, multiword phrases might equip one with a relevant advanced vocabulary for different linguistic scenarios.

Another reason multiword terms are beneficial in writing is that they facilitate understanding a text written by a non-native person. Millar (2011) found a connection between the written production of L2 learners and its processing by native speakers of English. Thus, if a learner used formulaic elements, L1 readers needed less time to analyse that text. Research on self-paced reading tasks also proved that FL phrases are processed more quickly than their non-formulaic analogues by L1 and L2 readers (Conklin & Schmidt, 2008). By including FL elements in texts, EFL learners have more chances to be understood correctly by readers, especially by native speakers.

Additionally, collocations significantly contribute to making texts more comprehensive in writing, according to Erman (2009, as cited in Khodadady & Shamsaee, 2012, p. 42). The researcher stated that the lack of fluency in writing stems from the inability to place collocations in a text correctly. Hence, according to the study, two elements should be considered when

studying and teaching these elements: 1) the meaning and the functions in a particular case and 2) the limited choice of the elements (e.g., *"make a choice"* instead of *"do a choice"*).

In speech, fluency might be associated with several factors: speaking rate, number of syllables, and words uttered between pauses (Wood, 2009, as cited in Khodadady & Shamsaee, 2012, p. 42). Wood's definition is the most common; however, there are disagreements (Wingate, 1984, as cited in Khodadady & Shamsaee, 2012, p. 42). Nevertheless, formulaic phrases impact fluency, as they help speakers save time on thinking and thus produce more flowy speech.

Finally, Pawley and Syder (1983, as cited in Khodadady & Shamsaee, 2012, p. 42) found a connection between fluent speech and FL. According to research, speakers of English opted to use multiword phrases instead of considering separate lexical and grammatical options. Therefore, incorporating formulaic sequences into language learning activities can be efficient for achieving fluency in English.

3.2 The Impact of Formulaic Language on Cognitive Skills

Referring to the point that multiword phrases save time, it was found that they can be more easily extracted from memory for later use, as Hay and Bresnan (2006, as cited in Khodadady & Shamsaee, 2012, p. 42) state. The results also demonstrated that FL might be coded and stored in memory along with their phonological attributes and thus be retrieved with less effort when speaking.

According to Carrol and Conklin (2019), the brain processes FL faster than non-formulaic phrases, supporting Conklin and Schmidt's idea from section 3.1. The researchers examined how people analyse idioms and collocations compared to regular expressions. This progress in pace stems from the participants' prior experience with those formulaic expressions. In the case of idioms, the reading pace depended on how familiar the phrase was. As for collocations, the predictability of a second word mattered how fast the participants could read the collocations. Therefore, familiarity and predictability explain why formulaic sequences are processed more

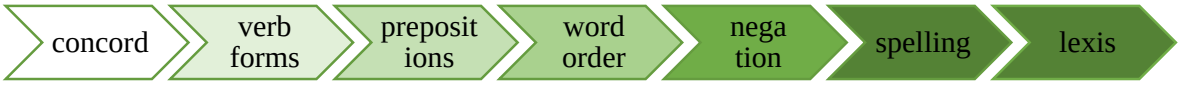
swiftly than non-formulaic ones. While there are variations in how different types of FL are analysed, these variances are minor.

Furthermore, using formulaic expressions affects the impression of EFL learners' language production. There is a proven connection between FL use and language proficiency, especially at advanced levels, since these sequences give students more confidence to express themselves in the language (Lundell, 2021). For instance, those who scored higher on an EFL proficiency test used more FL constructions (Ohlrogge, 2009). In addition, L2 speakers seemed more advanced when using multiword elements in writing (Boers et al., 2006). In summary, if a learner consciously includes more FL phrases in texts, that might indicate a higher English level; it can also make a better impression on a reader.

The lexical ones fall under the most severe category when referring to mistakes. McCretton and Rider (1993, as cited in James, 1998, p. 229) provided the hierarchy, where concord and lexis are the least and most severe, respectively. FL falls under the "lexis" category:

Figure 5

Hierarchy of mistakes related to formulaic languages (James, 1998)



One can create a message using the wrong verb form or word order in a sentence; however, errors in choosing a word might lead to a complete misunderstanding, and the same applies to formulaic phrases.

4 Benefits of Formulaic Language in Language Teaching

Even though FL is ubiquitous in the language, it appears extremely difficult to master and is time-consuming (Lundell, 2021). Additionally, there were pedagogical and accessibility issues ten years ago. Firstly, most EFL teachers and learners thought about vocabulary in terms of single words and not multiword elements (Schmitt, 2010), although they perform essential functions. An incomplete picture of vocabulary elements and a lack of emphasis on FL may lead to significant issues for EFL learners. Secondly, although this concept of language needed to be taken into proper attention, virtually no classroom materials capitalising on FL elements were available for commercial purposes (Wood, 2002).

However, after realising the contradiction between the importance of FL and the lack of appropriate teaching and learning methods, some research regarding pedagogical implications has been done. For example, Meunier (2012) compared the situation in English language teaching a decade later, after Wood's findings. It became clear that education took a step forward: digital sources now provide access to authentic materials and thus FL. The author also referred to another research in which new activities promoting the role of multiword phrases were suggested – *shadowing* and *dictogloss*. *Shadowing* is reading out loud a text simultaneously with a speaker. *Dictogloss* is a task during which learners collect keywords from a dictated text and then work in groups aiming to reconstruct its gist (Ellis, 2008). These activities help students find separate units of multiword expressions and then try to create whole phrases.

Erman and Lewis (2022) found that the type of educational activities significantly impacted FL production. They compared groups of English speakers – living in England and studying English in Sweden – in role-playing and retelling tasks. In the second group, the learners' mother tongue was Swedish. Students sounded like native-level speakers in the retelling activity, focusing on more typical phrases. In the role-playing one requiring collaboration, those who lived in

England performed better than those who did not live abroad. It follows that exposure to English is exceptionally beneficial for FL acquisition.

The next point of the paper is that students who recognised the exact formulaic phrases in Swedish tended to use them more often in English speech. Hence, finding connections between a native language and English might be fruitful for learning multiword phrases. In conclusion, the study showcases the importance of formulaic sequences in spoken English and the role of native environment exposure, task type, and similarities with L1.

The age at which people arrive in a new country may not influence the acquisition of FL in English; instead, knowledge of FL phrases in L1 depends on it. Vaynshteyn and Van Lancker Sidtis (2019) analysed how the age at which people immigrate impacts their ability to learn formulaic phrases. The researchers examined individuals who arrived from Russia to the USA early or late and compared them to English speakers. Regardless of arrival time, both bilingual groups performed less proficiently than English speakers in tasks assessing their use of formulaic command. Thus, the importance of the age when one migrated and its impact on FL acquisition might be overestimated.

Another point of the paper is that those who acquired English later performed better in tasks evaluating FL proficiency in Russian. This means the age at which one moves to another country could impact how well someone retains knowledge of formulaic phrases in L1; however, they lack FL skills in L2. In conclusion, teaching formulaic expression as early as possible might benefit proficiency in the second language later.

Discussing FL use on different levels, high-level Japanese speakers of EFL, like speakers of other languages, rely on many formulaic sequences, especially for describing objects, actions, time, and location, according to Culter (2017). However, recognising these phrases can be complicated because of differences in fluency (fillers, pauses, and hesitation in speech) and cultural gaps. Despite limitations, Culter (2017) states that EFL learners with a higher fluency use and repeat formulaic sequences more often, especially those for holding discussions and starting

conversations. Hence, instructions focusing on particular formulaic sequences can help advanced English language learners improve their speaking skills.

Even though FL phrases will be highly favourable at advanced levels, it is also vital to incorporate them at lower ones. In their study, Orta (2020) focused on this language group and analysed how low-level EFL students might benefit from memorising FL units. The results showed that the approach significantly contributed to FL acquisition. For instance, the participants could extract multiword phrases and recall their meanings a month later without practice. Also, memorising allowed the participants to integrate formulaic sequences in their active vocabulary, even when the initial understanding was incorrect. Thus, according to the results, repetition in real-life situations can improve FL comprehension. Common phrases like greetings or requests *("I am fine, thank you"* or *"Can I go to the toilet?"*) showcased the participants' progress. Memorisation might be a milestone to progress, and repeated practice in authentic conditions can demonstrate how those phrases work.

4.1 The Advantages of Formulaic Language for Writing and Speaking Skills

Using FL improves speaking skills, according to McGuire and Larson-Hall (2017. They tested the theory of whether teaching multiword sequences in class can improve fluency in EFL by comparing a treatment group (who received instructions on FL) and a control group (who did not receive them). The first group improved their speech flow, measured by speaking rate and mean run length, and used formulaic sequences more often than the second group. Nevertheless, the study has a few limitations that reduce generalisation: the method for identifying FL elements might have been subjective and tested on a specific cohort – students going abroad. Despite this, the paper proves that teaching formulaic phrases can improve fluency in English.

In terms of the connection between FL use and writing skills, Kilich (2015) researched the exam conditions in which the students needed to craft an opinion essay. The findings suggested no significant relationship between FL use and students' coherence and total writing scores. It is possible that other essential linguistic elements, such as robust content grammar and vocabulary,

impact the quality of writing along with FL. However, research does not diminish the importance of multiword phrases; the point was to highlight that FL should not be considered separately from the rest of linguistic objects.

Turning to the pedagogical implications listed in Kilich's article (2015), most importantly, the administration of educational organisations should carefully approach the choice of learning materials. First, these materials should include many multiword expressions that EFL students can use in classes and proficiency tests. Then, FL should be incorporated into the curriculum along with grammar and vocabulary as they complement each other. There should also be writing models provided where it is noticeable that FL is an excellent addition to an already advanced level of English rather than its basis. Lastly, material developers should offer more activities to help students apply the elements to grammar and meaningful context. In conclusion, this article presents a new perspective on the approach to learning multiword sequences: they should be studied along with other elements of English.

Vocabulary-based learning of academic FL would enhance students' skills in using those expressions when writing, as Peters and Pauwels (2015) state. It follows that the suggested approach helps recognise, recall, and use multiword phrases later for performing essays. In addition, the participants utilised the new phrases in their final projects and incorporated them into the active vocabulary. Peters and Pauwels (2015) also compared different methods of teaching FL and claimed that activities aimed at recalling expressions might be more productive than those aimed at recognising the expressions. Finally, students who received instructions on FL utilised more of them and wrote essays more formally. Generally, the study highlights the importance of explicit instruction on FL for academic writing.

4.2 The Advantages of Separate Formulaic Language Categories for Language Skills

Phrasal verbs are one of the most popular FL categories, and an effective strategy to learn them was identified – item grouping technique with a particle list. Spring (2018) analysed the efficiency of teaching English phrasal verbs (PV) to Japanese speakers by combining item groups

(motion, change of state, etc.) and using a particle list (up, down, in, etc.). This approach significantly increased students' knowledge of PV compared to the whole-unit method (everything grouped, without detailing topics and particles). Those who used the grouping method showed better results than other participants and liked English classes more. Therefore, the grouping approach is more effective. To sum up, Spring's research showcases the positive strategy of using the item grouping technique with a particle list to enhance learners' understanding of PV compared to the whole-unit one. A more robust command of PV will eventually enrich FL vocabulary.

Integrating idioms into English lessons is exceptionally beneficial for improving speaking skills (De, 2009). The study showed that workshops focusing on idioms were advantageous for EFL students. For example, the new lessons helped them enhance their English skills, broaden vocabulary, and better understand the subtle meanings embedded in music and movies native speakers enjoy. The positive outcomes demonstrate that including expressions in the curriculum can significantly improve fluency in English and raise cultural awareness, which is vital for communication and developing language skills.

Lui (2003) identified drawbacks in how spoken American English idioms are taught at educational institutions and proposed solutions. The corpus-based analysis showcased that current teaching materials often fall short of reflecting real-life cases. To address this, Lui (2003) suggested incorporating data from natural speech, providing clear explanations with examples for each idiom, and emphasising the contexts in which they are used. Additionally, the researcher recommended prioritising frequently encountered idioms for beginners. Overall, this study proves the applicability of corpus research in creating teaching materials that cover idioms and explain how idioms function in real-world communication.

4.3 The Advantages of FL for Exam Preparation and Studying Abroad

Standardised test (IELTS, TOEFL, state exams) preparation is a milestone for EFL learners, and focusing on pragmatic formulaic phrases might be practical in this case (Khodadady & Shamsaee, 2012). As the authors state, traditional test preparation might include memorising

vast numbers of formulaic elements, and this strategy is not efficient. The study suggests that students should focus on FL's actual function instead of its frequency or randomness. For example, it would be better for learners to memorise and be able to utilise smaller sets of high-impact multiword phrases, such as transitions and markers.

Another point of the paper is that EFL test preparation should cover FL in real situations. An effective FL learning strategy would also include exercises that improve grammar, vocabulary, and pronunciation because this approach affects various English skills. In summary, efficient EFL test preparation should emphasise the practical application of formulaic phrases over memorisation. Memorisation is beneficial for beginners but not in the case of standardised tests.

According to Goncharov (2019), who analysed a short-term IELTS speaking course and its advantages, formulaic phrases can influence speaking performance and even improve the band score. Five learners participated in this course to estimate their progress after. Thanks to it, the students enhanced their FL vocabulary in English, eventually enhancing fluency (there were fewer pauses later). Moreover, three participants improved their IELTS speaking band score. The findings suggest that courses covering FL phrases can enhance fluency and increase speaking scores on the IELTS test, but further research is needed to confirm the correlation.

It is claimed that pre-instructions on FL can help students going overseas to respond to social situations better. Wang and Halenko (2022) examined how pre-learning and exposure abroad impact learners' skills to use multiword phrases in various cases (asking for a pen or an excuse to use the bathroom). Those who were pre-instructed on formulaic phrases significantly enhanced accuracy compared to those who were not. Overall, the results showcase that preparation in advance helps students understand social contexts where formulaic phrases are used better. However, while exposure improved the FL skill, some students used formulaic sequences inaccurately, and the researchers suggested emphasising appropriateness. Hence, pre-instructions on FL could benefit EFL learners, but accurate usage should be a priority.

4.4 The Role of New Technologies in Formulaic Language Learning

Speaking about cutting-edge technologies in applied linguistics, Lin (2021) introduced a new tool for FL learning from YouTube videos – IdiomsTube. The point is that the software automatically generates tasks based on the captions of videos. Nowadays, the program is facing severe challenges:

- It cannot yet tell the difference between metaphorical and literal use of idioms.
- YouTube's captions are not always accurate and do not always align with the actual video, which impacts IdiomsTube's ability to create tasks.
- The number of idioms to include in the tasks varies from video to video. Shorter videos usually have fewer idioms.

Nevertheless, the software developers are improving the new tool. Their main goal is teaching AI to prioritise idioms and distinguish metaphorical and literal ones. However, only some things depend on the developers; they also hope that YouTube will improve the quality of video captions shortly.

Regarding another study in educational technologies, Cobb (2018) found that computer-assisted language learning (CALL) might become an efficient tool for FL acquisition, but there are still ongoing discussions. The program highlights words and phrases in different contexts (games, books, videos) and provides their translation. While CALL has successfully detected singular words, it has neglected formulaic sequences. The study raised many questions and arguments, such as:

- Should formulaic phrases be taught before or in combination with the single words?
- Can small screens allow teaching FL?
- Is there a need for learners to perform corpus analysis themselves?

However, the researchers claim that while many multiword phrases exist, EFL learners do not need to learn them all. Furthermore, single words and FL can be covered together. The chapter concludes that CALL programs should incorporate some awareness of FL alongside a focus on

singular words. Hence, corpus technology can be a helpful tool for exposing EFL learners to formulaic expressions, but more research is needed on integrating formula work into CALL best.

5 Conclusion

This literature review aimed to prove FL's relevance for English language learners – multiword constructions stored in memory as single units (Wood, 2020, p. 30). To achieve this goal, a profound literature review focusing on FL's classification, vitality, and applicability for language teaching was conducted.

The paper showcased that incorporating formulaic phrases into vocabulary can benefit EFL learners in several ways. Firstly, multiword expressions can enhance the ability to express themselves confidently, ensure successful communication in various social situations, and improve writing skills. Secondly, FL can contribute to more comprehensive speech delivery and faster information recall thanks to the efficient storage of these linguistic elements in memory.

However, the paper also showed that further research is needed to find efficient methods for teaching formulaic sequences and identifying the most essential phrases for learners at various proficiency levels. As referred to in the literature review, corpora-based and machine learning-based software are now being tested for integrating FL phrases. Many concerns and questions must be addressed and resolved soon, but educational technologies are progressing in successfully incorporating new learning tools.

References

Altenberg, B. (1998). On the phraseology of spoken English: The evidence of recurrent word-combinations. In A. P. Cowie (Ed.), *Phraseology: Theory, analysis, and applications*, (pp. 101–122). Oxford University Press.

Arnaud, P. S. L. (1993). Nattinger, J. R. & Jeanette S. Decarrio (1992). *Lexical Phrases and Language Teaching*, XVI + 218. In S.L. Arnaud (Ed.), *Meta: Translators' Journal, 38*(3), (pp. 565–568). https://doi.org/10.7202/001902ar

Biber, D. (2006). *University language: A corpus-based study of spoken and written registers*. John Benjamins.

Biber, D., Conrad, S., & Cortes, V. (2004). If you look at…: Lexical bundles in university teaching and textbooks. *Applied Linguistics, 25(*3), 371–405. https://doi.org/10.1093/applin/25.3.371

Biber, D., Johansson, S., Leech, G., Conrad, S., & Finegan, E. (1999). *Longman grammar of spoken and written English*. Pearson Education.

Boers, F., Eyckmans, J., Kappel, J., Stengers, H., & Demecheleer, M. (2006). Formulaic sequences and perceived oral proficiency: Putting a lexical approach to the test. *Language Teaching Research, 10*(3), 245–261. https://doi.org/10.1191/1362168806lr195oa

Carrol, G., & Conklin, K. (2019). Is all formulaic language created equal? Unpacking the processing advantage for different types of formulaic sequences. *Language and Speech, 63*(1), 95–122. https://doi.org/10.1177/0023830918823230

Cobb T. (2018). FROM CORPUS TO CALL: The use of technology in teaching and learning formulaic language. In A. Siyanova-Chanturia & A. Pellicer-Sanchez (Eds.) *Understanding Formulaic Language: A Second Language Acquisition Perspective* (pp. 192–211). https://doi.org/10.4324/9781315206615

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and non-native speakers? *Applied Linguistics*, *29*(1), 72–89. Oxford University Press. https://doi.org/10.1093/applin/amm022

Cutler, S.F. (2017). The use of psycholinguistic formulaic language in the speech of higher-level Japanese speakers of English. *Vocabulary Learning and Instruction*, *6*(1), 48–60. https://doi.org/10.7820/vli.v06.1.cutler

Dakhi, S., & Fitria, T. N. (2019). The Principles and the Teaching of English Vocabulary: A Review. *Journal of English Teaching, 5*(1), 15–25. https://doi.org/10.33541/jet.v5i1.956

De, E. E. R. (2009). The advantages and importance of learning and using idioms in English. *Cuadernos De Lingüística Hispánica*, *14*, 121–136. https://dialnet.unirioja.es/descarga/articulo/3618851.pdf

Denning, K. M., Kessler, B., & Leben, W. R. (2007). *English Vocabulary Elements*. Oxford University Press.

Dóczi, B. (2024). A szókincs szerepe a nyelvtanulásban. In *Akadémiai Kiadó eBooks*. https://doi.org/10.1556/9789636640125

Egamova, J. & Sharofova, N. (2022). The importance of vocabulary in learning English language. *Central Asian Research Journal For Interdisciplinary Studies (CARJIS), 2*(11), 187–190.

Ellis, N. C. (2008). Phraseology: The periphery and the heart of language. In F. Meunier & S. Granger (Eds.), *Phraseology in foreign language learning and teaching*, (pp. 1–13). John Benjamins.

Erman, B. (2009). Formulaic language from a learner perspective: What the learner needs to know. In R. Corrigan, E. Moravcisk, H. Ouali, & K. Wheatley (Eds.), *Formulaic language, Vol. 2: Acquisition, loss, psychological reality, and functional explanations*, (pp. 324–346). John Benjamins.

Erman, B., & Lewis, M. (2022). Formulaic language in L1 and advanced L2 English speech: multiword structures in the speech of two Swedish groups compared to a group of L1 English speakers. *Text & Talk*, *44*(2), 175–196. https://doi.org/10.1515/text-2021-0090

Erman, B., & Warren, B. (2000). The Idiom Principle and the Open Choice Principle. *Text, 20*(1), 29–62. https://doi.org/10.1515/text.1.2000.20.1.29

Fernando, C. (1996). *Idioms and idiomaticity*. Oxford University Press.

Goncharov, G. M. (2019). The effect of direct instruction in formulaic sequences on IELTS students' speaking performance. *Advanced Education*, *6*(11), 30–39. https://doi.org/10.20535/2410-8286.132528

Hay, J., & Bresnan, J. (2006). Spoken syntax: The phonetics of giving a hand in New Zealand English. *The Linguistic Review 23*, 321–349. http://doi.org/10.1515/TLR.2006.013

James, C. (1998). Errors in Language Learning and Use: Exploring Error Analysis. *Applied Linguistics and Language Study*. Routledge.

Kecskés, I. (2007). Formulaic language in English Lingua Franca. In *De Gruyter eBooks*, 191–218. https://doi.org/10.1515/9783110198843.3.191

Khodadady, E., & Shamsaee, S. (2012). Formulaic Sequences and Their Relationship with Speaking and Listening Abilities. *English Language Teaching*, *5*(2), 39–49. https://doi.org/10.5539/elt.v5n2p39

Kilich, S. (2015). The use of formulaic language by English as a foreign language (EFL) learners in writing proficiency exams. The Graduate School of Education of İhsan Doğramacı Bilkent University. https://core.ac.uk/download/pdf/52927434.pdf

Lin, P. M. (2021). Developing an intelligent tool for computer-assisted formulaic language learning from YouTube videos. *ReCALL (Hull)*, *34*(2), 185–200. https://doi.org/10.1017/s0958344021000252

Liu, D. (2003). The most Frequently used spoken American English Idioms: A Corpus Analysis and its Implications. *TESOL Quarterly*, *37*(4), 671–700. https://doi.org/10.2307/3588217

Lundell, F. (2021). Formulaicity. In N. Tracy-Ventura & M. Paquot (Eds.), *The Routledge Handbook of Second Language Acquisition and Corpora*, (pp. 370–381). Routledge.

McCretton, E. & N. Rider (1993). Error gravity and error hierarchies. *IRAL*. *31*(3), 177–188.

McGuire, M. P., & Larson-Hall, J. (2017). Teaching Formulaic Sequences in the Classroom: Effects on Spoken Fluency. *TESL Canada Journal*, *34*(3), 1–25. https://doi.org/10.18806/tesl.v34i3.1271

Meunier, F. (2012). Formulaic Language and Language Teaching. *Annual Review of Applied Linguistics*, 111–129. https://doi.org/10.1017/S0267190512000128

Millar, N. (2011). The processing of malformed formulaic language. *Applied Linguistics, 32*(2), 129–148. https://doi.org/10.1093/applin/amq035

Nation, I. S. P. (1989). Improving speaking fluency. *System, 17*(3), 377–384. https://doi.org/10.1016/0346-251x(89)90010-9

Ohlrogge, A. (2009). Formulaic expressions in intermediate EFL writing assessment. In R. Corrigan (Ed.), *Formulaic language*, (pp. 375–386). https://doi.org/10.1075/tsl.83.07ohl

Orta, J. M. (2020). Fostering speaking in the EFL classroom: the role of formulaic language in low proficiency L2 learners of English. http://dspace.uvic.cat/xmlui/handle/10854/6378

Pawley, A., & Syder, F. H. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards, & R. W. Schmidt (Eds.), *Language and communication*, (pp. 191–226). Longman.

Peters, E., & Pauwels, P. (2015). Learning academic formulaic sequences. *Journal of English for Academic Purposes*, 28–39. https://doi.org/10.1016/j.jeap.2015.04.002

Rafieyan, V. (2018). Knowledge of formulaic sequences as a predictor of language proficiency. *International Journal of Applied Linguistics and English Literature*, *7*(2), 64–69. https://doi.org/10.7575/aiac.ijalel.v.7n.2p.64

Rogošić, G. D. (2014). Lexical Collocations as a Building Block in Teaching ESP. *Conference: International Scientific and Professional Conference, Contemporary Issues in Economy and Technology, CIET 2014*, 171–180. https://www.bib.irb.hr/701905

Shin, D. (2007). The high-frequency collocations of spoken and written English. *English Teaching, 62*(1), 199–218. https://doi.org/10.15858/engtea.62.1.200703.199

Sullivan, R. A. & Alba, J. O. (2010). Criteria for EFL course books' vocabulary selection: Does it have any practical consequences? *Universidad de Las Palmas de Gran Canaria*, 197–210. https://ojsspdc.ulpgc.es/ojs/index.php/ElGuiniguada/article/viewFile/431/370

Schmitt, N. (2010). Researching Vocabulary: A Vocabulary Research Manual. *Palgrave Macmillan*, 8–12. http://dx.doi.org/10.1057/9780230293977

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences: Acquisition, processing and use,* (pp. 1–23). John Benjamins.

Schmitt, N., & Martinez, R. (2004). A Phrasal Expressions List. *Applied linguistics, 33*(3), 299–320. Oxford University Press.

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

Spring, R. (2018). Teaching phrasal verbs more efficiently: using corpus studies and cognitive linguistics to create a particle list. *Advances in Language and Literary Studies, 9*(5), 121-135. https://doi.org/10.7575/aiac.alls.v.9n.5p.121

Talmy, L. (1985). Lexicalization patterns. Semantic structure in lexical form. In T. Shopen (Ed.) *Language typology and syntactic description,* (pp. 36–149). CUP. https://doi.org/10.1017/cbo9780511618437.002

Talmy, L. (2000). Toward a cognitive semantics: Typology and process in concept structuring. *Vol. 2: Language, speech, and communication series,* viii+565. MIT Press.

Thim, S. (2012). Phrasal Verbs: The English Verb-Particle Construction and its History. *English Language and Linguistics English Language and Linguistics*, *18*(3), 572–586. https://ci.nii.ac.jp/ncid/BB10611511

Vaynshteyn, I., & Van Lancker Sidtis, D. (2019). Effects of age of arrival on acquiring formulaic expressions in the second language. *Study in English Language Teaching*, *7*(4), 391–418. https://doi.org/10.22158/selt.v7n4p391

Wang, J., & Halenko, N. (2022). Developing the use of formulaic language for study abroad: a targeted instructional intervention. *Language Learning Journal*, *50*(4), 409–426. https://doi.org/10.1080/09571736.2022.2088446

Wingate, M. E. (1984). Fluency, disfluency, dysfluency, and stuttering. *Journal of Fluency Disorders*, 163–168. https://doi.org/10.1016/0094-730x(84)90033-0

Wood, D. (2002). Formulaic language acquisition and production: Implications for teaching. *TESL Canada Journal*, *20*(1), 1–15. https://doi.org/10.18806/tesl.v20i1.935

Wood, D. (2009). Effects of focused instruction of formulaic sequences on fluent expression in second language narratives: A case study. *ACLA*, *1(*12), 39–57.

Wood, D. (2010). *Formulaic language and second language speech fluency: Background, evidence, and classroom applications.* Continuum.

Wood, J. (2020). Classifying and Identifying Formulaic Language. In S. Webb (Ed.) *The Routledge Handbook of Vocabulary Studies*, (pp. 30–46). Routledge.

Wray, A. (1999). Formulaic language in learners and native speakers. *Language Teaching*, *32*(4), 213–231. https://doi.org/10.1017/s0261444800014154

Wray, A., & Perkins, M. R. (2000). The functions of formulaic language: An integrated model. *Language and Communication*, *20*, 1–28. https://doi.org/10.1016/S0271-5309(99)00015-4

Zhi-Hong, B. (2018). An analysis of English vocabulary learning strategies. *Journal of Language Teaching and Research*, *9*(4), 849–855. https://doi.org/10.17507/jltr.0904.24