Formulaic Sequences in Second Language Acquisition

Supakorn Phoocharoensil
yhee143@gmail.com
Thammasat University

Abstract
This article is aimed at exploring the characteristics of formulaic sequences, as well as how they are acquired by second language learners. The main focus is on the vital role formulaic sequences are currently playing in English vocabulary use and instruction, namely their prevalence in spoken and written English, and their contribution to learners' significant L2 fluency development. As regards L2 formulaic sequence acquisition, two major different viewpoints exist. On the one hand, it is believed that vocabulary should be learned as a whole chunk rather than as units segmented from a string of words. Some scholars, on the other, argue that learners tend to learn vocabulary through component word analysis. The final topic to which this paper is devoted concerns formulaic sequence teaching, which evidently depends upon two principal factors: repetition and degree of lexical knowledge.

Keywords: formulaic sequence, formulaic language, chunk, English vocabulary language learning and teaching, second language acquisition

Introduction

English vocabulary instruction has primarily focused on single words since they are normally regarded as the basic unit of lexis. To many teachers, individual words seem to be convenient units to base instruction upon and incorporate into materials. As noted by Schmitt (2010, p. 9), "The main vocabulary reference source, dictionaries, are set up around individual headwords". It, accordingly, is an unsurprising fact that most teachers and learners view English vocabulary as a set of single words. Several researchers, likewise, also place an emphasis on individual words.

However, it is becoming increasingly accepted that mere knowledge of words in isolation appears inadequate for English vocabulary mastery. To successfully acquire L2 English lexical knowledge, one also has to take into consideration formulaic language as "an important element of language learning and use" (Schmitt, 2010, p. 9). Formulaic language concerns a chunk or sequence defined as "a sequence, continuous or discontinuous, of words or other elements, which is, or appears to be, prefabricated; that is, stored or retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar" (Wray, 2002, p. 9).

What do formulaic sequences look like?

Applied linguists have been employing a variety of related terms to refer to the concept expressed by formulaic language, as observed by Wray (2002), e.g. chunks, multiword units, collocations, clusters, lexical bundles, ready-made utterances, conventionalized forms, etc. Schmitt and Carter (2004, p. 3) remarked that formulaic language can be long as an idiom, e.g. *you can lead a horse to water, but you can't make him drink*, or short, such as an exclamation *Oh no!*, or anything in between.

Formulaic sequences are used for various purposes (Schmitt, 2010, p. 10). For instance, they can express a message or idea, e.g. the idiomatic expression *The early bird catches the worm*, interactional functions, such as *(I'm) just looking (thanks)*, which involves declining an offer of assistance from a shopkeeper, or social solidarity, e.g. *I know what you mean*, to express an agreement with an interlocutor. In terms of their patterns or occurrences, formulaic sequences can be entirely fixed, e.g. *Ladies and gentlemen*, in which the order of the components which *and* is between cannot be switched, or they sometimes have slots which is fillable with suitable words or groups of words, e.g. [someone/thing, usually with authority] made it plain that [something as yet unrealized was intended or desired] (Schmitt & Carter, 2004, p. 3).

Carter (2012, p. 3) maintained that fixedness is a major characteristics of formulaic language, "allowing such language to be memorized and used as wholes, rather than being newly created for each use". Some formulaic sequences are totally fixed, such as idioms. Carter exemplifies this notion using the fixed idiom *to have forty winks*, meaning 'to have a short sleep', as opposed to variations like **to have thirty nine winks* or **to have a short wink*. Still, plenty of formulaic language is not fixed but allows for flexibility. For example, verbs can vary in chunks like *cost/pay/spend/charge an arm and a leg* (Moon, 1997, 2010).

Why should formulaic language be given a prominent place in vocabulary use and pedagogy?

Formulaic language is of growing importance in English vocabulary use as well as instruction for at least two main reasons:

1. Formulaic language is very common in both spoken and written discourse. Erman and Warren (2000) discovered that 52-58% of the L1 English contains formulaic sequences. In a similar vein, Foster (2001), using different research methods and criteria, reported that 32% of L1 English contains formulaic language. In Howarth (1998)'s study of frequent verbs in a social science-academic corpus, approximately 31-40% of formulaicity was found in restricted collocations or idioms. It is worth noting that "a large proportion of the language we meet and produce is made up of multiword units" (Nation and Webb, 2011, p. 175). Likewise, according to O'Keeffe, McCarthy and Carter (2007), lexical chunks occur very frequently in spoken English, as indicated in CANCODE (The Cambridge and Nottingham Corpus of Discourse in English), which is a collection of spoken English recorded at hundreds of locations across the British Isles in a wide variety of situations. The British spoken English corpus-based evidence has demonstrated that several high-frequency chunks, e.g. *you know*, "are more frequent and more central to communication than even very frequent words" (O'Keeffe, 2007, p. 69). The prevalence of formulaic language, e.g. lexical chuncks, clearly necessitates its inclusion in a modern English syllabus, as well as in structure of other languages', since formulaic language is also ubiquitous in a range of languages, including French, Italian, Spanish, German, Swedish, Polish, Arabic, Hebrew, Greek, and Chinese (Conklin and Schmitt, 2008).

2. Making use of formulaic language increases English speakers' fluency and accuracy. Pawley and Syder (1983) revealed that even though native speakers have rather limited language processing speed, it is evident that their language production ability is beyond the processing capacity. The reason for their higher fluency in language production primarily deals with formulaic clusters which have already been memorized as ready-made prefabricated units. Put differently, these formulaic sequences are "…stored as single wholes and… instantly available for use without the cognitive load of having to assemble them on-

line as one speaks (Schmitt, 2010, p. 11). Similarly, Kuiper (2004) found support for Pawley and Syder in that speakers employ a large amount of formulaic language in their speech, under severe time constraints.

Not only does formulaic language facilitate speakers' speech production, but it also plays a vital role in language processing. Ellis (2006) and Conklin and Schmitt (2008) have provided evidence that formulaic sequences, e.g. making requests, making apologies, responding to compliments, refusing, complaining, etc. These speech acts are realized by conventionalized language, e.g. *I'm (very) sorry to hear about* _____, to express sympathy, or *I'd be happy to* _____ in response to a request (Nattinger and Decarrico, 1992). These formulaic sequences are viewed as ready-made chunks that enable speakers to achieve the relevant speech act in a quick, reliable manner.

McCarthy and Carter (1997), moreover, suggested that certain formulaic sequences are a key element of informal spoken discourse. These concern chunks that are commonly used in social interaction to maintain interpersonal relationship between interlocutors, as in Small talks. Schmitt (2010) proposed that such formulaic language allows speakers to express social solidarity, e.g. comments about the weather, e.g. *Nice weather today*, agreeing with interlocutors, e.g. *Oh, I see what you mean*, or providing positive feedback to another speaker, e.g. *How interesting*.

How is formulaic language learned?

Two distinct approaches to formulaic language learning are available: formulaic and non-formulaic models, to be discussed in detail below.

A formulaic approach to chunk learning

The proponents of this approach take a view that vocabulary should be learned as a whole chunk rather than as analyzed units segmented from a string of words (e.g. Lewis, 1997, 2000; Durrant and Schmitt, 2009, 2010; Ellis, 2001; Kuiper, Columbus, and Schmitt, 2009). Lewis (1997, 2000), in what he proposed as The Lexical Approach, takes a view that vocabulary and grammar as inseparable, essential components which constitute a lexical chunk or phrase should be highlighted in English instruction. It is claimed that vocabulary needs to be memorized and retained as chunks or a grammaticalized lexis ready for future use, as opposed to single words. In support of Lewis, Ellis (2001) postulated that language learners have a tendency to conceptualize formulaic language as the product of chunking. He observed that learners can enormously increase the amount of information which their memory can handle through recording vocabulary as clusters. When learners encounter two or more words that co-occur with very high frequency, it is highly likely that they will record these words together as a single unit, storing them in their memory so that this particular chunk will be ready for future production. Such a prefabricated unit is believed to help learners communicate effectively in L2 English. In addition, when they meet this chunk again, the time spent on processing it is expected to be much shorter, which means acquisition of vocabulary as a chunk seems to expedite human language processing, apart from production.

Kuiper et al. (2009), in a similar way, also indicated that formulaic sequences offer processing efficiency because single memorized lexical units consisting of word sequences are found to be processed at higher speed and with greater ease than a string of words generated creatively. Furthermore, Conklin and Schmitt (2008) corroborated Ellis (2001) and

gave support to Kuiper et al. (2009), exhibiting findings that non-native speakers read formulaic units more quickly than non-formulaic ones. This apparently shows that formulaic language is a factor contributing to clearly effective language use.

Saliency is also considered a key factor in acquiring formulaic sequences. As Ellis (2001) suggested, salient sequences are more likely to be learned than those that are less salient. That is, learners tend to acquire chunks to which they are more exposed. Durrant and Schmitt (2009) accorded with Ellis (2001) in that they reported learning collocations, i.e., a common form of formulaic language, is closely connected with the degree of saliency. Simply put, adult non-native English learners extensively used L2 English collocations which appear to be salient, i.e., occurring with high frequency, such as *good example* and *long way*, as opposed to less-salient collocations, e.g. densely populated and preconceived notions. In brief, it is clearly seen that English learners acquire collocations to which they have more exposure.

Durrant and Schmitt (2010)'s research was in line with their study in 2009. The more recent work investigated adult L2 learners' collocational knowledge retention through exposure. Insufficient collocation input was a major cause of learners' low collocational competence, which confirms how important exposure is in collocation acquisition. Having been exposed to prepared adjective + noun collocations in various training conditions, the participants achieved collocation retention in the verbatim and varied repetition conditions, with the former being slightly more effective. The researchers revealed, "…fluency-oriented repetition of individual sentence contexts has a greater impact on collocation learning than does exposure to the same collocations in different contexts" (Durrant and Schmitt, 2010, p. 182). Both studies by Durrant and Schmitt provided supportive evidence that formulaic language, e.g. collocations, is evidently learned as a single, prefabricated chunk instead of individual words. Nonetheless, there also exists certain research indicative of formulaic sequence acquisition through chunk division for analysis, to be elaborated on in the next subsection.

A non-formulaic approach to chunk learning

Whereas the aforementioned studies subscribe to acquiring chunks as a single unit, others seem to provide a counter argument against such a formulaic approach. It has been claimed that adults L2 learners access formulaic language in a different way from what L1 child acquirers do.

Wray (2002) postulated that adult L2 learners mainly notice and remember individual component words, as opposed to the entire chunk. To be precise, the way L1 child acquires formulaic language differs from how L2 learners do. For example, in L1 acquisition, once encountering a collocation like *major catastrophe*, L1 learners tend to view it as a single unit to be memorized as a cluster meaning 'large disaster'. In contrast, adult L2 learners are inclined to analyze this lexical bundle, dividing it into two words, one of which means 'big' and the other means 'disaster', before keeping the two words separately in their memory. When they meet this word combination again, they will have no memory of *major catastrophe* as a chunk but instead as two separate words (Wray, 2002, pp. 206-209). Obviously, L2 learners do not seem to remember a collocation as a single bunch of words. In fact, they focus more on words than sequences since managing single words is much easier and gives learners a feeling of control over the language.

There appear two principal reasons for the non-formulaic approach in L2 vocabulary acquisition. With respect to social factors, unlike L1 learners, hardly do adult L2 learners have pressure to memorize useful prefabricated clusters for everyday communication. This is

because in many cases the surrounding social pressure could stem from the learners'L1, rather than the L2 community. Therefore, L2 learners' adoption of native-like formulaic sequences can be discouraged. In terms of cognitive factors, there is a strong likelihood that learners will feel uneasy or uncomfortable not knowing how sequences are broken into individual words. With an attempt to perform chunk analysis, learners fail to store formulaic language as wholes.

Research has presented supportive evidence of non-formulaic model of L2 chunk acquisition. As demonstrated by Siyanova and Schmitt (2008), frequency of occurrences of formulaic language was a major determinant of L2 learners' collocation processibility. More specifically, the collocations occurring with high frequency seemed to be learned more easily than infrequent chunks the components of which were less familiar to them.

Liu (2010) confirmed necessity of a cognitive analysis of collocations in learners' formulaic language acquisition. Liu insisted that not all collocations are unpredictable or semantically unmotivated as the advocates of formulaic-learning approach often claim. To put it another way, whilst some formulaic-language researchers proposed that a chunk is acquired as a single unit and then kept for future use (Ellis, 2001), it can be clearly seen that some common collocations allow for learning by analysis. Being analyzable, collocations are claimed to be acquired through learners' cognitive exploration process. This learning process enables them to look at L2 collocations with an analytical mind, resulting in effective collocation retention.

How should formulaic sequences be taught?

How should English teachers present the notion of formulaic sequences to students? This seems to be what teachers have wondered for decades. As research indicates, there are as great a number of occurrences, if not greater, of formulaic language in English as there are of single words (Jackendoff, 1995). It is, for this reason, important for learners to acquire a number of formulaic sequences, in addition to single words, so as to have an acceptable level of L2-English lexical competence (Phoocharoensil, 2013). Some may claim that incidental vocabulary learning from language input results in advanced lexical L2 knowledge. Schmitt (2008), nevertheless, found that learning vocabulary incidentally is apparently slow, requiring a lot of exposures. Moreover, this incidental learning, as noted by Schmitt, does not always guarantee a productive vocabulary mastery level. This implies the significance of explicit instruction of vocabulary items, including formulaic sequences.

There has existed so far limited research on direct teaching of formulaic language. Jones and Haywood (2004) were among a group of very few researchers who examined formulaic language teaching. They discovered that highlighted formulaic sequences during a 10-week course evidently raised students' awareness of chunks, despite the fact that this awareness did not lead to any substantial increment in the use of the presented chunks in students' output. The researchers also suggested that instruction could have more impact on the accuracy and appropriateness of formulaic language use.

Even though it is troublesome in deciding on which teaching method serves as the best way to introduce formulaic sequences to students, there appear two major factors influencing students' achievement of chunks in relation to vocabulary pedagogy: repetition and degree of vocabulary knowledge.

As it holds true for individual word acquisition, repetition is considered to be conducive to learning formulaic sequences as well. That is, learners' need for recycling and multiple exposures to target vocabulary items is undeniable and outstanding (Nation, 2001).

In other words, new words learned must be recycled for students' retention of them. According to Wood (2002), repetition should be taken into account as a way to promote learners' formulaic sequence acquisition. This idea gains support from Alali and Schmitt (2012), who revealed Kuwati EFL students benefited from repetition in learning formulaic language when it came to form recognition. Additionally, they suggested "…repetition is certainly beneficial in pushing vocabulary knowledge up to the level where it can be useful in real-life contexts" (Alali and Schmitt, 167). With regard to types of repetition, written repetition was relatively more effective than its oral counterpart. Overall, the teaching methods used for single word instruction were found to be also effective in formulaic sequence teaching, although the chunk learning turned out to be somewhat less effective when applied to chunk learning as opposed to individual word learning.

Furthermore, the degree of vocabulary knowledge was also viewed as another key to lexical competence, as borne out by many studies. There is a point of immense importance for differentiation of the degree of vocabulary knowledge. Alali and Schmitt (2012) pointed out little difference between any of the treatments, namely no review, oral review, and written review. This was because almost all the participants received maximum scores. For a recall level of knowledge, the lexical learning appeared to be far less retained. The researchers posited that teachers are expected to carefully consider what levels of lexical knowledge they really want to enhance and then use the most appropriate method(s) to address these particular levels.

References

Alali, F.A. & Schmitt, N. (2012). Teaching formulaic sequences: The same as or different from teaching single words. *TESOL Journal, 3*(2), 153-180.

Carter, R. (2012). *Vocabulary. Applied linguistic perspectives*. London: Routledge.

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

Durrant, P. & Schmitt, N. (2009). To what extent do native and non-native writers make use of collocations? *International Review of Applied Linguistics, 47*, 157-177.

Durrant, P. & Schmitt, N. (2010). Adult learners' retention of collocations from exposure. *Second Language Research, 26*(2), 163-188.

Ellis, N. C. (2001). Memory for language. In Robinson, P. (Ed.), *Cognition and second language instruction*. Cambridge: Cambridge University Press.

Ellis, N. C. (2006). Language acquisition as rational contingency learning. *Applied Linguistics, 24*, 143-188.

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

Foster, P. (2001). Rules and routines: A consideration of their role in the task-based language production of native and non-native speakers. In M. Bygate, P. Skehan & M. Swain (Eds.), *Researching pedagogic tasks: Second language learning, teaching, and testing* (pp. 75-94). Harlow: Longman.

Howarth, P. (1998). Phraseology and second language proficiency. *Applied Linguistics, 19*, 45-72.

Jackendoff, R. (1995). The boundaries of the lexicon. In M. Everaert, E. van der Linden, A. Schenk, & R. Schreuder (Eds.), *Idioms: Structural and psychological perspectives* (pp. 133-166). Hillsdale, NJ: Lawrence Erlbaum.

Jones, M. & Haywood, S. (2004). Facilitating the acquisition of formulaic sequences: An exploratory study in an EAP context. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 269-292). Amsterdam, the Netherland: John Benjamins.

Kuiper, K. (2004). Formulaic performance in conventionalized varieties of speech. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 37-54). Amsterdam: John Benjamins.

Kuiper. K., Columbus, G., & Schmitt, N. (2009). The acquisition of phrasal vocabulary. In S. Foster-Cohen (Ed.), *Language acquisition* (pp. 216-240). Hamshire: Palgrave Macmillan.

Lewis, M. (1997). Pedagogical implications of the lexical approach. In J. Coady & T. Huckin (Eds.), *Second language vocabulary acquisition* (pp. 255-270). Cambridge University Press.

Lewis, M. (2000). *Teaching collocation: Further development in the lexical approach*. London: Commercial Colour Press Plc.

Liu, D. (2010). Going beyond patterns: Involving cognitive analysis in the learning of collocations. *TESOL Quarterly*, *44*(1), 4-30.

McCarthy, M. & Carter, R. (1997). Written and spoken vocabulary. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary. Description, acquisition, and pedagogy* (pp. 20-39). Cambridge: Cambridge University Press.

Moon, R. (1997). Vocabulary connections: Multi-word items in English. In N. Schmitt & M. McCarthy (Eds.), *Vocabulary. Description, acquisition, and pedagogy* (pp. 40-63). Cambridge: Cambridge University Press.

Moon, R. (2010). What can a corpus tell us about lexis? In A. O'Keeffe and M. McCarthy (Eds.), *The Routledge handbook of corpus linguistics* (pp. 197-211). London: Routledge.

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

Nation, I. S. P., & Webb, S. A. (2011). *Researching and analyzing vocabulary*. Heinle, Cengage Learning.

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

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

Pawley, A. & Syder, F. (1983). Two puzzles for linguistic theory: Nativelike selection and nativelike fluency. In J. C. Richards & N. Schmitt (Eds.), *Language and communication* (pp. 191-225). London: Longman.

Phoocharoensil, S. (2013). Cross-linguistic influence: Its impact on L2 English collocation production. *English Language Teaching, 6*(1), 1-10.

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

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

Schmitt, N. & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 1-12). Amsterdam, the Netherland: John Benjamins.

Siyanova, A. & Schmitt, N. (2008). L2 learner production and processing of collocation: A multi-study perspective. *Canadian Modern Language Review, 64*(3), 429-458.

Wood, D. (2002). Formulaic language in acquisition and production: Implications for teaching. *TESL Canada Journal, 20*, 1-15.

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