The Development of Formulaic Language in L2

奥　脇　奈津美

OKUWAKI Natsumi

Abstract

The prevalence of formulaic language in language use has been acknowledged for long, and in the past decade research around formulaic language seems to have increased considerably. This article reviews the studies which examined the acquisition and processing of formulaic sequences in recent years. Based on the observation that formulaic sequences were less successfully and used less by nonnative speakers than native speakers, I discuss the difficulties second language (L2) learners face in the acquisition of formulaic sequences and the factors that could explain these challenges. Intervention studies on formulaic language are also reviewed, examining the effectiveness of teaching formulaic sequences to learners. I argue that much can be done to promote the acquisition of formulaic sequences, but one should keep in mind that this is an area where L2 learners make slower progress than other L2 properties and which shows an incremental process that typically requires multiple encounters with target items.

1 ．Introduction

Formulaic language forms a considerable part of native speakers' discourse and helps L2 learners to appear more native-like. As Foster (2001) acknowledges, "the fluency and familiarity of native-like language can be explained by the fact that it is generally not composed of novel combination of words but use a lot of prefabricated sequences shared by everyone in the speech community" (p. 77). In the mental lexicon, it is generally acknowledged that representations of individual words are stored and retrieved online, and there is growing agreement that the lexicon also contains formulaic language. Ample evidence of the highly formulaic nature of language has been provided by recent research in psycholinguistics, phraseology, corpus linguistics, first language (L1) acquisition, L2 acquisition, computational linguistics, and others (see Wray, 2002).

Although mastering formulaic language contributes to L2 learners' successful and fluent use of the language, it is an area where they only very slowly develop their competence (Howarth, 1998; Kupier, Colunbus, & Schmitt, 2009; Nesselhauf, 2003, 2005). Even highly advanced learners may find formulaicity particularly challenging and in active usage they may tend to rely more than native speakers on a relatively narrow range of high-frequency

word strings (Durrant & Schmitt, 2009).

This article reviews the studies which examined the acquisition and processing of formulaic sequences. There is plenty of evidence that L2 learners are different from native speakers in many respects. The question is what remains challenging for very advanced learners to master L2 formulaic language. I present factors that could explain unsuccessful performances by L2 learners. Intervention studies in recent years are then reviewed, examining the effectiveness of teaching L2 formulaic sequences to learners.

２．Formulaic sequences and L2 proficiency

According to Erman & Warren (2000), a formulaic sequence is, as a phenomenon of human language use, "a combination of at least two words favored by native speakers in preference to an alternative combination which could have been equivalent had there been no conventionalization" (p. 31), which intends to leave some flexibility and variation in the word combinations. Boers & Lindstromberg (2012) consider formulaic sequences as word strings that "have become conventionalized in a given language as attested by native-speaker judgement and/or corpus data" (p. 83). This statement contains two important aspects in the identification of formulaic language. Firstly, we must rely on native-speaker intuitions to judge if the target combination is formulaic or not, and secondly, that must be attested by corpus data which provides valuable insights otherwise unnoticeable. Boers & Lindstromberg thus propose that native norms and frequency (provided by corpus-linguistics techniques) are two options for the basis of determination of formulaic sequences.

Formulaic language composes a considerable portion of one's vocabulary and enables the comprehension and expression of messages that might otherwise fail to get across. Boers & Lindstromberg (2012) listed the same functions which formulaic sequences fulfil as single words in many ways. For example, word strings function as content words (e.g., collocations: *blow your nose*). They are also helpful for conveying an evaluative stance (e.g., exclamations: *no kidding*). Schmitt & Carter (2004) listed the communicative functions which formulaic sequences fulfil; expressing a message of idea (e.g., *The early bird gets the worm*), realising functions (e.g., *I'm just looking*), expressing social solidarity (e.g., *Yeah it is*), transacting specific information in a precise and understandable way (e.g., *Cleared for takeoff*), and singling discourse organisation (e.g., *on the other hand*).

Vocabulary size has been found to be a predictor of general proficiency in L2 (Staeher, 2009), and some researchers have examined if L2 knowledge of formulaic language also correlates to learners' levels of proficiency. For example, Khodadady & Shamsaee (2012) investigated the relationship between the use of formulaic language, speaking ability, and speech fluency and found two types of formulaic sequences which can predict learners' speech fluency. Keshavarz & Salimi (2007) found a correlation between learners' knowledge of collocation and general proficiency; Hsu & Chiu (2008) reported a correlation between

collocation test scores and speaking test scores; Dai & Ding (2010) reported strong associations between the number of formulaic sequences produced by L2 learners in L2 writing and their scores provided by independent assessors. Laufer & Waldman (2011) also found that very advanced L2 learners, particularly who have benefited from a constant exposure to L2 in the L2 environment, seemed to present knowledge of formulaic language that is similar to that of native speakers.

However, these results could only be the case in terms of width of vocabulary (the quantity of L2 words which the learner is familiar with, without specifying how familiar they are). When it comes to depth of vocabulary (how common it is and in what contexts it is most likely to occur), L2 learners' intuitions about the frequency and usage patterns of formulaic language are likely to be less reliable (Siyonava, 2013; Siyonava & Schmitt, 2008).

The use of learner corpora is another method which can explore the production of L2 formulaic language. Since learner corpus data are in electronic format, it is possible to use quantitative methods to see what sequences learners use and identify types of chunks in phraseology. There is a growing body of evidence in learner corpora to suggest that the use of formulaic sequences is related to proficiency in L2. It is difficult to establish a clear link between language proficiency and L2 competence of formulaic sequences, but in general, high levels of proficiency is usually characterised by a higher rate of use of formulaic language such as collocations and phrasal verbs (Hsu, 2007; Laufer & Waldman, 2011; Lewis, 2009).

Learners in very advanced stages may be able to master L2 formulaic forms that resemble to native speakers in some way, but the development is relatively slow. Laufer & Waldman (2011) found a statistically significant difference between the number of verb-noun collocation used by advanced and elementary learners. There were no significant differences, however, between intermediate and either of the other two proficiency groups, suggesting that L2 development in this area is possibly slow. Similarly, a longitudinal data of Chinese learners given in Li & Schmitt (2011) did not show a major change in the production of adjective-noun collocations over the course of a year, a period that may seem long enough for other linguistic elements to improve. Thus, L2 learners only slowly develop the knowledge of formulaic language.

Another thing which should be noted is that "more does not necessarily mean better" (Paquot & Granger, 2012, p. 138). Thewissen (2013) has shown that collocations remain an error-inducing property at advanced proficiency level. Nesselhauf (2005) provided an interesting result that L2 learners who had learned L2 for 10 to 17 years produced a similar proportion of erroneous collocations as learners who only learned L2 for 5 to 10 years. Similarly, Laufer & Waldman (2011) reported learners produced deviant collocations in about a third of all cases, regardless of the learner's proficiency level, and furthermore, advanced and intermediate learners were found to produce significantly more erroneous collocations than elementary learners when computed as a proportion of total number of words. There seems a reverse relationship between proficiency and the accuracy of collocations. This may

seem a paradox.

　The question then is how this paradox can be reconciled. It may be true that more advanced learners make a larger number of collocation errors than less advanced learners, but that does not mean they are less proficient in its use. Despite more deviant use of collocations by more advanced learners, as Thewissen (2008) suggests, this may be a sign of increased formulaic richness; learners at higher proficiency levels are continuously trying a new and wider range of formulaic language in L2, but their performance may be less successful. If this is the case, more advanced learners are more promising as they produce "a large number of near hits (e.g., *to turn over a new leave* instead of *to turn over a new leaf*)", while less advanced learners are more likely to produces "a higher number of wide-of-the-mark errors ... such as *the time work's duration* instead of *working hours*" (Paquot & Granger , 2012, p. 138), which do not exist in L2.

　Recurrence is a notion in phraseology similar to formulaic sequences discussed so far, but they are different in that recurrent strings, often called as lexical bundles, are the repetition of continuous word sequences of a given length retrieved by taking a corpus-driven approach (Paquot & Granger, 2012), such as (e.g., *in the context of*). Collocations, which are referred to as co-occurrences as a phraseological pattern, are found to be used less by L2 learners than native speakers, but L2 learners tend to use more lexical bundles in writing compared to native speakers (Ping, 2009). All in all, the use of lexical bundles seems to decrease as proficiency in the language increases.

　This section reviewed the studies which examined how knowledge of formulaicity in L2 is related to L2 general proficiency. Many of the studies considered the relationship between knowledge of formulaic sequences in terms of size. It turned out that formulaic language is very slow to develop, with less success compared to other elements of L2. Even though very advanced learners are able to use a comparable number of formulaic sequences as native speakers, when it comes to its accuracy, they do not outperform less advanced learners. In the following sections, I discuss factors which make formulaic language so challenging for L2 leaners even in their final stages of development.

３．Difficulties L2 learners have in acquiring formulaic language

　Although a number of recent studies investigated learners' sensitivity to formulaic sequences (Arnon & Snider, 2010; Hilpert, 2008; Kapatsinski & Radicke, 2009; Tremblay, Derwing, Libben, & Wesbury, 2011), many reports, exemplified in such articles as Granger & Menuier (2008), Menuier & Granger (2008), and Wood (2010), revealed that L2 learners tend to have small inventory of formulaic sequences. L2 learners often overuse (as in the case of recurrences) and underuse word strings (as in the case of co-occurrences). In addition, leaners often do not show sensitivity to register differences (as they tend to lack knowledge in terms of depth of formulaic sequences). In cases where advanced learners do seem to have

mastered an inventory similar to native speakers in terms of width, this does not necessarily mean they have acquired the repertoire in terms of depth. All in all, the studies suggest that the development of formulaic sequences tends to be slow, incomplete, and inaccurate.

Frequency, L1 influence and processing are commonly cited as being the key elements to which the difficulties L2 learners face in learning formulaic language is attributable. In the following sections, the effects of these factors on the acquisition of L2 formulaic sequences are discussed.

3.1 Frequency

In order to examine the effect of frequency, Sosa & MacFarlane (2002) focused on *of* in two-word combinations varying in frequency (*sort of* and *kind of*) and investigated reaction times of *of* using an auditory word-mentioning task with native speakers. They found that reaction times of *of* in high-frequency combinations were significantly slower than those in low-frequency combinations. This means, very frequent combinations are stored as wholes, so reaction times of *of* on its own took longer than *of* in low-frequency cases. Sosa & MacFarlane argued that for native speakers frequently-used multiword units became chunked and were stored as a unit in the mental lexicon. Similarly, Arnon & Snider (2010) focused on the comprehension of compositional four-word phrases (e.g., *don't have to worry*) with native speakers and compared reading times for phrases, which differed in phrasal frequency. They found that the more frequent phrases were processed faster than the less frequent phrases and concluded that native speakers store frequency information about multi-word phrases as well as words. These findings suggest that for native speakers the more frequent a formulaic sequence is, the more likely it is to be represented as a whole in memory, which would ease the burden on working memory and processing.

Only a few studies investigated the role of frequency in formulaic language with nonnative speakers. An eye-tracking study by Siyanova-Chanturia, Concklin, & van Heuven (2011) investigated processing of formulaic sequences by nonnative speakers. They focused on formulaic sequences imbedded in sentences that differed in phrasal frequency (e.g., *bride and groom* vs. *groom and bride*). Nonnative speakers across a range of proficiency read more frequent sequences faster than less frequent ones, indicating that phrases are entrenched in memory. In a similar study, Siyanova-Chanturia, & Spina (2013) investigated how L2 learners (advanced and intermediate) judge collocation frequency. They used 80 noun-adjective collocations, which differed in frequency, and found that L2 learners were remarkably native-like in their intuitions of collocation frequency.

If this is on the right track, it is plausible to argue that nonnative speakers are sensitive to frequency information about formulaic sequences as native speakers. However, the issue is not so straightforward. In terms of a difference between native speakers and L2 learners, Ellis, Simpson-Vlach, & Maynard (2008) reported that processing speed of formulaic sequences by native speakers is positively influenced by the strength of the association between the words, rather than by the frequency of occurrences of the sequences per se.

This is because native speakers have had plenty of exposure to even relative low-frequency multiword units, and the effect of frequency has reached a ceiling. Many L2 learners, on the other hand, may not be able to meet low-frequency sequences enough for them to be firmly established in memory; and instead, unlike native speakers L2 learners are strongly influenced by the frequency in processing of formulaic sequences. Durrant & Schmitt (2009) showed that nonnative writers underuse less frequent, strongly associated collocations which are probably highly salient for native speakers, but they rely on high-frequency collocations. As they note, "learners are quick to pick up highly frequent collocations, but less common strongly associated items (e.g., *densely populated, bated breath, preconceived notions*) take longer to acquire" (p. 175). This combination of overuse of high-frequency collocations and underuse of less-frequency collocations tallies with results from learner-corpus-based research, such as as Kaszubski (2000). These suggest that it is not that nonnative speakers lack formulaic language, but they do not have exposure to L2 sufficiently enough to meet less frequent items to be represented as prefabricated sequences in memory.

3.2　L1 influence

In learner corpus research, the impact of the learner's L1 is most discussed as a factor of the phraseological production of L2 learners. Corpus-based studies of collocation errors often tend to support the claim that a large portion of L2 learners' deviant collocations can be attributed to the influence of L1 (Paquot & Granger, 2012).

Nesselhauf (2005) investigated the deviant verb-noun collocations produced by German leaners and found that about 50 per cent of them were potentially influenced by L1 phraseology. She also observed that majority of verbs that were likely to have been influenced by the L1 were used in a figurative sense in verb-noun collocations, though words with figurative meanings have been considered less transferable than those with nonfigurative meanings (Kellerman, 1978). Nesselhauf suggested that collocations seem to be the type of combination that is most susceptible to transfer. In a similar study, Laufer & Waldman (2011) investigated the use of English verb-noun collocations in the writing of native speakers of Hebrew at three proficiency levels. They found that L1 influence was evident in about half of the erroneous combinations produced by Hebrew learners at three levels of proficiency, and errors persist even at advanced levels. Gilquin & Granger (2011) focused on phrasal verbs produced by L2 learners of English and found a general underuse of phrasal verbs by L2 learners with Romance L1 backgrounds. This is consistent with a number of experimental studies which have reported underuse of English phrasal verbs by learners whose L1s do not have this type of construction (Laufer & Eliasson, 1993). Paquot (2010) also showed that L1 influence was evident in French learners' overuse of a number of discourse-oriented formulaic language in learner writing, which do not conform to English academic writing conventions.

Thus recent learner-corpus-based studies have documented a number of transfer effects which have not previously been much illustrated, and that the influence of L1 on learners'

use of formulaic language may be much more substantial than was previously thought.

3.3   Processing and exposure to L2

If formulaic sequences are well-entrenched in a language user's long term memory (stored as a holistic unit in the mental lexicon), these sequences can be retrieved from memory as a whole without the need to compose the sequences word by word online, which would make processing easier and faster than non-formulaic language (Wray, 2002, Conlkin & Schmitt, 2012). A range of investigations confirm that formulaic sequences have a privileged processing status; reaction time tasks (e.g., Arnon & Snider, 2009; Durrant & Schmitt, 2010; Ellis et al., 2008; Jiang & Nekrasova, 2007), self-paced reading (e.g., Conklin & Schmitt, 2008; Tremblay et al., 2011), silent reading as evidenced by eye-tracking (Columbus, 2010; Siyanova-Chanturia et al., 2011), and recall (Tremblay & Baayen, 2010). Thus, experiments with native speakers have shown that formulaic strings of various types are processed significantly faster than non-formulaic strings[1].

However, things are not straightforward for nonnative speakers. Some evidence shows that L2 learners less rely on phrases and sequences, focusing more on individual words (Foster, 2001) and do not fully enjoy the processing advantage afforded by formulaic sequences (Columbus, 2010; Conklin & Schmitt, 2008). Many studies found processing advantages by nonnative speakers (e.g., Siyanova-Chanturia et al., 2011), but often the participants were advanced learners, practically ones who have had experience in an immersion context.

In order for an L2 learner to receive the processing advantage afforded by a formulaic sequence, a learner needs to be familiar with it (Boers & Lindstromberg, 2012); that is, a learner must be exposed to the L2 required to reach such a level of familiarity with formulaic sequences[2]. Highly frequent phrases (e.g., *how are you?, it's lunch time*) are readily learnable by dint of being highly frequent[3], but those which do not frequently appear in L2 are not readily learnable and tend to be nontransparent in their interpretation. As Ellis (2008) and Ellis et al. (2008) suggest, learners require considerable language exposure before they encounter formulaic sequences, but many L2 learners may not be able to encounter low-frequency sequences sufficiently enough for them to be firmly established in memory. This is one of the reasons L2 learners are reported not to master native-like formulaic language.

4．Intervention studies

Having discussed the difficulties L2 learners have in acquiring formulaic sequences and the factors that could explain these challenges, this section considers teaching treatments that would improve phrasal competence of L2 learners. I focus on the effects of awareness-raising and input in L2 teaching, as these are the topics most widely investigated in the field. As has been discussed in previous sections, L2 learners can greatly benefit from mastering formulaic language, but in order to reap the full benefits of formulaic language, they must

achieve both width and depth of knowledge. In what follows, I discuss the studies which investigated if pedagogical treatments, such as awareness-raising and input flood, help learners meet that challenge.

Boers, Eyckmans, Kappel, Stengers, & Demecheleer (2006) investigated the effect of awareness-raising through text chunking. They found that the advanced L2 learners who engaged in text chunking activity in the course of a school year produced significantly more formulaic sequences in the narratives than those who engaged in other activities with the same text, when they were asked to orally retell the content of a new English text. However, it was also found that the learners in the text chunking group often recycled word strings verbatim from the new text. Stengers, Boers, Eyckmans, & Housen (2010) repeated the experiment using an input text in the learners' L1. They found no evidence of any differential learning of formulaic sequences between the groups which had engaged in text chunking and those which had not.

There are other ways to raise awareness to formulaic language; highlighting sequences in texts, discussing their usefulness for writing, and recycling the encountered sequences in writing tasks. The effects of these techniques were investigated by Jones & Haywood (2004). The L2 participants in their study who had engaged in these activities showed evidence of a greater awareness of formulaic language than a control group. They were able to underline word strings in a new text when asked to underline vocabulary which they would advise other learners to learn. When it comes to the retention of formulaic sequences, however, they did not produce more word strings in writing essays than the control group. As Boers & Lindstromberg (2012) suggest, the lack of surplus learning in awareness-raising activities is possibly attributable to the number of encounters with formulaic sequences; "taking notice of a given word sequence just once or twice is hardly enough to leave durable memory traces" (p. 89), which is what has repeatedly been found in the studies concerning single word learning (for a review, see Laufer, 2005).

Peters (2012) investigated the effect on retention of typographic enhancement (underlined and in bold font) and of glossing on formulaic sequences. In her study, L2 learners in the experimental group read a glossed German text paying attention to both formulaic sequences and single words and wrote down unfamiliar words, while the control group was instructed to pay attention to unfamiliar vocabulary in general. All the participants knew that a vocabulary posttest would follow the reading task. The result that typographic salience had an effect on participants' recall scores led Peters to suggest that typographic salience facilitates L2 learners' noticing and learning of unknown lexical items, and learning of formulaic sequences in particular. However, as Boers & Lindstromberg (2012) point out, it was conceivable that the learners made more of an effort to remember the highlighted items in the text than those which were not highlighted because they knew a vocabulary posttest would follow and the former were the most likely targets for the test. Peters (2012) would be consistent with the thesis that L2 learners are not likely to autonomously recognise and attend to formulaic sequences unless they are told to do so. In other words, formulaic

language attracts very little attention from L2 learners on its own.

Now consider the role of L2 input in the acquisition of formulaic language. As Meunier (2012) states, "in instructed contexts, authentic documents, textbooks, and teacher-talk are the key source of native(-like) input" (p. 113). Indeed a main source of target language input learners obtain is from textbooks teachers use in a classroom. However, there are discrepancies between written input included in teaching materials and corpus-based language (Meunier & Gouverneur, 2009). The lack of authenticity in textbooks does not help learners acquire formulaic language, and this in turn lessens the possibilities that future language teachers would become very fluent in L2 in countries where L2 teachers are mostly nonnative.

In addition, incidental learning of formulaic sequence is not likely to occur with a few encounters with the target phrase. Webb, Newton, & Chang (2012) found that it took 15 encounters for L2 learners to recall collocations correctly only half of the time tested. The study certainly shows that recurring collocations can be retained from reading, but it does not guarantee the rapid acquisition of collocation as a by-product of reading.

The next question in the intervention studies is whether L2 learners are able to retain formulaic sequences in memory if they are encouraged to try to remember the items. In Webb & Kagimoto (2011), L2 learners were told to learn unfamiliar adjective-noun collocations with the help of L1 translation in three minutes' study time. An immediate recall test followed. They reported L2 learners obtained the highest score when several collocations in the same set had the same adjective (e.g., *deep respect, deep sleep,* and *deep voice*), whereas the test score was the poorest when collocations in the same set consisted of different adjectives, including semantically related adjectives (e.g., *narrow escape* and *slim chance*). The negative effect of semantically related words in target items to be learned has also been found in the study of single words (Waring, 1997). The study by Webb & Kagimoto confirmed that the same is true for formulaic sequences.

In this section, I reviewed several intervention studies on formulaic sequences in L2, the area which has been rigorously studied in recent years. The rapid increase in the number of studies in formulaic sequences indicates our knowledge and understanding of the formulaic nature of language are increasingly refined, which clearly impacted L2 teaching. The availability of digital tools and corpora make L2 learners and teachers more accessible to the world of formulicity, which has been demonstrated to be indispensable for a competent language user.

5．Discussion

The prevalence of formulaic language in language use has been acknowledged for long, and in the past decade research activity around formulaic language seems to have increased considerably. The paradox in the acquisition of formulaic language presented by Ellis (2012)

is as follows: "Formulas seed language acquisition and yet learners typically do not achieve native-like formulaicity" (p. 18). Put it another way, there is "a missing link regarding why learners do not feel more empowered to harvest L2 input in larger chunks in the pursuit of painless routes to effective communication" (Wray, 2012, p. 236).

What is so challenging for L2 learners in the acquisition of formulaic language? Firstly, some formulaic sequences are easily learnable as it is highly frequent and prototypical in their functionality, which Ellis (2012) called "phrasal teddy bears" (p. 29). Even advanced L2 learners often overuse high-frequency basic words, avoiding going for less frequent but more appropriate words and making a word selection error. This leads to a consistent difficulty and unsuccessful performance in low frequent, rare, and nontransparent formulaic sequences, which are required for the native-like knowledge. Therefore, teaching time should be devoted to medium- and low-frequency formulaic sequences, as high-frequency ones stand the best chance of being picked up by learners implicitly (Boers, Deconinck, & Lindstromberg, 2010).

Secondly, L2 learners have difficulties deploying the knowledge of formulaic language and tend to resort to L1-based nonnative like items, and the impact of L1 typically continues into later stages of development. The impact of learner's L1 has been mostly discussed in learner corpus research, which revealed a number of transfer effects on learners' use of formulaic sequences. L1 impact has been explored long in L2 research, but learner corpus research revealed that it may be more substantial than was previously thought.

Processing formulaic language is also a difficult task for L2 learners. For some L2 learners it is possible to hold processing advantages similar to native speakers, but this may only be possible after the multiple encounters with the same form in input. The problem for L2 learners, however, is that they need to have considerable language experience in order to encounter formulaic sequences multiple times, which may not be always possible.

Next, consider width and depth of knowledge of L2 formulaic sequences. Depth of knowledge concerns such information as to commonness of a given formulaic sequences and contexts in which it is most likely to occur. As Boers & Lindstromberg (2012) suggest, deep knowledge of L2 formulaic language leads to processing advantages; that is, when a formulaic sequence is deep entrenched in a language user's long term memory, he/she can use the sequence automatically, and the user is able to predict other constituents comprising the text, which provides processing advantages.

Finally, it may be true that L2 learners in very advanced stages will be able to attain native-like knowledge of formulaic sequences in terms of size, when it comes to accuracy, however, they tend to use as many deviant forms as less proficient learners. It takes long to acquire accuracy of formulaic sequences in L2.

These findings suggest that frequency seems to be a key factor determining the success of formulaic sequences, and only proficient L2 learners who have had the time to amass sufficient language exposure over a long period begin processing formulaic sequences as quickly and automatically as native speakers. Considering adult L2 learners cannot avoid

analysing L2 in the process (Wray, 2002), it is worthwhile to provide pedagogical treatments involving cognitive engagement with formulaic sequences to channel that analytical processing along ways known to enhance retention.

6．Conclusion

　This article discussed the difficulties L2 learners have in acquiring formulaic sequences and the factors which could explain these challenges. I presented the differences observed between native and nonnative speakers in the representation and storage of prefabricated sequences in the mental lexicon, and the differences in the processing advantages of formulaic sequences. Based on the observation that formulaic sequences were unsuccessfully and not frequently used by nonnative speakers, I proposed much can be done to promote the acquisition of formulaic sequences in L2 teaching, keeping in mind that this is an area where L2 learners make slower progress than other L2 properties. They will show an incremental progress that typically requires multiple encounters with target items. As Ellis (2012) remarks, in the research of formulaic sequences, much remains to be done, needing triangulation across areas of research such as L1 acquisition, L2 acquisition, corpus linguistics, and psycholinguistics.

Notes

1　One of the complications in the processing studies is the great diversity in kinds of sequences, some of which intuitively seem more likely to be processed as unitary lexical units (e.g., idioms and strong collocations) than others. As Columbus (2010) showed, in the eye-tracking measures of speed of processing during silent reading, not all kinds of word strings offer the same degree of processing ease.

2　If language users are required to reach a certain level of familiarity with formulaic sequences so that the sequences are entrenched in memory, the differences of performance found between native and non-native speakers, and the differences among proficiency levels, are in some way considered to be revealing the amount of exposure to the language, especially in the case of not very frequent sequences. Some studies indeed indicated the impact of immersion-based L2 exposure on phraseological competence. An immersion study in Groom (2009) found a high correlation between the number of statistical co-occurrences and the amount of time spent in an English-speaking country. Similarly, Waibel (2008) found the length of time abroad is a significant predictor of the number of phrasal verb used.

3　Ellis (2012) called these familiar phrases "phrasal teddy bears" by analogy with "lexical teddy bears" by Hasselgren (1994) and "collocational teddy bears" by Nesselhauf

(2005).

Acknowledgement

References

Arnon, I., & Snider, N. (2010). More than words: Frequency effects for multi-word phrases. *Journal of Memory and Language, 62* (1), 67-82.

Boers, F., & Lindstromberg, S. (2009). *Optimizing a lexical approach to instructed second language acquisition.* Basingstroke: Palgrave Macmillan.

Boers, F., & Lindstromberg, S. (2012). Experimental and Intervention studies on formulaic sequences in a second language. *Annual Review of Applied Linguistics, 32,* 83-110.

Boers, F., Deconinck, J., & Lindstromberg, S. (2010). Choosing motivated chunks for teaching. In S. De Knop, F. Boers, & T. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 239-256). Berlin/New York: Mouton de Gruyter.

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,* 245-261.

Columbus, G. (2010). Processing MWUs: Are MWU subtypes psychologically real? In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 194-210). New York: Continuum.

Conklin, K., & Schmitt, N. (2008). Formulaic sequences: Are they processed more quickly than nonformulaic language by native and nonnative speakers? *Applied Linguistics, 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.

Ellis, N. C. (2008). Optimizing the input: Frequency and sampling in usage-based and form-focused learning. In M. H. Long & C. Doughty (Eds.), *Handbook of second and foreign language teaching.* Oxford: Blackwell.

Ellis, N. C. (2012). Formulaic language and second language acquisition: Zip and the phrasal teddy bear. *Annual Review of Applied Linguistics, 32,* 17-44.

Ellis, N.C., Simpson-V. R., & Maynard, C. (2008). Formulaic language in native and second-language speakers: psycholinguistics, corpus linguistics, and TESOL. *Tesol Quarterly, 42* (3), 375-396.

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

Foresberg, F. (2010). Using conventional sequences in L2 French. *International Review of Applied Linguistics in Language Teaching, 48,* 25-51.

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. 75-93). Harlow: Longman.

Gilquin, G., & Gaëtanelle, S. (2011). From EFL to ESL: Evidence from the international corpus of learner English. In M. Hundt & D. Mukherjee (Eds.), *Exploring second-language varieties of English and learner Englishes: Bridging a paradigm gap* (pp. 55-78). Amsterdam: John Benjamins.

Granger, S., & Menuier, F. (Eds.). (2008). *Phraseology: An interdisciplinary perspective.* Amsterdam: John Benjamins.

Groom, N. (2009). Effects of second language immersion on second language collocational development. In A. Barfield & H. Gyllstad (Eds.), *Researching collocations in another language* (pp. 21-33). Basingstoke: Palgrave Macmillan.

Hilpert, M. (2008). New evidence against the modularity of grammar: Constructions, collocations, and speech perception. *Cognitive Linguistics, 19,* 391-411.

Hoey, M. (2005). *Lexical priming: A new theory of words and language.* London: Routledge.

Howarth, P. (1998). The phraseology of learners' academic writing. In A. Cowie (Ed.), *Phraseology: Theory, Analysis and Applications* (pp. 155-175). Oxford: Oxford University Press.

Hsu, J. (2007). Lexical collocations and their relation to the online writing of Taiwanese college English majors and non-English majors. *Electronic Journal of Foreign Language Teaching, 4,* 192-209.

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

Kapatsinski, V., & Radicke, J. (2009). Frequency and the emergence of prefabs: Evidence from monitoring. In R. Corrigan, E. Moravcsik, H. Ouali, & K. Wheatley (Eds.), *Formulaic Language. Vol. II. Acquisition, loss, psychological reality, functional explanations* (pp. 499-520). Amsterdam: John Benjamins.

Kaszubski, P. (2000). *Selected aspects of lexicon, phraseology and style in the writing of Polish advanced learners of English: a contrastive, corpus-based approach.* Unpublished doctoral dissertation, Adam Mickiewicz University, Poznań, Poland.

Kellerman, E. (1978). Giving learners a break: Native language intuitions as a source of predictions about transferability. *Working papers on bilingualism, 15,* 59-92.

Khodadady, E., & Shamsaee, S. (2012). Formulaic sequences and their relationship with speaking and listening abilities. *English Language Teaching, 5* (2), 39-49.

Keshavarz, M. H., & Salimi, H. (2007). Collocational competence and cloze test performance:

A study of Iranian EFL learners. *International Journal of Applied Linguistics, 17,* 81-92.

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

Laufer, B. (2005). Focus on form in second language vocabulary acquisition. In S. Foster-Cohen (Ed.), *EUROSLA Yearbook 5* (pp. 223-250). Amsterdam: John Benjamnins.

Laufer, B., & Eliasson, S. (1993). What causes avoidance in second language learning: L1/L2 difference, L1/L2 similarity, or L2 complexity? *Studies in Second Language Acquisition, 15,* 35-48.

Laufer, B., & Waldman, T. (2011). Verb-noun collocations in second language writing: a corpus analysis of learners' English. *Language Learning, 61* (2), 647-672

Lewis, M. (2009). The idiom principle in L2 English: Assessing elusive formulaic sequences as indicators of idiomaticity, fluency, and proficiency. Saarbrucken: VDM Verlag Dr. Müller.

Meunier, F. (2012). Formulaic language and language teaching. *Annual Review of Applied Linguistics, 32,* 111-129.

Meunier, F., & Gouverneur, C. (2009). New types of corpora for new educational challenges: collecting, annotating and exploiting a corpus of textbook material. In K. Aijmer (Ed.), *Corpora and Language Teaching* (pp. 179-201). Amsterdam: Benjamins.

Menuier, F., & Granger, S. (2008). *Phraseology in foreign language learning and teaching.* Amsterdam: John Benjamins.

Nekrasova, T. M. (2009). English L1 and L2 speakers' knowledge of lexical bundles. *Language Learning, 59,* 647-686.

Nesselhauf, N. (2003). The use of collocations by advanced learners of English and some implication for teaching. *Applied Linguistics, 24* (2), 223-242.

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

Oppenheim, N. (2000). The importance of recurrent sequences for non-native speaker fluency and cognition. In H. Riggenbach (Ed.), *Perspectives on fluency* (pp. 220-240). Ann Arbor: University of Michigan Press.

Paquot, M., & Granger, S. (2012). Formulaic language in learner corpora. *Annual Review of Applied Linguistics, 32,* 130-149.

Pawley, A., & Syder, F. (1983). Two puzzles for linguistic theory: Native-like selection and native-like fluency. In J. Richards & R. Schmidt (Eds.). *Language and communication* (pp. 191-226). London: Longman,.

Peters, E. (2012). Learning German formulaic sequences: The effect of two attention-drawing techniques. *Language Learning Journal, 40* (1), 65-79.

Schmitt, N., & Carter, R. (2004). Formulaic sequences in action: An introduction. In N. Schmitt (Ed.), *Formulaic sequences* (pp. 1-22). Amsterdam: John Benjamins.

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

Siyanova-Chanturia, A., Conklin, K., & van Heuven, W. (2011). Seeing a phrase 'time and

again' matters: The role of phrasal frequency in the processing of multi-word sequences. *Journal of Experimental Psychology: Language, Memory, and Cognition, 37* (3), 776-784.

Siyanova-Chanturia, A., & Spina, S. (2013). Native and non-native speaker intuition of collocation frequency: A case of Italian. Paper presented at Eurosla 2013, Amsterdam, the Netherland.

Staeher, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition, 31,* 577-607.

Stengers, H., Boers, F., Eyckmans, F., & Housen, A. (2010). Does chunking foster chunk uptake? In S. De Knop, F. Boers, & A. De Rycker (Eds.), *Fostering language teaching efficiency through cognitive linguistics* (pp. 99-117). Berlin/New York: Mouton de Gruyter.

Thewissen, J. (2013). Capturing L2 accuracy developmental patterns: Insights from an error-tagged EFL learner corpus. *Modern Language Journal, 97,* 77-101.

Tremblay, A., & Baayen, H. (2010) Holistic processing of regular four-word sequences: A behavioral and ERP study of the effects of structure, frequency, and probability on immediate free recall. In D. Wood (Ed.), *Perspectives on formulaic language: Acquisition and communication* (pp. 151-173). London: Continuum.

Tremblay, A., Derwing, B., Libben, G., & Wesbury, C. (2011). Processing advantages of lexical bundles: Evidence from self-paced reading and sentence recall tasks. *Language Learning, 61,* 569-613.

Waibel, B. (2008). *Phrasal verbs: German and Italian learners of English compared.* Saarbrucken, VDM Verlag Dr. Müller.

Waring, R. (1997). The negative effects of learning words in semantic sets: A replication. *System, 25,* 261-274.

Webb, S., & Kagimoto, F. (2011). Learning collocations: Do the number of collocates, position of the node word, and synonymy affect learning? *Applied Linguistics, 32,* 259-276.

Webb, S., Newton, J., & Chang, A. (2013). Incidental learning of collocation. *Language Learning, 63* (1), 91-120.

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

Wray, A. (2012). What do we (think we) know about formulaic language? An evaluation of the current state of play. *Annual Review of Applied Linguistics, 32,* 231-254.