

**“COMFORTABLE FOSSILIZATION” CHINESE EFL LEARNER’S
ACQUISITION AND USE OF FORMULAIC SEQUENCES IN L2 WRITING**

Zhisheng Wen

Macao Polytechnic Institute, MACAU

RESUMO: O objetivo do presente estudo foi investigar o conhecimento e uso de sequências formulaicas na produção de textos em L2 por chineses aprendizes de inglês como língua estrangeira. A partir das definições de Alison Wray (Wray, 2000, p. 465), sequências formulaicas foram operacionalizadas como "expressões idiomáticas, colocações e enquadres sentenciais (incluindo conectivos)". Amostras autênticas produzidas por 16 aprendizes de nível avançado em inglês como L2 foram coletadas e analisadas com o objetivo de avaliar o conhecimento e uso de sequências formulaicas. Os resultados mostram que o conhecimento e uso de sequências formulaicas na escrita destes alunos foram fortemente afetados pela língua materna. Argumentamos que a dependência a sequências formulaicas incorretas, por parte do aprendiz, pode levar ao que Foster e Skehan (2001) chamam de "fossilização confortável". Discutimos também as implicações do presente estudo para o ensino da produção textual em inglês como língua estrangeira.

Palavras-chave: sequências formulaicas; agrupamentos lexicais; enquadres sentenciais; transferência da L1; fossilização confortável; escrita em L2.

ABSTRACT: The present study set out to investigate the knowledge and use of formulaic sequences in Chinese EFL learner’s L2 writing. Adopting Alison Wray’s definition (Wray, 2000, p.465), formulaic sequence was operationalized as “idioms, collocations and sentence frames (including connectives)”. Authentic samples written by 16 advanced EFL writers in China were collected and analyzed with a view to probing into their knowledge and use of formulaic sequences. The results indicated that a strong influence of these EFL learners’ native language (L1) affected their knowledge and use of formulaic sequences in L2 writing. The study argues that such over-reliance on these incorrect formulaic sequences is likely to become what Skehan and Foster (2001) have called “comfortable fossilization”. Implications of these results for EFL writing instructions are discussed.

Keywords: Formulaic sequences, lexical bundles, sentence frames, L1 transfer, comfortable fossilization, L2 writing

1. Introduction

In the field of second language acquisition (SLA), the term “formulaic sequence” is usually defined as:

‘a sequence, continuous or discontinuous, of words or other meaning elements, which is, or appears to be, prefabricated: that

is, stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar'. (Wray, 2000, p.465)

In many ways, Alison Wray's definition of formulaic sequences highlights the two key linguistic and psycholinguistic features of formulaic sequences (Schmitt and Carter 2004, p.3): a) they are sequences of lexis and b) these sequences are handled, or appear to be handled by the mind at some level of representation as wholes.

Since the mid-1980s, studies of formulaic sequences have come to the center stage of interests among many English as a Foreign Language (EFL) practitioners who are partly inspired by the concept of lexico-grammar of M. A. K. Halliday, especially after the emerging of pragmatics as an independent discipline (Granger, 1998). Furthermore, the rapid developments of corpus linguistics in recent years have not only greatly facilitated and modernized the research technique, but have also added great momentum to this line of investigation. So far, if we put into perspective these research efforts directed at the acquisition and use of formulaic sequences, the following major conclusions can be safely drawn. First, formulaic sequences are indeed ubiquitous in language use (Nattinger and DeCarrico, 1992; Erman and Warren, 2000; Foster, 2001; Pawley and Syder, 1983; Moon, 1998; Schmitt and Carter, 2004). Second, formulaic sequences serve many functions in language learning and so they do seem to play a very important role in the learning process (Weinert, 1995; Wray, 2000 and 2002; Wray and Perkins, 2000). Third, most major attempts to foreground formulaic sequences in teaching syllabuses (Willis, 1990; Nattinger and DeCarrico, 1992; Lewis, 1993), though succeeded in raising awareness among educationalists and applied linguists, have not made their way into the mainstream EFL curriculum (Wray, 2000; though see

Gatbonton and Segalowitz, 2005 for a new attempt). Fourth, most previous research into this phenomenon has been basically concerned with L1 learning, and consequently research into L2 formulaic sequences has been limited in number (Schmitt and Carter, 2004). Fifth, such restricted number of previous L2 research has been for the most part concerned with oral production, and empirical research into L2 writer's knowledge and use of formulaic sequences has been far and few between.

In terms of L2 written performance, as Cowie (1992) bluntly pointed out, an appropriate use of formulaic sequences (idiomaticity) is essential for L2 writing, "It is impossible to perform at a level acceptable to native users, in writing or in speech, without controlling an appropriate range of multiword units (p.10)". The use of formulaic sequences enables L2 writers, for example, to express technical ideas economically, to signal stages in their discourse and to display the necessary level of formality, while the absence of such features may result in a student's writing being judged as inadequate (Jones and Haywood, 2004). Moreover, some previous research has shown that L2 writers show little awareness of the speech vs. writing distinction and tend to include speech-like features in their writing (Granger and Rayson, 1998; Virtanen, 1998; Altenberg and Tapper, 1998). Other researchers have also found that L2 writers tend to either overuse, underuse or misuse the formulaic sequences in their writing (Cortes, 2004; Granger, 1998; De Cock, 2000; Milton and Freeman, 1996). Despite all these, empirical research into how EFL/L2 writers acquire or use formulaic sequences is relatively scarce. It is little wonder why Granger made the following claims that "We possess insufficient knowledge to decide what role they (formulaic sequences) should play in L2 teaching", simply because "when we consider how little we know about them, how they are acquired, what production difficulties they cause,

and how L1 and L2 prefabs interact, this is quite alarming (1998, p.159).” Hence the urgent need for more empirical work into this area.

2. The Present Study

2.1 Research design and methodology

In the present research, formulaic sequences were operationalized as including lexical bundles and discourse frames (sentence builders and connectives) even at the risk of simplicity. In order to probe Chinese EFL learners’ acquisition and use of these formulaic sequences in their written production, I collected authentic compositions written by college students (N=38) from a southern key university in Mainland China. To be more specific, the data contained two sets of compositions sampled from an intact class (16 out of the 38 students) in their final exams. More importantly, the data represents a contrast between the eight students at the lower quartile of the class (in terms of their English grades) and those eight others at the high quartile. Among these compositions, one set of data was collected from the participants’ final exam papers at the end of their first academic year, while the other set was collected from their final exams at the second academic year. The motivation for such a research design is to try to open a more fine-grained window on the underlying profile of EFL learners’ acquisition and use of formulaic sequences in their written performance, a picture that may be blurred by enormously huge learner corpora.

2.2 Results and Discussion

2.2.1 Lexical bundle: the case of “*instead of*” vs. “*take the place of/replace*”

A cursory analysis of the data already reveals one very striking use of the lexical bundle of “*instead of*” by many Chinese EFL learners in the first set of written data

when what they actually mean is “*to take the place of*” or simply “*to replace*”. For example:

(1). In their opinions, books made from wood is so important that no other kinds of books can *instead*. (Learner 2/Group 1)

(2). Recently, some people say that electronic books will *instead* paper books because of the development of computers. (Learner 5/1)

(3). But others consider that the paper-books can't be *instead of*. (Learner 8/Group1)

(4). I think electronic books at best supplement but not *instead* paper-made books as the carrier of knowledge, information, news and so on. (Learner2/Group 1)

Assuming that this phenomenon is pervasive not just for this group of learners, we need to go further so as to have a better idea how the other group of learners managed the meaning of “*to replace*”. Therefore, I carefully analyzed each composition in the first data set by the two groups and tried to identify each instance when they mentioned (meant) this. Most likely (though not always), the structure takes the form of “*NP1* (e.g. electronic books) *instead NP2* (e.g. paper books)”. Table 1 shows some examples of their actual use of all lexical bundles to mean “*to replace*” (including the word “replace” itself):

Table 1

Formulaic sequences meaning “to replace”

No	Group 1	Correct	Group 2	Correct
1	will take the place of (C) are take over (W) it isn't take the place of by (W)	1/3	will replace (C)	1/1
2	will take place (W) can instead (W)	0/2	will replace (C) can't be replaced (C) (to)supplement rather replace (C)	3/3
3	will take the place of (C) can't be taken place of (W) will take the place of (C)	2/3	will replace the other (C)	1/1
4	will take place of (W) won't be took place forever (W) can take place (W)	0/3	will replace (C) won't be replaced (C)	2/2
5	NP1 will instead NP2 (W)	0/1	would replace (C) were unreplacable (W) would replace (C)	3/3
6	will take the place of (C) can take over (W) can supplement but not instead (W)	1/3	take the place of (C)	1/1
7	would take the place of (C) won't be replaced (C)	2/2	will take the place of (C) can take the place of (C) to take the place of (C) will not take the place of (C)	4/4
8	can't be instead of (W) will be taken place (W)	0/2	will be replaced by (C)	1/1
Total	18	6/19	16	15/16

It is quite clear from Table 1 then, that the lower proficiency group suffered great difficulties with the lexical bundles meaning “*to replace*”. Out of the 19 instances, they were wrong on 13 occasions, which simply amounts to almost 70%, while the higher proficiency group (Group 2) had no problem with this. It seems to suggest that, in order to express the meaning of “*to replace*”, learners of lower proficiency either handle it casually, assuming that it has something to do with the verb phrase “*take + something*”, or simply resort to their L1 lexicon, which produces “*instead of*” immediately. One way or another, learners relied on these formulaic sequences initially as a quick means to be communicative (Schmitt and Carter, 2004), albeit in a limited and sometimes even in the wrong way. In terms of the causes underlying learners’ quick use of “*instead of*” in this particular case, it can be conceived that there are at least four factors that might contribute to this phenomenon. First, there is the L1 negative transfer which results in this word-class confusion (‘*instead of*’ is misused as a verb), simply because “*instead of*” in Chinese (*dai4qi4*) carries exactly the same meaning as “*to replace*”. Besides these, two other factors might also play a part: (a) learners had already acquired the phrase ‘*instead of*’ long ago (as early as when they were in primary school), that is the one that is readily available and (b) this phrase is frequently encountered which makes it easily accessible from the long-term memory.

Then, there is also the effect of individual differences that might also play a part. As we can see that, different learners got wrong in different ways, which might reflect the different acquisition process among them. To that regard, how individual differences affect the acquisition of formulaic sequences deserves further investigation. In terms of the theoretical account, there is still another issue that remains unresolved. That is, how can such errors be classified. Should they be classified as grammatical errors or alternatively as errors of formulaic sequences? It must be admitted that to categorize a

formulaic sequence has never been easy. As a matter of fact, it has been one of the thorniest issues inflicting applied linguists interested in this line of investigation (see Hunston, 2002; Read and Nation, 2004 for more details).

2.2.2 Formulae: sentence builders and connectives

This part is based on the analysis of two discourse frames: *sentence builders* and *connectives*. *Sentence builders* are those formulaic sequences which “function as macro-organizers in the text” (Granger, 1998, p. 154). One obvious example is the use of “*With + NP*” acting as a subordinate clause (for example, “With the development of science and technology, ...”). On the other hand, *connectives* are mainly those sequences that serve to connect sentences logically (e.g. firstly, secondly, on the one hand, on the other hand, etc.) and also those sequences that serve to express the writers’ viewpoints (in my opinion, in my view etc.). It must be admitted that the use of all these phrases and frames could be viewed as instances of what Dechert (1984: 227, cited in Granger, 1998:156) has called ‘islands of reliability’, i.e. “prefabricated formulaic stretches of verbal behavior whose linguistic and paralinguistic form and function need not be ‘worked upon’.” Based both on anecdotes and my own teaching experience, a typical example is the ‘*With + NP*’ structure and the ‘*As + NP+V-ing*’ structure that seem to be so favored by Chinese EFL learners to begin an essay (some of them used the two so often that they sometimes even got confused as to which was which). Participants in the first group (the lower proficiency level group) are no exception (though learner 5 did manage to avoid it!). The following are sentences directly copied from their writing (with grammatical mistakes from word forms unmarked):

- (1) With the developing of the computer technology, the computer is using in various of regions. (learner 1/Group 1)

- (2) With the development of science and technology, the technology of computer is developing rapidly. (learner 2/Group 1)
- (3) With the develop of the computers, some people think that the electronic books will take the place of the paper books. (learner 3/Group 1)
- (4) As the improvement of computer technology, electronic books will be use. (learner 4/Group 1)
- (5) With the development of computer technology, some people believe that electronic books will take the place of paper-made books. (learner 6/Group 1)
- (6) With the development of computer technology, some people thought that electronic books would take the place of nowadays's common books made by paper. (learner 7/Group 1)
- (7) As the development of computer going on, some people believe that electronic books will take the place of paper-books. (learner 8/Group 1)

Given the overwhelming number of the structure of “*With + NP*” and/or the “*As + NP + V-ing*” manipulated by participants of the study, it can be conceived that L1 negative transfer is felt at two levels for the lower proficiency EFL learners. First, using a structure of “*With+NP*” and/or “*As+NP+V-ing*” to begin a composition is absolutely not an uncommon phenomenon when Chinese EFL learners are writing in their L1 (i.e. Chinese). Second, the two words ‘with’ and ‘as’ carry more or less the same meaning in Chinese (*sui2zhe3*), which adds further confusion to these learners when they are applying this golden rule to begin an English essay. As for Group 2, the somewhat higher level group, the picture seems to change quite significantly, but still two out of the eight students used these structures to begin their composition:

- (1) As the development of the computer technology and PC, in the eyes of some people, the electronic books will replace the paper books. (learner 2/Group 2)
- (2) As the technology is developing with each passing day, aside from the paper making books, the electronic books have come into existence in our daily life. (learner 3/Group 2)

Such an effect of L1 negative transfer is also felt in these learners' use of sequences expressing viewpoints. A typical case would be the use of “*agree+something*”, simply because in Chinese, the word ‘agree’ (*tong2yi4*) does not need an obligatory preposition, you just “**agree something*” or “**agree someone's viewpoint*”. So, I identified all the instances for expressing viewpoints by these two groups of learners, both from the first data set and the second set.

Table 2

Formulaic sequences expressing viewpoints

No	Group 1	Correct	Group 2	Correct
1	I agreed the second view (W)	1/3	On contrary to the view, (W) In my opinion, (C) So from my standpoint (C?) So, I stand for the young to...(W)	3/6
			As I see it, (C) To the contrary, (W?) From my perspective, (C)	
2	In the eye of people (W)	2/6	In their opinions, (W)	2/3

	Someone didn't agree it. (W)		I think, ...(C)
	In their opinions (W)		
	I think it is not useful to...(C)		So from my standpoint, (C)
	Personally, (W)		
	So from my standpoint, (C?)		
3	Some people think that...(C)	$\frac{3}{4}$	But in my opinion, (C)
	In my opinion, (C)		And this is my opinion, do you agree with me? (C)
	I don't think...(C)		
	Contrary to it, (W?)		Different people has different perception about...(W)
			As I see it, (C)
			I don't think...(C)
4	Some people think that...(C)	$\frac{3}{4}$	There is no consensus of opinions of the topic...(C)
	But I don't think that...(C)		Some people hold that...(C)
			As to me, (C?)
	Some people say that..., others believe that... (C)		
	In my opinion, both ideas of them...(W)		Some people hold that...(C)
			Other people argue that...(C)
5	But I don't agree these opinion. (W)	$\frac{1}{2}$	Some said that,...(C)
			Some others disagreed. (C)
	In my opinion, (C)		In my opinion, (C)
			The question differs from one to another. (W)
6	What about my opinion? I think...(C)	$\frac{3}{3}$	Some people consider that...(C)
			Others believe that...(C)
	Some people say that...(C)		In my eyes, (C?)
	But I don't agree with the point of view. I believe that... (C)		As you see, (W?)
			Different people have different opinions.

			(W)	
			In my standpoint, (W?)	
			I don't agree that. (W)	
7	In their opinions, (W)	3/6	In some people's opinions, (W)	4/6
	What about my opinion? I stand up for the second mind. (W)		It is no doubt that,...(C)	
	In my thought, (W)		But in another people's view, (W)	
			In my eyes, (C?)	
	As I see it, (C)			
	From my standpoint, (C?)		From my standpoint, (C?)	
	Generally speaking, (C)		In my view, (C)	
8	This is my point of view. (C)	2/4	But I would like to say that why not believe the two can exist and develop together? (C)	5/5
	..., I think. (C)			
	So I directly believe that...(W?)		In fact, I would like... (C)	
	To the standpoint of mine, (W)		Different people have different ideas about...(C)	
			I think we should look at...(C)	
			In my opinion, (C)	
Total	32	18/32	16	29/41

3. Implications and Conclusion

In the present study, though it might have left more questions unanswered than it have solved and admittedly it represents all the limitations that can be readily solved by the modern means of learner corpora and native language corpora: It is small-scale, and the analysis is done manually, which seems to be so primitive in face with the enormously huge corpus today. However, despite all these limitations, I hope it has been able to accomplish its goal of pushing more empirical research into this wonderfully intriguing while relatively uncharted phenomenon of formulaic sequences

acquisition by EFL learners. Building on this, the following potential lines of investigation can be fruitfully pursued in further studies.

First, in terms of research orientation, I am calling for a combination of corpus linguistic research and SLA research. In the past, both sides seem to have worked quite separately, which is unfortunate indeed. With the enormously huge size of corpora, applied corpus linguists stand at a better position to depict a global picture of patterns and use of language by the learners, while SLA researchers can contribute their parts by providing insightful investigation into the acquisition and learning process of some particular cases of formulaic sequences. Such aligning efforts from both sides will achieve the synergy effect that shall be beneficial to the TEFL field as a whole.

Second, in the present paper, I have speculatively designated L1 negative transfer as a major cause accounting for L2 learners' difficulties with using formulaic sequences, a cause that have also been identified by other applied corpus linguists (e.g. Granger, 1998). To this regard, I am calling for a reconsideration of error analysis and contrastive analysis as informed by SLA research in terms of formulaic sequences. As Granger (1998) bluntly claims, "Given the essentially language-specific nature of prefabs, this is a major issue that must be addressed if we are serious about giving learners the most efficient learning aids" (p. 159). Granger has also emphasized the importance of introspective tests following corpus research. In the present study, a case study approach has been adopted. Building on it, further research can now expand the scope by seeking to investigate "how these formulaic sequences are acquired by Chinese EFL learners in other learning contexts"? For example, in terms of learner groups, we can expand it into the primary level, the secondary level, the college level (which includes the English majors and nonmajors), postgraduate level etc. With these data at hand, it will then become possible to chart out the developmental path of the

acquisition of formulaic sequences for participants of different proficiency levels. Such information will prove to be useful for educators and language researchers in designing syllabus or preparing for course materials.

Third, as already noted above, the effects of individual differences on formulaic sequence acquisition need to be further demystified with more rigid research designs. For one thing, such designs can take into account the latest theoretical constructs of individual difference research in the light of the information-processing based SLA view (Skehan, 1998). For example, in terms of foreign language aptitude, it should incorporate the memory component along with the language analytic component (Skehan, 1998, cf. Dornyei et al., 2004). Other individual differences (like learning style, learning strategies etc.) might also prove to be worthy of investigating as they may also play a part in the acquisition of formulaic sequences. More importantly, such designs can be combined to consummate with some well-established SLA research paradigms, for example, the task-based research that have really blossomed in the recent past. In this respect, Dornyei and colleagues (Dornyei, 2002; Dornyei and Kormos, 2000) have made profoundly insightful discoveries. Efforts heading towards this direction should produce fruitful results.

Last but not least, there are also a few caveats looming ahead of such research into the phenomenon of formulaic sequences that all researchers working towards this end should be cautious about or alternatively, should make concerted efforts to come up with possible solutions. First and foremost, there is the very basic yet intractable question of “what constitute formulaic sequences?” As formulaic sequences seem to exist in so many forms and have so many names (e.g. fifty in Wray’s 2002 list, p. 9) that it becomes insurmountably difficult even just to come up with a comprehensive definition of the phenomenon (Schmitt and Carter, 2004). Second, there is also the

frequency effect to be wary about. Though frequency definitely has its effect on learner's acquisition of formulaic sequences, its effect should not be overstated, simply because there are many other factors working to influence the input of formulaic sequences noticed by L2 learners (Skehan, 1998; Schmidt, 1990).

References:

- Altenberg, B., Tapper, M., 1998. **The use of adverbial connectors in advanced Swedish learners' written English**. In Granger, S. (ed.), *Learner English on computer*. Longman, London, pp. 80-93.
- Cortes, V., 2004. **Lexical bundles in published and student disciplinary writing: Examples from history and biology**. *English for Specif. Purposes* 23, 397-423.
- Cowie, A. P., 1992. **Multi-word lexical units and communicative language teaching**. In Anaud, P., Bejoint, H., (Eds.), *Vocabulary and applied linguistics*. Mcmillan, London.
- Ding Y., Qi Y., 2005. **Use of formulaic language as a predictor of L2 oral and written performance**. *J. of PLA Uni. of Foreign Lang.* 28, 49-53.
- Dechert, H., 1983. **How a story is done in a second language**. In Faerch, C., Kasper, G., (Eds.), *Strategies in interlanguage communication*. Longman, New York, pp.125-140.
- Dornyei, Z., 2002. **The motivational basis of language learning task**. In Robinson, P. (ed.), *Individual differences and instructed language learning*. John Benjamins, Amsterdam.
- Dornyei, Z., Kormos, J., 2000. **The role of individual and social variables in oral task performance**. *Lang. Teach. Research* 4, 275-300.
- Dornyei, Z., Durow, V., Zahran, K., 2004. **Individual differences and their effects on formulaic sequence acquisition**. In Schmitt, N. (Ed.), *Formulaic sequences*. John Benjamins, Amsterdam, pp. 87-106.
- 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 Bygate, M., Skehan, P., Swain, M., (Eds.), *Researching pedagogic tasks*. Longman, Harlow.
- Gatbonton, E., Segalowitz, N., 2005. **Rethinking communicative language teaching: A focus on Access to fluency**. *Canadian Modern Lang. Review* 61, 325-353.
- Granger, S., 1998. **Prefabricated patterns in advanced EFL writing: collocations and formulae**. In: Cowie, A. P., (Ed.), *Phraseology*. Clarendon Press, Oxford, pp. 145-160.
- Granger, S., Rayson, P., 1998. **Automatic profiling of learner texts**. In Granger, S. (ed.), *Learner English on computer*. Longman, London, pp. 119-131.
- Hunston, S., 2002. **Corpora in applied linguistics**. CUP, Cambridge.
- Jones, M. A., Haywood, S., 2004. **Facilitating the acquisition of formulaic sequences: An exploratory study**. In Schmitt, N., (ed.), *Formulaic sequences*. John Benjamins, Amsterdam.
- Lewis, M., 1993. **The lexical approach**. Teacher Training Publications, Hove.
- Moon, R., 1998. **Fixed expressions and idioms in English**. Clarendon Press, Oxford.
- Nattinger, J. R., DeCarrico, J. S., 1992. **Lexical phrases and language teaching**. OUP, Oxford.

- Pawley, A., Syder, F. H., 1983. **Two puzzles for linguistic theory: nativelike selection and nativelike fluency**. In: Richards, J. C., Schmidt, R.W., (Eds.), *Language and communication*. Longman, New York, pp. 191-226.
- Read J., Nation, P., 2004. **Measurement of formulaic sequences**. In Schmitt N., (Ed.), *Formulaic sequences*. John Benjamins, Amsterdam.
- Schmidt, R., 1990. **The role of consciousness in second language learning**. *Appli. Ling.* 11, 17-46.
- Schmitt N., Carter, R., 2004. **Formulaic sequences in action: An introduction**. In Schmitt N., (Ed.), *Formulaic sequences*. John Benjamins, Amsterdam.
- Skehan, P., 1998. **A cognitive approach to language learning**. OUP, Oxford.
- Skehan, P., and Foster, P., 2001. **Cognition and tasks**. In Robinson, P., (ed.) *Cognition and second language instruction*. CUP, Cambridge.
- Sinclair, J., 1991. **Corpus, concordance, collocation**. OUP, Oxford.
- Weinert, R., 1995. **The role of formulaic language in second language acquisition: a review**. *Appli. Ling.* 16, 180-205.
- Willis, D., 1990. **The lexical syllabus**. Harper Collins, London.
- Wray, A., 2000. **Formulaic sequences in second language teaching**. *Appli. Ling.* 21, 463-489.
- Wray, A., 2002. **Formulaic language and the lexicon**. CUP, California.
- Wray, A., Perkins, M., 2000. **The functions of formulaic sequences: An integrated model**. *Lang. and Comm.* 20, 1-28.