Formulaic Language: An Analytical Review

Siping Liu[1]

Abstract

This paper intends to give an analytical review of formulaic language. First, a brief review is made on various definitions and identification of formulaic language, a topic that has remained unsettled to date. Next, another topic --- transferability of unanalysed chunks to analytical competence will follow. As this topic has aroused heated discussion in the literature, one section will be spared to it for some detailed coverage. The last section, under the title of rationale and functions, is devoted to various features of formulaic language, which are elaborated in several subsections.

Keywords: formulaic language, linguistic functions, review

1. Introduction

Ever since the introduction of Chomsky's 'competence', research on second language acquisition has shifted its focused on the importance of creative, rule-governed process and on the systematic nature of learner language (Myles, Hooper & Mitchell, 1998). In contrast to this academic interest, formulaic language, though standing side by side with analytical system, has not been fully studied until the last 20 years.

Despite the negligence in the literature for many years, formulaic language has always occupied a vital place in human communication, especially in fluent and idiomatic production of spoken language (Biber, Conrad & Cortes, 2004).

---

[1] PhD in Teacher Education, graduated from University of Nevada Las Vegas, U.S.A.
E-mail: siping.liu@whu.edu.cn, Jianshe Dadao #613-36-2 Wuhan Hubei 430030 China.

From the protolanguage uttered by primitive man in chunks (Wray 1998) to the latest corpus calculation (Altenberg, 1998), formulaic language shares a great proportion of our speech. Chomskyan ideal speaker' (Chomsky, 1965) can hardly open his or her mouth for communication without the pragmatic aid of formulaic language. If a speaker could communicate only by applying grammatical rules, he or she would no doubt be a man of far fewer words (Wong-Fillmore, 1976).

It is true that human beings are the only creature on the earth that can create indefinitely meaningful messages by using acceptable grammatical rules. But it is also true that we usually utter ready-made expressions just like chimpanzees' 'warning signals' (McMahon, 1994), both of which are similar in one point of their utterance in that they are unanalytical.

2. Definitions

In terms of terminology, the concept of formulaic language covers a wide range of expressions in the literature (Wray & Perkins, 2000). In contrast to some of the definitions like 'fossilised forms' and 'frozen phrases', the terminology of formulaic language itself, ironically, is far from being 'fossilised' and has actually generated some ambiguity in literature. A mere glance at the literature generates some forty terms (Weinert, 1995) such as chunks, collocations, fixed expressions, holistic, lexicalised phrases, prefabricated routines, ready-made expressions, unanalysed chunks of speech, to mention only a few.

In the mess of terminology, however, one can still draw some consensus and find that 'researchers have very much the same phenomenon in mind' (Weinert, 1995: 128). Formulaic language, according to Weinert, is 'multi-word or multi-form strings which are produced or recalled as a whole chunk, much like an individual lexical item, rather than being generated from individual lexical items/forms with linguistic rules' (*ibid:* 182).

Based on a handful of reading my understanding is characterised by the following criteria:

✧ Phonologically coherent phrases memorised as a whole like one big word;
✧ Holistically retrievable without drawing on the analytical processing;
✧ Socially acceptable configuration and situational dependence with little novel information.

✧ Relatively invariably or frequently uttered by speakers or writers who are otherwise lower in grammatical competence.

As for the reasons of ambiguity in terminology, the nature of formulaic language itself has caused the 'chaos', for there is actually 'no boundary between an idiom and a collocation or between a collocation and a freely generated phrase' (Bolinger, 1977: 168). Another reason lies in the fact that formulaic language is observed from different perspectives and therefore has been subject to independent labelling (Wray, 2000). In a larger background, however, formulaic language can be viewed as a continuum between tightly idiomatic expressions and freely generated phrases (Weinert, 1995;Wray & Perkins, 2000) or non-formulaic strings (Wray, 2000). Along the line, various types of formulae settle in a proper point in accordance with the degree of unanalyticality or conventionality (Chenoweth, 1995).

Wray's two variables of formulaic language may shed more light on further understanding of this matter. Wray (2000) defines formulaic language as *compositionality* and *flexibility*. The former concerns itself with how the meaning of the whole reflects the sum of the parts. Non-compositional formulae are less understandable from their constituent words. As for flexibility, it reflects the similar idea of the continuum, with 'entirely fixed sequences at one extreme, to templates for unlimited paradigms of novel ones at the other' (*ibid*: 215).

From the conception of continuum, we may basically split formulaic language into two taxonomies: open and closed. The open formulae or open slots are lexicalised sentence stems, within which some variation can occur (Weinert, 1995). Wray & Perkin (2000) tie them near one end of the continuum, for their properties of transparency and flexibility. On the other hand, the closed formulae are exemplified by 'idioms and phrases that have taken on a semantic and syntactic life of their own' (Chafe, 1994: 113).

At the end of this section, from the puzzling definitions, one may be clarified to some degree just by one simplified statement: formulaic language is a 'phenomenon that encompasses various types of word strings which appear to be stored and retrieved whole from memory (Wray & Perkins, 2000: 1).

3. Analytical Competence vs. Unanalytical Chunks

L2 learners would certainly feel hopeful if the unanalysed chunks they have stored in mind could be gradually unpacked or broken down later for analytical or creative competence. But so far the few empirical studies concerning L2 have only drawn discouraging conclusions (DeCock, 2004; Foster, 2001), especially for adult learners (Wray, 2000), though formulaic language may contribute more positively to children. Given the different nature of the subjects in the researches, it is impossible to reach a uniform result in this issue in the literature.

Krashen and Scarcelle (1978) argue that formulaic language and rule-created speech are neurologically unrelated and there cannot be an interface between them. According to this view, learners do not linguistically unpack the linguistic information contained in formulae, but internalise L2 rules independently by attending to input. Bohn (1986) also concludes that formulae serve only 'short-term production tactics', and play no role in acquisition. Krashen (1981) explains that formulae are processed in the right hemisphere and are therefore not part of central language processes, which are located in the language-dominated left hemisphere.

These beliefs seem to be further justified by Chinese children's recitation of some of the classic poetry, a type of fixed formulae according to categorization (Wray, 2000). These highly rhymed poems that can be easily blurted out are far beyond the children's analytical competence and thus they can hardly be decoded naturally in later time. For real understanding and acquisition, the poetry, like any other formulae, has to be dismantled and taught in different ways (Hopper, 1987).

For some other linguists, formulaic language is the initial step leading to analyticality, especially for children learning L1 or L2. Wong-Fillmore (1976) claims that imitated utterances lay the foundation for later analysis, and the unpacked pieces become part of the learners' developing linguistic system.

Weinert (1995) also cites empirical evidence that previously unanalysed chunks feed into the developing rule formation.

Ellis (1994) explains that creative speech is promoted when learners armed with formulaic language come to realize what they initially use as whole consists of discrete constituents.

Myles et al (1998) believe that the rote-learning of formulae and the construction of rules are not independent processes, but interact and actively feed into one another. The discussion of the issue would be groundless if learners are not associated with a particular situation. From different perspectives, we may reach different conclusions.

In fact, formulae and rule-governed systems are complementary or supplementary to each other and co-exist in different hemispheres with a bridge linking them together or at least they are located in different layers, even if only one hemisphere dominates our language. Under different pragmatic scenarios we will alternate between creative expressions and routinized forms. For example, in real time conversations, because of peer's pressure, speakers are prone to resort to prefabricated units in order to reduce the processing load (Jiang and Nekrasova, 2007; Kecskes, 2007).

Weinert states that 'L2 acquisition data in general provide apparently ample evidence of creative construction processes which are independent of previously learned chunks of language' (1995:186). Accordingly, we cannot safely jump to the conclusion that 'learning always begins with large unanalysed units which are then broken down' (ibid: 186). Weinert's 'parallel and interactive process' (ibid) and Wray's two variables, compositionality and flexibility (2000), drop more hints on the understanding of this issue.

Since formulae differ in degree (Chenoweth, 1995), their contribution to the development of analytical competence may vary accordingly. Some formulae, especially proverbs, idioms or other semantically opaque and syntactically irregular ones, are learned in parallel with their segmented constituents; while some others, especially the semantically transparent and syntactically regular open slots, interact more positively with the growth of grammatical knowledge.

Whether they interact more or less, analytical competence and unanalytical chunks are stored separately in the mind. Under different scenarios, we exploit different resources, though both function simultaneously in mutually complementary manner.

In speaking, formulaic language is preferred because of 'too little on-line processing capacity to easily handle analytical language at all times' (Wray, 1998). This explanation may lead to the assumption that analytical language is deposited deeper than unanalytical chunks in the mind, besides the apparent reason that analytical language takes more time for speakers to re-assemble small pieces of lexical items syntactically. In writing it is the same story. When her students were time-constrained in writing, according to Chenoweth (1995), they naturally fell back on formulaic language for ready-made expressions. If the pressure was lifted and accuracy emphasised, it can be inferred, the students would be more creative and expressive in their compositions. Were the two blended together, speakers or writers might inevitably find trouble for quick access.

4. Rationale and Functions

4.1 Economicality

Psycholinguistically, human brains tend to function at the minimum analytical cost. Once the brain is familiar with a linguistic task, it by-passes the processing route that was used to learn it (McCrone, 1999) through a short-cut to a storehouse of large linguistic chunks retrievable holistically. Wray (1998) explains that we rely on holistic processing not because of the failure of analytical system but because of its expensive strategy, which we resort to compulsorily when confronted with emergencies. Becker says that in communication or social interaction, 'we don't have to go through the labour of generating an utterance all the way' (1975:17). In other words, if we depended exclusively on rule-based system, our mind would have been exhausted, for we would be constantly required of full access to all the language facilities (Wray, 1998).

It can be concluded from above that our brain is a great energy saver. We only consume our analytical 'energy' when pressed to do so and we are very skilful in keeping an economic balance between the expected and the unexpected in communication, which in analogy is just like driving a car along a formulaic-paved highway. Drivers running smoothly and comfortably along the way are only alerted when they make a sharp turn, or bump into an obstacle.

4.2 Acceptability

Formulaic language is not only communicatively necessary for both speakers and hearers to code and decode so that fluency is maintained for successful conversation. It also plays a vital role in socialinguistics. L2 learners, no matter how proficient they are in target language, will easily betray themselves, if they do not speak as idiomatically as the native counterparts do, for a person's group identity is established by using the right words or phrases (Wray, 1998). As formulaic language is socially-bound and culturally-laden, grammaticality is by no means acceptability. Although all grammatical sequences expressing the same idea are equally valid and equally like to occur, only some are preferred over others (ibid). Therefore, to answer the question of 'to be or not to be', the acceptable idiomatic expressions alone will mark even highly proficient non-native speakers out as non-natives (Pawley & Syder, 1983). Thus, an 'ideal speaker' is a non-affiliated proficient speaker unless s/he speaks in the way characteristic of a particular community.

4.3 Dependability

The unbelievable pervasiveness of formulaic language shows how much we have to depend on it, no matter in speaking or writing. Dechert (1984) calls it 'islands of reliability' that helps to construct and execute production plans (Ellis, 1994).

For L2 speakers, the frequent use of formulaic language, though inappropriate on some occasions, is survival strategies, which will help learners to deal with communicative situations unmanageable with their own competence (Yorio, 1980). Without rich resources of formulaic language, L2 speakers would feel tongue-twisted or frequently pause and hesitate in conversations (Schmidt, 1983). Bygate (1988) describes that speakers who are short of formulae tend to repeat fragments of their own, or their interlocutors' previous utterances.

This argument is also evidenced by Weinert's longitudinal and cross-sectional study (1987). According to her, the first year German learners relied heavily on routine and patterns in speech production.

The underlying reason for speakers' dependence on holistic retrieval is the limitation of human on-line processing capacities.

The interactional function (Wray 1998) pushes those involved in conversations back to the more accessible and reliable holistic system in the real-time communication, in which fluency is prioritized above anything else. Meanwhile, formulaic language also serves as a means of transition for speakers to generate novel information. In Weinert' words, prefabricated formulae indirectly make creative speech possible (1995).

In writing, formulaic language does the same job. Many formulae act as sentence-builders, or macro-organisers in the text (Granger, 1998). The result of Chenoweth (1995)'s research documented the organising role of formulaic language in achieving sentence coherence. Her subjects made extensive use of formulaic expressions. Liu also found formulaic language was so overused in the college English test in China that many composition raters took it as a sign of failure because they do not show students' real language ability (Liu, 2009). The reason, as Granger explains, is that 'learners' repertoires for introducing arguments and point of views are very restricted and they therefore cling on to certain fixed phrases and expressions which they feel confident in using (1998:156).

4.4 Difficulty

Judging from its nature, formulaic language is the first pair of crutches for the L2 learners to fall back on. However, with the shortage of such resources, they may more easily feel worn out in more analytical processing, compared with native counterparts, who accumulate formulaic language through the natural and everlasting exposure to their mother tongue.

The reason for the L2 speakers' linguistic fatigue is the fact that formulaic language is dynamically associated with pragmatics, psycholinguistics and socialinguistics (Chenoweth, 1995), or it is known as socialization in the second language (Ellis, 1994), which is certainly far less manipulatable than grammatical rules and is constrained in classroom teaching (Weinert, 1995). In this sense, many of formulae have to be acquired naturally than learned in a foreign soil. Considering that EFL learners have rare chance to immerse themselves in natural environments of the target language, many of formulae will always lie beyond their reach and therefore will remain a life-long target for most of us to approach. To bite this hard nut, EFL learners should have culturally sharpened teeth in a native-like dinning hall. Even if they swallow them without chewing, EFL learners can hardly retain and retrieve them without the frequent use in real-time situation (Wray, 2000).

Besides the situational dependence, the unanalyticality and irregularity are other constant problems for adult learners, who may skilfully assemble single words in perfect grammatical form. Since some formulae have relinquished their semantic compositional meaning in favour of a holistic one, they usually carry a metaphorical meaning, which, without substantial pragmatic or direct explicational content, is hardly possible for learners to understand (Nattinger and DeCarrico 1992). Pawley and Syder mention the same difficulty confronting L2 learners when they analyse lexicalized sentence stems, which have a 'grammar that is unique in that they are subject to an idiosyncratic range of phrase structure and transformational restrictions' (1983:192).

The same truth also holds in our memorization of telephone numbers. The frequently dialled numbers are clearly stored and easily retrieved, while the rarely used numbers have to be jotted down in a notebook for occasional reference. But with some help of analysis, many numbers may be easily and permanently stored in our memory. For example, in Chinese the homophony of some telephone numbers and words helps memorization. To illustrate, 881571 can be figuratively understood as 'Dad asks me to take medicines' in terms of homophony.

In short, formulaic language is not so easy to learn as it appears to be. In fact, it is the final and most difficult stumbling block for otherwise advanced non-native speakers (Lancker-Sidtis & Rallon, 2004).

Even for native speakers, research confirms that some formulaic language is not appropriately processed in the first language until late teenage (Wray, 2000).

5. Conclusion

This paper offers a glimpse of the formulaic language. In terminology, formulaic language has not been clearly defined because it has extended to a wide range of area such as psycholinguistics and socialinguistics. In view of this, the literature has to approach it from different perspectives. Another controversial issue addressed is the interplay between analytical competence and formulaic language. Again, debate has arisen because of the study of different subjects, children or adults, L1 or L2. For children learning in natural environments, formulaic sequences may be segmented and absorbed to promote construction of rules.

For adults, especially EFL learners learning in non-native English speaking environments, each seems to go their own way.

The features of formulaic language are also covered in some length. The reasons why people prefer formulaic language are closely associated with the working of the brain, the social membership and speaking strategies. It appears to be routinized but by nature it is rather unmanageable than grammatical rules. Given the necessary condition of natural exposure to target language for the acquisition, most of L2 learners' proficiency is doomed to be non-native like.

On account of the limitation of space and shortage of time, my coverage is certainly a condensed one and is far from being exclusive. I will feel greatly satisfied if my analyses of formulaic language can possible arouse any interest for more empirical studies.

References

Altenberg, B. (1998). 'On the Phraseology of Spoken English: the Evidence of Recurrent Word Combination' In: A. Cowie (ed.) Phraseology: Theory, Analysis and Applications Oxford: Oxford University Press.

Biber, D., Conrad, S. & Cortes, V. (2004). 'If you look at lexical bundles in university teaching and textbooks'. Applied Linguistics, 25 (3), 371-405.

Becker, J. (1975). The Phrasal Lexicons. Bolt Beranek and Newman Report No. 3081, AI Report No. 20.

Bohn, O. (1986). 'Formula, frame structure, and stereotypes in early syntactic development: Some new evidence from L2 acquisition'. Linguistics, 24, 185-202.

Bolinger, D. (1977). Idioms have relations Forum Linguisticum, 2, 157-169.

Bygate, M. 1988 Units of Oral Expressions and Language Learning Applied Linguistics 9(1)

Chafe, W. (1994). Discourse, Consciousness, and Time: The Flow and Displacement of Conscious Expectation in Speaking and Writing Chicago: University of Chicago Press.

Chenoweth, N.A. (1995). Formulacity in essay exam answers. Language Sciences, 17(3)

Chomsky, N. (1965). Aspects of the theory of syntax Mass: M.I.T. Press

Dechert, H. (1983). How a story is done in a second language. In: Faechard, C and G. Kasper, Casper (eds.), Strategies in Interlanguage Communication (pp. 175-195). London: Longman.

Ellis, R. (1994). The Study of Second Language Acquisition. Oxford: Oxford University Press.

DeCock, S., (2004). Preferred Sequences of Words in NS and NNS Speech. Belgian Journal of English Language and Literatures (BELL), 2, 225-246.

Foster, P., (2001). Rules & 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: Second Language Learning, Teaching and Testing (pp. 75-94). Longman, New York.

Granger, S. (1998). Prefabricated Patterns in Advanced EFL Writing: Collocation and Formulae In Cowie, A.P. (ed.), Phraseology: Theory, Analysis, and Applications (pp. 145-160). Oxford: Oxford University Press.

Hopper, P. (1987). Emergent Grammar Berkeley Linguistic Society, 13, 139-157.

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

Kecskes, I., (2007). Formulaic language in English Lingua Franca. In: Kecskes, Istvan, Horn, Laurence R. (Eds.), Explorations in Pragmatics (pp. 191-219). Mouton de Gruyter, Germany.

Krashen, S. (1981). Second Language Acquisition and Second Language Learning Oxford: Pergamon.

Krashen, S. & R. Scarcella (1978). On routines and patterns in language acquisition and performance. Language Learning, 28(2), 283-300.

Lancker-Sidtis, D. Van. & Rallon, G. (2004). Tracking the incidence of formulaic expressions in everyday speech: methods for classification and verification. Language & Communication, 24, 207–240.

Liu, S. (2009). The hidden side of EFL writing in China. Foreign Languages Teaching Journal. 13(6), 34-51.

McCrone, J. (1999). State of mind. New Scientist, 161(2178), 30-31.

McMahon, A. (1994). Understanding language change. Cambridge: Cambridge University Press.

Myles, F., Hooper, J and Mitchell, R. (1998). Rote or rule, exploring the role of formulaic language in classroom foreign language learning. Language Learning, 48(3), 323-363.

Nattinger, J. R. & DeCarrico J.S. (1992). Lexical Phrases and Language Teaching. Oxford: Oxford University Press.

Pawley, A. & Syder, F.H. (1983). Two puzzles for linguistic theory: Native selection and native-like fluency. In: Richard, J.C. & Schmidt, R. W. (eds.), Language and Communication (pp. 191-225). London: Longman.

Schmidt, R. W. (1983). Interaction acculturation, and the acquisition of communicative competence: A case study of an adult. In Wolfeson, N and Judd, E. (eds.), Sociolinguistics and language acquisition (pp. 137–174). Rowley, MA: Newbury House.

Weinert, R. (1987). Process in classroom second language development --- the acquisition of negation in German. In R. Ellis (ed.), Professional competence and quality assurance in the caring professions (pp. 34-56). London: Chapman and Hall.

Weinert, R. (1994). Some effects of a foreign language classroom on the development of German negation. Applied Linguistics, 15(1), 76-101.

Weinert, R. (1995). The role of formulaic language in second language acquisition: A review Applied Linguistics, 16(2), 180-205.

Wong-Fillmore, L. (1976). The second time around: Cognitive and social strategies in second language acquisition. Ph.D dissertation, Stanford University.

Wray, A. (1998). Protolanguage as a holistic system for social interaction. Language and Communication, 18(1), 47-67.

Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice Applied Linguistics, 21(4), 463–489.

Wray, A. & Perkins, M. R. (2000). The functions of formulaic language: An integrated model Language and Communication, 20(1), 1-28.

Yorio, C.A. (1980). Conventionalised language forms and the development of communicative competence. TESOL Quarterly,14(4), 433-442.