The Effects of English Academic Formulaic Sequence Processing on Chinese Graduate Students

Dongmei Meng[1,2], Lulu Jin[1,*]

[1]School of Foreign Languages, East China Jiaotong University, Nanchang, Jiangxi, China

[2]Research Center for Applied Translation of Transportation and Engineering, East China Jiaotong University, Nanchang, Jiangxi, China

*Corresponding author: Lulu Jin

Abstract: The processing of formulaic sequence is a hot discussion of research in the field of language acquisition, but the differences in experimental subjects and experimental methods have led to different point of view for many studies on programmatic language processing, and even different conclusions are often obtained on certain issues. The research on the influence of formulaic sequence processing has reached a high attention in recent years. By reviewing the literature on formulaic sequence processing in recent years, especially the review and summary of newer studies, hoping to sort out the influencing factors of formulaic sequence processing among Chinese graduate English students.

1. Intrdduction

The use of language is not completely arbitrary, and people tend to use some fixed expressions to convey certain meanings (Pawley&Syder, 1983). This fixed form is generally composed of several words glued together and reproduced in high frequency in the use of language, which is called formulaic sequence. In language use, the frequency of formulaic sequence is about 30% to 50%(Oppenheim, 2000; Foster, 2001), that is to say, people are accustomed to using some fixed expressions, which also appears in academic communication. Formulaic sequence is an important component of academic English discourse. The mastery and use of formulaic sequence in academic English is conducive to understanding and constructing discourse and to improving language fluency and accuracy. The study of formulaic sequence can be divided into three categories: the use of formulaic sequence, the acquisition of formulaic sequence and the processing of formulaic sequence. This paper reviews the recent research on formulaic sequence processing, and summarizes the influencing factors, so as to provide reference for future research.

2. Formulaic sequence

Formula sequence is an important part of vocabulary knowledge, a basic part of discourse construction, and a decisive indicator of learners' overall language ability and development (Ellis, 2012). Some scholars believe that after matching the functional meaning of the word with the form when remembering, the sequence formula sequence composed of multiple orthographic words or phonemic words is stored as a word memory. Under this definition, such languages may not have clear meanings or functions. The storage and retrieval methods of character strings vary from person to person.

In recent years, the cognitive processing of formula sequences has become a hot topic in linguistic research, focusing on whether formula sequences are stored and extracted in a mental dictionary. On this issue, Pawle & Sider (1983) proposed the "global representation" hypothesis in 1983, which holds that formula sequences have two cognitive modes: global representation and global processing. They believed that the former is the basis of the latter, and the latter is the development of the former. The British applied linguist Alison Wray (2002) also holds a similar view that formula sequence are

"extracted or stored from overall memory, rather than generated or analyzed through grammar". Psycholinguistics experiments show that the first language and second language learners process formula sequence much faster than non-formula sequence. Therefore, the formula sequence is regarded as a complete storage and extraction language unit in the brain, thus, it can be concluded that no matter from the perspective of linguistics and psycholinguistics, the processing of formula sequence is integrated, and its processing speed is higher than that of non-formula sequence.

3. Literature review

3.1 Research methodolgy

This paper takes CNKI as the platform, sets the retrieval mode as "advanced retrieval", and takes the keyword "English formulaic sequence processing" as the retrieval condition, and finds 47 relevant literatures. Although these literatures did not directly study academic formulaic sequence, their research on formulaic sequence not only examined the language level, language use experience, but also the relationship between language background, genre, discipline and the use of formulaic sequence, which is closely related to academic formulaic sequence.

3.2 Research on formula sequences processing

Foreign research on formulating sequence processing began in the 1970s. After experimental research, more consistent conclusions have been drawn on the processing advantages of formula sequences. Related research in my country started relatively late. Relevant data from CNKI showed that there were 70 literature search results for the keyword "formula sequence processing", of which the earliest publication date was 2010, and 12 related documents were published in 2018, indicating that domestic research on this issue has attracted widespread attention, ushered in its research climax.

From a linguistic point of view, the processing advantage of formula sequence comes from frequency effects. According to the frequency influence hypothesis, the more language items with higher frequency of contact, the stronger the mental representation, and the faster and more accurate the identification of language users. For high-level English learners, most domestic research results show that their formula sequence processing has a frequency effect, but for low-level English learners whether they are sensitive to the frequency of formula sequences, the existing research results are not consistent. In addition, some scholars are also concerned about the impact of semantic transparency on the processing of formulaic sequences. Studies have found that the chunk processing of Chinese English learners has a semantic transparency effect.

At present, researchers at home and abroad have applied a variety of research paradigms in order to empirically study the psychological processing of formulaic sequence from multiple perspectives and methods. In the literature mainly reviewed in this paper, reaction time paradigm was applied in all experiments, and the collected data were analyzed and sorted out.

4. Influencing factors

According to the literature data on CNKI, this paper temporarily divides the influencing factors into the following four categories: formula sequence frequency, learners' English level, semantic-related knowledge and task type. The following will sort out the literature from the experimental methods, experimental tools, experimental purposes and experimental results.

4.1 Formulaic sequence frequency

The frequency of the sequence of formulas refers to the number of times the special phrase appears in the language learning process. The higher the probability of occurrence and the higher the frequency, the learner is more familiar with the formula sequence; otherwise, the lower the frequency, the less familiar the learner is with the formula sequence.

Huang Jia (2019) used the online grammar evaluation task and E-prime software to record the judgment and reaction time of the subjects, measure the processing speed and accuracy and explore the differences in the frequency and semantic transparency of the formula sequence to affect the

participants' grammar and judgment of the formula sequence. In the form of a preliminary questionnaire, the author selected some formula sequences with high or low semantic transparency from the list of semantic transparency tables compiled by Martinez & Schmit, and investigated the frequency and transparency of these formula sequences among Chinese students, and based on the final questionnaire. The results showed that there was a significant frequency effect in the recipe sequence processing. The response time and error rate of the two groups are lower than the low-frequency formula sequence recognition, indicating that the high-frequency formula sequence has processing advantages.

Through the above experiments by Huang Jia (2019), it can be concluded that the frequency of the formula sequence has an impact on the processing. Words can be represented as a whole in the learner's second language mental vocabulary. This phenomenon may also be due to the frequent co-occurrence of the components, which makes the relationship between the components close and improves the processing efficiency of the formula sequence.

4.2 Semantics of formulaic sequence

Semantics in the narrow sense mainly includes the understanding of language knowledge and the grasp of language meaning. From the perspective of operability and practicality, this article mainly studies the impact of semantic transparency and lexical semantics on formula sequence processing.

Huang Jia (2019) studied the effects of three variables on formula sequence processing at the same time in his experiment, including semantic transparency. The program language is divided into high transparency and low transparency according to the survey results. Experimental results showed that semantic transparency affected language processing. For learners with high and low proficiency, compared with learners with high semantic transparency, the subjects can more accurately identify formula sequence with low semantic transparency in a short period of time in the grammatical judgment task, indicating formula sequence with low semantic transparency. It is easier for English learners to handle.

Cao Yu (2018) studies the relationship between semantics and formula sequence processing. He also used E-prime to record the reaction time and accuracy of the subjects, but unlike Huang Jia, he used the collocation semantic judgment task under the cross-language activation condition. In this experiment, the task of judging collocation words was performed, and English learners were required to distinguish between formula and non-formula sequence in the auditory channel.

Huang Jia (2019) and Cao Yu (2018) used visual online grammatical judgment tasks and auditory collocation semantic judgment tasks to obtain consistent experimental results, indicating that the semantics have an impact on formula sequence. The reason for these results may be that Chinese English learners need semantic assistance in formula sequence processing.

4.3 Language level

Language level refers to the level of language learners' mastery of the language. In practical experiments, in order to facilitate the control of variables, languages learning time or relevant domestic language test scores are often used as the basis for judging.

Luo Tingting (2020), like Huang Jia, used the DMDX program to investigate the mental representation and processing of high and low proficiency test subjects in the online grammar judgment task. In terms of language selection, he uses three or four words as Huang Jia did. There were two main findings from the experiment. First, the processing speed of formula sequence is significantly faster than that of non-formula sequence, and the error rate is significantly lower in both high level and low level. Second, there was no significant difference in the processing of formula sequence among different levels, which indicates that the second language level has no influence on the processing of formula sequence among Chinese English majors.

Luo tingting's experimental results were consistent with those of Zhong Zhiying (2016), who showed that the different language levels of second speakers did not significantly affect formula sequence processing. This discrepancy may be due to differences in materials used and subject selection. For example, Huang Jia chose non-English major students and divided them into groups based on cet-4 and CET-6 scores, while Luo Tingting chose English major subjects and divided them

into groups based on cloze test. This difference caused the difference in experimental results.

4.4 Types of tasks

Sheng Jie (2019) studied the impact of different input task types on the quality of Chinese English learners's formula sequence Acquisition, exploring the impact of input on the memory, application and learning of different types of formula sequence. First, different input tasks have a significant impact on formula sequence memory. Second, high-input tasks are more conducive to the improvement of language application ability. However, the sentence-making task has obvious advantages over the translation task under the same input, which means that learners need to process sentences in deeper semantic processing. Third, different input tasks had a significant impact on different types of formula sequence learning. Different input tasks had a significant impact on different types of formula sequence learning. In the case of the same amount of input, the learning effect of a fixed formula sequence was higher than that of a semi-fixed formula sequence.

The above experimental results may be due to different types of task learning, and learners have different requirements for the processing depth of the formula sequence. For example, memory tasks have a higher degree of mastery of formula sequence than sentence-making and translation tasks. Not only should we pay attention to the semantics of formula sequence, but also make adaptive choices based on context. Although the literature discusses the relationship between task types and formula sequence acquisition, formula sequence acquisition is the basic part of it's processing, so it can also be regarded as the processing of formula sequence.

5. Conclusion

By summarizing and analyzing the collected literature, the following results can be drawn: the frequency of formula sequence use, the learner's English proficiency, the semantics of the formula sequence and the task type all affect the processing of the formula sequence. The higher the frequency of use of the formula sequence, the better the processing effect. The clearer and more specific the semantics of a formula sequence, the easier it is to deal with. Although there are some differences in the research results of learners' English proficiency, for non-English majors, the higher the English proficiency, and the better the processing effect. Finally, the type and number of tasks are also important factors for formula sequence processing. The greater the number of tasks, the higher the degree of language mastery required by the task type, and the more conducive to program language processing.

According to the summary of the literature in this article, the frequency effect inspires us to pay attention to repetition in textbook compilation and teaching. For example, English textbooks should contain a large number of high frequency, real formulaic sequences to increase exposure. As Bybee (2006) said, the use of high frequency makes the formula sequence lose its internal structure, and finally forms an overall sequence that is automatically stored and retrieved (Bybee, 2008). This phenomenon is slowly formed in the repeated practice of high frequency. Secondly, semantic transparency is an important factor that affects learners' processing of corpus. Specifically, teachers should use different teaching strategies to teach these two formulaic sequences. In addition, the language proficiency effect indicates the importance of grammatical knowledge and language experience. While increasing English input, learners should also strengthen the learning of grammar knowledge. For language learners of different English proficiency levels, it is necessary to teach students in accordance with their aptitude. For example, for low-level students who are not majors in English, procedural language learning should be strengthened to help them learn English. Finally, for different types of tasks, English learners have different performances in procedural language processing. In English teaching, we should pay attention to multi-task collaborative teaching, and formulate special teaching tasks according to the specific requirements of different teaching difficulties and key points.

Acknowledgement

References

[1] Bybee, J. (2006).From usage to grammar: The mind's response to repetition. Language, 82.

[2] Bybee, J.(2008).Usage-based grammar and second language acquisition. *Handbook of Cognitive Linguistics and Second Language Acquisition.* New York: Taylor &Francis.

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

[4] Wray A. Formulaic language and the lexicon [M]. Cambridge, Cambridge University Press, 2002, pp.09.

[5] Cao Y. The influence of one-language semantic knowledge on Chinese English learners' English Formulaic language processing [J]. Foreign Language and Foreign Language Teaching, 2018(05):109-117+150.