

Research Progress on the Psychological Reality of Formulaic Sequences: Evidence from Multiple Research Paradigms

Cui Ma

School of Foreign Languages , Zhengzhou Normal University , Zhengzhou, Henan, 450000 China

Abstract: Psychological reality of formulaic sequences is that the processing and storage of formulas are represented in a holistic way in mental lexicon. This paper, based on psycholinguistic approach, combined with linguistics characteristics, including phonetics, semantics, components and the visual as well as auditory holistic processing, surveys empirical evidence for the psychological processing of formulas and different subtypes of the combination of multi-units (components and phrase-level) from the perspective of reaction time (RT), eye-tracking, E-prime, and ERP. The multi-disciplinary and multi-paradigm psychological reality research of formulas will verify the nature of holistic representation of formulaic sequences in a more comprehensive and systematic way.

Keywords: Formulaic Sequences; Mental Lexicon; Psychological Reality; Holistic Representation

Fund Project:

2023 Special thesis Research on General projects of Humanities and Social Sciences Research in colleges and universities in Henan Province.

Introduction

Formulaic sequences are one of the basic units of language processing, and are combinations of several words that language users have lexicalized during language use or learning (Zhan Hongwei, 2009). Scholars in applied linguistics have a broad consensus on this, that is, formulaic sequences or “formulaic” language sequences are conducive to fluent, complete and accurate use of language (Pawley and Syder 1983; Wray 2002). This basic assumption is that prefabricated formulaic sequences are stored as a whole in the brain, without grammatical analysis, and can be used as a unit in the mental lexicon with minimal cognitive processing, recognition, and extraction, promoting rapid and accurate use (Wray 2002).

1. Reaction Time Paradigm

The reaction time paradigm is one of the main paradigms in the experimental study of psychological representation of formulaic sequences, and also one of the most commonly used and important dependent variables in other paradigms of psycholinguistics. Underwood et al (2004) believed that the internal components of formulaic sequences and the reaction speed of words at the end of formulaic sequences can reflect the overall processing advantages of formulaic sequences. Carroll and Coklin (2004) pointed out in their study of the processing advantages of idioms that the ending noun is the core component of the internal components of the formulaic sequences, and the reading speed is the fastest among all stages of measurement. Therefore, based on the processing process between the constituent words of formulaic sequences, the processing of each position word in the formulaic sequences word string is examined, not just the reaction speed of the final word. This can further understand which component word or words have an activating effect on the formulaic sequences of the language block. Based on this assumption, Schmitt and Underwood (2004) designed a self-paced reading experiment using E-prime programming to examine the reading response time of the target word in both formulaic and non formulaic sequences of the target word (NS vs NNS). The results showed that there was no significant

difference in the RT of the word by word self paced experiment between NS and NNS for the target word block in both formulaic and non formulaic sequences contexts. A self paced reading experiment based on the processing process between the components of formulaic sequences and examining the processing of words at each position in the formulaic sequences word string did not demonstrate the processing advantage of the overall formulaic sequences. Conklin and Schmitt(2008) improved the presentation method of self paced reading tasks, replacing word by word display with line by line display. The test materials included formulaic sequences and control word strings. The research results validate the holistic characteristics of the psychological representation of formulaic sequences. The above self paced reading experimental design based on the reaction time paradigm and different presentation modes of the internal components of three formulaic sequences confirms the hypothesis of the overall processing advantage of formulaic sequences.

2. Online Grammar Paradigm

Jiang&Nekrasova(2007) used online grammar judgment tasks to investigate the overall visual processing of formulaic sequences in response to their grammatical judgments. Based on language processing theory, grammar judgment of a word string can be divided into two parts: 1) identifying words in the word string; 2) Syntactic analysis and judgment. The holistic representation of formulaic sequences suggests that if formulaic sequences are processed and stored as independent units, the recognition of the correct order of words between their components can activate the holistic representation of formulaic sequences, without the need for further syntactic recognition. If the assumption is correct, the processing speed of formulaic sequences is faster than that of control word strings. Experiment 1 examines the reaction time and error rate of NS and NNS on both formulaic sequences and non formulaic sequences (with the same length and word frequency). Based on the online grammar paradigm, the grammaticality and non grammaticality of word strings, as well as the processing response time and error rate of functional and non formulaic sequences, are examined, verifying that the fast processing speed of formulaic sequences is essential for elementary word level recognition and saving the assumption of syntactic analysis of formulaic sequences. Experiment 2 further verifies whether the processing speed of formulaic sequences is due to the saving of syntactic analysis on formulaic sequences, or is it caused by the visual form (familiarity) presented with formulaic sequences, in order to control the impact of the visual form of formulaic sequences on word recognition. All experimental materials are presented in capital letters. The experiment controls the visual form presented by formulaic sequences, ruling out the possible impact of word recognition advantages on online grammar judgment tasks, and proving the essence of psychological reality of formulaic sequences.

3. Eye Tracker Paradigm

Based on eye tracker research, the psychological reality of formulaic sequences is mainly revealed from the various components of the formulaic sequences (the ending words of formulaic sequences). Underwood et al(2004) argued that the reaction speed of words at the end of formulaic sequences can reflect the overall processing advantages of the blocks. The study examined the number and duration of fixations of NS and NNS on final words in both formulaic and non formulaic sequences environments using an eye tracker. The experimental results indicated that in terms of fixation frequency, NS had significantly less fixation frequency on the last word than NNS. From the perspective of fixation time, NS has a shorter fixation time for the target word of formulaic sequences compared to NNS. The data on the number of fixation times and time of final words in both formulaic and non formulaic sequences environments provide evidence for the psychological reality of the overall storage and processing of formulaic sequences. From the record of eye movements of final words, the overall processing of formulaic sequences can be inferred.

Carrol and Conklin(2019) studied and compared the reading patterns of three different types of formulaic sequences (idiom, binomial, collocation) under three reading conditions (formulaic sequence, controlled word string 1, controlled word string 2). Research has shown that in a series of eye tracking tests, three different types of formulaic phrases have varying degrees of formulaic sequences processing advantages (RT) compared to the corresponding two control phrases. Compared with the corresponding two types of control groups, the three types of variants of formulaic sequences show processing advantages in the dimension of word skipping at the end. The ending word is the core component of formulaic sequences. Word skipping is assumed to be at least partially recognized or processed by the visual fovea, so word components that are known as formulaic sequences are likely to be completely skipped or reduced in reading fixation time.

4. Research Paradigm of Event Related Potential(ERP)

In the study of multi word units, such as idioms, binomial, collocations and word combinations with corresponding creation and different semantic transparency, it is proved that N400 amplitude is related to easier semantic processing and integration. Siyanova et al(2009) provided evidence for the psychological reality of formulaic sequences from the ERP perspective by studying the processing of

binomial expressions. The experiment presents the subjects with word strings under three conditions, namely, binomial phrases, inverted word strings, and abnormal word strings. The experiment assumes that if the binomial phrase is used as an independent representation unit of formulaic sequences, then the N400 amplitude of the three word strings is different. Moreover, from binomial phrase to "inverted word string" and then to "abnormal word string", it gradually increases. The experiment verifies the hypothesis that the N400 amplitude of abnormal word strings is the largest, because the deviation between context and semantics is the largest, and the amplitude of inverted word strings is greater than that of binomial, indicating that the frequency of inverted word strings is low, cognitive processing is difficult, and it is difficult to integrate with context. Binomial phrase is regarded as a whole representation of memory unit, which has psychological reality. In the study of the psychological processing mechanism of multi word units, it has been proven that N400 amplitude is related to familiar and easy semantic processing and integration, but it is not the only ERP component that processes highly predictable information. Siyanova et al (2017) provided evidence for the processing advantages of formulaic sequences through the ERP component analysis of the addition process of binomial phrases. Three groups of word string materials were presented, including binomial phrase, low frequency and similar word string with connection strength, and semantic violation of word string. The results showed that the binomial triggered a larger P300 and a smaller N400, reflecting the activation of templates matching predictable information (P300) and easier semantic integration (N400).

5. Conclusions

The research on the psychological reality of formulaic sequences mainly focused on the phonological, semantic, component processing, and overall visual form processing of formulaic sequences. In terms of research methods, it involved reaction time paradigm, online grammar paradigm, eye tracker paradigm, and ERP paradigm. Most studies have demonstrated the psychological reality of holistic representation and processing of formulaic sequences (Carroll & Conklin 2019). The multi paradigm study of formulaic sequences collects multidimensional research evidence, and different methods promote and supplement each other, deepening a more comprehensive and in-depth understanding of the processing mechanism of formulaic sequences, thereby promoting the acquisition of multi word units in second language.

References:

- [1] Jiang, N. & T. Nekrasova. 2007. The processing of formulaic sequences by second language speakers. *The Modern Language Journal*, 91(3): 433 – 445.
- [2] Siyanova-Chanturia, A., Conklin, K., Caffarra, S., & Kaan, E. 2017. Representation and processing of multi-word expressions in the brain. *Brain and Language*, 175, 111-122.
- [3] Wray, A. 2002. *Formulaic language and the lexicon*. Cambridge: UK Cambridge University Press.