A Cognitive Study of College Students' English Vocabulary Based on Electroencephalogram

Zefen Zhang

ABSTRACT

In this study, a vocabulary cognition experiment is designed with cross-language startup paradigm from the cognitive perspective to analyze the time course of English vocabulary semantic accessibility of college students, the influence of words and speech information on English semantic accessibility, and the relationship between English learning and Chinese learning and their mutual influence on the cognitive process. The experiment uses event-related potential technology to record the sensory and cognitive processes of the brain online in a continuous and real time manner. The components of the scalp area are separated to determine the cortical areas activated by the mental activity, thus determining the internal nerve source of the explicit behavior. This study discusses English vocabulary learning from the perspective of cognitive linguistics, mainly to understand learners' cognitive characteristics and cognitive process of English vocabulary learning. In view of this phenomenon, it is proposed that learners should establish the cognitive schema of English vocabulary to form the semantic network of English vocabulary, overcome the negative transfer of mother tongue, and achieve vocabulary cognition through a great deal of English input and mastery of context knowledge. This study complements the research on vocabulary cognition of college students, and the results provide theoretical guidance for college students' English learning and language teaching.

Key Words: Electroencephalogram (EEG), Vocabulary cognition, Cognitive Linguistics, Event-related Potential

DOI Number: 10.14704/nq.2018.16.5.1299 **NeuroQuantology 2018; 16(5):363-368**

Introduction

To explore the mechanism of vocabulary and semantic change and the cognitive process of vocabulary cognition from the perspective of cognition is another new way to study vocabulary cognition, better deepening the understanding of acquired vocabulary knowledge (Davis *et al.*, 2017). The study analyzes the relationship between cognitive linguistics and vocabulary cognition, and uses cognitive schema theory to study Chinese English learners' cognitive process of English vocabulary learning, trying to find out how English learner perceive vocabulary based on cognitive schema, explain the influence of learners' mother tongue concept knowledge on English vocabulary cognition in vocabulary learning, seek to promote learners' understanding, memory and cognition of English vocabulary knowledge, and overcome the negative transfer of mother tongue, so as to improve learners' foreign language learning efficiency and language use ability (Lawrence *et al.*, 2012). Since modern times, with the rapid development of science and technology, a large number of techniques for measuring the brain without human injury have been born. The main brain imaging techniques include MEG, FMRI, ERP, PET and NIRS (Adeniyi and Lawal, 2012). Each of these techniques has its own advantages. In the actual experiment, the researchers use these techniques to complement each other and obtain great success. A lot of excellent research

Corresponding author: Zefen Zhang
Address: School of Foreign Languages, Henan University of Chinese Medicine, Zhengzhou 450008, China
e-mail ✉ eshui2010@163.com
Relevant conflicts of interest/financial disclosures: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
Received: 14 March 2018; **Accepted:** 13 April 2018

results have greatly promoted the research of brain cognition.

In this study, a vocabulary cognition experiment is designed with cross-language startup paradigm from the cognitive perspective to analyze the time course of English vocabulary semantic accessibility of college students, the influence of words and speech information on English semantic accessibility, and the relationship between English learning and Chinese learning and their mutual influence on the cognitive process. The experiment uses event-related potential technology to record the sensory and cognitive processes of the brain online in a continuous and real time manner. The components of the scalp area are separated to determine the cortical areas activated by the mental activity, thus determining the internal nerve source of the explicit behavior (Liu, 2011).

The language cognitive psychology model is shown in Figure 1.

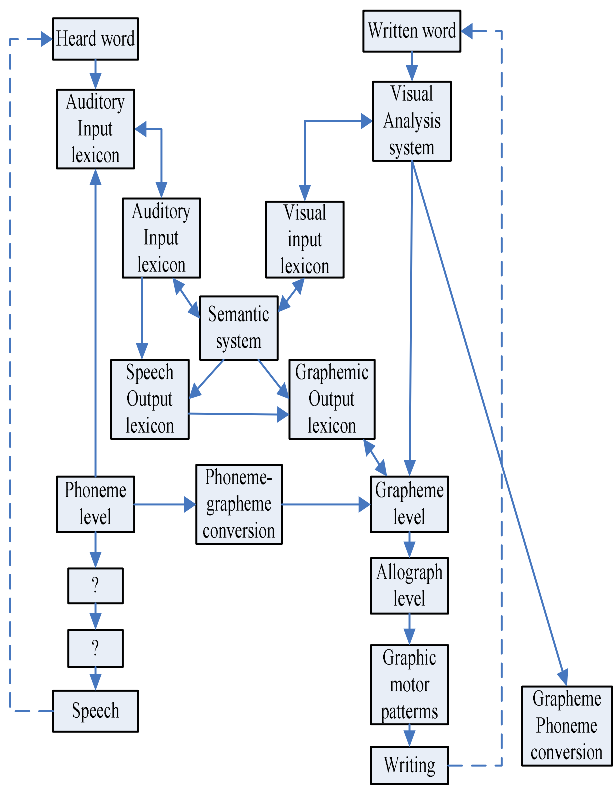

Figure 1. Linguistic cognitive psychology mode

The principle and experimental technique of ERP

Because the brain is a good conductor of electrical signals, the electrical activity of neurons in the brain can be transmitted to the cerebral cortex and captured by recording instruments. However, the signal strength of individually active neurons is too weak, and only the same electric field direction and the sum of synchronized neuronal activities can have sufficient signal strength to be captured and recorded. Therefore, it can also be said that electrical brain signals recorded on the scalp are in the electrical activity of the brain neurons with a certain distance apart, and mostly they are only part of the situation. The intensity of voltage collected on the scalp is influenced by the distance factor, that's, whether the position point of collecting the signals is consistent with or close to the neurons (Bobde *et al.,* 2018). Neurons in the brain are divided into two types by distance from the scalp: near-field sources and far-field sources. Among them, the potentials evoked by the primary somatosensory belong to the former, with a strong electric field due to near the scalp, mainly located in the central posterior gyrus, and the amplitude reaches the peak at the top. The latter type is represented by brainstem auditory evoked potentials. The amplitude of EEG signal that can be detected is very small because the endogenous signal of brain is far away from scalp and some are located in brain stem.

ERP is formed by the electrical activity of neurons in the brain caused by stimulation events, resulting in voltage changes. ERP can reflect the brain processing (brain mechanism), different ERP shows the different brain processing. However, if the same ERP is presented under different conditions, it does not mean that the brain mechanisms are the same, but it may be the result that the changes of neuron electric field direction or activity intensity offset each other. In addition, the ERP extracted by capture is represented by a potential waveform, and the same waveform can be formed by any number of endogenous brain sources (Rakshit *et al.,* 2017). The obvious advantage of using ERP for studying brain processing: first, the time resolution of ERP is millisecond, which is much better than fMRI and PET, which can be imaged in three dimensions; second, the spatial resolution of ERP technology is in centimeters. However, due to the development of high-density conductance, EEG cap can reach 128 or even 256 conductance, which greatly enhances spatial resolution accuracy. At the same time, dipole tracing technology also provides a good tool for localization (Wen *et al.,* 2015); third, the recording of ERP technology is multi-dimensional, which can describe brain activity from the aspects of polarity, amplitude, latency period, the range of scalp distribution and so on. These unique advantages of ERP provide technical support for the research and development of brain and cognition, and also establish the use state and good prospect of ERP technology (Khosrowabadi *et al.,* 2014). The processing steps of ERP experiment are shown in Figure 2 below:

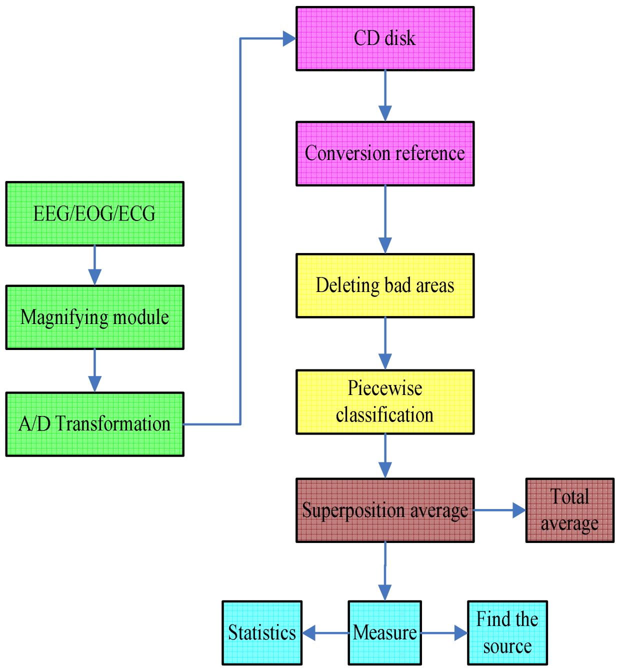

Figure 2. ERP experimental processing step

Methods

Subjects

There are 17 college students (9 males and 8 females) with an average age of 21.4 years and normal visual acuity or corrected visual acuity. Their native language is Chinese, they acquire English acquisition at the age of 12~13 years old, and all pass the College English Test Band 4.

Experimental materials

The experimental items include 40 pairs of Chinese-English translations, in which Chinese part are single words, and all of them are specific nouns that do not represent animals, such as "fire", with a frequency of 90-217 words per million. English is the translation of Chinese words, such as "fire", with a word length of 3 to 6 letters and high familiarity. In addition, 40 pairs of translation words representing animals are selected as complementary items.

Experimental design

This experiment is a single factor 3 level, and language type of target words: Chinese and English words are designed by subjects. 40 experimental items and 40 complementary stimuli in Chinese are set as one test group, and the stimuli in each test group are randomly presented. The three languages are presented in an identical manner, and the materials are presented separately, and the order of presentation is balanced among the subjects.

Experimental program and subjects' task

After the electrode installation, the subjects seat in front of the computer screen with a visual distance of about 80cm. The subjects are presented with a point of regard of 500 milliseconds, followed by a target word. English target words are of the bold 32-number Times New Roman and Chinese target words are the bold 32-number Song characters. The target words are presented in the center of the screen with a maximum viewing angle of no more than 2.5. The subjects judge whether a target word represents an animal: if yes, press the left mouse button; and if not, press the right mouse button.

Recording of behavior data and ERP data

The contact resistance between the electrodes and scalp is less than 5 Kilohms. The reference electrode is the center point. One electrode is arranged on the outside of both eyes to record the horizontal and vertical eye electricity. The filter band pass is 0.1 ~ 100Hz and the sampling frequency is 250Hz. The continuously recorded original EEG data is subjected to 30Hz low-pass filtering off-line, and then segmented according to the standard of 200ms before and 1000ms after the target word is presented, and after removing the data segments containing artifacts such as eye movements, the data segments under the same conditions are averaged. After baseline correction and whole brain average, the average waveform of each subject under each experimental condition is obtained.

Data analysis

Behavior data: Conduct statistical analysis of response time and accuracy data of target words. The time course of ERP data analysis is 1000 ms since the target word is presented, and the baseline is 200 ms before the target word is presented. The total average waveform of ERP is P120 (100~130ms), P220 (180~240ms) and N400.

Results

Behavior data results

Using SPSS13.0 statistical software, t-test is performed on the average response time and error rate of subjects. Prior to formal statistical processing, delete data with response time of less than 250 ms and greater than 2500 ms from raw data, replace data with data with response time of less than 250 ms with that of 300 ms, and replace data with data with response time of greater than 2,500 ms with data with the average response time of the subject plus 2 standard deviations. The average response time and error rate are shown in Table 1. Table 1 shows the response time (milliseconds) and error rate (%) of the subjects' target word semantic classification task, and the standard deviation of the mean value of the second behavior.

By paired t-test on the response time data, the difference between Chinese and English words is significant ($t(18)=-11.87$, $P<0.001$. Paired t test is carried out for the error rate of data, and the difference among the three is extremely significant, $t(18)=3.93$, $P<0.001$. According to the analysis of the behavior data, we can see that in terms of the response time Chinese words are shorter than English words and in terms of the error rate Chinese words are lower than English words. From the behavior data, it can be seen that there are obvious differences among the three. In order to identify the sources of these differences, we further analyzed the ERP data.

Table 1. Reaction time and error rate table

Language	Chinese	English
Reaction time	522	859
	20.6	36.9
Error rate	1.61	6.24
	0.69	1.03

Table 2. Cz point wave amplitude index

Language	Chinese	English
P120	2.64	3.88
P220	3.16	4.94
P400	-4.9	0.46

EEG data results

Table 2 is the average amplitude of Cz points P120 (100~130 milliseconds), P220 (180~240 milliseconds) and N400 (300~500 milliseconds) components. Figure 1 is the total mean ERP waveform of three language types of Chinese and English target words Fz, Cz, and Pz points. Figure 2 is the topographic map of three language types of Chinese and English target words in 120, 220, and 400 milliseconds, respectively. The average amplitude of the above 3 components on 19 electrodes is analyzed.

In the experimental task, the response time of subjects and the accuracy of judgement are the two key reference indicators of behavioral data. They indicate whether subjects understand stimuli and the extent of understanding in terms of language cognition. In this experiment, the direct results can be presented in Figure 3 and Figure 4.

It can be seen from the figure that, firstly, comparing the differences between different languages in judging the consistency of word meaning, we can see that the average response time of English and Chinese words is about 600 milliseconds and it can be preliminarily concluded that the way of learning English for college students is to rely on the Chinese cognitive way though there are many similarities between English and Chinese, which belong to phonemic characters; secondly, comparing the different types of stimulus materials, it is found that the response time of the two languages with equivalent semantic is shorter than the semantic unequivalence. In particular, by averaging the data, it is found that the response time of subjects to the experimental materials of English and Chinese translation words is about 100 milliseconds less than that of the non-translation words. This shows that when the semantic relations between the two languages are not consistent, and the difficulty of judgment will increase, so the subjects need more time to think and judge. According to the obtained data of standard deviation, the maximum difference is the judgment of English and Chinese materials with semantic equivalence. The standard deviation shows the difference in time of judgment among subjects, and the smaller the numerical value is, the more stable it is. In terms of the error rate of the judgment, the concrete sample data of ten subjects are visually displayed from the graph. By comparing the differences between different languages, the average judgment error rate of English and Chinese is 4.83%. From the error rate, this is consistent with the response time result. Secondly, by comparing different types of stimulus materials, it is found that the error rate of semantic equivalence in both languages is slightly less than that of semantic unequivalence. In particular, the error rates of the translated words and non-translated words of the experimental materials are 4.67% and 5.29% respectively, that's, it is less difficult to process and understand the translated words. Compared with the standard deviation data, the minimum difference is 0.68 and the maximum is 1.57. The results are similar and the numerical difference is not large, which shows a good stability.

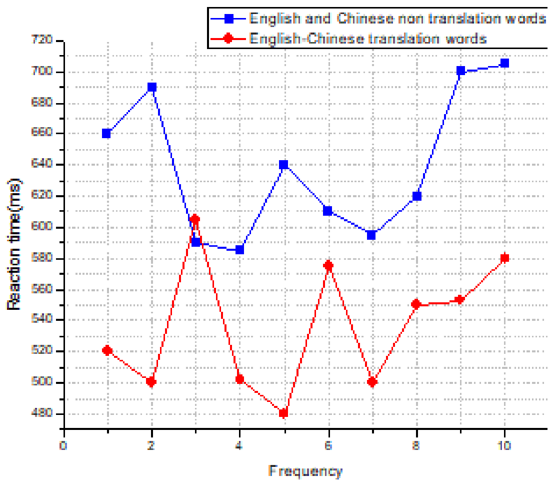

Figure 3. The response time of the word judgment task

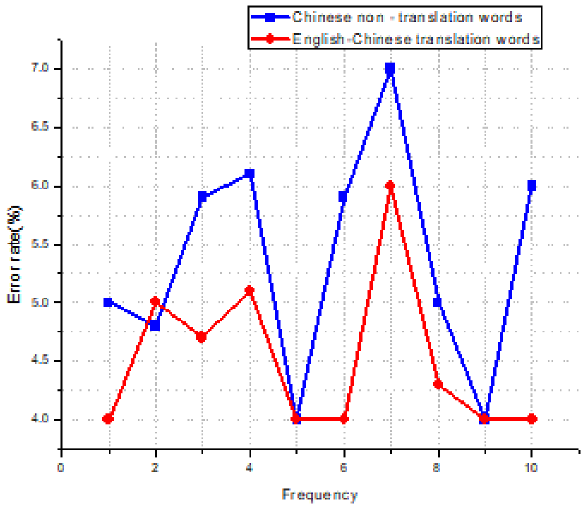

Figure 4. Error rate of word judgment task

Traceability analysis

LORETA is used to analyze the source location of ERP data. It divides the standard spherical head model into 239 pixel x10 millimeter deep area. In this range, the algorithm of seeking the smoothest solution is used to find the inverse operator and calculate the source of the skin electric activity in the brain. The total average waveform of ERP is input into LORETA, and the sampling period is 100-500 milliseconds after the start of stimulation presentation. The interval is taken from 100 milliseconds after the start of stimulation. Figure 3 is an activation diagram of three language types of target words. Table 3 is the maximum active area of different target words in the 100-500 millisecond time window. As can be seen from Table 3, in the 100-500 millisecond time window, the main activation areas of English target words are BA17, BA18, and BA19, and the main activation areas of Chinese target words are BA17, BA18, BA19, and BA6; the main activation areas of English target words are BA17 and BA18.

Table 3. Activation area list

Time	Chinese	English
100ms	Left BA18	Left BA17
200ms	Left BA18	Right BA18
300ms	Right BA19	Right BA18
400ms	Left BA16	Right BA18
500ms	Left BA17	Right BA18

Table 4. Comparison of reaction time length and error recognition rate

Language	English	Chinese
Average value	589.6ms	487.36ms
	4.93%	3.86%
Standard deviation	46.33ms	14.51ms
	0.93%	0.89%

This experiment uses the task of semantic classification to investigate the time course and brain activation of Chinese and English nouns with the same mental lexicon by Chinese-English bilinguals of the middle proficiency (Davis and Fan, 2016). The results show that the English processing speed of such proficient bilinguals is obviously slower with higher error rate, and their Chinese processing speed is middle with a small gap among them, which is proportional to the time of language mastery. From the perspective of language cognition, the direct results of examining whether or not the subjects understand the stimuli and the degree of understanding is shown in Table 4 below. According to the total average waveform of ERP, the waveform of English target words is more negative than that of Chinese target words, and especially in the N400 component, the average amplitude of English is significantly larger than that of Chinese target words. ERP records not only provide information about the time course of word processing, but also provide general spatial information. Combining with LORETA, it can provide an internal source of electrical activity of the brain in a specific period of time, which to some extent makes up for the low temporal resolution of fMRI and PET. However, the result of this experiment reveals to a certain extent that the language type of the target words varies over time. In the early stage of visual processing, English words produce more left occipital activation than Chinese words, and during the whole course of vocabulary recognition, English words produce more right cerebral activation than Chinese words (Alipoor and Jadidi, 2016). The results show that the brain areas activated by English words is very similar to the brain areas activated by Chinese characters in both tasks, but the left hemisphere is more strongly activated when Chinese native speakers process Chinese, while the right hemisphere is more strongly activated when they process English, which suggests that the right hemisphere has a greater role in processing a second language.

Conclusions

Chinese-speaking English learners are subconsciously influenced by Chinese vocabulary knowledge. When English vocabulary concepts are input into the brain, they interact with the concepts of native language words that originally exist in the brain of the learners. From the students' choice of words in translation, the concept of native language vocabulary in the brain of learners has a certain degree of solidification. To overcome this phenomenon, English learners should continuously activate and expand the original mother tongue semantic network and turn to English vocabulary semantic network. The new semantic information is input into the existing semantic network, and it is recognized, analyzed and edited by the brain, and connected with all elements in the network to form a more comprehensive lexical network. The structure of vocabulary network can not only connect and activate related words, but also facilitate the deep processing of words. In other words, after the English learners are familiar with a word, they can form a series of vocabulary and meaning related to the word, and construct the English vocabulary schema. Therefore, in English vocabulary cognition, English learners should not only pay attention to the literal meaning of a word, but should carry on the diffusion learning of the word. After activating the word, all aspects related to the word should be included in the scope of learning. English learners' native language meanings are recognized after a long period of imperceptible use. Therefore, English vocabulary meanings should also be experienced in an appropriate context, which requires time, situation, and a process of gradual recognition, a process of gradually establishing complex cognitive networks from the initial simple cognition. Along with the comprehensive understanding and application of English vocabulary knowledge, English learners will gradually develop semantic network and finally realize the cognition of English vocabulary.

References

Adeniyi FO, Lawal RA. Comparative effects of multisensory and metacognitive instructional approaches on english vocabulary achievement of underachieving nigerian secondary school students. International Education Studies 2012; 5(1): 18-27.

Alipoor R, Jadidi E. The Relationship between Iranian EFL Teachers · Reflection and their Cognition about Vocabulary Teaching Style. Procedia-Social and Behavioral Sciences 2016; 232: 769-75.

Bobde SP, Mantri ST, Patil DD, Wadhai V. Cognitive Depression Detection Methodology Using EEG Signal Analysis. InIntelligent Computing and Information and Communication Springer, Singapore, 2018: 557-66.

Davis GM, Fan W. English vocabulary acquisition through songs in chinese kindergarten students. Chinese Journal of Applied Linguistics 2016; 39(1): 59-71.

Davis JJ, Lin CT, Gillett G, Kozma R. An integrative approach to analyze EEG signals and human brain dynamics in different cognitive states. Journal of Artificial Intelligence & Soft Computing Research 2017; 7(4): 287-99.

Khosrowabadi R, Quek C, Ang KK, Wahab A. ERNN: A biologically inspired feedforward neural network to discriminate emotion from EEG signal. IEEE transactions on neural networks and learning systems 2014; 25(3): 609-20.

Lawrence JF, Capotosto L, Branummartin L, White C, Snow CE. Language proficiency, home-language status, and english vocabulary development: a longitudinal follow-up of the word generation program. Bilingualism Language & Cognition 2012; 15(3): 437-51.

Liu H. Vocabulary Learning Strategies For Communication Promotion-On English Vocabulary Learning Related to Listening, Speaking, Writing Capability/Lernstrategie der Englischwörter für die Funktion des Sprachkontakts-Fähigkeit für Hören, Mündlich, Schreiben und Wörterlernen. Studies in Literature and Language 2011; 3(2): 129.

Rakshit A, Banerjee A, Mazumder A, Ghosh P, Dey A, Tibarewala DN. Fractal analysis of EEG signals for studying the effect of cognitive stress on brain. International Journal of Biomedical Engineering and Technology 2017;25(2-4): 336-69.

Wen D, Zhou Y, Li X. A critical review: coupling and synchronization analysis methods of EEG signal with mild cognitive impairment. Frontiers in Aging Neuroscience 2015; 7(54): 54-64.