Morphological Decomposition in Second Language Word Processing

Guihua Cao
School of Foreign Languages, Hubei Engineering University, Xiaogan, Hubei Province, 432000, China

Abstract—This paper made a masked priming experiment to examine the Chinese learners’ online processing of English derivatives or derived words so as to find out the effect of morphological structures on second language (L2) word processing. 39 Chinese freshmen of non-English majors participated in the experiment. Results indicated that derivatives had significant priming effect on the identification of their stems and that this kind of priming effect was not affected by the brain’s familiarity with it, which meant that like L1, L2 derivative processing was also affected by the morphological structures and that morphological decomposition did exist. This conclusion supports the view of Diependaele et al. that the characteristics of the target words determined their processing mechanism. In addition, the main factors that affect the complicated processing mechanism prove to be the features which are composed of the morphemes instead of the features of the whole words.

Index Terms—L2 processing, morphological decomposition, masked priming, derived words

I. INTRODUCTION

“Decomposition” and “composition” is an essential characteristic of a language (Johnson 2004; Werning et al. 2012), which produces effect on various levels of a language, from the inner structures of vocabularies to the whole organization of texts. Due to this characteristic, languages are filled with regularities so that people can understand and use it, and at the same time it turns out to be the important means for language creations and universally exists in human languages, but is revealed in different ways in different languages. For example, at the vocabulary level, the composition of derivatives or inflectional morphemes which are composed of free morphemes or bound morphemes are quite popular in English, but seldom used in Chinese. It has become the mutual focus of the second language acquisition and psychological linguistic academic field that whether the differences of word composition between different languages affect L2 vocabulary learning and whether there is “decomposition” and “composition” in the processing of L2 complicated words (Basnight-brown et al. 2007; Silva & Clahsen 2008; Zhang 2014.)

Most of the researchers discussed about this topic via online behavior experiments, of which masked priming was employed most frequently. Such experiment pattern revealed the following typical characteristics. The time for the presentation of the priming words was short (within 80ms), and there were a series of masked symbols (e.g.#) before the presentation, which ensured the fact that the participants processed the words unconsciously and avoided any impact of intentional processing strategies and effectively revealed the characteristics of automatic processing at the early stage of the priming words presentation. A great number of observations of L1 learners which employed this mode found out that affected by the inner structures of the words, there was morphological decomposition in the processing of complicated words and this decomposition occurred at the early stage of the word processing, proved to be an automatic processing mechanism for complex words and not affected by the factors such as word familiarity (Rastle & Davis 2008; McCormick et al. 2009).

Compared with L1, L2 processing proved to be more complicated. On the one hand, the two processing mechanisms might be different in nature. On the other hand, the differences between learners’ L1 morphological structures and those of L2 might also produce effect on the processing. As for the processing of inflectional words, some researchers pointed out that L2 processing mainly depended on declarative memory and seldom used procedural memory. In other words, L2 processing was made via whole words and it was almost impossible for morphological decomposition to occur (Clahsen et al. 2010). This view was confirmed in some empirical researches (Silva & Clahsen 2008; Neubauer & Clahsen 2009; Clahsen & Neubauer 2010). Nevertheless, there were also some researchers which arrived at contrastive conclusions. For example, Basnight-Brown et al. (2007) observed Chinese EFL learners, Gor & Jackson (2013) observed the British learners of Russian and found morphological decomposition in both of them.

Up to now there have been few researches with regards to the processing of L2 derivatives and there are also differences between their conclusions. Silva & Clahsen (2008) found out that although there was no morphological decomposition in L2 learners’ processing of inflectional words, there was morphological decomposition in their processing of derived words. In addition, there were no significant differences between Chinese EFL learners and German EFL learners and their performance in the processing of both inflectional words and derived ones proved to be the same. Diependaele et al (2011) also found morphological decomposition in the processing of derived words for the Spanish and Netherlands EFL learners and the significant level proved to be equal to L1 processing. Hence whether the target language was learners’ L1 or L2 and what was learners’ L1 did not affect their processing of derived words. What
did affect it should be mainly the characteristics of the target language itself, such as the morphological regularities and semantic transparency (Diependaele et al. 2011). Nevertheless, Neubauer & Clahsen (2009), Clahsen & Neubauer (2010) arrived at the opposite conclusions; there was no morphological decomposition in Polish learners’ processing of German derived words and inflectional words, indicating that L2 learners seemed not to be able to make morphological analysis of the complicated words and that their processing had to be made merely via whole words (Clahsen et al. 2010).

Chinese researchers have also begun to conduct behavior experiments to explore the online processing of learners’ L2 complicated words. Chen (2009) and Chen et al. (2007) employed word naming task to examine learners’ online processing of complicated words and found different ways of processing for different categories of complicated words. In other words, derived words and compound words were processed via decomposition and inflectional words via whole words. Zhang’s (2014) word-judgment task found that Chinese EFL learners processed the derived words via dual routes-decomposition processing and whole words and that the choice of these two ways was affected by the relative familiarity (whole word familiarity/stem familiarity). The above researches were conducive to the understanding of the learners’ EFL online processing of derived words but failed to indicate the inner mechanism of the complicated words processing. Did the morphological decomposition in the processing of L2 complicated words occur automatically or intentionally? Would there be morphological decomposition in the processing of complicated words via whole words? Up to now there have been no answers to the two questions. Within China only Yue et al. (2012) used masked priming experiment to examine Chinese learners’ EFL processing mechanism. Yue et al. simultaneously examined learners’ processing of inflectional words and derived words and found no differences between the two, but both of them were affected by stem frequency. In another word, there occurred morphological decomposition in both inflectional words and derived words with high frequency stems and there was no morphological decomposition in inflectional words and derived words with low frequency stems.

The above mere one research was not sufficient to explain Chinese EFL learners’ inner processing mechanism of English complicated words. In addition, in the aspect of research design, Yue et al. (2012) fixed the time interval (SOA) between the priming words and the target words as 49 millisecond (MS). There was no doubt that such short period of time could ensure the fact that the participants processed the priming words unintentionally. But there was also some other possibility that due to the effect of word frequency and also the fact that L2 was processed slower than L1 (Clahsen et al. 2010), learners failed to perceive the existence of the words with low frequency within a short period of time and did not trigger the decomposition processing mechanism. Therefore if SOA was appropriately increased and the participants were ensured not to perceive the appearance of the priming words, the characteristics of L2 learners’ EFL processing would be examined in a more all-around way. Secondly, as for the familiarity, most of the previous researches employed the word frequency data from L1 corpus as reference index for familiarity, which seemed to be not fair for L2 learners, in particular those with low LFL proficiency (Zhang, 2014). Nevertheless, merely either whole word frequency or stem word frequency was taken into account instead of the effect of both. The previous researches had confirmed that compared with the whole word or stem familiarity, the effect of the comparative value for both of them (comparative familiarity) proved to be more effective (Marcolini et al. 2011). Of the researches within China, only Zhang’s (2014) employed participants’ evaluation to obtain the familiarity index for the experiment materials and simultaneously took the effect of whole word familiarity and stem familiarity into account. In view of that, this research intends to use Zhang’s materials and the masked priming experiment with 80 millisecond SOA in order to examine the morphological decomposition in Chinese EFL learners’ processing of derived words and further reveal the effect of morphological structures on L2 word processing. The research questions to be answered include the following two: (1) Is there morphological decomposition that automatically occur in the Chinese EFL learners’ processing of derived words? (2) Is the decomposition affected by the comparative familiarity of the derived words and their stems if the answer to the first question is yes?

II. METHODS

A. Participants

The participants in this research came from the same learner group as Zhang’s (2014) who were the non-English majors in their first year from a university in Wuhan, Hubei province, China (with 23 male students and 16 female ones). They had studied English for 7 or 8 years and were enrolled into the same class according to their English proficiency and there were no significant differences between the individuals (p>0.05). Their eyesight or CVA (corrected visual acuity) was normal and all of them participated in the experiment of their own accord.

B. Design

This research took simple words as the target ones and fixed three different priming conditions for each of the target word: derived priming (derived words with priming words as target words, e.g. speaker-speak), repeated priming (the priming words are the same as the target ones, e.g. speak-speak) and irrelevant priming (the priming words are not related to the target words in terms of both morphology and meaning, e.g. tomorrow-speak). The research would reveal whether there was morphological decomposition in the processing of EFL derived words via comparison between the facilitation effect of derived priming conditions and that of the other two priming conditions on the target
words identification, and examined the reaction time (RT) and the accuracy rate (AR) of the target words as variables. The hypothesis for this research is that if the derived priming conditions produced significant effect on the target words and the effect was the same as that of the repeated priming, namely RT \( \text{derived priming} < \text{RT} \) \( \text{repeated priming} \), AR \( \text{derived priming} < \text{AR} \) \( \text{repeated priming} \), the derived words had the same priming effect as their stems, called full priming. Full priming indicated that in the processing of the derived words the psychological representations of their stems were activated and proved to be the more convincing evidence for the morphological decomposition. If the derived priming had some priming effect at certain degree but could not reach the level of the repeated priming, or RT \( \text{derived priming} < \text{RT} \) \( \text{repeated priming} \), AR \( \text{derived priming} < \text{AR} \) \( \text{repeated priming} \), it revealed that the appearance of the derived words did have some facilitating effect on the processing of their stems, called partial priming. The existence of the partial priming effect also indicated that there was morphological decomposition in the processing of the derived words. Nevertheless, it could not reach the level of the repeated priming since L2 learners were not familiar with the derived words and their automation proficiency was low, or RT \( \text{derived priming} = \text{RT} \) \( \text{repeated priming} \), AR \( \text{derived priming} = \text{AR} \) \( \text{repeated priming} \), indicating that there was no morphological decomposition in the processing of the derived words.

According to the comparative familiarity, this research re-divided the derived words into high ones (the whole words had low familiarity and the stems had high familiarity) and low ones (the whole words had high familiarity and the stems had high familiarity) and examined them respectively. Therefore this research used the design of 2 (familiarity: high and low) \( \times 3 \) (priming conditions: derived priming, repeated priming and irrelevant priming).

C. Materials

This research also used the same experiment words as Zhang’s (2014), of which the derived words were used as the priming words for the derived priming, the simple words correspondent to the derived words as the priming words for the irrelevant priming and the stems of the derived words as the target words. These derived words were made via the combination of the stems and the highly productive suffixes such as \(-\text{er}, -\text{ful}\) and \(-\text{able}\), and the morphology and pronunciation of the stems did not change in the course of the derivation. In addition, as the participants in this research came from the same learner group as Zhang’s (2014), Zhang’s result for the familiarity evaluation could be used once again.

This research also controlled the semantic relation and the morphological similarity between the priming words and the target words, and semantic relation was obtained via the previous semantic relation evaluation before the experiment. The experts who participated in the evaluation were familiar with all the experiment words. The evaluation instrument was 5-level Likert scale (one point for no relation, two points for little relation, three points for relation, four points for strong relation and five points for fairly strong relation). The evaluation result indicated that there was strong semantic relation between the derived priming words and the target words, but there was almost no relation between the irrelevant priming words and the target words. The morphological similarity was obtained by means of the letter coincidence rate between calculating priming words and the target words. Result indicated that there was significant similarity between the derived priming words and the morphology of the target words and there was no significant similarity between the irrelevant priming words and the morphology of the target words.

The researcher divided all the experiment items into three groups and each participant used merely one of the group in the experiment. Each target word appeared only once in each group and the numbers of appearance for the different priming types were the same in each group. At the same time 60 pairs of distractors for irrelevant priming were added in each group. Hence the proportion for the repeated priming items and the derived priming items was lower than 14% in each group, and greatly decreased the possibility that the participants formed the correspondent processing strategies. In addition, this experiment also consisted of 100 pairs of coined (invented) distractors (the priming words as the real words and the target words as the false or coined words). These coined words were made by deleting a letter of the real words and were the same as the target words for the real items in terms of the length and the familiarity. The coined distractors were also divided into the derived priming, repeated priming and the irrelevant priming and their composition proportion were the same as those of the real items. Accordingly each group consisted of 200 pairs of experiment items. The researcher equally divided them into 4 sections, and within each of the sections the experiment items were arbitrarily ordered with no disturbance between the sections.

D. Procedures

The experiment was conducted in a quiet room and only one participant was allowed to enter the room each time and the whole experiment was finished on the same computer. The experiment procedures used DMDX (Forster & Forster, 2003). During the experiment, the background color of the computer was light grey, the stimulus were black with word size of 24 and font Arial. To avoid the effect that the morphological similarity might cause, the priming words were presented in lowercase and the target words in capital letters. Before the formal experiment, the participants did the exercises for 15 items and the exercises might be repeated until the participants had been familiar with the correct ways and the procedures. When the exercises ended, the participants pushed “Enter” and began the experiment. The order for the formal experiment was as the following: fixation point “+” 250 millisecond, masked symbol (12 “#”) 500 millisecond, priming words 80 millisecond and the target words 600 milliseconds. The participants were required to
make the right or wrong spelling judgment for the target words within the shortest time and push the key (“L” for right and “A” for wrong). When the judgment ended, the next item began. If the participants failed to make the judgment within 5 seconds, the next item began automatically. The interval between the two experiments was 3 seconds. During the experiment, there were 15 adapted items before each section began and the participants might take a rest by themselves after each section was finished.

The masked priming experiment required the participants to process the priming words unconsciously or the participants were not perceptive of the priming words. To ensure the fact that this experiment could satisfy the above requirement, each participant had a visual test of the experiment materials after the experiment was finished. The test was composed of the 15 priming words and target words for the real filling items in the experiment which ranked arbitrarily and the participants were required to answer whether they had seen these words or not during the experiment.

E. Results

Before the data analysis, the researcher deleted the data for the participants and items with high error rate (consisting of 4 participants with more than 20% of the error rate and 8 items with more than 30% of the error rate), the extreme data (5.4%) with more than ±2 of standard deviations for the reaction time (RT) and reserved the data with correct response. Then the researcher made a standard logarithmic transformation of the data before the analysis so as to decrease the effect that might be caused by the data distribution skewness.

The data analysis used the mixed effect model (Baayen et al. 2008), and the participants and the items were added to the model as random variables. In view of the fact that the materials might be distributed unevenly, the researcher introduced the material groups and the item presentation order into the model when it was initially created. Yet its effect seemed to be not significant, could be neglected and was accordingly deleted. The later model took the priming type, the relative familiarity and their interactivity as the main effect. Analysis of the model found that the interactive effect between the priming type and the relative familiarity was not significant (repeated priming x relative familiarity was low: t=0.34, p=0.73; irrelevant priming x relative familiarity was low: t=0.21, p=0.83) and was deleted.

Finally the researcher merely reserved the priming type as the main effect for the analysis and found significant effect. Analysis of the data for reaction time indicated that there were significant differences between the repeated priming and the irrelevant priming, between the derived priming and the irrelevant priming (repeated priming minus irrelevant priming: t=5.36, p<0.001; derived priming minus irrelevant priming: t=4.44, p=0.00), but no significant differences between the derived priming and the repeated priming (t=1.03, p=0.30). The result for the analysis of the accuracy rate was similar to the above but the differences were not that significant. Merely the differences between the repeated priming and the irrelevant priming were significant at above 5% (Z=2.23, p=0.02). The accuracy rate for the derived priming was only a bit higher than that for the irrelevant priming and not significant (Z=0.75, p=0.46). The difference between the derived priming and the repeated priming was not significant either (Z=1.56, p=0.12). The above result indicated that the derived priming obtained the same full priming effect as the repeated priming, meaning that there was morphological decomposition in the Chinese EFL learners’ processing of the derived words.

The result for the participants’ visual test of the experiment materials after the experiment indicated that 77% of the target words were believed to have appeared in the experiment. Such a high rate revealed that the participants did have seen these words and that they had processed them consciously. More 9.2% of the priming words were believed to have appeared. Such low rate indicated that the participants did not really find them. And these priming words were chosen absolutely due to the participants’ guessing. Hence the SOA in this research met the requirements for the masked priming experiment.

III. DISCUSSION

This research used masked priming experiment to examine the morphological decomposition in the Chinese EFL
learners’ processing of the derived words. Result indicated that the derived words with both high and low comparative familiarity produced significant priming effect on the processing of the stems and the strength was similar to the repeated priming, called full priming. This finding has provided convincing evidence for the morphological decomposition in the Chinese EFL learners’ processing of the derived words and further confirmed the effect of morphological structures in the processing of L2 complicated words.

It should be noted that in the priming experiment, the derived priming effect was totally different from the priming effect caused by morphological similarity or semantic relations. On the one hand, the derived priming was affected by its morphological structures in the processing of the derived words and thus decomposed, or the psychological representations of its stem was activated. In other words, the processing of the derived priming words and the target words was an understanding of the same representation unit, which was exactly the same as the repeated priming in nature. Hence it could achieve the full priming effect which was equal to that of the repeated priming, see the test result for the reaction time data in this research. However, the priming caused by morphological similarity or semantic relations was due to the close relations between the two in the psychological lexicon network and the understanding of the representation unit for the priming words increased the activation level of the representation unit for the target words, and made them easier to be identified. Nevertheless, as the processing of the priming words and the target words on this priming condition was an understanding of the different representation units, it was impossible for the priming effect to reach the strength of the repeated priming. On the other hand, a great number of researches on L1 have found that morphological decomposition was a processing mechanism for the complicated words that occurred automatically. This mechanism was activated at the early automatic stage of word identification and would occur even if the participants were unconscious of it (Rastle & Davis 2008), while the priming effect caused by morphological similarity or semantic relations would occur only at the conscious processing stage and it was hard for it to happen in the masked priming experiment when the priming words were presented in a very short period of time (Rastle et al. 2004). In this research the visual test of the materials used after the experiment indicated that the participants did not perceive the existence of the priming words, and therefore there would be no morphological or meaningful priming. Accordingly the derived priming effect found in this research should only be the result for the morphological decomposition of the priming words. However, the measurement result for the target words’ accuracy rate in this research failed to reveal their significant priming effect that was similar to the data for the reaction time (RT) and even the repeated priming effect was not that significant. Hence the researcher believed that it was due to the presentation of the target words in capital letters. After the experiment, the participants confirmed that they had not been adaptive to the reading of the words in capital letters and committed errors in their judgment during the experiment so that the accuracy rate was decreased. In spite of that, the data for the reaction time (RT) indicated that there was morphological decomposition in learners’ processing of the derived words.

The above conclusion coincides with those from Silva & Clahsen (2008), Diependaele et al. (2011) and Yue et al. (2012), revealing that L2 learners could acquire the decomposition processing mechanism for the derived words and made an online analysis of their morphological structures. This conclusion has provided future evidence for the view of Diependaele et al. (2011) that the characteristics of the target words determined their processing mechanism. It also indicates that the familiarity had no significant effect on the priming effect, and that the derived words with both high and low relative familiarity revealed their full priming effect. Previously L1 research by McCormick et al. (2009) and L2 investigation by Otto (cited in Zhang & Zhu 2013) also arrived at similar conclusions. The author of this research also observed that it was decided by the nature of the morphological decomposition. As an automatic processing mechanism for the derived words, the morphological decomposition was also activated when the brain perceived the morphological structures (Rastle & Davis 2008). Some researches found that this decomposition even occurred in the processing of the simple words (e.g. corner) which consisted of the similar morphological structures, while Yue et al. (2012) found that the morphological decomposition merely existed in the processing of the complicated words with high frequency stems instead of those with low frequency stems, which might be simultaneously affected by both word frequency and SOA. The higher the word frequency was, the higher level the words were activated and the easier they would be identified and vice versa. Hence when SOA was short, the high frequency stems would be also identified and activated the decomposition processing mechanism, while when the activation level for the low frequency stems was low, it took more time to identify them and the short SOA might not be sufficient for the participants to recognize them among the whole words and unable to activate the decomposition processing mechanism. This research appropriately increased SOA and made it possible for the stems with low familiarity to be identified and accordingly found the popularity of the morphological decomposition.

Although the derived words and their stems had different familiarity, their suffixes had high level of productivity. Therefore no matter how the learners were familiar with the derived whole words and their stems, they could identify the letter combination at the end of the words as suffixes, which also promoted the morphological decomposition at certain degree. In might be concluded that whether there was decomposition in the processing of complicated words depended on the features of the decomposition morphemes instead of those of the whole words, which coincided with Yue’s (2012) view of the stem frequency effect: the higher frequency of the stems was, the more possible they would be identified as the morphemes and the more easier the decomposition processing mechanism would be activated; in contrast, the lower frequency of the stems was, the less possible they would be identified as the morphemes and the
more possible the decomposition processing mechanism would be activated.

The words used in this research were the same as those in Zhang’s (2014). Zhang found that the derived words with low familiarity of stems and high familiarity of whole words were processed via whole words, while this research indicated that there was morphological decomposition in the processing of these words. These two conclusions were not contradictory since Zhang (2014) investigated the whole process of the processing while this research examined the early stage of the processing. The possibility might be that there was morphological decomposition in these derived words at the early stage of the processing and provided decomposition route for their processing, but the whole word representation for these words existed in them and the processing route for the whole words could also be used, which made it possible for the competition between the whole word processing route and the decomposition processing route. As the whole words were superior in terms of familiarity, the whole word processing route won in the competition, which was the whole word processing found in Zhang’s (2014) research and also coincided with the view of double processing routes proposed by Schreuder & Baayen (1995).

The morphological decomposition mechanism in the processing of L2 complicated words might be the result affected by the decomposition and combination which was popular in language processing. Although there are not many derived words in Chinese, there is a great number of compound words among them. The composition mechanism of the two is very much similar and most of the Chinese derived words are composed of two or more morphemes, which means that the Chinese learners might have acquired the meta-cognitive ability of morphological decomposition in learning and using Chinese (L1) (Clark et al. 2011). When they learn L2 (EFL), this ability might be transferred and used in the analysis of L2 derived words. The other possible explanation might be that the learners used the word-building knowledge a great deal in the explicit learning and the explicit knowledge was gradually internalized and formed the implicit processing mechanism of morphological decomposition and combination.

IV. Conclusion

This research employed masked priming experiment to examine the Chinese learners’ processing of EFL derived words and found that the derived words produced significant priming effect on their stems and the effect was not affected by the level of familiarity, indicating that like L1 processing, the processing of L2 derived words might be also affected by the morphological structures and according morphological decomposition occurred, which coincided with the view of Diependael et al. (2011) that the characteristics of the target words determined their processing mechanism. This research further points out that what mainly affects the processing mechanism for the derived words prove to be the characteristics of the complicated word composition, not the features of the whole words.

There might be some limitations for this research for various reasons. Firstly, the materials or words used in this research were all derived words with semantically transparent suffixes which had high productivity of suffixes. These features might have provided conditions for the decomposition of the derived words. Then what would happen to the derived words with other suffixes or derived words with prefixes and no semantic transparency? How would be the processing mechanism for them? Secondly, the participants’ EFL proficiency was limited and the number of the derived words fit for this research was also limited. Hence this research merely examined the two group of materials: the derived words with high familiarity of whole words but low familiarity of stems and those with low familiarity of whole words but high familiarity of stems. Then what would happen to the derived words with both high familiarity of whole words and high familiarity of stems and also those with low familiarity of whole words and low familiarity of stems? In addition, this focused on the Chinese EFL learners’ processing of the derived words. Would it be different from the processing of the EFL inflectional words due to the limitation of the morphological system for the Chinese inflectional words? At the same time cross-linguistic contrastive researches might be the new hot topic to be further discussed. For example, Zhang & Zhu (2013) examined the Chinese learners’ processing of the Japanese complicated words. Would the differences in morphological structures between Chinese and Japanese affect the learners’ processing of Japanese words? To sum up, it proves to be only a beginning for the researches about L2 complicated words processing. It is expected that the researchers in the field of linguistics, psychology and neurology cooperate with each other, employ different methods and discussed about the unknown field from different perspectives.

REFERENCES

21-43.


Guihua Cao was born in Jinzhou, China in 1979. She received her master degree in linguistics from Central China Normal University in 2007. She is currently an associate professor in the School of Foreign Languages, Hubei Engineering University, Xiaogan, Hubei, China. Her research interests include linguistics and EFL teaching.