Neuropsychologia 50 (2012) 1918–1928

Contents lists available at SciVerse ScienceDirect

Neuropsychologia

journal homepage: www.elsevier.com/locate/neuropsychologia

Multiple semantic processes at different levels of syntactic hierarchy: Does the higher-level process proceed in face of a lower-level failure?

Xiaoming Jiang [a], Xiaolin Zhou [a,b,*]

[a] Center for Brain and Cognitive Sciences and Department of Psychology, Peking University, Beijing 100871, China
[b] Key Laboratory of Machine Perception and Key Laboratory of Computational Linguistics (Ministry of Education), Peking University, Beijing 100871, China

ARTICLE INFO

Article history:
Received 18 January 2012
Received in revised form
16 April 2012
Accepted 19 April 2012
Available online 26 April 2012

Keywords:
Hierarchical structure
Semantic process
Local priority
Coordination process
ERP
N400
P600
Late negativity

ABSTRACT

Humans have special abilities in processing hierarchical, recursive structures. Here we investigated how an upcoming word embedded in a hierarchical structure is semantically integrated into the prior representation during sentence comprehension. Participants read Chinese sentences with a complex verb argument structure "subject noun+verb+numeral+classifier+object noun", in which the object noun was constrained by the classifier in a local structure and by the verb in a higher-level structure. The semantic congruence between the classifier and the noun, between the verb and the noun, and between the verb and the classifier was manipulated individually or simultaneously to create a local mismatch (i.e., with classifier–noun mismatch), a sequential mismatch (with verb–classifier and classifier–noun mismatches) or a triple mismatch (with verb–classifier, classifier–noun, and verb–classifier mismatches) condition. Event-related potentials locked to the object noun showed increased N400 and late negativity responses over the local mismatch, the sequential mismatch and the triple mismatch conditions. The local mismatch additionally elicited a posterior positivity effect on the object noun. The verb–classifier mismatch elicited a right N400-like effect followed by a posterior positivity (P600) effect on the classifier. The N400 effects demonstrate that the semantic process at a higher syntactic level can proceed in face of the failure of semantic processes at lower levels when no structural re-interpretation is available, and that the semantic congruence between earlier sentence constituents can affect the integration of the upcoming word in the hierarchical structure. The P600 effects suggest the immediate triggering of a co-ordination process across syntactic levels whereas the late anterior negativity effects suggest the initiation of a second-pass semantic re-interpretation process.

1. Introduction

The extraction of semantic relations between words depends upon the syntactic structure of the sentence into which the words are embedded (Jackendoff & Pinker, 2005). During sentence comprehension, readers/listeners need to integrate constituent words constrained at different levels of syntactic hierarchy to construct a coherent sentential representation (Filik & Leuthold, 2008; Hald, Steenbeek-Planting, & Hagoort, 2007; Zhang, Jiang, Saalbach, & Zhou, 2011; Zhou et al., 2010; Nieuwland and Van Berkum, 2006). Recent studies suggest that semantic processes at different syntactic levels are probably supported by differential neuro-cognitive mechanisms (Kemmerer, 2000; Kemmerer, Tranel, & Zdanczyk, 2009; Zhou et al., 2010). It is therefore of particular interest to investigate how these mechanisms might interact and how multiple semantic constraints in a hierarchical structure might impact the integration of upcoming words during sentence comprehension.

One issue is how semantic processing in a local phrase is affected by semantic processing in a hierarchical structure, and vice versa. It seems that there are only two studies so far that have directly examined this issue. These two studies, one on German (Zhang et al., 2011) and one on Chinese (Zhou et al., 2010), nevertheless, provided contradictory data concerning the interplay of semantic processes at different syntactic levels. Although these two studies consistently showed that the semantic process at the local level can proceed when the semantic process at the higher level fails, they differed on whether the semantic process at the higher-level proceeds (Zhang et al., 2011) or not (Zhou et al., 2010) in face of the local failure.

In an event-related potential (ERP) study, Zhou et al. (2010) asked participants to read Chinese sentences with the hierarchical structure "subject noun+verb+numeral+determiner/classifier+object noun", in which the object noun was constrained by the determiner/classifier

* Corresponding author at: Center for Brain and Cognitive Sciences and Department of Psychology, Peking University, Beijing 100871, China. Fax: +86 10 62761081.
E-mail address: xz104@pku.edu.cn (X. Zhou).

0028-3932/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.neuropsychologia.2012.04.016

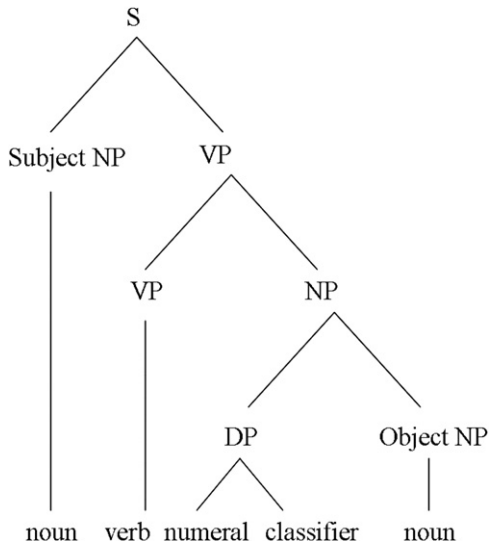

Fig. 1. The hierarchical structure of the sentence in the form of "subject noun+verb+numeral+classifier+object noun". S=sentence; NP=noun phrase, VP=verb phrase; DP=determiner phrase. The classifier (DP) and the object noun form a local, lower-level phrase whereas the verb and the object noun form a higher-level structure.

phrase (CP) at the lower level and by the verb at the higher level (see Fig. 1 for the tree structure). The semantic congruence between the classifier and the noun, between the verb and the noun, and/or between the verb and the classifier, was manipulated. Compared with the correct sentences, sentences with either classifier–noun or verb–noun mismatch elicited increased N400 responses followed by an anteriorly-maximized negativity effect on the object noun, suggesting that the mismatch at either level cannot be completely overridden by the construction of a coherent representation at the other level. Moreover, the N400 effect was larger for classifier–noun mismatch than for verb–noun mismatch, and was as large for double mismatch (i.e., the noun mismatching the classifier and the verb simultaneously) as for classifier–noun mismatch. These results suggested that the local constraint between the classifier and the noun may have processing priority over the higher-level constraint between the verb and its object, and that the local semantic process proceeds in face of the failure of the semantic process at the higher level. Furthermore, a posterior positivity (P600) effect was observed on the object noun in both the verb–noun and the double mismatch conditions, and on the classifier mismatching the verb in the triple mismatch condition. This effect was interpreted as reflecting a cross-level coordination process for multiple semantic constraints in a hierarchical structure. Attention or processing focus is shifted from the semantic process at one level to the process at another level in order to search for a possible way of constructing a coherent representation in face of the failure at a particular level.

In contrast, when Zhang et al. (2011) asked native speakers of German to read German sentences with a hierarchical structure "subject noun+verb+article/determiner+adjective+object noun+prepositional phrase", they observed a pattern of ERP responses on the object noun different from that reported for the Chinese. Here both the local mismatch between the adjective and the noun and the higher-level mismatch between the verb and the noun elicited comparable N400 effects, which were both smaller than the N400 effect for double mismatch. Moreover, all the types of mismatch elicited increased P600 responses on the object noun, with the positivity effect larger for the double mismatch than for either of the single mismatches. While the difference in N400 between the double mismatch and the verb–noun mismatch is consistent with the effect for Chinese, suggesting that the local semantic process proceeds in face of the failure of the higher-level semantic process, the difference between the double mismatch and the adjective–noun mismatch indicates that the semantic process at the higher-level, unlike in the study on Chinese, proceeds in face of the failure of the semantic process at the lower-level.

We suspect that the discrepancy between the two studies may be due to the availability of an alternative structural interpretation initiated by a local mismatch in Chinese. Given that Chinese is a non-case marked language, the construction of sentence representation relies heavily on word order and semantic properties encoded in lexical items (Bornkessel-Schlesewsky & Schlesewsky, 2009; Phillip, Bornkessel-Schlesewsky, Bisang, & Schlesewsky, 2008; Wang, Schlesewsky, Bickel, & Bornkessel-Schlesewsky, 2009; Ye, Luo, Friederici, & Zhou, 2006; Zhou et al., 2010). In Chinese, an NP at the object position can be assigned with a different thematic role, such that it becomes a modifier in a pre-nominal relative clause. For example, a sentence with a local classifier–noun mismatch (1) can be easily revised by dropping the classifier and the determiner, becoming (2):

(1) ∗*Laoshi tuijian yi shou jiaocai.*
∗*The teacher recommended one [song-classifier] textbook. (The teacher recommended one textbook.)*
(2) Laoshi tuijian jiaocai.
The teacher recommended a textbook.
Alternatively, it can be revised by treating the object noun as a modifier in relative clause (3):
(3) Laoshi tuijian yi shou jiaocai tidao de gequ.
The teacher recommended one [song-classifier] textbook mentioned DE song. (The teacher recommended the song that the textbook mentioned.)

This way of revision becomes more prominent when both the local constraint between the classifier and the noun and the higher-level constraint between the verb and the noun are mismatched (e.g. the incorrect sentence (4) becoming (5)):

(4) ∗*Xiaozhao xiuli yitai xinzhi.*
∗*Xiaozhao repaired one [electric appliance-classifier] letter. (Xiaozhao repaired one letter.)*
(5) Xiaozhao xiuli yitai xinzhi fugai de fuyinji.
Xiaozhao repaired one [electric appliance-classifier] letter covered DE copy machine. (Xiaozhao repaired a copy machine which was covered by letters.)

For these double mismatch sentences, this relative clause strategy could be the only viable way to repair the sentences, as other strategies such as dropping the determiner or replacing the verb or noun would be either unworkable or too costly. Since the pre-nominal clauses are commonly used, native speakers of Chinese are accustomed to this temporary ambiguity during the unfolding of sentence input and they may even actively use the classifier-noun incongruence as a cue for the upcoming relative clause (Wu, Kaiser, & Andersen, 2009). In processing a hierarchical structure with mismatches on the object noun, this relative clause strategy may make the integration process at the higher-level redundant or unnecessary in face of failure at the lower-level since the alternative, relative clause strategy provides an easy way to reinterpret the sentence.

However, in a case-marked language like German, the construction of sentence representation relies heavily on case marking (Bornkessel, Schlesewsky, & Friederici, 2002). When an adjective mismatches the object noun, as in (6), it is impossible, given the use of the accusative case for the object noun, for the comprehender to continue this sentence and construct an structure alternative; instead, the comprehender may focus on the integration between the verb and the object noun at the higher level, ignoring or replacing the mismatching adjective. This

Table 1
Experimental conditions and exemplar sentences with the structure of "structure+verb+numeral+classifier+noun". The selectional restrictions of the classifiers are noted in the brackets. The match or mismatch of semantic constraints between verb and classifier and between classifier and noun are marked in the right columns, with "✓" indicating a semantic match and "×" indicating a semantic mismatch.

Condition	Sentence example					Verb–classifier	Classifier–noun	Verb–noun
Baseline	小赵	修理	—	张	长椅	✓	✓	✓
	Zhao	Repaired	One	Zhang (chair-classifier)	Chair			
	Zhao repaired a chair							
Classifier–noun mismatch	小赵	修理	—	台	长椅	✓	×	✓
	Zhao	Repaired	One	Tai (electric appliance-classifier)	Chair			
	Zhao repaired a chair							
Sequential mismatch	小赵	修理	—	棵	长椅	×	×	✓
	Zhao	Repaired	One	Ke (tree-classifier)	Chair			
	Zhao repaired a chair							
Triple mismatch	小赵	修理	—	棵	长椅	×	×	×
	Zhao	Sewed	One	Ke (tree-classifier)	Chair			
	Zhao sewed a chair							

higher-level processing is hence responsible for the enlarged N400 responses in the double mismatch condition and for the appearance of the late positivity effects in all the three mismatch conditions.

(6) ∗*Felix erkannte den appetitlichen Punkt nach der Lektüre.*
∗*Felix recognized the appetizing point after reading.*

One way to test the above account for the discrepancy between the ERP patterns in the Chinese and German studies and to test the local priority in sentence comprehension is to create Chinese sentences that prevent the comprehender from using the relative clause strategy in face of local mismatch between the classifier and the object noun (Table 1). ERP responses to the critical words in such sentences should then be more similar to the pattern observed in German (Zhang et al., 2011) than to the pattern in Chinese (Zhou et al., 2010). To achieve this aim, we manipulated the semantic congruence between the verb and the classifier (7):

(7) ∗*Zhao fengbu yike yizi.*
∗*Zhao sewed one [tree-classifier] chair.*

Because the verb *sew* is not normally used together with an object noun modified by the classifier for trees, the mismatch between the verb and the classifier prevents the comprehender from treating the object noun *chair* as a modifier in a relative clause. Thus the system may instead focus on the integration at the higher level between the verb and the object noun. If so, the noun mismatching the classifier and the verb should behave more like the object noun in a German double-mismatch sentence than in a Chinese double-mismatch sentence, i.e., exhibiting increased N400 responses in this condition (with triple mismatch) than in the classifier–noun mismatch condition.

The above reasoning implicitly assumes that the verb–classifier mismatch by itself has no effect upon the ERP responses to the subsequent object noun. However, previous studies showed that earlier mismatch between sentence constituents may affect the semantic integration of upcoming words and the ERP responses to such words (Hagoort, 2003; Jiang, Tan, & Zhou, 2009; Zhou et al., 2010). In Zhou et al. (2010), for example, the comparison between the verb–classifier incongruent condition (in triple mismatch) and the verb–classifier congruent condition (in double mismatch) revealed an N400 effect on the object noun (in addition to an N400 effect on the classifier). Therefore, to obtain a "pure" effect for the high-level congruence in face of local mismatch, we need to compare the triple mismatch condition against another condition in which both verb–classifier and classifier–noun are mismatched but verb–noun is matched (i.e., sequential mismatch). Any differences in this comparison would be attributed to the higher-level semantic congruence between the verb and the noun, rather than to the potential confounds by the congruence between the verb and the classifier.

Moreover, comparing the sequential mismatch condition with the local mismatch and the baseline conditions could provide additional evidence concerning whether an earlier mismatch on the verb–classifier would affect the integration of the subsequent object noun into the prior context. Although Zhou et al. (2010) provided a positive answer to this question by comparing the triple mismatch with the double mismatch and by observing a larger N400 effect for the former, the verb and the noun in those conditions were mismatched. In the present design, the verb and the noun were congruent in both the sequential mismatch and the local mismatch conditions.

Therefore, this experiment had four conditions: (1) triple mismatch; (2) sequential mismatch; (3) local mismatch; and (4) the correct baseline in which all the constraints between sentence constituents were intact (see Table 1). As in Zhou et al. (2010), we asked participants to read Chinese sentences with the hierarchical structure "subject noun+verb+numeral+classifier+object noun" and measured ERP responses to the object noun and to the classifier. On the basis of Zhou et al. (2010) and other studies, we expected to observe an N400 effect on the object noun when the local mismatch condition was compared with the baseline and an even larger N400 effect on the object noun when the sequential mismatch condition was compared with the local mismatch (and the baseline) condition. We also expected to find an N400 effect on the mismatching classifier in the sequential mismatch condition.

Importantly, by comparing the triple mismatch with the local mismatch and the sequential mismatch, we would see clearly whether the brain responses to the object noun would exhibit a larger N400 effect than other mismatches, as in the previous study on German (Zhang et al., 2011), or a similar N400 effect as in the previous study on Chinese (Zhou et al., 2010). If the processing system focuses on the higher-level semantic integration process between the verb and the noun when the local process between the classifier and the noun fails and when there is no alternative approach to reverse the local mismatch (i.e., using the relative clause strategy), then we should observe increased N400 responses to the object noun, as in the study on German. If, however, the difference in the patterns of the N400 effect between Zhang et al. (2011) and Zhou et al. (2010) was

Table 2
Mean scores and standard deviations in the four pretests. The classifier–noun and the verb–classifier phrase acceptability rating and the sentence acceptability rating used five-point Likert scales, with 5 representing "totally acceptable" and 1 representing "totally unacceptable". The listed scores for the cloze probability test are for the target object nouns used in the sentence. The sentence completion possibility was tested on the basis of the sentence fragment without the target object noun.

Experimental condition	Classifier–noun acceptability		Verb–classifier acceptability		Sentence acceptability		Cloze probability of the target noun (%)		Sentence completion possibility (%)	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Correct	4.93	0.14	4.12	0.91	4.80	0.21	17.60	0.23	96.20	0.12
Classifier–noun mismatch	1.29	0.18	3.94	0.96	1.72	0.23	0.20	0.01	94.30	0.13
Sequential mismatch	1.21	0.11	1.83	1.04	1.68	0.20	–	–	20.10	0.20
Triple mismatch	1.21	0.11	1.76	0.89	1.16	0.14	–	–	19.02	0.24

simply due to some unexplained differences between the two languages, then we should observe no difference in the N400 responses to the noun, as in the study on Chinese.

2. Method

2.1. Participants

Twenty right-handed undergraduate and graduate students (11 females, age ranging between 18 and 26 years) at Peking University participated in the experiment. All of them were native Chinese speakers and had normal or corrected-to-normal vision. None of them suffered from psychiatric or neurological disorders. ERP data from three participants (including two females) were excluded due to excessive artifacts. This study was approved by the Ethics Committee of the Department of Psychology, Peking University.

2.2. Materials

Sixty quartets of sentences were selected on the basis of 5 pretests (see below). Incorrect sentences were created based on the correct sentences by replacing the classifiers with new ones that violated the semantic constraint between the classifier and the object noun (for the local mismatch) or that violated the constraints between the verb and the classifier and between the classifier and the noun (for the sequential mismatch). For the triple mismatch, verbs in the sequential mismatch sentences were also replaced with new ones such that a verb mismatched both the object noun and the classifier (Table 1). Thus, the verb was the same for the correct (baseline), local mismatch and sequential mismatch conditions in a quartet, and was a new one for the triple mismatch condition. The classifier was the same for the sequential and triple mismatch conditions. All the object nouns were two-character, two-syllable concrete words denoting inanimate objects. All the classifiers were selected from the Dictionary of Classifier Usages in Contemporary Chinese (Guo, 2002) and were all one-character, one-syllable words that were commonly used in the language. The classifiers in different conditions were matched in frequency (93.6 per million for the baseline condition, 111.7 per million for the local mismatch condition, 88.3 per million for the sequential and the triple mismatch conditions, according to Cai & Brysbaet, 2010) and visual complexity (in terms of the mean number of stroke, 8.9, 8.0, 8.8, and 8.8 per character for the four conditions, respectively). For these measures, no significant differences were found between the conditions, p's > 0.1. The numeral preceding the classifiers were always "—" (*one*). All the subject nouns were two- or three-character animate nouns denoting human names and/or their occupations.

One hundred and eighty filler sentences were constructed with essentially the same structure as the critical ones. In order to make the critical object nouns less predictable at the sentence-final position, all the object nouns in filler sentences were in the middle of the sentences. Among them, 90 were correct, 30 were incorrect with classifier–noun mismatch, 30 were incorrect with sequential verb–classifier and classifier–noun mismatches, and 30 were incorrect with triple mismatches.

2.3. Pretests

Five pre-tests, including three acceptability ratings, one cloze probability test and one sentence completion test, were carried out in order to select the critical stimuli (Table 2). The sentence acceptability rating was conducted to ensure that sentences with various types of mismatches were indeed unacceptable. The phrase acceptability rating was conducted to ensure that the classifier–noun congruence was violated to the same extent in the three mismatch conditions and that the verb–classifier congruence was violated to the same extent in the sequential and triple mismatch conditions. A five-point Likert Scale was used for either of the ratings, with 20, 20, and 12 participants for the sentence, the classifier–noun, and the verb–classifier rating, respectively. Four hundred and eighty sentences were included for the sentence acceptability rating, and 360 phrases with the structure of "numeral + classifier + noun" (all taken from the 480 sentences) were included for the phrase acceptability rating.

Mean scores for the finally selected critical stimuli are shown in Table 2 as a function of experimental condition. Clearly, relative to the correct sentences, sentences with classifier–noun mismatch and/or verb–classifier mismatch had much lower acceptability rating, p's < 0.001. Sentences with triple mismatches had the lowest acceptability rating, with the mean differing from the means in the other three conditions, p's < 0.001. Moreover, the mean rating for sentences with the local classifier–noun mismatch did not differ from the mean rating for sentences with sequential mismatches, $p > 0.1$. On the other hand, the local phrase acceptability test for the classifier and the noun showed significant differences between the baseline condition and the other three conditions, which did not differ between themselves.

To determine the cloze probability of a word at the object noun position, forty participants were instructed to complete the sentence fragments (i.e., without the final object nouns) of sentences in the baseline and the local classifier–noun mismatch conditions. Results showed that the average cloze probability was 17.6% for the target nouns used in the correct sentences and was approximately zero for the target nouns used in the classifier–noun mismatch sentences. The average cloze probability for the most produced words (which were generally not the ones used in the actual stimuli) was 41.1% for sentence fragments in the correct condition and 39.2% for sentence fragments in the classifier–noun mismatch condition.

To make sure that the classifier was congruent or incongruent with the verb in each sentence, another 16 participants were instructed to complete the sentence fragments of "subject + verb + numeral + classifier" with any word or phrase that made sense and to skip fragments that were impossible to continue. It is clear from Table 2 that the sentence fragments containing the verb–classifier mismatch in the sequential mismatch condition had a very low possibility of completion, compared with fragments in which the classifiers were congruent with the preceding verbs in other two conditions, p's < 0.001. For the fragments in the sequential and triple mismatch conditions, although some participants did provide grammatical continuations, none of the completions was a single object noun; rather, they were relative clauses.

To ensure that the degree of semantic congruence between the verb and the classifier was the same for the verb–classifier congruent (i.e., the correct and classifier–noun mismatch) conditions and for the verb–classifier incongruent conditions (i.e., the sequential and triple mismatch) conditions, a group of 12 participants were asked to rate on the acceptability of the verb–classifier phrase. Results showed that the average acceptability rating was much lower for the incongruent verb–classifier combinations (1.83 and 1.76 for the sequential and triple mismatch conditions, respectively) than for the congruent verb–classifier combinations (4.12 and 3.94 for the correct and classifier–noun mismatch conditions, respectively), p's < 0.001. Importantly, the rating was equally low for the sequential and for the triple mismatch conditions, $p > 0.1$.

2.4. Procedure

Participants were seated in a sound-attenuating and electrically shielded chamber. They were instructed to move as little as possible and to keep their eyes fixated on a sign at the center of the computer screen. This fixation sign was at eye-level and was approximately 1 m away. After this fixation presented for 700 ms, sentences were presented word-by-word in a serial visual presentation mode at the center of the screen, with a visual angle of less than 1°. Each word was presented for 400 ms, with a blank screen lasting 400 ms between the two consecutive frames. The numeral and the classifier were presented separately. This was to make sure that the measurement of ERP responses to the classifiers could have appropriate pre-stimulus baselines. A question mark appeared 800 ms after the end of each sentence and lasted for 1000 ms. Participants were asked to judge whether the sentence was semantically plausible by pressing buttons with their right or left index fingers. The assignment of response

buttons was counter-balanced across participants. The following trial started 1000 ms after the participants made their responses.

Twenty different test sequences were created with the restriction that (1) no more than three sentences from the same condition were presented consecutively; (2) sentences from the same critical set were separated by at least 30 other sentences (see also Zhou et al., 2010). In this way, any effects due to the repeated use of verbs, object nouns in different conditions were minimized. Each participant read 420 sentences in total, with 60 sentences from each experimental condition. Sentences were divided into six blocks after pseudo-randomization. There were 18 practice trials prior to the formal test.

2.5. EEG recording

The EEGs were recorded from 30 electrodes in an secured elastic cap (Electrocap International) located at the following positions: FP1, FP2, F7, F3, Fz, F4, F8, FT7, FC3, FCz, FC4, FT8, T7, C3, Cz, C4, T8, TP7, CP3, CPz, CP4, TP8, P7, P3, Pz, P4, P8, O1, Oz and O2. The vertical electro-oculogram (VEOG) was recorded from electrodes placed above and below the left eye. The horizontal EOG (HEOG) was recorded from electrodes placed at the outer canthus of each eye. The EEG was referenced online to the left mastoid and was re-referenced to the linked bilateral mastoids offline. Electrode impedance was kept below 5 kΩ. The signals were amplified with a band pass between 0.05 Hz and 70 Hz. The 50 Hz notch filter was on in order to eliminate the powerline interference. The EEG and EOG were digitized on-line with a sampling frequency of 500 Hz.

2.6. Data analysis

The raw EEG signals were offline filtered with a low-pass 30 Hz filter, based on which the following analyses were performed. Incorrectly judged sentences and sentences contaminated by EEG artifacts (with potentials greater than $\pm 70\,\mu V$) were rejected before the EEG averaging procedure, resulting in on average 85.0% of the artifact-free trials for the experiment (86.7% out of 60 sentences in the correct condition, 83.3% in the classifier–noun mismatch condition, 85.0% in the sequential mismatch condition, and 83.3% in the triple mismatch condition). ERPs were computed separately for each participant and each experimental condition, from −200 ms before to 800 ms after the onset of the critical classifiers or the object nouns. For classifiers, ERPs in the 200 ms pre-stimulus interval were used for baseline correction; for object nouns, ERPs in the first 100 ms after stimulus onset were used for baseline correction, given that ERP responses to the preceding classifiers could be different (due to sequential mismatch). Nevertheless, essentially the same pattern of effects was obtained when we used the pre-stimulus 200 ms period for baseline correction. An extra high-pass 10 Hz filter (24 db/oct) was implemented for visual demonstration of ERP effects.

Based on visual inspection of the grand averages and on findings in previous studies, two time windows were selected for the object nouns: 250–500 ms for the N400 component, and 500–800 ms for the late positivity and the late negativity; and for the classifiers: 350–450 ms for the negative component (N400) and 450–800 ms for the late positivity. For ERP responses to the object nouns, repeated-measures ANOVA were conducted with sentence types (baseline, local classifier–noun mismatch, sequential mismatch, triple mismatch) and topographic factors as within-participant variables. Topographic factors were included for midline and lateral analysis. The midline analysis had two factors: sentence type and electrode (Fz, FCz, Cz, CPz, and Pz). The lateral analysis had four factors: sentence type, region (frontal, fronto-central, central, centro-parietal and parietal), hemisphere (left vs. right) and electrode. The hemisphere and the region were crossed, forming ten regions of interest (ROIs), each of which was represented by two electrodes: F3, F7 for the left frontal, FC3, FT7 for the left fronto-central, C3, T7 for the left central, CP3, TP7 for the left centro-parietal, P3, P7 for the parietal, F4, F8 for the right frontal, FC4, FT8 for the right fronto-central, C4, T8 for the right central, CP4, TP8 for the right centro-parietal, P4, P8 for the right parietal. Pairwise comparisons between each mismatch condition and baseline or between mismatch conditions were planned and reported (see Section 3 and see also Kuperberg, Kreher, Sitnikova, Caplan, & Holcomb, 2007; Zhou et al., 2010). For ERP responses to the classifiers, trials in the baseline and the local classifier–noun conditions were combined to form a verb–classifier congruent condition while trials in the sequential and the triple mismatch conditions formed a verb–classifier incongruent condition. ANOVAs with the verb–classifier congruence and topographic factors were conducted to determine the ERP effects of the verb–classifier congruence in the two time windows defined above. Greenhouse–Geisser correction was applied when appropriate (Greenhouse & Geisser, 1959).

3. Results

3.1. Behavioral results

The response accuracy of the semantic plausibility judgment was 92.5% for the correct condition, 92.1% for the classifier–noun mismatch, 97.5% for the sequential mismatch, and 97.7% for the triple mismatch. Pair-wise comparisons showed that the "no" responses in the sequential and triple mismatch conditions had higher accuracy than the "yes" responses in the baseline condition and the "no" responses in the classifier–noun mismatch condition, p's < 0.001. Nevertheless, the high accuracy overall indicates that the participants paid attention to the experimental stimuli.

3.2. ERP results

As shown in Figs. 1 and 2, compared with the baseline, all the mismatch conditions elicited an N400 effect on the object noun, with the triple mismatch eliciting the largest N400 responses; compared with the local classifier–noun mismatch condition, the sequential mismatch condition elicited enhanced N400 responses at the central and posterior sites. Moreover, in the late time window, the local and sequential mismatch conditions elicited differential ERP responses on the object noun: compared with the baseline, the local mismatch condition showed an increased posterior positivity and an increased anterior negativity, but the sequential mismatch elicited only an anterior negativity effect, although this effect was larger for the sequential mismatch than for the local mismatch condition. Again, the triple mismatch elicited the largest negativity effect. These observations were confirmed by the statistical analysis.

3.3. The 250–500 ms time window on the object noun

Repeated-measures ANOVA revealed a significant main effect of sentence type on the midline and lateral electrodes, $F(3, 48)=16.60$, $p<0.001$, and $F(3, 48)=16.54$, $p<0.001$, respectively. This effect interacted with electrode in the midline analysis, $F(12, 192)=3.23$, $p<0.05$.

Further comparisons between the local classifier–noun mismatch and the baseline conditions revealed a significant effect of sentence type on the midline, $F(1, 16)=5.78$, $p<0.05$, and on the lateral, $F(1, 16)=4.49$, $p<0.05$, suggesting that the classifier–noun mismatch elicited more negative N400 responses ($-1.12\,\mu V$ on the midline; $-0.86\,\mu V$ on the lateral) as compared with the baseline. Similarly, the comparison between the sequential mismatch and the baseline also revealed a significant effect of sentence type on the midline, $F(1, 16)=15.96$, $p<0.005$, and the lateral electrodes, $F(1, 16)=14.33$, $p<0.005$, suggesting that the sequential mismatch ($-1.78\,\mu V$ on the midline; $-1.30\,\mu V$ on the lateral) elicited more negative N400 responses. The comparison between the triple mismatch and the baseline revealed a significant effect of sentence type, $F(1, 16)=5.78$, $p<0.05$ and $F(1, 16)=40.03$, $p<0.001$, respectively on the midline and the lateral electrodes, indicating that the triple mismatch elicited more negative N400 responses ($-1.99\,\mu V$ on the midline; $-0.94\,\mu V$ on the lateral). The sentence type interacted with hemisphere on the lateral electrodes, $F(1, 16)=4.90$, $p<0.05$, with the N400 effect being larger in the right, than in the left hemisphere.

Interestingly, the comparison between the sequential mismatch and the local classifier–noun mismatch revealed a significant two-way interaction between sentence type and region on the lateral sites, $F(4, 64)=3.73$, $p<0.05$, although there was no main effect of sentence type. Separate analysis for each region revealed an effect of sentence type on the central, $F(1, 17)=3.21$, $p<0.05$, the central parietal, $F(1, 17)=3.31$, $p<0.05$, and the parietal electrodes, $F(1, 17)=3.05$, $0.05<p<0.1$, suggesting increased N400 responses ($-0.78\,\mu V$ on average) in the sequential mismatch condition than in the classifier–noun mismatch condition. This finding suggests that the difficulty in integrating the object noun into the

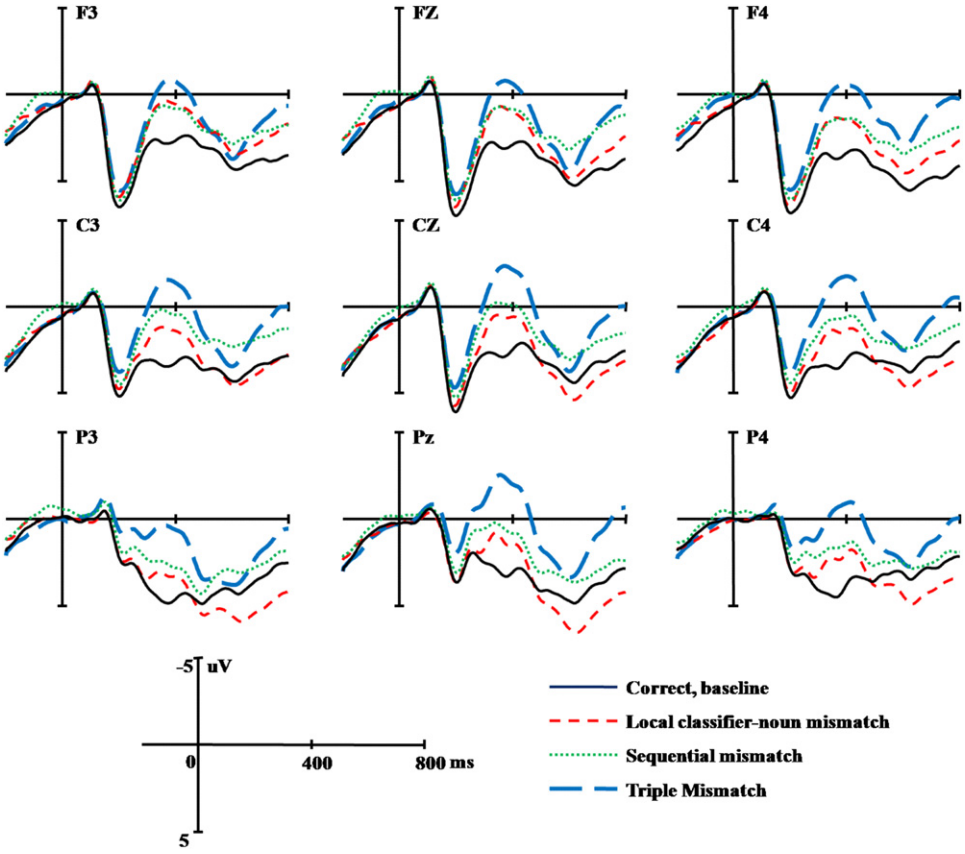

Fig. 2. Grand average waveforms epoched from 200 ms before to 800 ms after the onset of the object noun at nine exemplar electrodes.

local determiner phrase was increased when the constraint between the preceding verb and the classifier was violated.

Importantly, the comparison between the triple and the sequential mismatch conditions also revealed a main effect of sentence type on the midline, $F(1, 16)=14.00$, $p<0.005$, and the lateral electrodes, $F(1, 16)=14.39$, $p<0.005$, indicating that the triple mismatch elicited enlarged N400 responses as compared with the sequential mismatch. This finding suggests that the semantic process between the verb and the object noun continues to proceed when the processes between the verb and the classifier and between the classifier and the noun fail.

3.4. The 500–800 ms time window on the object noun

Repeated-measures ANOVA revealed a significant effect of sentence type on the midline, $F(3, 48)=6.62$, $p<0.005$, and on the lateral, $F(3, 48)=3.83$, $p<0.05$, and also a significant two-way interaction between sentence type and electrode on the midline, $F(3, 48)=6.62$, $p<0.005$, and between sentence type and region on the lateral, $F(12, 192)=3.90$, $p<0.05$.

Further comparisons between the classifier–noun mismatch and the baseline revealed a significant interaction between sentence type and electrode on the midline, $F(4, 64)=6.43$, $p<0.005$, and between sentence type and region on the lateral, $F(4, 64)=5.17$, $p<0.05$. It is clear from Fig. 3 that this interaction was due to the appearance of the frontal negativity and the posterior positivity. Separate analysis showed that the negative effect (-1.26 μV) over the frontal electrodes (Fz, F3, F7, F4, F8) and the positivity effect (1.71 μV) over the posterior electrodes (CPz, Pz, CP3, CP4, TP7, TP8, P3, P4, P7 and P8) were all significant, $F(1, 16)=5.15$, $p<0.05$, and $F(1, 16)=4.22$, $p<0.05$, respectively.

The comparison between the sequential mismatch and the baseline demonstrated a significant effect of sentence type on the midline electrodes, $F(1, 16)=5.47$, $p<0.05$, indicating that the sequential mismatch condition elicited a larger negativity (-1.22 μV) as compared with the baseline. The sentence type interacted with electrode in the midline analysis, $F(4, 64)=3.41$, $p<0.05$, and with region in the lateral analysis, $F(4, 64)=5.92$, $p<0.01$. It is clear from Fig. 3 that the negativity effect (-0.81 μV) was most evident on the frontal and frontocentral electrodes (Fz, F3, F7, F4, F8, FCz, FC3, FT7, FC4, FT8), $F(1, 16)=6.02$, $p<0.05$,

The comparison between the triple mismatch and the baseline revealed a significant effect of sentence type on the midline, $F(1, 16)=7.99$, $p<0.05$, and on the lateral electrodes, $F(1, 16)=6.49$, $p<0.05$, suggesting that the triple mismatch elicited a larger negativity as compared with the baseline (-1.84 μV on the midline and -1.28 μV on the lateral). The sentence type interacted with hemisphere and with region in the lateral analysis, $F(1, 16)=7.15$, $p<0.05$, and $F(1, 16)=3.76$, $p<0.05$, respectively. It is clear from Fig. 3 that the negativity effect was most evident on the right hemisphere and in the anterior regions. Detailed statistical analysis confirmed this observation.

The comparison between the sequential mismatch and the local classifier–noun mismatch found also a significant effect of sentence type on the midline, $F(1, 16)=15.88$, $p<0.005$, and on the lateral, $F(1, 16)=5.34$, $p<0.05$, suggesting that the sequential mismatch in general elicited more negative responses than the local mismatch (-1.94 μV on the midline; -1.08 μV on the lateral). In particular, when only the ERP responses on the frontal and frontocentral electrodes (on which the ERP responses were negative compared with the baseline; see Fig. 2) were compared for the two conditions, the negativity effect (-0.74 μV) was still

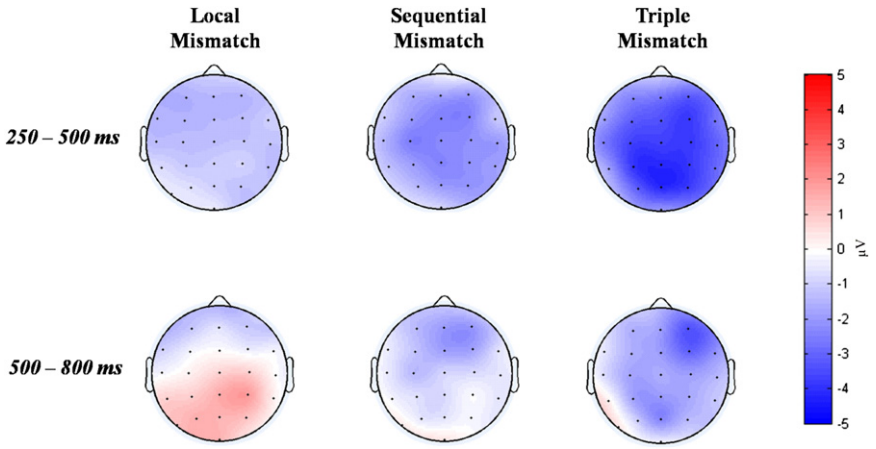

Fig. 3. Topographic maps for the contrast between each mismatch conditions and the baseline condition, with the upper panel demonstrating the mean differences in 250–500 ms time window and the lower panel demonstrating the mean differences in the 500–800 ms time window.

significant, $F(1, 16)=3.29$, $p<0.05$, suggesting that the anterior negativity effect, as compared with the baseline, was larger for the sequential than for the local mismatch.

It is evident in Fig. 2 that compared with the sequential mismatch, the triple mismatch in general had more negative ERP responses in the 500–800 ms window, particularly on the right hemisphere. Statistical analysis revealed a significant two-way interaction between sentence type and electrode on the midline, $F(1, 16)=3.89$, $p<0.05$, and between sentence type and hemisphere on the lateral electrodes, $F(1, 16)=10.59$, $p<0.01$, indicating that the negativity effect, as compared with the baseline, was indeed larger for the triple mismatch than for the sequential mismatch on the right hemisphere.

It is also evident in Fig. 2 that, compared with the sequential mismatch, the triple mismatch elicited a small positivity effect in the 550–650 ms window on the left hemisphere. The lateral analysis revealed a significant two-way interaction between sentence type and hemisphere, $F(1, 16)=15.39$, $p<0.005$, with the late positivity effect (0.72 μV) being significant on the left hemisphere, $F(1, 16)=3.34$, $p<0.05$, but not on the right hemisphere, $F<1$.

3.5. The 350–450 ms time window on the classifier

Fig. 4 illustrates the grand average waveforms on the classifier for the verb–classifier congruent condition (i.e. ERP responses to the classifiers in the baseline and the local classifier–noun mismatch conditions) and the verb–classifier incongruent condition (i.e. ERP responses to the classifiers in the sequential and the triple mismatch conditions). Apparently the verb–classifier incongruent condition elicited a right-lateralized, N400-like negativity, followed by a large late positivity, as compared with the congruent condition.

Although repeated-measures ANOVA with experimental condition and topographic factors as within participant variables did not show a significant main effect of condition on the midline or lateral electrodes, $Fs<1$, it did reveal a significant interaction between condition and hemisphere in the lateral analysis, $F(1, 16)=3.84$, $p<0.05$. Further analysis found a significant effect of condition on the right hemisphere, $F(1, 16)=4.14$, $p<0.05$, but not on the left hemisphere, $F<1$, suggesting that the incongruent classifiers elicited enhanced more negative-going responses (-0.71 μV) than the congruent ones on the right hemisphere.

3.6. The 450–800 ms time window on classifiers

ANOVA revealed a significant effect of condition on the midline and lateral electrodes, $F(1, 16)=8.40$, $p<0.05$, and $F(1, 16)=4.25$, $p<0.05$, respectively, suggesting that incongruent classifiers elicited more positive responses than congruent ones (1.19 μV on the midline and 0.63 μV on the lateral). This effect interacted with electrode on the midline, $F(4, 64)=4.24$, $p<0.05$, and with region on the lateral, $F(4, 64)=9.52$, $p<0.005$, suggesting that the positivity effect was the largest on the central and posterior electrodes.

4. Discussion

This ERP study investigated the brain responses to semantic processes between constituent words embedded within a hierarchical structure. Crucially, it removed the possibility of structural reanalysis of the sentences which was present in the previous Chinese study (Zhou et al., 2010). In a complex verb argument structure, three types of semantic mismatch were created by manipulating the congruence between the classifier and the object noun, between the verb and the classifier, and between the verb and the noun. ERP responses to the object noun showed that all types of mismatch elicited increased N400 followed by enhanced anterior negativity, as compared with the baseline condition. Moreover, the N400 effect was larger in the sequential than in the local classifier–noun mismatch condition and was larger in the triple than in the sequential mismatch condition. The classifier–noun mismatch elicited an additional late posterior positivity (P600) on the object noun. Furthermore, on the classifier, the verb–classifier incongruent condition elicited a right N400-like effect followed by a late posterior positivity effect on the classifier, as compared with the verb–classifier congruent condition. These findings suggest a parallel processing of multiple semantic constraints in a hierarchical structure. In the following paragraphs, we focus on two issues: (1) the modulation of the N400 by the number/type of mismatch; and (2) the functional significance of the late posterior positivity and anterior negativity for semantic processes in a hierarchical structure.

4.1. The larger N400 effect for the triple than for the sequential mismatch

Consistent with our hypothesis, compared with the baseline, the triple mismatch elicited a larger N400 effect on the object

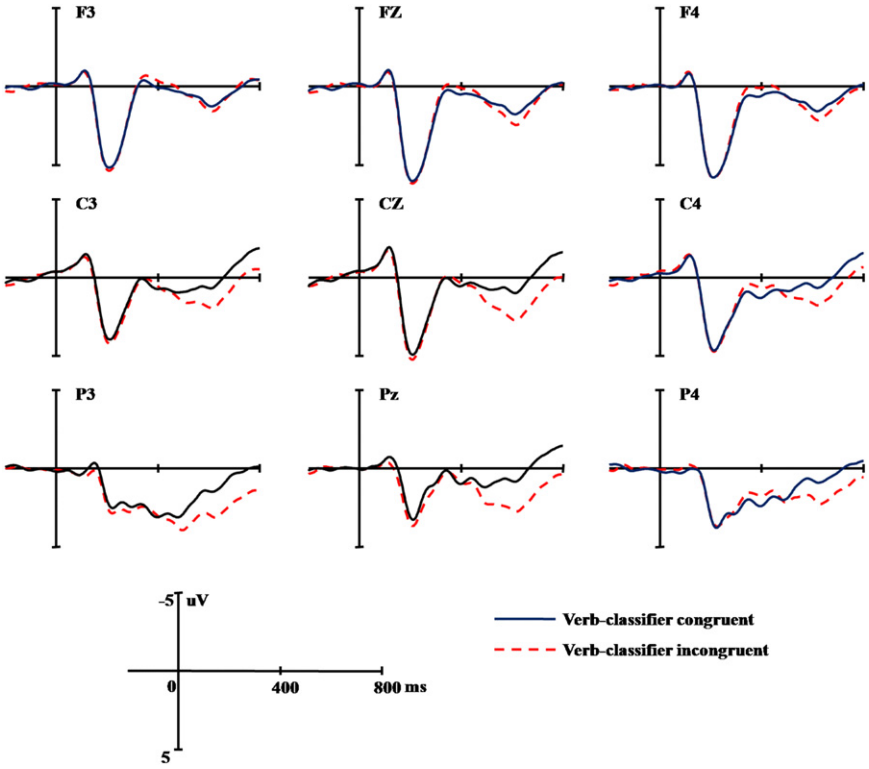

Fig. 4. Grand average waveforms for the verb–classifier congruent and verb–classifier incongruent sentences at nine exemplar electrodes, epoched from 200 ms before to 800 ms after the onset of the classifier.

noun than the sequential mismatch. In both conditions, the constraints between the verb and the classifier and between the classifier and the noun were violated, and the degree of violation was essentially the same for the two conditions (see Section 2.3). The only difference between the two conditions was that the semantic constraint between the verb and the noun was violated in the triple mismatch condition but was intact in the sequential mismatch condition. This finding of a larger N400 effect was consistent with Zhang et al. (2011) on German but was inconsistent with Zhou et al. (2010) on Chinese.

As we hypothesized in the Introduction, when facing the failure of the local semantic process between the adjective and the object noun in German (Zhang et al., 2011), the comprehender may make stronger effort to construct a meaningful sentence representation based on the verb and the object noun, because the case system in German prevents the comprehender from readily using other strategies (such as the bi-clause strategy) to deal with the mismatch information. In Chinese, however, when the constraint between the classifier and the object noun is violated, the comprehender may treat the mismatching noun as the onset of a relative clause (i.e., as a modifier). This relative clause strategy is helped by the intact constraint between the verb and the classifier. Thus the system may not focus the processing resources on the higher-level process between the verb and the noun in the same way as for the German sentences (Zhou et al., 2010). The present study suggests that, when the relative clause strategy is not available (because of the earlier mismatch between the verb and the classifier), the system *can* make stronger effort to construct a meaningful sentence representation based on the verb and the object noun, rendering the pattern of the N400 effects more similar to the one observed for German than the one observed for the Chinese.

It might be suggested that the differential N400 effect between the triple and the sequential mismatch may reflect the difference in making sense of the sentence, as suggested by the lower acceptability rating for the triple mismatch than for the sequential mismatch sentences. However, the sequential mismatch elicited enhanced N400 responses as compared with the local mismatch even though these two conditions had equally low acceptability ratings, suggesting that the differential N400 effects between conditions cannot simply be attributed to the difficulty in making sense of the incorrect sentence in different conditions. Instead, we suggest that this "making sense" process is reflected by the late negativity, rather than the N400 (see Section 4).

4.2. The larger N400 effect for the sequential than for the local mismatch

Interestingly, this study revealed a larger N400 effect for the sequential than for the local mismatch. In Zhou et al. (2010), when sentences with the triple mismatch were compared with sentences with the double mismatch, a larger N400 effect was also observed on the object noun, suggesting that the earlier mismatch between the verb and the classifier influenced the semantic integration of the upcoming object noun into prior sentence context. This finding, however, may be open to alternative accounts because the acceptability for the triple mismatch sentences was generally lower than that for the double mismatch sentences. In the present study, the pattern of the N400 effects was not confounded by the overall acceptability because sentences with sequential mismatch was rated as being equally unacceptable as sentences with local mismatch.

One account for the increased N400 responses for the sequential than for the local mismatch is to assume that the semantic integration difficulty between earlier, critical sentence constituents accumulates and affects the semantic process for the upcoming word. Integration of the object noun into the prior sentence representation would be more difficult when this representation by itself is not coherent (as in the sequential mismatch condition) than when it is coherent (as in the local mismatch condition). The

same rationale can be used to explain why the N400 responses were more negative in the triple mismatch condition than in the double mismatch (Zhou et al., 2010).

Alternatively, one might suggest that when the comprehender comes across a mismatching classifier, he/she might be "coerced" to interpret the classifier as standing for a correct one or simply to ignore the classifier (cf., Molinaro, Kim, Vespignani, & Job, 2008). Indeed, the right-lateralized N400-like effect and the late posterior positivity effect observed on the mismatching classifier might be taken as evidence for this "coercion". However, the "new" classifier may or may not be congruent with the upcoming object noun and in any case we should not expect to observe a *larger* N400 effect for the sequential than for the local mismatch on the object noun. One might also suggest that the "coercion" results in an reinterpretation of the verb, rather than the classifier, and this "new" verb may or may not be congruent with the object noun. This "partial mismatch" may enlarge the N400 effect caused by the mismatch between the classifier and the noun. However, by this account, the mismatch between the verb and the classifier in the triple mismatch condition would disappear due to the "coercion" and the triple mismatch would somehow resemble the double mismatch. It is not clear how this "new" double mismatch would elicit stronger N400 responses to the object noun than the real double mismatch condition (Zhou et al., 2010).

It should be noted that the absence of coercion here was for the semantic processes. Coercion does take place for syntactic processes. Molinaro et al. (2008) asked participants read sentences with a hierarchical co-referential structure (e.g. *The famous dancer was nervously preparing herself to face the crowd*), in which the reflexive pronoun (*herself*) was constrained either by a local verb (*was*) or by a higher-level subject noun (*dancer*). The number agreement was manipulated between the subject noun and the verb (e.g. *The famous dancer was/*were nervously preparing*…) and the reflexive pronoun agreed either with the verb (*themselves*) or with the subject noun (*himself*). When the subject and the verb disagreed, a P600 effect, typically associated with the grammatical number mismatch, was elicited on the pronoun when it agreed with the subject noun but disagreed with the verb. However, no such effect was observed on the pronoun when it agreed with the verb but disagreed with the subject noun. These findings suggested that, in order to interpret the pronoun, the erroneous representation for the subject noun is coerced to match the verb. Molinaro, Vespignani, Zamparelli, and Job (2011) extended this finding by showing that it is the difficulty of altering the morpho-syntactic features of constituent words, rather than distance between two critical constituents, that determines the direction of coercion. It is possible that absence of coercion for semantic processes here might be due to the enhanced difficulty in altering semantic representations, as compared with altering morpho-syntactic representations.

4.3. The late positivity and coordination of semantic processes across syntactic levels

We have interpreted the P600 as reflecting the coordination of multiple semantic processes at different levels of syntactic hierarchy (Zhang et al., 2011; Zhou et al., 2010). When the semantic process at one level meets difficulty, the processing system may initiate a process redeploying the attentional or processing focus from this level to the semantic process at another level in order to mitigate the difficulty in constructing a sentence representation. Thus the appearance of the P600 indicates the initiation of the coordination process within a hierarchical structure.

This account of the functional significance of the P600 fits well with the findings in this study. We observed a P600 effect for the local mismatch, as in our German study (Zhang et al., 2011). When the object noun mismatches the preceding classifier, the processing system may start to focus more on the higher-level semantic process between the verb and the noun. Similarly, the appearance the P600 effect on the mismatching classifiers can be taken as evidence that the system initiates a process in which the processing focus shifts from the process between the verb and the classifier to other processes between sentence constituents that might ease the difficulty in constructing a coherent sentence representation.

The absence of the P600 effect on the object noun in the sequential and triple mismatch conditions may be due to overlap with the late (anterior) negativity. As we will argue in the next section, the late anterior negativity may reflect a sentence re-interpretation process in which inappropriate input information is replaced or inhibited for the buildup of a coherent representation. For example, in the sequential mismatch condition, the processing system might attempt to inhibit the mismatching classifier or to replace it with an appropriate one when the object noun is met; this re-interpretation process might become stronger when more constraints between sentence constituents are broken, as in the triple mismatch condition, resulting in a stronger late anterior negativity (see Fig. 3). It is plausible that this late negativity may overshadow the late positivity that otherwise would be seen in the sequential and triple mismatch conditions.

The coordination account can explain why the P600 effect has been observed for semantically mismatching words embedded in a hierarchical, complex structure, such as between adjectives forming a hierarchical adjective sequence, e.g. **Jennifer rode a gray huge elephant* (Kemmerer, Weber-Fox, Price, Zdanczyk, & Way, 2007), between the noun and the classifier in long-distance dependency, e.g. **Haibao Qingfeng Zhao kajian yi liang benzhuode/ *Seal, Qingfeng Zhao saw one liang$_{car\text{-}classifier}$ clumsy* (Zhang, Zhang, & Min, 2012), between the noun and the verb separated by an adverbial clause (Gunter, Stowe, & Mulder, 1997) or by a prepositional phrase (Hoeks, Stowe, & Doedens, 2004), but not for mismatch words in a simple local phrase, e.g. between the verb and the noun (Li, Shu, Liu, & Li, 2006; Jiang et al., 2009; Ye et al., 2006) or between the adjective and the noun (Prior & Bentin, 2006).

Before conducting further studies to test the above suggestions, we may need to rule out three alternative accounts for the P600 effect on the object noun. The first account takes it as reflecting a syntactic reanalysis process (Friederici, 1995) in which the object noun in the local mismatch condition is reinterpreted as causing temporary ambiguity: it could be the (mismatching) object noun in a subject–verb–object structure or it could be reinterpreted as a modifier in a relative clause. The use of the relative clause strategy in face of the local-mismatch between the classifier and the noun leads to the posterior P600 effect in the local mismatch condition. On the other hand, the verb–classifier mismatch in the sequential and triple mismatch condition blocks such structural reinterpretation, leading to the null P600 effect. However, the appearance of the P600 in the German study (Zhang et al., 2011) in which no alterative structural analysis was possible allows us to conclude that the above structural account of the P600 for semantic processes in a hierarchical structure cannot be a general theory.

Another account assumes that the P600 effect (preceded by an N400 effect) due to semantic manipulations reflects the categorization of sentence well-formedness in a certain experimental environment (Bornkessel-Schlesewsky et al., 2011; Bornkessel-Schlesewsky & Schlesewsky, 2008; Frenzel, Schlesewsky, & Bornkessel-Schlesewsky, 2011). This P600 effect usually appears in an "error-detection" task (such as the acceptability judgment used in this study) in which the well-formedness of a sentence

needs to be verified. However, this account would predict that the P600 effect appears on the object noun in all the mismatching conditions, especially for the sequential and the triple mismatch, which had equal or lower acceptability relative to the local mismatch and were easier to be categorized as "ill-formed".

The third account assumes that the P600 reflects a conflict monitoring process for potential errors in face of a processing failure (Kolk & Chwilla, 2007; Van de Meerendonk, Kolk, Vissers, & Chwilla, 2010; Vissers, Chiwilla, & Kolk, 2007; Vissers et al., 2008). In this way, the P600 effect is consequential of a general executive control mechanism (Ye & Zhou, 2009), which is involved in resolving conflicts or in-determinacy in language perception. The P600 effect has been observed on words violating the plausibility of an event (e.g. the monophasic P600 in semantically reversible sentences with competing sentence representations; Kolk, Chwilla, Van Herten, & Oor, 2003; Hoeks et al., 2004; Van Herten, Chwilla, & Kolk, 2006; Vissers et al., 2007) and on words violating the expectancy built upon the preceding context (e.g. the biphasic N400–P600 for conflicts between prediction-based words and the actual input; Nieuwland & Van Berkum, 2005; Van de Meerendonk et al., 2010). It is suggested that the strength of the un-expectedness determines the presence of a P600 effect. Highly unexpected words trigger a monitoring process for potential error and elicit a biphasic N400–P600 effect; moderately unexpected words do not produce a strong conflict and elicit only the N400 effect (Federmeier, Wlotko, Ochoa-Dewald, & Kutas, 2007; Van de Meerendonk et al., 2010).

In the present study, the context words preceding the object noun (i.e., subject–verb–numeral–classifier) in the local classifier–noun mismatch condition formed a coherent representation and promoted a moderate expectancy toward a particular noun (39.2% for sentence fragments in the classifier–noun mismatch condition). In contrast, the context words (i.e., subject–verb–numeral–classifier) in the sequential and the triple mismatch conditions could not form a coherent representation and promote a prediction towards the object noun. Thus, the executive control account seems to fit the above observations. However, this account has difficulties in explaining the left-lateralized P600 effect for the triple mismatch, relative to the sequential mismatch (Fig. 2). In both conditions, no particular expectation could be formed towards an object noun.

4.4. The late negativity and second-pass semantic re-interpretation

The anterior late negativity on the object noun replicated our previous results concerning semantic mismatch in a hierarchical structure (Zhou et al., 2010). In Zhou et al. (2010), in addition to the increased N400 responses, broadly-distributed negativities (with anterior maxima) in a later time window (550–800 ms) were found in the local, higher-level, double and triple mismatch conditions, as compared with correct sentences. The negativity effect, increasing in size as a function of the degree of mismatch in both Zhou et al. (2010) and the current study, might be interpreted as reflecting the triggering of a re-interpretation process after the initial detection of semantic failure. This re-interpretation process could take the form of replacing the object noun or the classifier with a plausible one (Jiang et al., 2009; Zhou et al., 2010) or inhibiting irrelevant, incongruent representations in order to construct a coherent one (Politzer-Ahles, submitted for publication; Ye and Zhou, 2008). Note that, although the object noun could not be surely predicted to be the sentence-final word in a critical sentence (because the object noun could be the modifier of relative clause in a filler sentence), a sentence-final wrap-up process (Hagoort, 2003) may nevertheless take place on the encounter of the object noun and the late anterior negativity may reflect this process. However, this wrap-up process functions essentially in the same way as what we suggest for the re-interpretation process.

5. Conclusion

This study investigated how multiple constraints in a hierarchical structure may affect the semantic integration processes at different levels of syntactic hierarchy during sentence comprehension. The semantic congruence between the classifier and the object noun, between the verb and the noun, and between the verb and the classifier was manipulated for Chinese sentences with the complex argument structure "subject noun+verb+numeral+classifier+object noun". Compared with the correct sentences, a larger N400 followed by a larger anterior negativity was found for all the types of mismatch, with both the N400 effect and the late anterior negativity effect being larger for the sequential than for the local mismatch condition and being larger for the triple mismatch than for the sequential mismatch condition. A posterior positivity (P600) effect was also found on the object noun in the local mismatch condition and on the classifier in the sequential and the triple mismatch conditions. Extending Zhou et al. (2010), this study demonstrates that semantic processing in a hierarchical structure follows a parallel constraint-satisfaction principle and the semantic process at the higher-level can proceed in face of the failure of semantic process at the local level.

Acknowledgments

This study was supported by Grants from the Natural Science Foundation of China (30970889; 30110972) and from National Basic Research Program of China (973 Program: 2010CB833904) to Xiaolin Zhou, and Grants from China Postdoctoral Science Foundation (20100480150) to Xiaoming Jiang. We thank Mr. Steve Politzer-Ahles and two anonymous reviewers for their insightful comments on an earlier version of the manuscript.

References

Bornkessel-Schlesewsky, I., Kretzschmar, F., Tune, S., Wang, L., Genc, S., & Philipp, M. (2011). Think globally: Cross-linguistic variation in electrophysiological activity during sentence comprehension. *Brain and Language*, *117*, 133–152.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2008). An alternative perspective on "semantic P600" effects in language comprehension. *Brain Research Review*, *59*, 55–73.

Bornkessel-Schlesewsky, I., & Schlesewsky, M. (2009). The role of prominence information in the real-time comprehension of transitive constructions: A cross-linguistic approach. *Language and Linguistic Compass*, *3/1*, 19–58.

Bornkessel, I., Schlesewsky, M., & Friederici, A. (2002). Beyond syntax: Language-related positivities reflect the revision of hierarchies. *Neuroreport*, *13*, 361.

Cai, Q., & Brysbaet, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PLOS One*, *5*, http://dx.doi.org/10.1371/journal.pone.0010729.t001.

Federmeier, K., Wlotko, E., Ochoa-Dewald, E., & Kutas, M. (2007). Multiple effects of sentential constraint on word processing. *Brain Research*, *1146*, 75–84.

Filik, R., & Leuthold, H. (2008). Processing local pragmatic anomalies in fictional contexts: Evidence from the N400. *Psychophysiology*, *45*, 554–558.

Frenzel, S., Schlesewsky, M., & Bornkessel-Schlesewsky, I. (2011). Conflicts in language processing: A new perspective on the N400–P600 distinction. *Neuropsychologia*, *49*, 574–579.

Friederici, A. (1995). The time course of syntactic activation during language processing: A model based on neuropsychological and neurophysiological data. *Brain and Language*, *50*, 259–281.

Greenhouse, S. W., & Geisser, S. (1959). On methods in the analysis of profile data. *Psychometrika*, *24*, 95–112.

Gunter, T. C., Stowe, L. A., & Mulder, G. (1997). When syntax meets semantics. *Psychophysiology*, *34*, 660–676.

Guo, X. (2002). *Dictionary of classifier usages in contemporary Chinese. Xiandai hanyu liangci yongfa cidian*. Beijing: Yuwen Press.

Hagoort, P. (2003). Interplay between syntax and semantics during sentence comprehension: ERP effects of combining syntactic and semantic violations. *Journal of Cognitive Neuroscience*, *15*, 883–899.

Hald, L. A., Steenbeek-Planting, E. G., & Hagoort, P. (2007). The interaction of discourse context and world knowledge in online sentence comprehension: Evidence from the N400. *Brain Research*, *1146*, 210–218.

Hoeks, J. C. J., Stowe, L. A., & Doedens, G. (2004). Seeing words in context: The interaction of lexical and sentence level information during reading. *Cognitive Brain Research*, *19*, 59–73.

Jackendoff, R., & Pinker, S. (2005). The nature of the language faculty and its implications for evolution of language. *Cognition*, *97*, 211–225.

Jiang, X., Tan, Y., & Zhou, X. (2009). Processing the universal quantifier during sentence comprehension: ERP evidence. *Neuropsychologia*, *47*, 1799–1815.

Kemmerer, D. (2000). Selective impairment of knowledge underlying pronominal adjective order: Evidence for the autonomy of grammatical semantics. *Journal of Neurolinguistics*, *13*, 57–82.

Kemmerer, D., Tranel, D., & Zdanczyk, C. (2009). Knowledge of the semantic constraints on adjective order can be selectively impaired. *Journal of Neurolinguistics*, *22*, 91–108.

Kemmerer, D., Weber-Fox, C., Price, K., Zdanczyk, C., & Way, H. (2007). Big brain dog or brown big dog? An electrophysiological study of semantic constraints on prenominal adjective order. *Brain and Language*, *100*, 238–256.

Kolk, H. H. J., & Chwilla, D. J. (2007). *Late positivities in unusual situations. Brain and Language*, *100*, 257–261.

Kolk, H., Chwilla, D., Van Herten, M., & Oor, P. (2003). Structure and limited capacity in verbal working memory: A study with event-related potentials. *Brain and Language*, *85*, 1–36.

Kuperberg, G., Kreher, D., Sitnikova, T., Caplan, D., & Holcomb, P. (2007). The role of animacy and thematic relationships in processing active English sentences: Evidence from event-related potentials. *Brain and Language*, *100*, 223–237.

Li, X., Shu, H., Liu, Y., & Li, P. (2006). Mental representation of verb meaning: Behavioral and electrophysiological evidence. *Journal of Cognitive Neuroscience*, *10*, 1774–1787.

Molinaro, N., Kim, A., Vespignani, F., & Job, R. (2008). Anaphoric agreement violation: An ERP analysis of its interpretation. *Cognition*, *106*, 963–974.

Molinaro, N., Vespignani, F., Zamparelli, R., & Job, R. (2011). Why *brother and sister* are not just *siblings*: Repair processes in agreement computation. *Journal of Memory and Language*, *64*, 211–232.

Nieuwland, M., & Van Berkum, J. (2005). Testing the limits of the semantic illusion phenomenon: ERPs reveal temporary semantic change deafness in discourse comprehension. *Cognitive Brain Research*, *24*, 691–701.

Nieuwland, M., & Van Berkum, J. (2006). When peanuts fall in love: N400 evidence for the power of discourse. *Journal of Cognitive Neuroscience*, *18*, 1098–1111.

Philipp, M., Bornkessel-Schlesewsky, I., Bisang, W., & Schlesewsky, M. (2008). The role of animacy in the real time comprehension of Mandarin Chinese: Evidence from auditory event-related brain potentials. *Brain and Language*, *105*, 112–133.

Politzer-Ahles, S., Fiorentino, R., Jiang, X., & Zhou, X. Event-related potential investigation of scalar implicature processing in Chinese using picture-sentence verification, submitted for publication.

Prior, A., & Bentin, S. (2006). Differential integration efforts of mandatory and optional sentence constituents. *Psychophysiology*, *43*, 440–449.

Van de Meerendonk, N., Kolk, H., Vissers, C., & Chwilla, D. (2010). Monitoring in language perception: Mild and strong conflicts elicit different ERP patterns. *Journal of Cognitive Neuroscience*, *22*, 67–82.

Van Herten, M., Chiwilla, D. J., & Kolk, H. J. (2006). When heuristics clash with parsing routines: ERP evidence for conflict monitoring in sentence perception. *Journal of Cognitive Neuroscience*, *18*, 1181–1197.

Vissers, C. Th. W. M., Chiwilla, D. J., & Kolk, H. J. (2007). The interplay of heuristics and parsing routines in sentence comprehension: Evidence from ERPs and reaction times. *Biological Psychology*, *75*, 8–18.

Vissers, C. Th. W. M., Kolk, H. H. J., Van de Meerendonk, N., & Chiwilla, D. J. (2008). Monitoring in language perception: Evidence from ERPs in a picture-sentence matching task. *Neuropsychologia*, *46*, 967–982.

Wang, L., Schlesewsky, M., Bickel, B., & Bornkessel-Schlesewsky, I. (2009). Exploring the nature of the 'subject'-preference: Evidence from the online comprehension of simple sentences in Mandarin Chinese. *Language and Cognitive Processes*, *24*, 1180–1226.

Wu, F., Kaiser, E., & Andersen, E. (2009). The effect of classifiers in predicting Chinese relative clauses. *Paper presented at western conference on linguistics*.

Ye, Z., Luo, Y., Friederici, A., & Zhou, X. (2006). Semantic and syntactic processing in Chinese sentence comprehension: Evidence from event-related potentials. *Brain Research*, *1071*, 186–196.

Ye, Z., & Zhou, X. (2008). Involvement of cognitive control in sentence comprehension: Evidence from ERPs. *Brain Research*, *1203*, 103–115.

Ye, Z., & Zhou, X. (2009). Executive control in language processing. *Neuroscience and Biobehavioral Reviews*, *33*, 1168–1177.

Zhang, Y., Jiang, X., Saalbach, H., & Zhou, X. (2011). Multiple constraints on semantic integration in a hierarchical structure: An ERP study on German. *Brain Research*, *1410*, 89–100.

Zhang, Y., Zhang, J., & Min, B. (2012). Neural dynamics of animacy processing in language comprehension: ERP evidence from the interpretation of classifier–noun combinations. *Brain and Language*, *120*, 321–331.

Zhou, X., Jiang, X., Ye, Z., Zhang, Y., Lou, K., & Zhan, W. (2010). Semantic integration processes at different levels of syntactic hierarchy during sentence comprehension: An ERP study. *Neuropsychologia*, *48*, 1551–1562.