Can L2 sentence processing strategies be native-like? Evidence from English speakers' L2 processing of Chinese base-generated-topic sentences

Abstract

This article reports on an empirical study examining English speakers' L2 processing of Chinese base-generated-topic (BGT) sentences. Forty-four highly proficient English-speaking L2 learners of Chinese and 23 native Chinese speakers were involved in the study. Results of a self-paced reading task reveal that both native Chinese speakers' and L2 Chinese learners' processing of Chinese BGT sentences is syntactically induced in a top-down manner. English speakers are sensitive to and are able to make use of syntactic cues as well as semantic information in their processing of Chinese BGT sentences. The study provides disconfirming evidence against the Shallow Structure Hypothesis (Clahsen and Felser, 2006a, b), which predicts that unlike native speakers, L2 learners do not rely on structure-based processing strategies when solving ambiguities in L2 sentence processing.

Kay words

second language, sentence processing, Chinese base-generated topics, native-like strategies

1. Introduction

It is widely observed that children generally learn their mother tongues rapidly and successfully, but few adults can have native-like mastery of the target language in their

acquisition of a second language (L2). One of the accounts for this contrast is the Shallow Structure Hypothesis (SSH) by Clahsen and Felser (2006a, b), which states that during real-time language comprehension, L2 learners can only construct shallow structure representations that contain basic argument-predicate relations but lack detailed syntactic information, and therefore their comprehension relies almost exclusively on lexical-semantic and pragmatic information. The SSH has brought many researchers' attention to the mechanism that native (L1) speakers and L2 speakers utilize in sentence processing. However, most studies that Clahsen and Felser (2006a) refer to in support of their SSH focus on filler-gap dependencies in processing L2 wh-questions or relative clauses, and it is not clear from studies in the L2 processing literature whether the SSH can be confirmed in any "gapless" structure in L2 sentence processing. In this article, we will report an empirical study investigating L2 processing of the Chinese base-generated-topic sentence, which we hope can provide useful evidence about how "gapless" structures are processed in L2 as well as L1 sentence processing.

In Mandarin Chinese (henceforth Chinese), it is common to have sentences like (1), where the topic *Shuiguo* "fruits" is a base-generated topic and is not a constituent derived from inside the sentence. There is no gap in the sentence and all positions in the argument structure are phonetically and lexically filled. Since the Chinese base-generated-topic sentence has a "gapless" structure, it would be interesting to see whether L1 and L2 parsers would initially process the first two NPs, i.e. *Shuiguo* "fruits" and *wo* "I", as the topic and the subject of the sentence respectively, whether any restructuring of the initial analysis would have to take place, and how the subcategorization need of the verb *chi*

48 "eat" is satisfied in the sentence processing. Semantic constraints of the Chinese base-
49 generated-topic sentence will be examined as well.
50
51 (1) *Shuiguo wo zui ai chi xiangjiao.*
52 fruit I most love eat banana
53 As for fruits, I like to eat bananas the most.
54

55 ## 2. Base-generated-topic Sentences in Chinese

56 Chinese has been considered a topic-prominent language in the literature, in contrast to
57 English, which is claimed to be a subject-prominent language (cf. Li and Thompson,
58 1976, 1981; Huang, 1984a,b; Xu, 2006; Xu and Langendoen,1985; Huang, Li, and Li,
59 2009; among many others). In Chinese, it is common to have a topic at the sentence-
60 initial position, followed by a sentence, which serves as a comment about the topic. This
61 can be exemplified in (1), in which the topic S*huiguo* "Fruits" has no syntactic relation
62 with any constituent in the comment and there is no gap in the comment either. This
63 "gapless" topic structure suggests that the sentence-initial topic is base-generated in the
64 left periphery and is not a result of movement. Sentences like the one in (1) are what
65 Gundel (1988) calls the topic-comment construction and are also known in the literature
66 as a "Chinese-style" topic structure, a term which originated in Chafe (1976). English
67 does not allow sentences with a base-generated topic, and for the topic in the "Chinese-
68 style" topic structure to be acceptable in English, it is usually encoded into a prepositional
69 phrase like *as for…*, *of…*, or *speaking of…*, as can be seen in the English translation of
70 the topic in (1).

Li and Thompson (1976) suggest that the notion of topic in Chinese is as basic as that of subject in general grammar descriptions and that the topic in Chinese cannot be viewed as derived by movement from some argument position in the sentence. They point out that an important characteristic of the topic in Chinese is that it is independent of the verb and need not be an argument of a predicative constituent in the sentence. From the sentence in (1), we can see that the topic *Shuiguo* "Fruits" is not determined by the verb, and sentences of this type provide clear evidence that the topic leaves no "gap" in the sentence and that no process of movement is involved.[1, 2]

Huang (1984a) argues that topic-comment sentences in Chinese "must count as basic forms in that they cannot be plausibly derived from other 'more basic' forms" (p. 550), and this view is also shared by Xu (1986) and Cole (1987). In this article, we assume that the topic in the "Chinese-style" topic structure is base-generated in the Specifier of the Topic Projection (TopP) in the left periphery of the sentence, in the sense of Rizzi (1997).

It should be pointed out that although base-generated-topic (henceforth BGT) sentences are common in Chinese, Chinese also allows topic structures in which the topic is a result of movement, as indicated in the sentence in (2), where the topic *Zhe ben shu* "this book" is originally based-generated as the object of the verb *xihuan* "like" before it

[1] The topic in (1) should not be treated as being the same as the left dislocated NP, *John*, in the following example, because the left dislocated NP in English, although also base-generated, has to be co-indexed with a constituent in the sentence, as shown by the co-indexation between *John* and the pronoun *him* in the following example.

(i) John$_i$ I don't trust him$_i$.

[2] Shi (2000) argues that every topic must be syntactically licensed and that it cannot be merely semantically related to the comment as a whole. However, his argument has been challenged by many linguists, including Pan and Hu (2001, 2002), who provide counter-evidence to Shi's analysis. As pointed out by Xu (2006), if Shi were correct, there would be no significant structural difference between topic-prominent languages and other languages.

is topicalized to the Specifier of TopP. While this kind of Chinese topic sentences are not the focus of the study, their existence in Chinese is likely to affect both native Chinese speakers' and L2 Chinese learners' processing of Chinese BGT sentences, as will be shown in our empirical study.

(2) *Zhe ben shu$_i$ wo bu xihuan t_i.*
this CL book I not like
This book, I don't like.

3. Semantic Constraint on Base-generated Topics in Chinese

The base-generation of the topic is subject to various pragmatic and semantic constraints. In (1), the topic *Shuiguo* "Fruits" on the one hand, and the NP *xiangjiao* "banana" in the comment on the other, form a hyponymy relationship, with the topic being the superordinate and the NP in the comment being its hyponym. The Chinese-style topic structure would not be felicitous if the hyponymy relationship is violated even if the topic is base-generated in Spec TopP. As can be seen in (3a), the topic *Xiangjiao* "bananas" is a hyponym while the NP *shuiguo* "fruit" in the comment is the superordinate. This reversed hyponymy relationship leads to the infelicity of the sentence. Similarly, the sentence would be unacceptable if the base-generated topic forms a sisterhood relationship with the NP in the comment, as shown between *xiangjiao* "bananas" and *pingguo* "apples" in (3b), or has no hyponymy relationship with the NP in the comment, as shown between *Shuiguo* "Fruits" and *binggan* "biscuits" in (3c).

(3)a. **Xiangjiao wo zui ai chi shuiguo*.

banana I most love eat fruit

*Bananas, I like to eat fruits the most.

b. **Xiangjiao wo zui ai chi pingguo.*

banana I most love eat apples

*Bananas, I like to eat apples the most.

c. **Shuiguo wo zui ai chi binggan.*

fruit I most love eat biscuit

*Fruits, I like to eat biscuits the most.

The relation between the topic and the comment is commonly characterized as "aboutness" in the literature, and according to Gundel's (1985) formulation of "aboutness", "an entity, E, is the pragmatic topic of a sentence, S, iff S is intended to increase the addressee's knowledge about, request information about or otherwise get the addressee to act with respect to E" (p. 86). Takami and Kamio (1994) also point out that the topic must be characterized by the rest of the sentence. Based on the formulations of "aboutness", we can argue that the infelicity of the sentences in (3) is due to the violation of the aboutness condition.

4. Studies of L2 Topic Structures

Since the stimulating paper by Li and Thompson (1976), topic-prominence as a linguistic phenomenon has attracted much attention, not only from scholars working on language typology and linguistic theory, but also from researchers in L2 acquisition, particularly

those working with special reference to target languages such as Chinese. An interesting question that people ask is whether native speakers of a subject-prominent language such as English are able to acquire features of a topic-prominent language like Chinese.

Jin (1994) conducted a L2 study examining the behaviours of adult native English speakers acquiring Chinese as a topic-prominent language. She used three production tasks, oral interviews, story retelling and free compositions, to elicit data from English speakers' L2 Chinese. The results indicate that English speakers go through a process of systematically transferring subject-prominence features to their L2 Chinese at early stages. When their Chinese proficiency is limited, they tend to rely on the subject-prominent structure of English in their L2 Chinese, which Jin argues is evidence of typological transfer from a L1 subject-prominent language to a L2 topic-prominent language. When learners have reached what Jin calls a requisite proficiency, they become sensitive to syntactic features of topic-prominence in Chinese and start to use base-generated topics. Similar results are also reported in Jung's (2004) study of L2 acquisition of Korean, a topic-prominent language, by English speaker. In Jung's study, English speakers are found to be able to use base-generated topics in their L2 Korean writing at an advanced level and there is evidence of L1 transfer of subject-prominence to English speakers' L2 Korean at earlier stages. Both Jin's and Jung's studies demonstrate that base-generated topics are acquirable by speakers of a subject-prominent language. However, it is not clear from these studies in what way Chinese or Korean BGT sentences are processed in real time by L2 learners and whether the L2 syntax of base-generated topics is governed by the semantic constraint discussed in the previous section.

AUTHOR (1995) carried out a study specifically investigating the acquisition of base-generated topics in Chinese by English-speaking learners. Over 100 English-speakers were involved in the study and the results of an acceptability judgement test indicate that although English-speaking learners of Chinese at earlier or intermediate stages had difficulty accepting sentences with a base-generated topic like (4), there is clear evidence that the base-generated topic in Chinese can be eventually acquired by English-speaking learners. However, AUTHOR's study, like Jin's and Jung's, only indicates that the base-generated topic can be established in English speakers' L2 Chinese syntax, and it does not provide us with any information about how Chinese BGT sentences are processed in real time and whether these sentences are regulated by the semantic constraint in L2 Chinese.

(4) *Ta jia li de ren wo zhi jian-guo ta mama*
his family in DE people I only meet EXP his mother
*People in his family, I have only met his mother.
(=(6) in AUTHOR (1995))

Another L2 study of Chinese as a topic-prominent language was conducted by Cao, Yang, Huang, Gao and Cui (2006), in which native speakers of Japanese, Korean and English were included in order to examine whether speakers of topic-prominent languages like Japanese and Korean have advantages over speakers of English in their L2 acquisition of Chinese topic structures. Their results suggest that the topic-prominence in learners' L1 can facilitate the acquisition of topic-prominence in their L2 because evidence was found

in an acceptability judgment task that base-generated topics were accepted by Japanese- and Korean-speaking learners but not by English-speaking learners. However, learners in this study were all at "intermediate and high-intermediate levels" and no learner at an advanced level was included. As shown in Jin's (1994) and AUTHOR's (1995) studies above, English-speaking learners at advanced levels are able to acquire base-generated topics, like *Na ke shu* "that tree" in (5), in their L2 Chinese.

(5) *Na ke shu, yezi hen da.*
that CL tree leaf very big
That tree has big leaves.

Studies reviewed above show that the syntax of base-generated topics is acquirable by English speakers. However, no evidence is provided in the literature as to whether L2 learners can process Chinese BGT sentences in the same way as native Chinese speakers and whether the BGT structure in L2 Chinese is governed by the relevant semantic constraint.

5. L2 Sentence Processing and the Shallow Structure Hypothesis

In recent years, an increasing number of researchers have paid attention to the mechanism that native speakers and L2 speakers utilize in sentence processing. Some have argued that the lack of native-like ultimate attainment in adult L2 acquisition can at least partially be attributable to adult L2 sentence processing problems, which include problems that adult L2 learners may have in integrating different information sources in

real-time L2 sentence processing. For example, Marinis, Roberts, Felser and Clashen (2005) carried out a self-paced reading task with four groups of L2 learners of English whose L1s were Chinese, Japanese, German and Greek, as well as a group of native English speakers. Their study focuses on sentences involving long distance wh-dependencies in sentences like (6a) and (6b).

(6)a. The manager *who* the consultant claimed _____ that the new proposal had pleased _____ will hire five workers tomorrow. (intermediate gap)

b. The manager who the consultant's claim about the new proposal had pleased _____ will hire five workers tomorrow. (no intermediate gap)

As sentences like (6a) involve wh-extraction from a complement clause, an intermediate gap is assumed to be present at the intervening clause boundary, which breaks the long dependency up into two shorter ones. However, no such intermediate gap is present in sentences like (6b) which involve extraction across a complex NP. In the study, it is assumed that although the linear distance between the filler, i.e. *who*, and its subcategorizer, i.e. *pleased*, is the same in both (6a) and (6b), integrating the filler with its subcategorizing verb should be facilitated by the availability of an intermediate gap at the clause boundary if the parser consults a mental representation of the filler at this point during processing. In the study, longer reading times were observed in the native English speaker group, but not in L2 groups, at the intervening clause boundary in the extraction-VP condition, as in (6a), compared to the corresponding nonextraction condition, as in (6b). The results also show that only native English speakers', but not L2 groups',

reading times at the segment containing the subcategorizing verb are significantly shorter for sentences that contain an intermediate gap, as in (6a), than for those that do not, as in (6b). The interpretation by the authors is that native English speakers associate the filler with an intermediate gap when processing sentences involving wh-extraction from an embedded clause, which facilitates filler integration later on. In contrast, there is no such interaction or intermediate gap effect in L2 processing although L2 learners, like native speakers, are able to integrate the filler with its subcategorizing verb in their sentence processing. The authors conclude that L2 learners "do not use native-like, phrase-structure-based processing mechanisms … during online comprehension." (p. 72)

On the basis of results of this type (also results from Felser, Roberts, Marinis, and Gross, 2003; Papadopoulou and Clahsen, 2003; among others), Clashen and Felser (2006a) propose the Shallow Structure Hypothesis that adult L2 learners are guided by lexical-semantic cues in their sentence processing in the same way as native speakers, but L2 learners' sensitivity to syntactic information is restricted and therefore their syntactic representations in sentence processing are shallower than those of native speakers. According to Clahsen and Felser, the mental processes involved in L2 learners' sentence processing are qualitatively different from those used in native speakers' L1 processing; unlike native speakers, L2 learners do not rely on structure-based processing strategies when solving ambiguities in the L2. Instead, they process L2 sentences primarily on the basis of lexical-semantic and pragmatic information.[3]

However, the SSH has been challenged by an increasing number L2 sentence processing studies in the literature. Omaki and Schulz (2011) conducted a self-paced

[3] Some other studies in the literature, e.g. Frenck-Mestre and Pynte (1997), Juffs (1998), Dussias and Piñar (2010), Dinçtopal-Deniz (2010), do provide evidence in support of the SSH.

reading experiment, comparing the extent to which advanced Spanish-speaking L2 learners of English and native English speakers make use of the relative clause (RC) island constraint in constructing filler-gap dependencies like (7).

(7)a. The murder case$_i$ that the law students [RC who learned about the constitution] discussed ____$_i$ was going to be on the exam.

b. *The murder case$_i$ that the law students [RC who learned about ____$_i$] discussed ____$_i$ was going to be on the exam.

In the grammatical condition (7a), the dependency between the filler *the murder case* and the verb *discussed* does not cross the RC island.[4] The ungrammatical counterpart in (7b) is constructed by taking the sentence in (7a) and deleting the object of an obligatorily transitive preposition inside the RC, such that the dependency between *the murder case* and the preposition *about* crosses the RC island. The results show that L2 learners pattern with native speakers in postulating a gap in non-island conditions but not in island conditions, suggesting that syntactic island constraints successfully blocked ungrammatical long-distance dependency formation in both native and non-native speakers' sentence processing, contra to the prediction of the SSH.

Williams (2006) also conducted a study of L2 processing of wh-dependencies, which was based on an earlier study by Williams, Mobius and Kim (2001). In the study, participants read sentences like (8a) and (8b) word-by-word in a self-paced fashion and pressed a button as soon as the sentence stopped making sense to them.

[4] It was acknowledged by the authors that the acceptability of (7a) is somewhat degraded due to the large processing cost incurred by the presence of more than one temporarily incomplete clause,

(8) a. Which car$_i$ did the tourist buy the really expensive radio for e_i two months ago?

b. Which friend$_i$ did the tourist buy the really expensive radio for e_i two months ago?

It was hypothesized that increased RTs should be found in the region after the determiner and prior to the noun, i.e. the region of *really expensive*, if L2 learners process the syntactic cues similarly to native speakers. On the other hand, if native English speakers start the reanalysis process based on syntactic cues from the determiner which informs the parser that an NP follows, while L2 learners do the reanalysis only based on lexical information after encountering the noun *radio*, this would indicate that L2 learners ignore the syntactic cue from the determiner and do the reanalysis on the basis of the noun, i.e. *radio*. The results showed that both native English speakers and L2 learners had longer RTs before the noun, indicating that the reanalysis started after the determiner and before the noun. This suggests that both native speakers' and L2 learners' sentence processing can be structurally driven, which is not in accordance with the SSH.[5]

In a more recent study, Pliatsikas and Marinis (2013) investigated the effect of naturalistic exposure in processing wh-dependencies. They used the same experimental sentences as in Marinis, Roberts, Felser and Clashen (2005), and examined the processing of sentences involving intermediate gaps like those in (6a) and (6b). The participants included 26 advanced Greek-speaking learners of L2 English with an average 9 years of naturalistic exposure, 30 with classroom exposure and 30 native English speakers. Results from a self-paced reading task show that L2 learners with naturalistic exposure

[5] See Aldwayan, Fiorentino and Gabriele (2010) for another L2 study of wh-movement processing, which also disconfirms the SSH.

are able to have native-like processing of the intermediate gaps like those in (6a) and (6b). That is, the naturalistic exposure L2 group converged with the native English group in revealing facilitation in processing the final gap when an intermediate gap was present. This suggests that extended immersion in naturalistic target language environments can lead to native-like abstract syntactic processing in L2, a case not predicted by the SSH but a case confirming VanPatten and Jegerski's (2010) prediction that differences in populations can be a factor affecting native-like abstract syntactic processing in L2.

As indicated above, the majority of L2 sentence processing studies in the literature, whether in support of or against the SSH, use filler-gap dependencies in either relative clauses or wh-questions in their investigations of L2 sentence processing, and there seems to be a lack of structural varieties in examining L2 sentence processing. In this aspect, the BGT sentence in L2 Chinese provides a good alternative for the investigation of L2 sentence processing. It has several advantages. Firstly, it has a "gapless" structure in a sense that neither its syntactic structure nor its argument structure contains any empty category as no movement of any constituent takes place from inside the sentence, thus no "gap" in the sentence. Secondly, unlike English wh-questions or relative clauses, in which the fronted wh-word, when it is processed, can reveal its "filler" status because of the *wh*- marking on the wh-word, the word in the topic position in the Chinese topic sentence does not have any overt marking whatsoever, and therefore it does not give away any information as to whether it is a potential filler or whether there is a gap in the sentence.

6. Research Questions

Based on the analyses above, we had the following research questions for the empirical study:

1. In what way would the topic and the subject, i.e. the first NP and then the second NP, in the Chinese topic sentence be processed initially and subsequently in L1 and L2 processing? According to the SSH, there would be differences between native speakers and L2 learners when solving syntactic ambiguities, because L2 learners, unlike native speakers, do not rely on structure-based processing strategies when solving syntactic ambiguities in L2 processing.
2. Since the topic in Chinese can be either base-generated in the sentence-initial position or derived from movement of a constituent from inside the sentence, and given that there is no overt marking at all on the topic NP as a potential filler, would the syntactic (re-)analysis of the first NP as the topic of the sentence assign to the topic NP a role of potential filler in L1 and L2 BGT sentence processing?
3. Would L2 learners be sensitive to the semantic requirement, i.e. the hyponymy relationship between the topic and a relevant item in Chinese BGT sentences, an ability predicted to be available in L2 processing by the SSH?

Unlike the wh-word in English wh-questions and relative clauses, which reveals itself as a potential filler by the overt wh-marking attached to it, any possible acquisition of the topic role by the first NP in Chinese topic sentences would be triggered by the syntactic (re-)analysis of the first NP (and probably also the second NP) in the processing of

Chinese BGT sentences. This would be triggered syntactically, rather than lexical-semantically or morphologically.

7. Experiment

Four tasks were included in the experiment: a) a self-paced reading task, which was used to examine the on-line processing of Chinese BGT sentences by English-speaking learners of Chinese as well as native Chinese speakers; b) a grammaticality judgment task, which was to check whether participants in the experiment had the grammatical knowledge of the Chinese BGT sentences, which is believed to be a prerequisite for processing this type of sentences; c) a cloze test for measuring participants' Chinese language proficiency; d) a questionnaire to collect information about participants' Chinese language learning background and the self-evaluation of their own Chinese language proficiency. The tasks were presented in the above order.

7.1 Participants

A total of 44 English-speaking learners of L2 Chinese and 23 native Chinese speakers participated in the experiment. The learners were highly proficient in Chinese and they were diplomats and business people working and living in China as well as English-speaking academics and students teaching or studying in universities in the U.K. or China.[6] The native Chinese speakers were graduate students, academics at universities in the U.K. or China, or office workers in China. Moderate payment was given to each

[6] It had originally been designed to divide the English-speaking learners of Chinese into two Chinese language proficiency groups, but it was decided to collapse the two groups into one because of similar results of the two groups in the experiment.

participant as a token of thanks for their participation in the experiment. Before data analyses were conducted, each participant's data underwent a screening check on the basis of their performance in the grammaticality judgment task, which was designed to identify participants who can demonstrate clear knowledge of the BGT structure in Chinese. Ensuring that participants could handle the Chinese BGT structure is important in order for us to rule out the possibility that any possible problem in their sentence processing in the experiment is due to the lack of grammatical competence in this area. As a result of this screening check, 10 participants from the Learner Group (L-Group) and 4 from the Native Group (N-Group)[7] were excluded because of their failure to pass the screening check (see below for detailed information about the screening check). Participants in the learner group were also asked to do a self-evaluation of their L2 Chinese with regard to speaking, listening, reading and writing. Table 1 provides information about participants included in the study. All participants had normal or corrected-to-normal vision at the time of the experiment.

Table 1. Information about the participants included in the study (standard deviations in parentheses)

	L-Group *(SD)*	N-Group *(SD)*
Number	44 - 10 = 34	23 - 4 = 19
Mean score of the cloze test (maxi =40)	35 *(2.7)*	38 *(1.6)*

[7] The four native speakers were excluded not because of their Chinese language competence but because of their carelessness in doing the experiment as they admitted afterwards that they were not paying careful attention while doing the experiment.

Mean age		31 *(15)*	28 *(5.6)*
Mean age starting Chinese		18.6 *(3.3)*	N/A
Years of learning Chinese		12.25 *(13.17)*	N/A
Years of residence in China/Taiwan		6.3 *(9.29)*[a]	N/A
Self-evaluation of Chinese:[b]	Speaking	4.6 *(0.78)*	N/A
	Listening	4.5 *(0.83)*	N/A
	Reading	4.6 *(0.89)*	N/A
	Writing	3.9 *(0.84)*	N/A

[a] The range is 1 year to 34 years.

[b] On a 1-6 scale where 1 = beginner level, 2 = post-beginner level, 3 = intermediate level, 4 = post-intermediate level, 5 = advanced level, 6 = very advanced level.

7.2 Cloze Test

A cloze test with 40 blanks was administered to all participants to assess their Chinese language proficiency. Although the result of an independent-sample *t*-test indicates that the native Chinese group performed significantly better than the learner group in the cloze test ($t(51)=5.544$, $p<0.001$), participants in the learner group are considered to be advanced learners of Chinese, given the information about the average number of years of their residence in Chinese-speaking environments (6.3 years), the average number of years of their studying Chinese (12.25 years) and their mean score in the cloze test (35), as shown in Table 1.

7.3 Self-paced Reading Task

In total, there were 28 types of stimuli in the self-paced reading task, and each type had 6 test sentences in it. Altogether, there were 168 sentences included in the self-paced reading task, out of which 3 counterbalanced presentation lists were constructed on the basis of a Latin square design, with each list containing 56 sentences. The stimuli relevant to the study reported in this article are the 4 types presented in Table 2, with each type having 6 test sentences (see the Appendix for all the 24 test sentences used in these 4 types). These 24 sentences were embedded in the other 144 sentences, which tested processing of other language structures (e.g. word orders of unaccusative and unergative verbs, etc.) in Chinese and are considered as suitable fillers for the examination of the 4 types relevant to this study.

Table 2. Sample set of experimental stimuli[8]

	Regions								
Types	1	2	3	4	5	6	7	8	9
A. BGT sentence with S-H	水果 fruit	我 I	最 most	爱吃 like eat	香蕉 banana	, 所以 so	我 I	经常 often	买香蕉。 buy banana
B.*BGT sentence with H-S	香蕉 banana	我 I	最 most	爱吃 like eat	水果 fruit	, 所以 so	我 I	经常 often	买水果。 buy fruit
C. *BGT sentence with sisterhood	苹果 apple	我 I	最 most	爱吃 like eat	香蕉 banana	, 所以 so	我 I	经常 often	买香蕉。 buy banana
D. Non-BGT sentence	以前 before	我 I	最 most	爱吃 like eat	香蕉 banana	, 所以 so	我 I	经常 often	买香蕉。 buy banana

The rationale for including these 4 types of sentences is that any locally increased processing efforts should be detected in longer reading times on a given region in comparison with the same region in a controlled sentence. For example, processing the pronoun *wo* "I" in Region 2 in Types A, B, and C is predicted to take longer times than

[8] The English gloss is given here only for the reader of this article, and it was not available in the experiment. Also, in Column 1, "S-H" stands for the superordinate-hyponym relationship between the topic and the relevant NP in the sentence, "H-S" stands for a hyponym-superordinate relationship, and "sisterhood" for a sisterhood relationship.

processing the same pronoun in Region 2 in Type D because the parser is likely to go back to the first NP (i.e. *shuiguo* "fruit", *xiangjiao* "banana" or *pingguo* "apple" in Region 1 in Types A, B, and C to correct the initial assignment of the first NP as the subject of the sentence and to re-analyze it as a topic. However, processing the pronoun *wo* "I" in Region 2 in Type D is unlikely to incur any extra cost or re-analysis because the first element that the parser processes is an adverb *yiqian* "before" in Region 1, which frequently appears at the beginning of the sentence in human languages, and the parser can easily integrate the pronoun *wo* "I" in Region 2 into the subject position without any reanalysis.[9]

Similarly, it is predicted that processing the NPs in Region 5 in Types A, B, and C (i.e. *xiangjiao* "banana" or *shuiguo* "fruit" will increase processing costs if the topic in Region 1 is temporarily stored in working memory as a topic resulted from topicalization. (Recall that in Chinese, a topic can be a result of topicalization of a constituent from inside the sentence, or a base-generated topic in the sentence initial position.) If the topic is stored in working memory as a potential filler, the parser may be looking for a gap in the sentence for the topic to fill as a result of the Filler-Driven Strategy (cf. Frazier and Clifton, 1989) or the Principle of Immediate Association (cf. Pickering and Barry, 1991).

[9] It should be noted that *wo* in Chinese does not have any case marking. That is, there is no morphological change whatsoever when *wo* is used as a subject pronoun, or as a (topicalized) object pronoun as in (i), or as a possessive pronoun as in (ii).

(i) Wo$_i$ ta bu renshi t_i.
Me he not know
*"Me, he doesn't know."

(ii) Shuiguo wo mama xihuan chi xiangjiao
Fruits my mother like eat bananas
"As for fruits, my mother likes to eat bananas."

When the verbal phrase *ai chi* "like to eat" in Region 4 is processed, the parser may treat it as the subcategorizer for the topic. However, once the NP in Region 5 in Types A, B and C is encountered, the parser will have to cancel its previous analysis and re-analyze the topic in working memory as a base-generated topic rather than a topic as a result of topicalization. Furthermore, the NPs in Region 5 in Types B and C are predicted to take even longer times for the parser to process than the NP in the same region in Type A because the former violate the hyponymy relationship between the base-generated topic in Region 1 and the NP in the object position in Region 5, with the topic being a hyponym of the object, i.e. a hyponym-superordinate (H-S) relationship, in Type B and with the topic and the NP in the object position forming a sisterhood relationship in Type C. Processing these semantic conflicts is likely to incur additional processing costs. However, this kind of processing delay is predicted not to occur in Type D because there is no topic in working memory for the NP *xiangjiao* "banana" in Region 5 in Type D to check against and therefore the parser can process the NP much faster.

Given the predictions made above, we treat Region 2 and Region 5 as *critical regions*, and as there may possibly be spill-over effects, we also consider Region 3 and Regions 6 and 7 as *post-critical regions* respectively.

In order to make sure that participants paid attention to the content of the test sentences, they were required, after reading each test sentence, to answer a true/false comprehension question about the sentence that they had just read. Below each comprehension question, there were two options on the screen, i.e. *dui* "true" and *bu dui* "false", with one on the left-hand side and the other on the right-hand side. Participants were instructed to press a designated key on the left half of the keyboard or a designated

key on the right half to answer the true/false comprehension question. For half of the comprehension questions, the correct answers appeared on the left-hand side of the screen, and for the other half, the correct answers on the right-hand side. Comprehension questions eliciting true or false answers were evenly distributed across all test sentence types. See the Appendix for comprehension questions for all test sentences, and the correct answers.

All sentences were presented in Chinese characters, and efforts were made to ensure that the number of characters in the same region across all the sentence types was the same, particularly in the critical regions and post-critical regions.

7.4 Grammaticality Judgment Task

As successful processing of the BGT sentence is dependent upon the availability of the relevant knowledge of the sentence structure, a grammaticality judgment task was designed to help to identify participants who had acquired the knowledge of the BGT sentence in Chinese and to exclude participants who lacked the relevant grammatical knowledge. All participants did the grammaticality judgment task after they had done the self-paced reading task, and this order was to try to minimize any possible effect of participants' awareness of the focuses of the experiment on the processing of similar structures in the self-paced reading task. Test sentences used in the grammaticality judgment task were exactly the same as those 168 sentences used in the self-paced reading task, except that irrelevant parts of the test sentences were deleted. That is, the words like those in Regions 6-8 in test sentences presented in Table 2 were deleted and the comma was replaced with a full stop. In the grammaticality judgment task, each

participant judged the 168 sentences in the same order as presented in the self-paced reading task. Test sentences were presented on the screen one at a time, and participants were instructed to judge whether the sentence was grammatically correct in Chinese or not. Below each sentence, there were two options on the screen, i.e. *zhengque* "correct" and *bu zhengque* "incorrect", with one on the left-hand side and the other on the right-hand side. Participants were instructed to press a designated key on the left or right half of the keyboard to judge the grammaticality of each sentence.

7.5 Procedures

In the self-paced reading task, the 24 sentences represented by the 4 types exemplified in Table 2 were embedded in 144 sentences, which were used to examine other linguistic phenomena in L2 Chinese. In the experiment, 3 counterbalanced presentation lists were constructed out of these 168 sentences, and one third of the participants did the 3 lists in the order to 1-2-3, one third in the order of 2-3-1 and one third in 3-1-2. The test sentences in each list were pseudo-randomized and mixed with the fillers. It took proximately 10 minutes for a participant to finish each list, and there was a break of minimally 10 minutes and maximally 4 days between any two lists.[10]

The main paradigm used in the experiment was a segment-by-segment non-cumulative self-paced moving windows task (c.f. Just, Carpenter, and Woolley, 1982), in which the participant read each sentence on a computer screen one segment at a time. Participants were aware that they were participating in a language experiment and that they would be reading sentences presented on a computer screen segment by segment.

[10] The majority of participants had a break of 10 minutes, and only a few had to have a longer break of 1 to 4 days due to their other commitments.

They were asked to read each segment as carefully and quickly as they could, and they were then prompted to answer a comprehension question when the last segment of the sentence disappeared from the screen. They were told that their reading time of each segment and their answer to the comprehension question would be recorded by the computer and would be used for the study. Each sentence began with an asterisk on the left edge of the screen, and participants were instructed to press the space bar on the key board to obtain the first segment. They then pressed the space bar for the next segment, which appeared to the right of the preceding segment after the preceding segment had disappeared. They continued doing this until they saw a segment followed by a full stop, which indicated the end of the sentence. When they pressed the space bar again at this moment, a true/false comprehension question appeared on the screen, and they had to press an appropriate key to indicate "true" or "false". The presentation of the sentences and the collection of the data were done with DMDX presentation software (Forster and Forster, 2003).

Participants were tested individually and the experiment was conducted in a quiet room in various cities in the UK and China. Efforts were made to include only daily-life vocabulary in the test sentences. A short list of relatively less common words was sent to each participant a few days before the experiment, and each English-speaking participant had to orally translate the words on the list into English at the very beginning of the experiment. This was to ensure that participants had no problem understanding the words used in the experiment, and none of the participants had problems with the vocabulary list. The participant received both written and oral instructions on how to do the tasks, and the

514 self-paced reading task was preceded by 6 practice sentences to familiarize participants
515 with the procedure.

516 The self-paced reading task was followed by the grammaticality judgement task, and
517 after both the self-paced reading task and grammaticality judgment task were completed,
518 each participant also did a language background questionnaire for biographical
519 information and the cloze test.

520 As the availability of the relevant grammatical knowledge of Chinese BGT sentences
521 is a prerequisite for successful processing of BGT sentences in Chinese and incomplete
522 knowledge of the target language can affect processing behaviours, a rather stringent
523 criterion was used to identify participants who showed clear knowledge of the BGT
524 sentence in Chinese. That is, to be included in the study, a participant must correctly
525 judge at least 10 of the 12 grammatical sentences in Types A and D (83% accuracy) and
526 the same participant must correctly reject at least 10 of the 12 ungrammatical sentences in
527 Types B and C (83% accuracy). As a result of this screening procedure, 10 participants
528 from the learner group and 4 from the native Chinese group were excluded from the study,
529 as indicated in Table 1. The high percentages of the learner group and the native Chinese
530 group in correctly judging the grammatical and ungrammatical sentences, as shown in
531 Table 3, suggest that both groups had knowledge of the BGT sentence in Chinese.

532

533 **Table 3.** Percentage of each group in correctly judging the grammatical and
534 ungrammatical sentences in the grammaticality judgment task

	Grammatical	Ungrammatical
Learner group	93%	95%
Native Chinese group	93%	98%

535

536 **7.6 Results**

Recall that in the self-paced reading task, each test sentence was followed by a true/false comprehension question to make sure that participants paid attention to the content of test sentences. The native Chinese group's percentage in correctly answering the comprehension questions related to the 4 types of sentences is 96.1%, and the learner group's is 90.2%. This indicates that both the native Chinese speakers and L2 Chinese learners were, in general, paying attention to the contents of test sentences in the self-paced reading task. Test items for which comprehension questions were not answered correctly were excluded from the reading time (RT) analyses.

Before analyzing the RT data, we also dealt with RT outliers. Any RT longer than 2000ms was eliminated, and any RT that was 2 standard deviations from the relevant cell mean of the relevant participant was also eliminated. The percentage of data thus affected was 5.3% in the native Chinese group and 9.8% in the learner group.[11]

7.6.1 Native Chinese Speakers

As we can see from the second column of Table 4 and Region 1 in Figure 1, native Chinese speakers' mean RTs for the first region of all the 4 sentence types are similar and no significant difference is found in a one-way ANOVA between any first region of the four types of sentences, $F(3, 432) = 0.346$, $p = 0.792$. This is unsurprising because

[11] A disproportionate number of outliers were found in the last region, i.e. Region 9, in both learner group's data and native Chinese group's data. These outliers are mainly RTs longer than 2000ms. This is likely to be due to the appearance of the full stop in the last region, which triggered participants to start to anticipate the comprehension question even before they were prompted. It could also be due to the increased number of characters in Region 9, i.e. 3 characters, or to the fact that this was the last region. If Region 9 is to be excluded, the total percentage of data thus affected would be 4.0% in the native Chinese group and 6.1% in the learner group.

555 theoretically no reanalysis or restructuring is expected in Region 1 of any of the 4

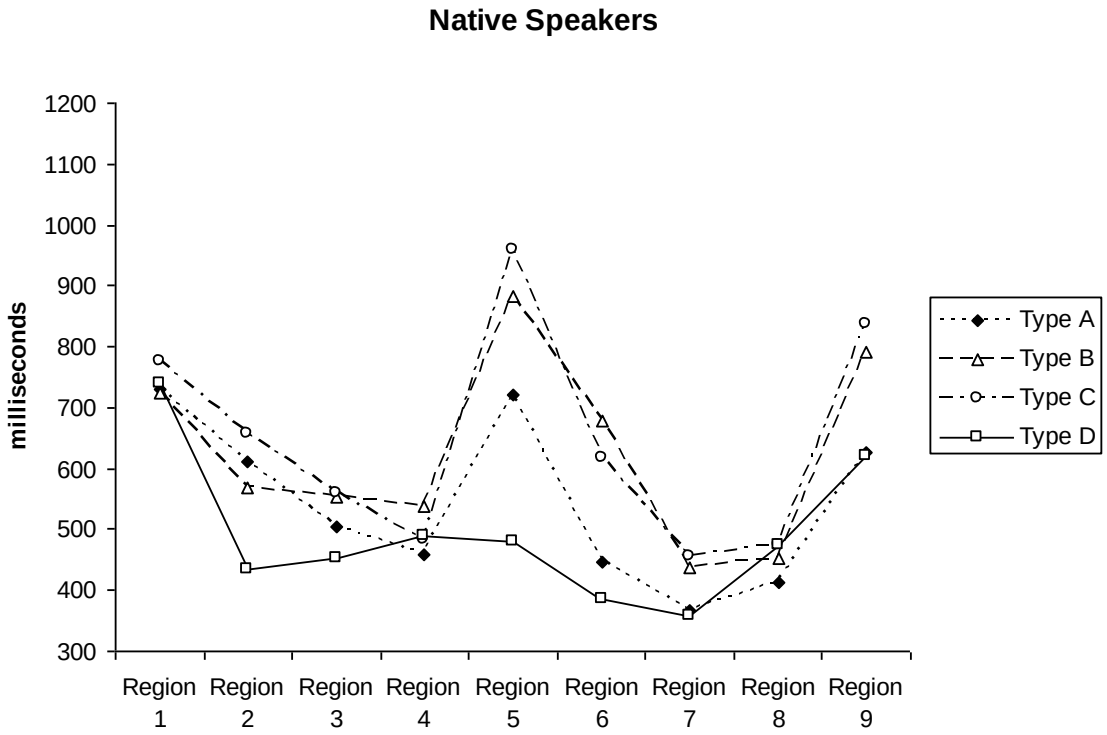
556 sentence types.

557

558 **Table 4.** Native Chinese group's mean reading times (in milliseconds) and standard
559 deviations (in parentheses) for each of the regions of the 4 types of test sentences

	Regions								
Types	1	2	3	4	5	6	7	8	9
A. BGT sentence	729	612	504	459	722	446	368	412	625
with S-H	(409)	(334)	(322)	(206)	(431)	(206)	(200)	(246)	(320)
B.*BGT sentence	723	570	555	537	884	681	437	453	792
with H-S	(355)	(358)	(316)	(307)	(461)	(496)	(271)	(210)	(458)
C. *BGT sentence	777	657	561	484	960	616	456	474	838
with sisterhood	(539)	(455)	(317)	(192)	(420)	(385)	(396)	(284)	(595)
D. Non-BGT	739	433	453	491	480	385	357	474	621
sentence	(368)	(235)	(276)	(228)	(221)	(187)	(201)	(244)	(242)

560



561

562 **Figure 1.** Native speakers' mean reading times for the regions in the 4 sentence types

563

However, significant differences are found in Region 2 in native Chinese speakers' RTs, $F(3, 430) = 8.121$, $p < 0.001$, in spite of the fact that Region 2 of all the four types of sentences includes the same word, i.e. the first-person pronoun *wo* "I" (see Table 2). Post-hoc Scheffé tests indicate that native Chinese speakers' RTs for Region 2 of Type C sentences are significantly longer than their RTs of the same region of Type D sentences. This is believed to be due to the fact that restructuring takes place in Region 2 of Type C sentences, where Region 1, which was originally processed and stored in working memory as a subject of the sentence, is revised and re-assigned to the topic position and Region 2 is then analyzed as the subject of the sentence. This restructuring results in longer RTs when native speakers read Region 2 of Type C sentences. The restructuring does not take place in Region 2 of Type D sentences because Region 1, i.e. *yiqian* "before", was processed and is stored as an adverbial of time, and when Region 2, *wo* "I", in Type D sentences is processed, it is stored in working memory as the subject of the sentence, without triggering any restructuring, and therefore it takes shorter RTs than Region 2 in Type C sentences. As the possible restructuring may not necessarily take place immediately after a relevant region is processed, we decide to combine Region 2, i.e. *wo* "I", and Region 3, i.e. *zui* "most", to see whether there is any spill-over effect (cf. Pearlmutter, Garnsey, and Bock, 1999; Sharkey and Sharkey, 1987; Warren and Gibson, 2002; Jiang 2013) of the restructuring. A one-way ANOVA indicates that there is a significant difference in native speakers' RTs of Regions 2 and 3 combined, $F(3, 423) = 9.405$, $p < 0.001$, and post-hoc Scheffé tests suggest that native speakers' RTs of the two regions in Types A, B and C are significantly longer than their RTs of the two regions in Type D sentences. This confirms our analysis above. That is, restructuring takes place

when Region 2 of Type C sentences is processed and this results in longer RTs. In addition, our data show that the effect of the restructuring is spilled over to Region 3 of Types A, B and C sentences. No restructuring is necessary in Region 2 of Type D sentences and as a result, it leads to shorter RTs.

No significant difference is found in native Chinese speakers' RTs for Region 4 of the four types of sentences, $F(3, 437) = 2.139$, $p = 0.095$. This is expected because the verbs like *ai chi* "like to eat" in Region 4 have the same function across all the four types of sentences, and therefore result in similar RTs. However, significant differences are found in native Chinese speakers' RTs for Region 5 of the four types of sentences, $F(3, 396) = 30.784$, $p < 0.001$. Post-hoc Scheffé tests reveal that native Chinese speakers' RTs for Region 5 in Types B and C are significantly longer than that in Type A, which is also found to be significantly longer than that in Type D. The shortest RT of Region 5 in Type D is believed to be due to the fact that Region 5, i.e. *xiangjiao* "banana", is the object of the verbal phrase in Region 4 *ai chi* "like to eat" and no restructuring is needed here. However, processing sentences of Types A, B and C is different. Recall that a topic is stored in working memory after Regions 1, 2 and probably also 3 of Types A, B and C sentences are processed, and also recall that a topic in Chinese can be base-generated in the sentence initial position or derived from topicalization of a constituent from inside the sentence. In the latter case, the parser will look for a gap in the sentence where the topic is originally derived, as predicted by the Filler-Driven Strategy (cf. Frazier and Clifton, 1989). After the verbs in Region 4 are processed, the parser is likely to expect a gap in the object position in Region 5. However, the encounter of *xiangjiao* "banana" in Region 5 in Types A, B and C forces the parser to revise its previous analysis and re-analyse the

topic in working memory as a base-generated topic rather than a topic derived from inside the sentence. This re-analysis is costly and increases the RTs of Region 5 in Types A, B and C.

Chinese BGT sentences are subject to semantic constraints, and one of the constraints is that the base-generated topic and its related NP in the sentence are required to have a hyponymy relationship. However, the topic in Region 1 and the NP in Region 5 in Types B and C violate such a requirement, with the topic being a hyponym of the NP in Region 5 in Type B, and the topic in Type C having a sisterhood relationship with the NP in Region 5. The longer RTs of Region 5 in Types B and C are considered to represent native Chinese speakers' sensitivity to the violation of the required semantic relationship involved in Chinese BGT sentences. Obviously, detecting such a semantic violation will further prolong the RTs of Region 5 in Types B and C, which are found to be significantly longer than the RT of the same region in Type A, where the topic in Region 1 is a superordinate of the NP in Region 5, meeting the requirement of the semantic relationship for Chinese BGT sentences. To check the spill-over effects, the RTs in Region 5 and 6 are combined, and then those in Regions 5, 6 and 7 are also combined. The data in the two combinations reveal that the effects of re-analyzing the topic in working memory as a base-generated topic rather than a topic derived from inside the sentence and detecting the violation of the required semantic relationship between the base-generated topic and the relevant NP in the sentence are spilled over, not only to Region 6 but also to Region 7; native Chinese speakers' RTs of the combination of Regions 5 and 6 and their RTs of the combination of Regions 5, 6 and 7 for Types B and C are significantly longer than that in Type A, which in turn is significantly longer than

633 that in Type D. No significant difference is found in native Chinese speakers' RTs
634 between Types B and C with regard to Region 5, or the combination of Regions 5 and 6,
635 or the combination of Regions 5, 6 and 7.

636 There is no significant difference in native Chinese speakers' RTs of Region 8 (F(3,
637 447) =1.590, p=0191) and Region 9 (F(3, 410) =0.706, p=0.549) between any of the four
638 sentence types.

639

640 *7.6.2 L2 Chinese Learner Group*

641 Table 5 and Figure 2 provide the learner group's mean RTs for each of the regions of the
642 four types of test sentences. As we can see from the second column of Table 5 and
643 Region 1 in Figure 2, the learner group's mean RTs for the first regions of all the 4
644 sentence types are similar, and no significant difference is found in a one-way ANOVA
645 between any of the first regions of the four types of sentences, F(3, 730) =0.214, p=0.887.
646 This is expected because no restructuring is predicted in our theoretical analysis of
647 Region 1 in any of the 4 sentence types.

648 In Region 2, however, significant differences are found (F(3, 635) =11.229, p<0.001),
649 and post-hoc Scheffé tests reveal that the learner group's RTs of Region 2 in Types B and
650 C are significantly longer than their RTs in Type D. This is in spite of the fact that the
651 second regions across all sentence types are the same word, i.e. the first-person pronoun
652 *wo* "I". The learner group's longer RTs of Region 2 in Types B and C are believed to be
653 due to some restructuring that takes places at this point, which is similar to what occurs in
654 native Chinese speakers' processing of Region 2. That is, unlike Region 1 in Type D,
655 which is an adverbial of time, the first regions in Types B and C are initially processed as

656 the subject of the sentence. When the first-person pronoun *wo* "I" is processed in Region
657 2 in Types B and C, the parser has to restructure its initial analysis, re-analyze Region 1
658 and store it as a topic rather than a subject in working memory. In the restructuring and
659 re-analysis, the subject position is vacated and this makes it possible for the first-person
660 pronoun *wo* "I" in Region 2 to be assigned to it. There is no significant difference in the
661 learner group's RTs between Region 2 in Type A and the same region in Type D, and the
662 restructuring and re-analysis scenario above apparently does not work for Region 2 in
663 Type A although the first four regions in Types A, B and C share the same grammatical
664 structure. However, we cannot rule out the possibility that the restructuring and re-
665 analysis may be slightly delayed. With this in mind, we calculate the data for Regions 2
666 and 3 combined across all the four sentence types. A one-way ANOVA reveals that the
667 learner group has significantly different RTs of Regions 2 and 3 combined between the
668 four sentence types, $F(3, 787) = 9.948$, $p < 0.001$, and post-hoc Scheffé tests indicate that
669 learners' RTs of the combination of Regions 2 and 3 in Types A, B and C are
670 significantly longer than their RTs of the same regions in Type D, which suggests that the
671 restructuring and re-analysis take place when they process Regions 2 and 3 in Types A, B
672 and C, but not in Type D. This finding is similar to what we have found in native Chinese
673 speakers' processing of Regions 2 and 3 of the four sentence types.[12]
674
675 **Table 5.** Learner group's mean reading times (in milliseconds) and standard deviations
676 (in parentheses) for each of the regions of the 4 types of test sentences

[12] The longer RTs in L2 learners' and native Chinese speakers' processing of Regions 2 and 3 of Types A, B and C sentences could also be due to the ambiguity of *wo* in Region 2 as a subject pronoun, or a possessive pronoun, or a (topicalized) object pronoun (see Note 8 above). This would provide further evidence that L2 learners are sensitive to syntactic cues and structure-based information.

	Regions								
Types	1	2	3	4	5	6	7	8	9
A. BGT sentence	889	552	529	610	922	583	411	532	954
with S-H	(395)	(344)	(299)	(315)	(398)	(312)	(198)	(334)	(479)
B.*BGT sentence	918	572	600	647	1139	804	527	611	874
with H-S	(402)	(340)	(362)	(264)	(480)	(442)	(355)	(378)	(582)
C. *BGT sentence	893	562	504	614	1125	749	472	568	959
with sisterhood	(353)	(307)	(232)	(264)	(556)	(405)	(316)	(319)	(496)
D. Non-BGT	906	405	483	636	558	454	387	564	947
sentence	(371)	(314)	(425)	(264)	(269)	(331)	(234)	(244)	(682)

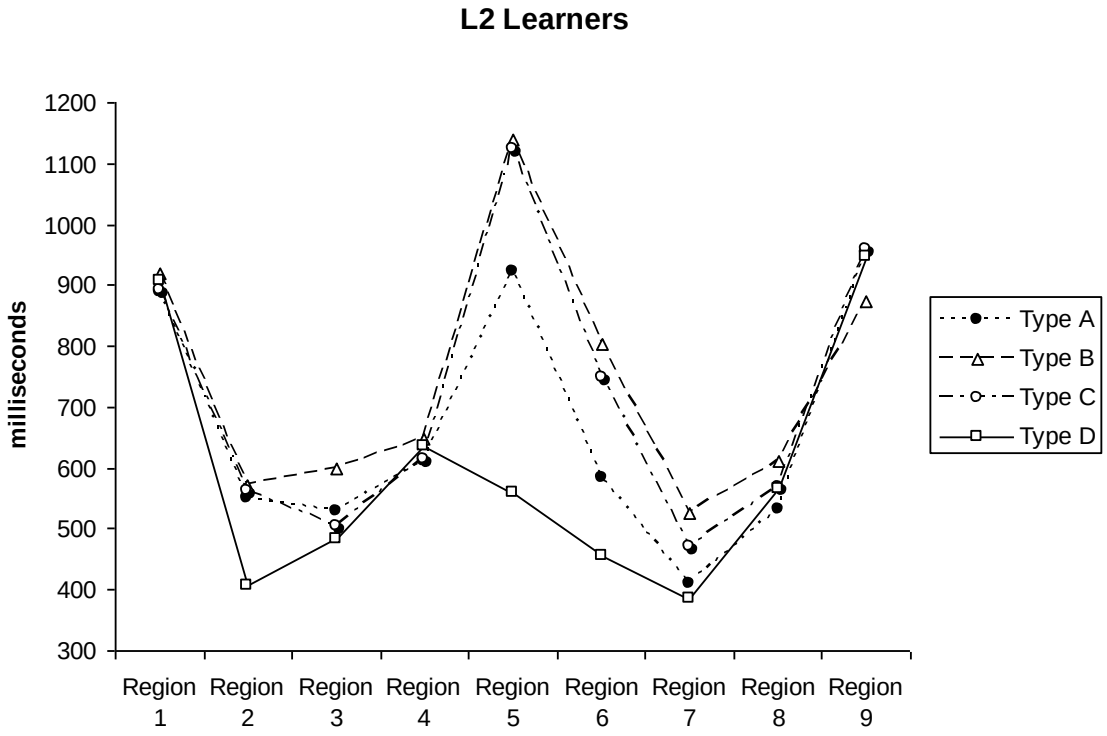


Figure 2. L2 learners' mean reading times for the regions in the 4 sentence types

The learner group's RTs for Region 4 of the four sentence types show no significant difference, $F(3, 779) = 0.782$, $p = 0.504$, and this is, to a large extent, expected as verbs in this region like *ai chi* "love to eat" have the same functions in all the test sentences and therefore, similar RTs of Region 4 are expected across all sentence types. However, the learner group's RTs of Region 5 are found significantly different between each of the

686 sentence types, $F(3, 639)=70.942$, $p<0.001$. Post-hoc Scheffé tests reveal that the learner
687 group's RTs of Region 5 are significantly different between each sentence type, except
688 for between Types B and C. The learner group's RTs of Region 5 in Types B and C are
689 significantly longer than the same region in Type A, which in turn is significantly longer
690 than that in Type D. It seems that our analysis of native Chinese speakers' data of Region
691 5 can also be used to account for the variation in the learner group's processing of the
692 same region in different types of sentences. That is, processing Region 5 in Type D
693 requires no restructuring or re-analysis, but restructuring and re-analysis have to take
694 place when learners process the same region in Types A, B and C, with more complicated
695 processing in Types B and C. Recall that when Regions 2 and 3 in Types A, B and C are
696 processed, the first-person pronoun *wo* "I" in Region 2 will trigger the re-analysis of
697 Region 1 as a topic rather than the subject of the sentence, as initially analyzed, and this
698 topic will be stored in working memory. As a topic in Chinese can be base-generated or
699 derived from topicalization of a constituent from inside the sentence, the parser, while
700 processing the rest of the sentence, is likely to look for a gap from which the topic in
701 working memory is originally derived, a similar strategy as we described for native
702 Chinese speakers above. When the transitive verbal phrase *ai chi* "love to eat" in Region
703 4 is processed, it could be taken by the parser as the subcategorizer of the topic, and if
704 this occurs, the parser would expect a gap in Region 5. The parser is forced to revise its
705 earlier analysis when the NP in Region 5, e.g. *xiangjiao* "banana", is processed, re-
706 analyzing the topic in working memory as a base-generated topic rather than a topic
707 derived from inside the sentence. The restructuring and re-analysis obviously lengthen
708 the RTs of Region 5 in Types A, B and C. What makes the learner group's RTs of Region

5 in Types B and C even longer is believed to be due to the learner group's native-like sensitivity to the violation of the hyponymy relationship required in the Chinese BGT sentence. In Type B sentences, the topic is the hyponym of the NP in Region 5, and in Type C sentences, the topic has a sisterhood relationship with the NP in Region 5. Both of these two sentences violate the constraint that the base-generated topic should be the superordinate of the relevant NP in the Chinese BGT sentence. The learner group's longer RTs of Region 5 in Types B and C reflect their detection of the violation of the semantic requirement in these two types of sentences, in addition to the re-analysis of the topic in the working memory as a base-generated topic rather than a topic derived from inside the sentence. The re-analysis and the checking of the semantic requirement obviously do not apply to Type D sentences, which do not have a topic, thus the learner group's shortest RTs of Region 5 in Type D. Our data also show that the effects of the learner group's re-analysis and their sensitivity to the semantic violations have spilled over to Regions 6 and 7; in both the combination of Regions 5 and 6 ($F(3, 686) = 91.748$, $p<0.001$), and the combination of Regions 5, 6 and 7 ($F(3, 686) = 88.395$, $p<0.001$), the learner groups' RTs in Types B and C are significantly longer than their RTs in Type A, which are in turn significantly longer than their RTs in Type D.

No significant difference is found in the learner group's RTs of Region 8 ($F(3, 788) = 1.969$, $p=0.117$) or Region 9 ($F(3, 604)=0.681$, $p=0.564$) between any of the four sentence types. On average, the RTs of Region 9 are longer than the RTs of many other regions, and this is probably due to what is called the sentence *wrap-up* effect in sentence processing (Just and Carpenter 1980).

8. Discussion

The data in our grammaticality judgment task show that English-speaking learners of L2 Chinese can acquire the explicit knowledge of Chinese BGT sentences as they are able to accept grammatical BGT sentences and reject those violating the semantic constraint of the hyponymy relationship. This is in conformity with the findings reported in the literature (Jin, 1994; AUTHOR, 1995; Jung, 2004; Cao, Yang, Huang, Gao, and Cui, 2006) that speakers of a subject-prominent language like English are able to acquire the knowledge of the BGT structure in their L2 acquisition of a topic-prominent language. There is plenty of positive evidence of BGT sentences in their L2 input, which is likely to enable L2 learners to be aware of the existence of BGT sentences in the target language.

We have seen in Tables 4 and 5 as well as Figures 1-2 that native speakers' RTs of test sentences are, in general, faster than those of L2 Chinese learners. This is expected as native speakers are fast and efficient in sentence processing while L2 learners lack automaticity in L2 processing (cf. Segalowitz, 2003; Dekydtspotter and Miller, 2013). In this sense, native speakers and L2 learners are expected to be different when automaticity in sentence processing is considered. What is more, the orthographic difference between English, which uses a romanization spelling system, and Chinese, which adopts a character script system, is likely to make English speakers' processing of Chinese sentences even slower and less automatic than native Chinese speakers'.

Although L2 learners are found to be generally slower than native Chinese speakers in processing test sentences, they pattern with native Chinese speakers in processing all critical regions and post-critical regions of the test sentences. Recall that unlike English, which is a subject prominent language, Chinese is a topic-prominent language and allows

a topic to appear at the sentence initial position preceding a subject NP. At the same time, it is common for Chinese sentences to have just one preverbal NP, as exemplified by *niurou* "beef" in (8), which can be analyzed as the subject of the sentence[13].

(8) *Niurou zai zheli feichang gui.*
beef at here very expensive
Beef is very expensive here.

How does the parser deal with input that is compatible with more than one grammatical analysis? When the first NP of a BGT sentence is processed, there is no information available that the parser can refer to in processing it as the topic or the subject of the sentence. If the processing default of the parser is to analyze the first NP as the topic of the sentence, no reanalysis of the BGT sentence is necessary when the second NP is processed as the subject of the sentence. However, if the default is for the parser to process the first NP as the subject of the sentence, syntactic reanalysis will have to take place when the second NP in the BGT sentence is processed. That is, the first NP, which was originally analyzed as the subject of the sentence, has to vacate the subject position,

[13] It is also possible to assume that the single NP in the preverbal position in Chinese sentences like (8) is the topic of the sentence with an empty subject, as in (i), or the subject of the sentence with an empty topic, as show in (ii). Interpretations of this type of sentences depend on appropriate contexts, and we will not go into these details, as what was included in our experiment was individual sentences without contexts.

(i) Niurou (zhe zhong dongxi) zai zheli feichang gui.
beef this type thing at here very expensive
"Beef (*this kind of thing) is very expensive here."

(ii) (chide dongxi) niurou zai zheli feichang gui.
edible thing beef at here very expensive.
"(*Food) beef is very expensive here.

772 be processed as the topic and be assigned to the topic position. The subject position thus
773 vacated is then filled with the second NP. The parser will obviously have to make efforts
774 to do the syntactic reanalyses and restructuring, and as a result, it takes longer RTs for the
775 parser to process the second NP, and possibly the following region as well because of the
776 spill-over effect. This is what the native Chinese speakers' and L2 Chinese learners' data
777 of Regions 2 and 3 in Types A, B and C sentences have suggested, which form a striking
778 contrast with their data of the same regions of Type D sentences, where the NP following
779 the adverbial of time is analyzed by the parser as the subject of the sentence and no
780 reanalysis or restructuring is required afterwards. Consequently, it takes shorter RTs for
781 the parser to process Regions 2 and 3 of Type D sentences than the same regions of
782 Types A, B and C sentences. The data also implicate that for both native Chinese
783 speakers and English-speaking learners of L2 Chinese, the processing default of Chinese
784 BGT sentences is to initially analyze the first NP as the subject of the sentence, rather
785 than as a topic. While it is not clear whether L2 Chinese learners' initial decision of
786 processing the first NP in Types A, B and C sentences as the subject of the sentence is
787 due to the transfer of the subject-prominence of their L1 English into their L2 sentence
788 processing, it seems possible to account for the default in both native Chinese speakers'
789 and L2 Chinese learners' initial analysis of the first NP as the subject on the basis of the
790 "least effort" principle by Frazier (1978, 1987). That is, the parser prefers the structurally
791 simplest analysis. Obviously, the S-V-O structure is simpler than the Topic-S-V-O
792 structure, which has a Topic in addition to the S-V-O structure.

793 Recall that Clahsen and Felser (2006a, b) argue in their SSH that adult L2 learners are
794 guided by lexical-semantic cues in their sentence processing in the same way as native

speakers, but L2 learners' sensitivity to syntactic information is restricted and therefore their syntactic representations in sentence processing are shallower than those of native speakers. According to the SSH, unlike native speakers, L2 learners do not rely on structure-based processing strategies when solving ambiguities in L2 sentence processing, and instead, they process L2 sentences primarily on the basis of lexical-semantic and pragmatic information. However, the SSH is not supported by our data here. As we discussed above, L2 Chinese learners, like native Chinese speakers, are sensitive to syntactic cues in solving ambiguities in processing Chinese BGT sentences. Obviously, there are no semantic or pragmatic cues in the reanalyses of the first two NPs of the BGT sentences, and the disambiguation has to be solved by structure-based strategies. Our data clearly demonstrate that L2 Chinese learners, like native Chinese speakers, are sensitive to syntactic information in dealing with ambiguities in processing BGT sentences.

The majority of studies in L2 sentence processing literature use filler-gap dependencies in either English relative clauses (e.g. Juffs, 1998; Papadopoulou and Clahsen, 2003; Marinis, Roberts, Felser and Clashen, 2005; Felser and Roberts, 2007; Dinçtopal-Deniz, 2010; Omaki and Schulz, 2011) or English wh-questions (e.g. William, Möbius and Kim, 2001; William, 2006; Dussias and Piñar, 2010; Aldwayan, Fiorentino and Gabriele, 2010) in their investigations of L2 sentence processing, where the fronted wh-word is identified as a potential filler because of the morphological marking of -wh on the wh-word, and there is a gap which can potentially trigger trace-based antecedent reactivation in processing because of the subcategorization requirement. However, the topic in Chinese BGT sentences is syntactically identified, and there is no gap in the sentence. An interesting question is whether the syntactically identified topic will be

processed by the parser as a potential filler or a structurally displaced constituent, in spite of the fact that there is no gap in the BGT sentence and that all subcategorization requirements in the sentence are met. That is, Types A, B and C sentences in Table 2 would be complete and grammatical sentences, even with the topic deleted. Our data suggest that the syntactically identified topic in Chinese BGT sentences is indeed stored as a potential filler or a structurally displaced constituent in working memory in both native Chinese speakers' and L2 Chinese learners' processing, and that the transitive verbal phrase *ai chi* "love to eat" in Region 4 is initially processed by both native Chinese speakers and L2 Chinese learners as the subcategorizer of the topic. More specifically, the parser seems to postulate a gap in working memory and immediately analyze the topic as the object of the verbal phrase as soon as the verbal phrase is processed, i.e. before the object of the verbal phrase in Region 5 is processed. This can be accounted for by a processing principle that requires the parser to complete grammatical dependencies as soon as possible (de Vincenzi, 1991; Frazier, 1987; Pritchett, 1992) or on the basis of the need to reduce the cost of retaining the filler in memory (Gibson, 1998). When the object (i.e. Region 5) of the verbal phrase *ai chi* "love to eat" in Types A, B and C is processed, the parser is forced to revise its earlier analysis and re-analyze the topic in working memory as a base-generated topic rather than a topic derived from inside the sentence. The revision and re-analysis obviously require extra efforts, which explains the longer RTs of native Chinese speakers' and L2 Chinese learners' processing of Region 5 (and also Regions 6 and 7 because of the spill-over effect) in Types A, B and C sentences than their RTs of the same regions in Type D sentences, where no topic is stored in working memory. Recall that while there are BGT sentences in Chinese, it is also

common in Chinese to have topic structures in which the topic is a result of movement, as exemplified in the sentence in (2), repeated in (9), where the topic *Zhe ben shu* "this book" is originally based-generated as the object of the verb *xihuan* "like" in the sentence before it is topicalized to the Specifier of TopP at the initial position of the sentence.

(9) *Zhe ben shu$_i$ wo bu xihuan t_i.*
this CL book I not like
This book, I don't like.

Positive evidence like the sentence in (9) is likely to set a default in both native Chinese speakers' and L2 Chinese learners' processing, treating the syntactically identified topic as a topic derived from inside the sentence and store it as such in working memory until contradicting information is processed, as in the case of Region 5 in Types A, B and C sentences. Note that unlike the filler-gap dependencies in processing English relative clauses or English wh-questions, which can be morphologically and semantically triggered in a bottom-up fashion, processing the Chinese topic by both native Chinese speakers and L2 Chinese learners as a potential filler is syntactically induced in a top-down manner, and no semantic, morphological or pragmatic cues are available that the parser could rely on in processing the topic as a potential filler and store it as such in working memory. This top-down structure-based processing strategy provides further counter-evidence against the SSH because our data demonstrate that both native Chinese speakers and L2 Chinese learners are similarly sensitive to syntactic information in processing Chinese BGT sentences.

There is strong evidence in the literature that L2 proficiency is an important factor for L2 syntactic processing (Frenck-Mestre, 2002; Hahne, 2001; Hopp, 2006; Jackson, 2008). Given that L2 learners in our study were all very proficient L2 speakers of Chinese, it is highly likely that L2 learners' native-like sentence processing is positively co-related with their proficiency of the target language, as proposed by Mendés, Farmer and Slabakova (2014). In addition, L2 learners in our study had an average stay of 6.3 years in China/Taiwan, as shown in Table 1, and in accordance with the suggestions by Frenck-Mestre (2002), more than five years of exposure to the target language can lead to the use of native-like processing strategies by L2 learners. This helps to account for English speakers' native-like structure-based processing of Chinese BGT sentences. The extended periods of immersion in naturalist Chinese environments are expected to play a facilitating and crucial role in L2 learners' ability to process Chinese BGT sentences in a native-like manner. Our findings here are also in conformity with the finding in Pliatsikas and Marinis (2013), where Greek-speaking L2 learners of English with an average of 9 years of immersion in English environments are found to be able to have native-like processing of intermediate traces in long distance wh-dependencies in English. This suggests that differences in populations do indeed play a role in assessing L2 processing, as pointed out by VanPatten and Jegerski (2010). BGT sentences are not allowed in English, but English-speaking L2 Chinese learners' structure-based processing of Chinese BGT sentences provides us supporting evidence that native-like L2 processing is achievable, even for L2 features that do not have equivalents in the L1 (Foucart and Frenck-Mestre, 2012).

Recall that the SSH predicts that L2 learners, like native speakers, are able to make use of semantic and pragmatic information available in sentence processing, and this part of the SSH is confirmed by data in our study. Chinese BGT sentences are subject to semantic constraints, one of which is that the base-generated topic and its related NP in the sentence are to have a hyponymy relationship. However, the topic in Region 1 and the NP in Region 5 in Types B and C sentences violate such a requirement, with the topic being a hyponym of the NP in Region 5 in Type B, and the topic in Type C having a sisterhood relationship with the NP in Region 5. The longer RTs of Region 5 in native Chinese speakers' and L2 Chinese learners' processing of Types B and C sentences are believed to reflect their sensitivity to the violation of the required semantic relationship involved in Chinese BGT sentences, as detecting such a semantic relationship violation will, obviously, further prolong the RTs of Region 5 in Types B and C, which are found to be significantly longer than the RT of the same region in Type A, where the topic in Region 1 is a superordinate of the NP in Region 5, meeting the requirement of the semantic relationship for Chinese BGT sentences. This implicates that the semantic information of the hyponymy relationship is stored together with the syntactically identified topic in working memory and can be made use of in both native Chinese speakers' and L2 Chinese learners' sentence processing to check against the corresponding semantic information on the relevant NP in the sentence.

9. Conclusion

No evidence is found in our study which shows that L2 structures are shallow in sentence processing, disconfirming the prediction by the SSH. Data in our study demonstrate that

like native speakers, L2 learners are sensitive to and are able to make use of syntactic cues as well as semantic information in their L2 sentence processing. As participants in this study were highly proficient L2 learners of Chinese with long periods of immersion in Chinese-speaking environments, we cannot rule out the possibility that L2 learners at earlier stages or with only classroom exposure are not able to rely on syntactic cues in their L2 sentence processing. However, our data do suggest that native-like structure-based processing of L2 Chinese BGT sentences is possible, at least for highly proficient L2 Chinese learners with extended periods of immersion in naturalistic Chinese environments.

References

Aldwayan, S., Fiorentino, R. and Gabriele, A. (2010). Evidence of syntactic constraints in the processing of wh-movement: A study of Najdi Arabic learners of English. In B. VanPatten and J. Jegerski (Eds.), *Research in Second Language Processing and Parsing* (pp. 65-86). Amsterdam: John Benjamins.

Cao, X.-L., Yang, S.-Y., Huang, Y.-Y., Gao, L.-Q. and Cui, X.-L. (2006). The L2 acquisition of Chinese topic constructions. *World Chinese Language Teaching,* 77, 86-102.

Chafe, W. (1976). Givenness, Contrastiveness, Definiteness, Subject, Topics and Point of View. In C. Li (Ed.), *Subject and Topic* (pp. 27-55). New York: Academic Press.

Cole, P. (1987). Null object in Universal Grammar. *Linguistic Inquiry,* 18, 597-612.

Clahsen, H. and Felser, C. (2006a). Grammatical processing in language learners. *Applied Psycholinguistics,* 27, 3-42.

Clahsen, H, and Felser, C. (2006b). Continuity and shallow structures in language processing. *Applied Psycholinguistics,* 27, 107-126.

Dekydtspotter, L. and Miller, A. K. (2013). Inhibitive and facilitative priming induced by traces in the processing of wh-dependencies in second language. *Second Language Research,* 29, 345-372.

de Vincenzi, M. (1991). *Syntactic parsing strategies in Italian*. Dordrecht: Kluwer.

Dinçtopal-Deniz, N. (2010) Relative clause attachment preferences of Turkish L2 speakers of English: Shallow parsing in the L2? In B. VanPatten and J. Jegerski (Eds.), *Research in Second Language Processing and Parsing* (pp. 27-63). Amsterdam: John Benjamins.

Dussias, P. E. and Piñar, P. (2010). Effects of reading span and plausibility in the reanalysis of wh-gaps by Chinese-English second language speakers. *Second Language Research,* 26, 443-472.

Felser, C., Roberts, L., Marinis, T. and Gross, R. (2003). The processing of ambiguous sentences by first and second language learners of English. *Applied Psycholinguistics,* 24, 453-489.

Felser, C and Roberts, L. (2007). Processing wh-dependencies in a second language: a cross-modal priming study. *Second Language Research,* 23, 9-23.

Forster, K. I. and Forster, J. C. (2003). DMDX: A windows display program with millisecond accuracy. *Behavior Research Methods, Instruments and Computers,* 35, 116-124.

Frazier, L. (1978). *On Comprehending Sentences: Syntactic parsing strategies.* Unpublished PhD dissertation. University of Connecticut, Storrs.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory, 5*, 519-559.

Frazier, L. and Clifton, C. (1989). Identifying gaps in English sentences. *Language and Cognitive Processes,* 4, 93-126.

Frenck-Mestre, C. (2002). An on-line look at sentence processing in the second language. In J. Altarriba and R. Herridia (Eds.), *Bilingual Sentence Processing* (pp. 218–236). Amsterdam: Elsevier.

Frenck-Mestre, C and Pynte, J. (1997). Syntactic ambiguity resolution while reading in second and native language. *Quarterly Journal of Experimental Psychology,* 50A, 119-148.

Foucart, A. and Frenck-Mestre, C. (2012). Can late L2 learners acquire new grammatical features? Evidence from ERPs and eye-tracking. *Journal of Memory and Language,* 66, 226–248.

Gibson, E. (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition,* 68, 1-76.

Gundel, J. K. (1988). Universals of topic-comment structure. In M. Hammond, E. Moravacsik, and J. Wirth (Eds.), *Studies in Syntactic Typology* (pp. 209-239). Amsterdam: John Benjamins.

Hahne, A. (2001). What's different in second-language processing? Evidence from event-related brain potentials. *Journal of Psycholinguistic Research*, 30, 251–266.

Hopp, H. (2006), Syntactic features and reanalysis in near-native processing. *Second Language Research*, 22, 369–397.

977 Huang, C.-T. J. (1984a). On the distribution and reference of empty pronouns. *Linguistic*
978 *Inquiry, 15*, 531-574.
979 Huang, C.-T. J. (1984b). On the typology of zero anaphora. *Language Research, 20*, 85-
980 105.
981 Huang, C.-T. J., Li, Y.-H. A. and Li, X.-G. (2009). *The Syntax of Chinese*. Cambridge:
982 Cambridge University Press.
983 Jackson, C. N. (2008). Proficiency level and the interaction of lexical and
984 morphosyntactic information during L2 sentence processing. *Language Learning,* 58,
985 875–909.
986 Jiang, N. (2013). Conducting Reaction Time Research in Second Language Studies.
987 London: Routledge.
988 Jin, H.-G. (1994). Topic-prominence and subject-prominence in L2 acquisition: Evidence
989 of English-to-Chinese typological transfer. *Language Learning,* 44, 101-122.
990 Jung, E.-H. (2004). Topic and subject prominence in interlanguage development.
991 *Language Learning,* 54, 713–738.
992 Juffs, A. (1998). Main verb versus reduced relative clause ambiguity resolution in L2
993 sentence processing. *Language Learning*, 48, 107-147.
994 Just, M. A. and Carpenter, P.A. (1980). A theory of reading: From eye fixation to
995 comprehension. *Psychological Review*, 87, 329-354.
996 Just, M. A., Carpenter, P. A., and Woolley, J. D. (1982). Paradigms and processes in
997 reading comprehension. *Journal of Experimental Psychology: General,* 111, 228-238.
998 Li, C. and Thompson, S. (1976). Subject and Topic: A new typology of languages. In C.
999 Li (Eds.), *Subject and Topic* (pp. 457-489). New York: Academic Press.

1000 Marinis, T., Roberts, L. and Felser, C. (2005). Gaps in Second language sentence
1001 processing. *Studies in Second Language Acquisition,* 27, 53-78.
1002 Mendés, T. L., Farmer, T. A. and Slabakova, R. (2014). A same-system view of L2
1003 processing: Evidence from long-distance syntactic dependencies in L2 Spanish. In W.
1004 Orman, and M. J. Valleau (Eds.), *Proceedings of the 38th Annual Boston University*
1005 *Conference on Language Development* (pp. 266-278). Somerville, MA: Cascadilla
1006 Press.
1007 Omaki, A., and Schulz, B. (2011). Filler-gap dependencies and island constraints in
1008 second language sentence processing. *Studies in Second Language Acquisition*, 33,
1009 563-588.
1010 Pan, H., and Hu, J. (2001). Deriving topic-comment constructions in Chinese. Paper
1011 presented at the 1st International Conference on Formal Linguistics, Changsha, Hunan,
1012 China.
1013 Pan, H., and Hu, J. (2002). Liscensing dangling topics in Chinese. Paper presented at the
1014 76th Annual Meeting of the Linguistic Society of America, San Francisco.
1015 Papadopoulou, D., and Clashen, H. (2003). Parsing strategies in L1 and L2 sentence
1016 processing: A study of relative attachment in Greek. *Studies in Second Language*
1017 *Acquisition*, 25, 501-528.
1018 Pearlmutter, N. J., Garnsey, S. M., and Bock, K. (1999). Agreement processes in sentence
1019 comprehension. *Cognition*, 64, 9-17.
1020 Pickering, M. J., and Barry, G. (1991). Sentence processing without empty categories.
1021 *Language and Cognitive Processes*, 6, 229-259.

1022 Pliatsikas, C., and Marinis, T. (2013). Processing empty categories in a second language:
1023 When naturalistic exposure fills the (intermediate) gap. *Bilingualism: Language and*
1024 *Cognition*, 16, 167-182.
1025 Pritchett, B. L. (1992). Parsing with grammar: Islands, heads, and garden paths. In H.
1026 Goodluck, and M. Rochemont (Eds.), *Island Constraints: Theory, acquisition, and*
1027 *processing* (pp. 321–349). Dordrecht: Kluwer Academic Publishers.
1028 Rizzy, L. (1997). The fine structure of the left periphery. In L. Haegeman, (Ed.),
1029 *Elements of Grammar* (pp.281-337). Dordrecht: Kluwer.
1030 Segalowitz, N. (2003). Automaticity and second languages. In C. Doughty and M. Long,
1031 (Eds.), *The Handbook of Second Language Acquisition* (pp. 382-388). Oxford:
1032 Blackwell.
1033 Sharkey, N. E., and Sharkey, A. J. C. (1987). What is the point of integration? The loci of
1034 knowledge-based facilitation in sentence processing. *Journal of Memory and*
1035 *Language*, 26, 255-276.
1036 Shi, D. (2000). Topic and topic-comment constructions in Mandarin Chinese. *Language*,
1037 76, 383-408.
1038 Takami, K-I., and Kamio, A. (1994). Topicalization and subjectivization in Japanese:
1039 Characterizational and indentificational information. *Lingua*, 99, 207-235.
1040 VanPatten, B. and Jegerski, J. (2010). Second language processing and parsing: The
1041 issues. In B. VanPatten and J. Jegerski (Eds.), *Research in Second Language*
1042 *Processing and Parsing* (pp. 3-23). Amsterdam: John Benjamins.
1043 Warren, T., and Gibson, E. (2002). The referential processing on sentence complexity.
1044 *Cognition*, 85, 79-112.

Williams, J. N. (2006). Incremental interpretation in second language sentence processing. *Bilingualism: Language and Cognition,* 9, 71-88.

Williams, J.N., Möbius, P. and Kim, C. (2001). Native and non-native processing of English wh-questions: Parsing strategies and plausibility constraint. *Applied Psycholinguistics,* 22, 509-540.

Xu, L. (1986). Free empty category. *Linguistic Inquiry,* 17, 75-93.

Xu, L. (2006). Topicalization in Asian Languages. In M. Everaert, M. and H. van Riemsdijk (Eds.), *The Blackwell Companion to Syntax*, Volume V, (pp. 137-174). Oxford: Blackwell Publishing.

Xu, L., and Langendoen, D. T. (1985). Topic structures in Chinese. *Language,* 61, 1-27.

Author(1995).XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX

APPENDIX: Experimental stimuli (Only Chinese characters were used in the experiment. Pinyin, English gloss and English translation are provided here for readers of this article. Comprehension questions are in parentheses; √=true, ×=false)

Type A. BGT sentences with a superordinate NP as the base-generated topic

水果我最爱吃香蕉，所以我经常买香蕉。（“我”喜欢水果。√）

shuiguo wo zui ai chi xiangjiao, suoyi wo jingchang mai xiangjiao ("wo" xihuan shuiguo. √)

fruit I most like eat bananas, so I often buy bananas ("I" like fruit √)

"As for fruits, I like to eat bananas most. Therefore I often buy bananas. ("I" like fruits. √)"

动物我最喜欢小狗，所以我有两只小狗。（“我”家里没有动物。×）

dongwu wo zui xihuan xiao gou, suoyi wo you liangzhi xiao gou. ("wo" jiali mei you dongwu. ×)

animal I most like little dog, so I have two little dog. ("my" home not have animal. ×)

"As for animals, I like little dogs most. Therefore I have two little dogs. ("I" have no animal at home. ×) "

外语他只会说法语，所以他常常去法国。（他常常去国外。√）

waiyu ta zhi hui shuo Fayu, suoyi ta changchang qu Faguo. (ta changchang qu guowai√)

foreign language he only can speak French, so he often go to France (he often go abroad√)

"As for foreign languages, he can only speak French. Therefore he often goes to France. (He often goes abroad. √)"

中国她只去过上海，可是她没去过别的城市。（她去过北京。×）

Zhongguo ta zhi qu guo Shanghai, but ta mei qu guo biede chengshi. (ta qu guo Beijing. ×)

China she only go EXP Shanghai, but she not go EXP other city. (She go EXP Beijing. ×)

As for China, she has only been to Shanghai, but she has not been to other cities. (She has been to Beijing. ×)

海鲜我最爱吃大虾，所以我常常买大虾。（"我"经常吃海鲜。√）

haixian wo zui ai chi daxia, suoyi wo changchang mai daxia. ("wo" jingchang chi haixian. √)

seafood I most like eat prawn, so I often buy prawn. ("I" often eat seafood. √)

"As for seafood, I like to eat prawns most. Therefore I often buy prawns. ("I" often eat seafood. √)"

体育他最喜欢足球，所以他经常踢足球。（他讨厌体育。×）

tiyu ta zui xihuan zuqiu, suoyi ta jingchang ti zuqiu. (ta taoyan tiyu. ×)

sport he most like football, so he often play football. (he hate sport. ×)

"As for sports, he likes football most. Therefore he often plays football. (He hates sports. ×)"

Type B. *BGT sentences with a hyponym NP as the base-generated topic

*香蕉我最爱吃水果，所以我经常买水果。（“我”讨厌水果。×）

xiangjiao wo zui ai chi shuiguo, suoyi wo jingchang mai shuiguo. (“wo” taoyan shuiguo. ×)

banana I most like eat fruit, so I often buy fruit. (“I” hate fruit. ×)

*“As for bananas, I like to eat fruits most. Therefore I often buy fruits. (“I” hate fruits. ×)”

*小狗我最喜欢动物，所以我有两只小动物。（“我”害怕动物。×）

xiao gou wo zui xihuan dongwu, suoyi wo you liangzhi xiao dongwu. (“wo” haipa dongwu.)

little dog I most like animal, so I have two small animal. (“I” am afraid of animal.)

*“As for little dogs, I like animals most. Therefore, I have two small animals. (“I” am afraid of animals.)

*法语他只会说外语，所以他外语很好。（他学过外语。√）

Fayu ta zhi hui shuo waiyu, suoyi ta waiyu hen hao. (ta xue guo waiyu. √)

French he only can speak foreign language, so he foreign language very good. (he study EXP foreign language. √)

*“As for French, he can only speak foreign languages. Therefore his foreign languages are very good. (He has studied foreign language before.√)

*上海她只去过中国，可是她没去过别的国家。（她去过法国。×）

Shanghai ta zhi qu guo Zhongguo, keshi ta mei qu guo biede guojia. (ta qu guo Faguo. ×)

Shanghai she only go EXP China, but she not go EXP other country. (she go EXP France. ×)

*“As for Shanghai, she has only been to China, but she has not been to any other country. (She has been to France. ×)”

*大虾我最爱吃海鲜，所以我常常买海鲜。（“我”吃过海鲜。√）

daxia wo zui ai chi haixian, suoyi wo changchang mai haixian. (“wo” chi guo haixian. √)

prawns I most like eat seafood, so I often buy seafood. (“I” eat EXP seafood. √)

*“As for prawns, I like to eat seafood most. Therefore I often buy seafood. (“I” have eaten seafood before. √)”

*足球他最喜欢体育，所以他经常做体育活动。（他喜欢运动。√）

zuqiu ta zui xihuan tiyu, suoyi ta jingchang zuo tiyu huodong. (ta xihuan yundong.)

football he most like sports, so he often do sport exercises. (he likes exercise.)

*"As for football, he likes sports most. Therefore he often does sport exercises. (He likes exercises.)"

Type C. *BGT sentences with a sisterhood relationship

*苹果我最爱吃香蕉，所以我经常买香蕉。（"我"不常常买水果。×）

pingguo wo zui ai chi xiangjiao, suoyi wo jingchang mai xiangjiao.("wo" bu changchang mai shuiguo.×)

apple I most like eat banana, so I often buy banana. ("I" not often buy fruit. ×)

*"As for applies, I like to eat bananas most. Therefore I often buy bananas. ("I" do not often buy fruits. ×)"

*小猫我最喜欢小狗，所以我有两只小狗。（"我"家里有动物。√）

xiao mao wo zui xihuan xiao gou, suoyi wo you liangzhi xiao gou. ("wo" jiali you dongwu. √)

little cat I most like little dog, so I have two little dog. ("my" home have animal. √)

*"As for little cats, I like little dogs most. Therefore I have two little dogs. ("I" have animals at home. √)"

*日语他只会说法语，所以他常常去法国。（他很少去外国。×）

Riyu ta zhi hui shuo Fayu, suoyi ta changchang qu Faguo. (ta hen shao qu waiguo. ×)

Japanese he only can speak French, so he often go France. (he rarely go abroad. ×)

*"As for Japanese, he can only speak French. Therefore he often goes to France. (He rarely goes abroad.×)"

*北京她只去过上海，可是她没去过别的城市。（她去过中国。√）

Beijing ta zhi qu guo Shanghai, keshi ta mei qu guo biede chengshi. (ta qu guo Zhongguo. √)

Beijing she only go EXP Shanghai, but she not go EXP other city (she go EXP China. √)

*"As for Beijing, she has only been to Shanghai, but she has not been to any other cities. (She has been to China. √)"

*海鱼我最爱吃大虾，所以我常常买大虾。（“我”喜欢海鲜。√）

hai yu wo zui hai chi daxia, suoyi wo changchang mai daxia. (“wo” xihuan haixian. √)

sea fish I most like eat prawn, so I often buy prawn. (“I” like seafood. √)

*“As for sea fish, I like to eat prawns most. Therefore I often buy prawns. (“I” like seafood. √)”

*篮球他最喜欢足球，所以他经常踢足球。（他讨厌体育。×）

lanqiu ta zui xihuan zuqiu, suoyi ta jingchang ti zuqiu. (ta taoyan tiyu. ×)

basketball he most like football, so he often play football. (he hates sports. ×)

*“As for basketball, he likes football most. Therefore he often plays football. (He hates sports. ×)”

Type C. Non-BGT sentences

以前我最爱吃香蕉，所以我经常买香蕉。（过去“我”很少吃水果。×）

yiqian wo zui ai chi xiangqiao, suoyi wo jingchang mai xiangjiao. (guoqu “wo” hen shao chi shuiguo. ×)

before I most like eat banana, so I often buy banana. (in the past “I” rarely eat fruit. ×)

“Before I liked to eat bananas most. Therefore I often bought bananas. (In the past “I” rarely ate fruits. ×)”

过去我最喜欢小狗，所以我有两只小狗。（“我”养过动物。√）

guoqu wo zui xihuan xiao gou, suoyi wo you liangzhi xiao gou. (“wo” yang guo dongwu. √)

past I most like little dog, so I have two little dog. (“I” raise EXP animal. √)

“In the past I like little dogs most. Therefore I had two dogs. (“I” once had animals. √)”

那时他只会说法语，所以他常常去法国。（那时候他不会说别的外语。√）

na shi ta zhi hui shuo Fayu, suoyi ta changchang qu Faguo. (na shihou ta bu hui shuo biede waiyu.)

that time he only can speak French, so he often go France. (that time he not can speak other foreign language.)

"At that time he could only speak French. Therefore he often went to France. (At that time he could not speak other foreign languages.)"

那时候她只去过上海，可是她没去过别的城市。（那时候她没去过北京。√）

na shihou ta zhi qu guo Shanghai, keshi ta mei qu guo biede chengshi. (na shihou ta mei qu guo Beijing.√)

that time she only go EXP Shanghai, but she not go EXP other city. (that time she not go EXP Beijing.√)

"At that time, she had only been to Shanghai, but she had not been to other cities. (At that time, she had not been to Beijing. √)"

从前我最爱吃大虾，所以我常常买大虾。（"我"没买过海鲜。×）

congqian wo zui ai chi daxia, suoyi wo changchang mai daxia. ("wo" mei mai guo haixian. ×)

before I most like eat prawn, so I often buy prawn. ("I" not buy EXP seafood. ×)

"Before I liked to eat prawns most. Therefore I often bought prawns. ("I" have never bought seafood. ×)"

小时候他最喜欢足球，所以他经常踢足球。（他没踢过足球。×）

xiao shihou ta zui xihuan zuqiu, suoyi ta jingchang ti zuqiu. (ta mei ti guo zuqiu. ×)

small time he most like football, so he often play football. (he not play EXP football. ×)

"When he was small he liked football most. Therefore he often played football. (He has never played football. ×)"