Constructing Filler-Gap Dependencies in Chinese Possessor Relative Clauses

Chien-Jer Charles Lin
Department of Linguistics,
The University of Arizona
Tucson, AZ 85721-0028, USA
clin@u.arizona.edu

Sandiway Fong
Department of Linguistics,
The University of Arizona
Tucson, AZ 85721-0028, USA
sandiway@u.arizona.edu

Thomas G. Bever
Department of Linguistics,
The University of Arizona
Tucson, AZ 85721-0028, USA
tgb@u.arizona.edu

Abstract

This article explores the construction of filler-gap dependencies in Chinese possessor relative clauses (PRCs), which are different from typical relative clauses (RCs) considered in the literature because Chinese PRCs contain no overt missing arguments (i.e. gaps). As Chinese RCs are prenominal, the gaps precede the head noun fillers. It has been suggested that when the gaps are close to the filler, the dependency is easier to construct; there is, thus, a processing advantage for object RCs over subject RCs (Hsiao & Gibson, 2003). The PRC data presented show that even in Chinese, a language with RCs that are head-final, it is possible to have a subject gap preference (over object) despite the fact that the subject is further away from the filler. Three experiments confirmed this subject preference with respect to naturalness and grammaticality ratings (Experiment 1), paraphrasing tasks (Experiment 2), and self-paced reading tasks (Experiment 3). The results support a theory of gap-searching which operates top down. Issues regarding locality and canonicity will also be discussed.

1. Introduction

As human languages are replete with dependency relations within and across sentences, one crucial task of the language parser concerns the efficient recovery of such relations and their correct interpretation. This process involves various factors, including the required on-line processing load (often discussed in terms of working memory), the complexity of the processed materials, and human syntactic knowledge, etc.[1] Most previous research focused on the first two factors. When a sentence consumes more processing resources, it is assumed to be more difficult and thus takes longer to understand. Similarly, when a sentence is more complicated or less usual, it requires longer processing time. However, the third factor, i.e. the structural properties of different syntactic positions in strategic on-line processing, has been less studied. Studies of focus positions and their salience for processing are both examples of this vein of inquiry (see for example Birch et al., 2000, and Frazier et al., 2005). In this paper, we present data that supports structural knowledge as a prominent factor in sentence processing.

By structural or syntactic knowledge, we mean knowledge about the function of specific syntactic positions. This knowledge allows the human parser to recover structure in efficient fashions. For example, in a probe-goal model such as Chomsky (2001), this might mean specific and direct access to certain syntactic positions (e.g. a probe to spec-TP), followed by more general top-down search for matching goals. As the experimental data with possessor relative clauses (PRCs) will suggest, surface subject positions are the most likely candidates for direct gap-probing. In what follows, section 2 reviews findings on processing Chinese relative clauses (RCs), and classical effects such as locality and canonicity. Section 3 introduces Chinese PRCs and issues regarding filler-gap dependencies. Section 4 presents three experiments on sentence comprehension. Section 5 discusses the implications of the

[1] This is not meant to underestimate the importance of semantic and pragmatic knowledge in processing dependencies. However, in this paper, we mainly focus on the syntactic aspect of dependent relations.

experimental results and proposes a structure-based theory of gap-searching. Crosslinguistic evidence from Japanese and Turkish is also discussed.

2. Locality, canonicity, and processing Chinese relative clauses

2.1.Locality

Locality-based theories of sentence comprehension predict that when two dependent elements are farther away from each other, the dependency is cognitively more consuming to construct. Frazier and Flores d'Arcais (1989) call this the Active Filler Strategy defined in (1) (see also Crain & Fodor 1985; Frazier & Clifton 1989; Frazier et al. 1983).

(1) Active Filler Strategy: Assign an identified filler as soon as possible; i.e., rank the option of a gap above the option of a lexical noun phrase within the domain of an identified filler. (Frazier & Flores d'Arcais, 1989)

This strategy accounts for Stowe's (1986) observation that (2b) is more likely to lead to garden-pathing than (2a). In on-line word-by-word reading, as soon as a filler is detected, the filler is expected to fill a gap at each potential point (indicated by the underlined spaces).

(2) a. My brother wanted to know *if* Ruth will bring us home to Mom at Christmas.
b. My brother wanted to know *who* Ruth will bring _ us _ home to _ at Christmas.

Gibson (1998) extends this strategy by focusing on the cognitive resources required during on-line sentence processing. His Syntactic Prediction Locality Theory (SPLT) (a.k.a. Dependency Locality Theory) has two components: integration cost and memory cost, both of which are affected by locality:

(3) Syntactic Prediction Locality Theory / Dependency Locality Theory (Gibson, 1998: 1): 1) the longer a predicted category must be kept in memory before being encountered, the greater is the cost for maintaining that prediction; and 2) the greater the distance between an incoming word and the most local head or dependent to which it attaches, the greater the integration cost.

Both theories predict that sentences with locally dependent elements are easier to understand than those with distant relations. These locality-based theories have been able to account for the asymmetry in subject and object relative clauses. For example, an English relative clause where the subject is extracted ((4), subject RC) is easier than its object counterpart ((5), object RC) in self-paced reading tasks (King & Just 1991; Gibson et al. 2005) and eye-movement monitoring reading tasks (Traxler et al. 2002). It is argued that this is due to a shorter filler-gap distance in subject RCs over object RCs.

(4) Subject Relative Clause in English:
The lady who _ laughed at John has a loud voice.
(5) Object Relative Clause in English:
The lady who John laughed at _ has a loud voice.

Consistent with this account, the opposite direction of preference is observed in Chinese. Chinese object RCs are easier than subject RCs because of a shorter filler-gap distance in the former (Hsiao & Gibson 2003). This preference follows from the fact that Chinese relative clauses are prenominal, and thus an object gap appears closer to the head noun.[2] However, the

[2] This theory has been challenged by Japanese relative clauses where subject RCs are preferred to object RCs while the distance between the filler and the gap is longer in subject RCs.

preferences in both Chinese and English are ambiguous for a locality-based theory. A confound for such an account is that the preferred RCs in both languages follow the typical agent-verb-patient order, while the dispreferred do not. This preference can be accounted for by preferred direction of thematic assignment in a language.

(6) Subject Relative Clause in Chinese:
_ chaoxiao zhangsan de **nyuren** sangmen hen da
_ laugh at Zhangsan DE woman voice very big
'The woman who laughed at Zhangsan has a loud voice.'

(7) Object Relative Clause in Chinese:
zhangsan chaoxiao _ de **nyuren** sangmen hen da
Zhangsan laugh at _ DE woman voice very big
'The woman who Zhangsan laughed at has a loud voice.'

Another difference between English and Chinese RCs is the relative position of the filler and the gap. In English, the filler precedes that gap. It is reasonable to assume that after the relativizer *who* is encountered, the parser starts to look for a gap. With Chinese RCs, however, the gap precedes the filler. Once a gap is identified by the parser, a relative clause construction may be assumed. The parser then expects to fill the gap with the head noun after the relativizer *de*, rather than trying to fill it with every upcoming NP in the clause. The filler-gap dependencies in Chinese and English are therefore quite different. In order to study whether Chinese uses a gap-filling strategy similar to English, we study PRCs where no overt gaps exist, and the gap filling process does not start until the relativizer and the head noun are reached.

2.2.Canonicity

The effect of canonicity predicts that sentences that undergo more complex syntactic derivations are more difficult than sentences that resemble base-generated word orders. The assumption is that in order to understand a sentence, the parser has to recover the base-generated structure; a sentence that has all elements in situ is easier than a sentence with elements displaced, in which the dependency between the moved word and its trace must be recovered for interpretation. This complexity is also due to more complex thematic assignment in sentences with movement. If we assume that the direction of thematic assignment is determined by underlying predicate-argument order, then sentences with displacement are more difficult because theta assignment has to be redirected to the surface positions.

In English, this theory has been supported by evidence that passives are more difficult to comprehend than actives. Ferreira (2003: 164) asked participants to identify thematic roles in aurally presented sentences and found that people adopted "simple processing heuristics" that rely on canonical word order for shallow processing. This produces "good enough" interpretations which may be inaccurate. Passives with the thematic roles appearing at non-typical positions induce more errors in identifying thematic roles than their active counterparts.

In languages with scrambling, it has also been reported that sentences with scrambling tend to be harder than sentences with all words in situ. This has been confirmed in Japanese, where scrambled sentences take longer both to read (Mazuka et al., 2002) and to be judged as correct sentences (Tamaoka et al. 2003, 2005). Tamaoka et al. (2005), in particular, specified that this canonicity effect is based on grammatical functions (such as subject, indirect object and direct object) rather than the order of thematic roles or case particles.

Both locality and canonicity are well-attested. However, we will demonstrate that in Chinese PRCs, sentences with distant filler-gap dependencies and more complex syntactic structures (e.g. passives) are preferred. These results suggest a gap-searching mechanism that recognizes structural positions and performs top-down search starting at the subject position.

3. Chinese possessor relative clauses

Possessor relative clauses (PRCs) are relative clauses in which the head noun serves as the possessor of an NP within the relative clause. In some languages, PRCs are overtly marked by a possessor relativizer. In (8) for example, *whose* marks the antecedent head noun (*the lady*) as the possessor of the NP (*daughter*).

(8) Possessor Relative Clause in English:
The lady **whose** daughter spilled some water has a loud voice.

In Chinese, there is no special relativizer that distinguishes PRCs from regular RCs. Nevertheless, the possessive relationship between the head noun and the appropriate nominal element is established when the relativizor *de* and the head noun are reached. In (9), the head noun *nyushi* 'lady' serves as the possessor of *nyuer* 'daughter', the subject of the RC. Chinese PRCs are different from the RCs typically considered in the literature because all argument slots are filled while a possessor-possessee relationship exists between the head and an argument. Because of this property, the part of a PRC prior to the relativizer may be taken as a thematically-complete sentence without the need for gap-filling. Gaps are not recognized prior to the fillers (like in typical Chinese RCs). Only when the relativizer and the head noun are reached is gap-filling initiated.

(9) Possessor Relative Clause in Chinese:
_ nyuer dafan shui de nawei **nyushi** sangmen hen da
daughter spill water DE that lady voice very loud
'The lady whose daughter spilled the water has a loud voice.'

In this paper, we focus on object PRCs, where the object NP is the possessee of the head noun. Objects in these PRCs can be displaced with different constructions, thus allowing us to manipulate the distances between the possessee (i.e. the gap) and the head noun (i.e. the filler). Sentences (10)-(12) demonstrate increasing filler-gap distances over three constructions. Sentence (10) is a PRC with the canonical order of agent-verb-patient. Sentence (11) is the *ba* variant with the patient preposed to the pre-verbal position. Sentence (12) is the passive *bei* construction, where the patient is displaced all the way to the initial subject position.

(10) Chinese possessor relative clause with canonical order:
huairen bangjia **_ laopo** de **zongcai** jueding baojing
bad guys kidnap wife DE chairman decide call police
'The chairman whose wife some bad guys kidnapped decided to call the police.'

(11) Chinese possessor relative clause with *BA* (agent BA patient V):
huairen ba **_ laopo** bangjia de **zongcai** jueding baojing
bad guys BA wife kidnap DE chairman decide call police
'The chairman whose wife some bad guys kidnapped decided to call the police.'

(12) Chinese possessor relative clause with *BEI* in the passive construction:
_ laopo bei huairen bangjia de **zongcai** jueding baojing
wife BEI bad guys kidnap DE chairman decide call police
'The chairman whose wife was kidnapped by some bad guys decided to call the police.'

As far as locality is concerned, the linear distance between the filler and the gap is longest in passives, shorter in *ba* sentences, and shortest in canonical sentences. A theory based on (linear) locality would predict (12) > (11) > (10) in terms of difficulty. In such an approach, the parser probes the NPs that appear most recently first, following stack or push-down automaton order. Canonicity theory, on the other hand, would also predict (11) and (12) to be more difficult than (10) because passives and BA sentences involve more complex syntactic derivations and trace relations. Three experiments were conducted to test the validity of these predictions.

4. Experiments

4.1. Experiment 1: Naturalness and grammaticality ratings

Participants Fifty-seven undergraduate students (11 males, 42 females, 4 unidentified) from National Cheng-Chi University participated in this experiment. The participants were all native speakers of Mandarin Chinese between the age of 19 and 22.

Materials In addition to PRCs, we collected the ratings of three sets of baseline sentences. The purpose of these baseline sentences was to establish the overall pattern of sentences without filler-gap dependencies, so that the particularity of PRCs can be revealed in comparison. These baseline sentences included (a) simple sentences, (b) sentences with relative clauses, and (c) sentences with adjunct relative clauses. The canonical, *ba* and *bei* (passive) versions of these baseline sentences were all rated. Examples of these sentences in their canonical versions are given in (13)-(15).

(13) **Baseline simple sentences (canonical version)**:
kexuejia jiejue le xuduo miti
scientist solve ASP many puzzle
'Scientists have solved many puzzles.'

(14) **Baseline sentences with relative clauses (canonical version)**:
bianju xinshang de yanyuan dianran le lazhu
playwright like DE actor light ASP candle
'The actor that the playwright likes lit the candle.'

(15) **Baseline sentences with adjunct relative clauses (canonical version)**:
liwei shanchu yusuan de liyou hen nan rang ren jieshou
legislator cut budget DE reason very hard make person accept
'The reason why legislators cut the budgets was difficult to accept.'

Each set of the experimental materials contained 24 sentences. Three different questionnaires were created, each of which contained 24 simples sentences (Baseline I), 24 sentences with RCs (Baseline II), 24 sentences with adjunct RCs (Baseline III), 24 PRCs (experimental group), and 77 filler sentences of various sentence types and grammaticality status. For each set of the baseline and experimental sentences, 8 sentences were in the canonical variant, 8 with the *ba* construction, and 8 with the passive construction. Each sentence was presented only in one variant (conical, *ba*, or passive) in each questionnaire. The participants rated a total of 173 sentences in an average of 25 minutes.

Procedure Each participant was randomly given one questionnaire, which started with questions about their linguistic background, followed by a set of instructions and examples. Participants were instructed to rate the naturalness of each sentence on a scale of 1 (very natural) to 6 (very unnatural). Immediately following each naturalness rating, they were asked to rate the grammaticality of each sentence. Three choices were given: grammatical,

ungrammatical, and unsure (selected only when the participants were unsure of the grammatical status of the sentence).

Results The results of two participants were excluded due to apparent carelessness in their responses. The average of all naturalness ratings (including filler sentences) is 3.12 (SD = 1.98). The means and SDs (given in parentheses) are given in Table 1.

Table 1: Naturalness scores

	Simple Sentences	**Sentences with RCs**	**Sentences with Adjunct RCs**	**Sentences with PRCs**
Canonical	1.41 (0.95)	1.64 (1.09)	1.75 (1.16)	4.48 (1.53)
BA	1.52 (1.06)	1.84 (1.27)	2.51 (1.65)	4.92 (1.29)
BEI	1.79 (1.25)	2.27 (1.53)	3.03 (1.76)	2.87 (1.79)

ANOVAs with repeated measures showed significant differences in main effects across sentence types ($F1(3, 162) = 502.30$, $p < 0.000$) and construction variants ($F1(2, 108) = 23.53$, $p < 0.000$) by subject analysis. The different construction variants within each sentence type were also significantly different from each other by item analysis (simple sentences: $F2(2, 46) = 5.17$, $p < 0.05$; sentences with RCs: $F2(2, 46) = 11.91$, $p < 0.001$; sentences with adjunct RCs: $F2(2, 46) = 21.34$, $p < 0.001$; sentences with PRCs: $F2(2, 46) = 102.47$, $p < 0.001$). Further paired-samples *t*-tests showed that the construction variants within each sentence type were significantly different from each other by subject analysis ($p < 0.05$). By item analysis, passives were rated less natural than both canonical and *ba* variants for simple sentences ($t(23) = 2.90$, $p < 0.01$; $t(23) = 2.27$, $p < 0.05$), sentences with RCs ($t(23) = 6.10$, $p < 0.001$; $t(23) = 2.96$, $p < 0.01$), and sentences with adjunct RCs ($t(23) = 6.51$, $p < 0.001$; $t(23) = 2.14$, $p < 0.05$), but more natural than both the canonical and BA variants for PRCs ($t(23) = 9.56$, $p < 0.001$; $t(23) = 12.37$, $p < 0.001$). A bar chart with the means of each group is given in Figure 1.

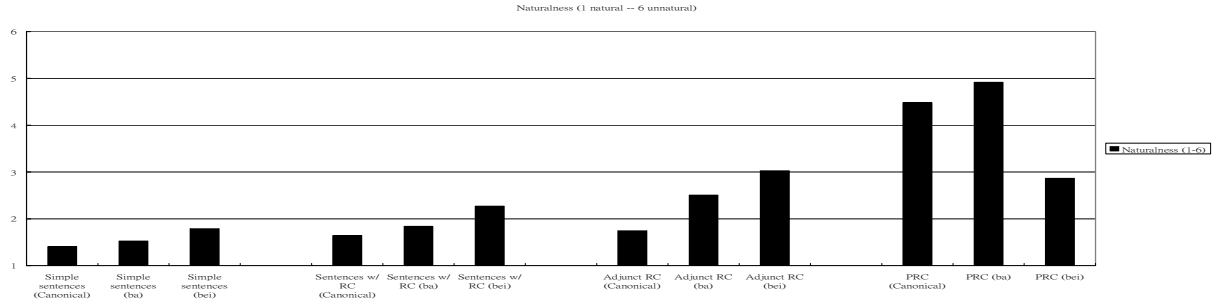


Figure 1: Bar chart for the naturalness scores

Similar patterns were found with the grammaticality judgments. Table 2 gives the percentage of responses in each category. We only focus on the percentages of grammatical judgments for simplicity; the ungrammatical percentages follow exactly same patterns in the opposite direction. ANOVAs with repeated measures showed significant differences in main effects across sentence types ($F1(3, 162) = 509.22$, $p < 0.000$) and construction variants ($F1(2, 108) = 19.96$, $p < 0.000$) by subject analysis. The different construction variants within each sentence type were also significantly different from each other by item analysis (simple sentences: $F2(2, 46) = 5.92$, $p < 0.01$; sentences with RCs: $F2(2, 46) = 10.50$, $p < 0.001$; sentences with adjunct RCs: $F2(2, 46) = 14.30$, $p < 0.001$; sentences with PRCs: $F2(2, 46) = 111.68$, $p < 0.001$). Further paired-samples *t*-tests showed that the construction variants within each sentence type were significantly different from each other by subject analysis ($p < 0.05$). By item analysis, passives were rated less grammatical than canonical variants for simple sentences ($t(23) = 3.07$, $p < 0.01$), less grammatical than both canonical and *ba* variants for sentences with RCs ($t(23) = 5.68$, $p < 0.001$; $t(23) = 2.27$, $p < 0.05$), and less grammatical

than canonical variants for sentences with adjunct RCs ($t(23) = 4.92$, $p < 0.001$). Passives were, however, rated more grammatical than both the canonical and BA variants for PRCs ($t(23) = 11.31$, $p < 0.001$; $t(23) = 11.43$, $p < 0.001$). Figure 2 illustrates the patterns of these grammaticality ratings.

Table 2: Percentage of grammaticality judgments.

	Simple Sentences			**Sentences w/ RC**			**Adjunct RC**			**Possessor RC**		
	gram.	ungram.	unsure	gram.	ungram.	unsure	gram.	ungram.	unsure	gram.	ungram.	unsure
Canonical	96.10	3.21	0.68	92.89	5.75	1.36	88.56	7.95	3.48	18.06	77.62	4.32
BA	92.01	6.85	1.14	85.68	13.18	1.14	71.07	24.84	4.09	10.39	83.86	5.75
BEI	86.69	10.81	2.50	77.14	18.96	3.90	58.57	37.27	4.16	65.02	31.96	3.02

(gram. = grammatical, ungram. = ungrammatical)

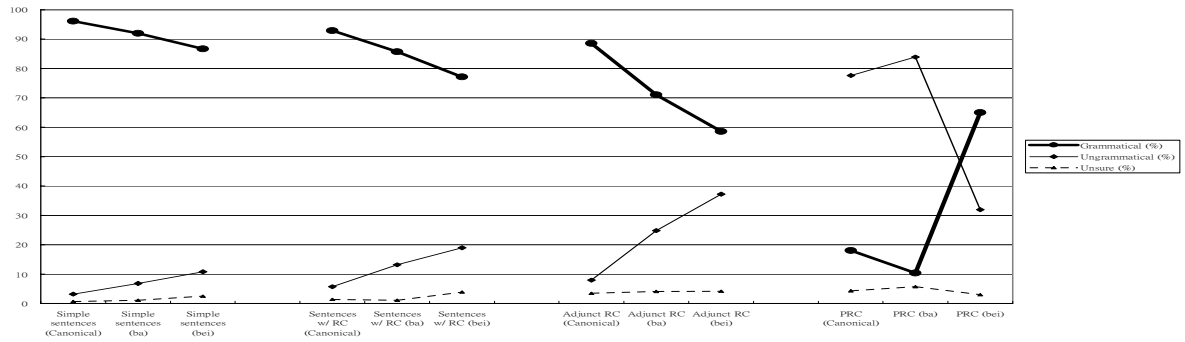


Figure 2: Percentages of grammaticality ratings.

Discussion The results for the baseline sentences showed a very similar trend. Canonical variants were considered more natural than the *ba* variants. Passives were considered most unnatural and most ungrammatical. This confirms the canonicity effect. However, such a pattern is only partially observed in PRCs. PRCs with *ba* were rated more difficult and less grammatical than their canonical counterparts. Passive PRCs were considered more natural and more grammatical than both the canonical and BA variants. These results suggest that the filler-gap relations constructed in PRCs produced an advantage for passives than for the canonical and *ba* sentences.

4.2. Experiment 2: Paraphrasing tasks

Since these PRCs are uncommon and predominantly rated as unnatural, we collected paraphrases in Experiment 2 to get at how they were understood, and whether the possessive relations between the head nouns and the objects were perceived.

Participants Eighty-three undergraduate students (14 males, 62 females, 7 unidentified) from National Cheng-Chi University, National Taipei Teachers' College and National Hsin-Chu Teacher's College participated in this experiment. The participants were all native speakers of Mandarin Chinese between the age of 19 and 22. None of them participated in Experiment 1.

Materials Twelve sets of sentences with PRCs were selected from those used in Experiment 1. These sentences were quasi-randomly placed into 4 questionnaires. Each questionnaire contained 9 sentences that were randomly ordered. Each participant only paraphrased a variant of each sentence.

Procedure Participants were instructed to carefully read each sentence and then use their own words to paraphrase them. They were also instructed to cross out a sentence if they did not understand it. The whole questionnaire took about 10 minutes to finish.

Results and discussion Among all the 741 collected paraphrases, participants were able to paraphrase 638 correctly. The overall accuracy of correct paraphrases was 86.1%. This shows

that most participants were able to understand and interpret PRCs correctly. Among these correct paraphrases, 331 responses (51.88%) overtly specified the possessive relation between the head noun and the possessed NP. Among the sentences that were *not* understood correctly, only 18% were passive PRCs; 38% were canonical PRCs; 44% were PRCs with BAs. This further confirms the findings in Experiment 1 that passive PRCs are easier to comprehend. The raw frequencies and percentages of accurate paraphrases are given in Table 3. In addition to the accuracy of interpretations, the percentage of correct paraphrases that overtly mentioned possessive relations was also calculated. Passive PRCs most frequently induce overt mentioning of possessive relations in paraphrasing. These are reported in Table 4.

Table 3: Frequency and percentage of correct and incorrect paraphrases among PRCs.

Possessor RCs	Correct Paraphrasing		Incorrect Paraphrasing	
	N	%	N	%
Canonical	207	83.81	40	16.19
BA	202	81.78	45	18.22
BEI	228	92.31	19	7.69

Table 4: Frequency and percentage of possessive relations identified in the correct paraphrases.

Possessor RCs	Possessive Relations Identified	
	N	%
Canonical	101	48.79
BA	94	46.53
BEI	135	59.21

The results of this experiment confirmed the findings of Experiment 1, showing that passive PRCs were most likely to be correctly interpreted. They also established a stronger possessive relation between the filler and the gap, as was revealed by a higher percentage of possessive relations being reported in the paraphrases of passive PRCs.[3]

4.3. Experiment 3: Self-paced reading tasks

Experiment 3 collected self-paced reading data to show the real-time process in constructing filler-gap dependencies.

Participants Twenty-six undergraduate students (8 m, 18 f) from National Cheng-Chi University participated. The participants were all native speakers of Mandarin Chinese between the age of 19 and 22. None of them participated in previous experiments.

Materials Experimental materials included the three sets of baseline sentences and the experimental sentences with PRCs used in Experiment 1. These sentences were arranged into three different lists by a Latin-square design. One hundred and four filler sentences of various types were added to each list, which were altogether randomly presented to the participants.

Procedure This experiment was run on a notebook computer (SONY PCG-VX89) using Linger 2.94 developed by Doug Rohde at MIT. All experimental materials were presented phrase by phrase on a single line. Sentences were presented as dashes (with each dash covering one Chinese character) without spaces between words and phrases. Participants were instructed to carefully read the sentences by pressing the space bar for each phrase to appear. A comprehension question (either a true/false question or a multiple-choice question) follows each sentence. After each incorrect response, a visual feedback appeared on the screen,

[3] These numbers are under-estimates of the actual perception of possessive relations, since not all possessive relations perceived were reported. They were, nevertheless, indicative of the different strengths of possessive relations these different constructions conveyed.

indicating that the response was incorrect. No feedback was given to correct responses. The whole experiment took about 30 minutes to complete.

Results The data of one participant were excluded due to technical problems during the experiment. The average comprehension accuracy for all other participants was 94.46%. All the participants were able to answer at least 90% of the comprehension questions correctly. Reading time data were analyzed by collapsing words and phrases into regions. For simplification, only the RTs of the adjunct RCs and the PRCs are compared since these two sentence types match best in their surface word orders, differing only in filler-gap dependency. The average reading times per Chinese character for the whole sentences were calculated for comparison. In a 2 by 3 ANOVA analysis of repeated measures, we found main effects in both sentence types (adjunct RCs versus PRCs) ($F(1, 24) = 7.80$, $p < 0.05$) and construction variants (canonical, *ba*, and passives) ($F1(2, 48) = 5.64$, $p < 0.01$; $F2(2, 46) = 3.56$, $p < 0.05$ for adjunct RCs; $F2(2, 46) = 11.24$, $p < 0.001$ for PRCs). Paired-samples *t*-tests showed that among the adjunct RCs, the passives took longer to comprehend than both the canonical variants ($t1(24) = 3.59$, $p < 0.01$; $t2(23) = 2.54$, $p < 0.05$) and the *ba* variants ($t1(24) = 2.24$, $p < 0.05$; $t2(23) = 1.54$, $p = 0.14$). RTs for the canonical adjunct RCs were shorter than those for the *ba* variants but the difference did not reach significance ($t1(24) = 2.04$, $p = 0.05$; $t2(23) = 1.14$, $p = 0.27$). Among the PRCs, passives were read faster than both the canonical ($t1(24) = 5.56$, $p < 0.001$; $t2(23) = 4.31$, $p < 0.001$) and *ba* variants ($t1(24) = 5.48$, $p < 0.001$; $t2(23) = 5.13$, $p < 0.001$). RTs for the canonical and *ba* PRCs were not significantly different ($t1(24) = 0.08$, $p = 0.94$; $t2(23) = 0.34$, $p = 0.74$). Comparing between adjunct RCs and PRCs within the same construction types, we found PRCs to be longer than adjunct RCs for both the canonical ($t(24) = 4.80$, $p < 0.001$) and the *ba* variants ($t(24) = 3.76$, $p < 0.01$), but shorter for the passives ($t(24) = 3.13$, $p < 0.01$). The mean RTs per character are given in Table 3.

Table 3: Mean RTs (ms) per Chinese character of adjunct RCs and PRCs across construction variants. (SDs are in parentheses.)

	Sentences with Adjunct RCs	**Sentences with PRCs**
Canonical	332.80 (80.47)	412.74 (137.18)
BA	347.95 (80.11)	414.00 (141.21)
BEI	368.08 (89.88)	329.54 (89.07)

By-region reading time analyses are provided in Figures 3 and 4. We compared the RTs of the regions after the relativizor DE where the relative clauses were recognized and the process of filler-gap identification was initiated. Both the RTs across sentence types and between construction variants were compared. Within the sentences with adjunct RCs, the main effect of construction variants was found at N3 ($F(2, 48) = 3.62$, $p < 0.05$) and V2 ($F(2, 48) = 21.42$, $p < 0.001$), but not with DE ($F(2, 48) = 2.23$, $p = 0.12$). RTs of passive adjunct RCs were longer than the canonical variants at N3 ($t(24) = 2.13$, $p < 0.05$) and V2 ($t(24) = 6.13$, $p < 0.001$), and longer than the *ba* variants at V2 ($t(24) = 3.61$, $p < 0.01$). RTs of adjunct RCs with *ba* were longer than the canonical variants at V2 ($t(24) = 3.16$, $p < 0.01$).

For sentences with PRCs, the main effect of construction variants was also found with N3 ($F(2, 48) = 10.63$, $p < 0.001$) and V2 ($F(2, 48) = 19.01$, $p < 0.001$), but not with DE ($F(2, 48) = 2.41$, $p = 0.10$). RTs of passive PRCs were *shorter* than the canonical variants at N3 ($t(24) = 3.53$, $p < 0.01$) and V2 ($t(24) = 4.64$, $p < 0.001$), and *shorter* than the *ba* variants at N3 ($t(24) = 4.12$, $p < 0.001$) and V2 ($t(24) = 6.12$, $p < 0.001$). RTs of the canonical and *ba* variants were not significantly different either at N3 ($t(24) = 1.98$, $p = 0.60$) or at V2 ($t(24) = 0.58$, $p = 0.57$).

Comparing between adjunct RCs and PRCs within the same construction variants, RTs were longer for PRCs than for adjunct RCs in the canonical variants (N3: $t(24) = 3.90$, $p < 0.01$; V2:

t(24) = 6.73, $p < 0.001$) and the *ba* variants (DE: t(24) = 3.13, $p < 0.01$; N3: t(24) = 3.59, $p < 0.01$; V2: t(24) = 6.06, $p < 0.001$). With the passive variants, however, PRCs and adjunct RCs were not significantly different except at V2, where the RTs of PRCs were *shorter* than those of adjunct RCs (t(24) = 2.99, $p < 0.01$).

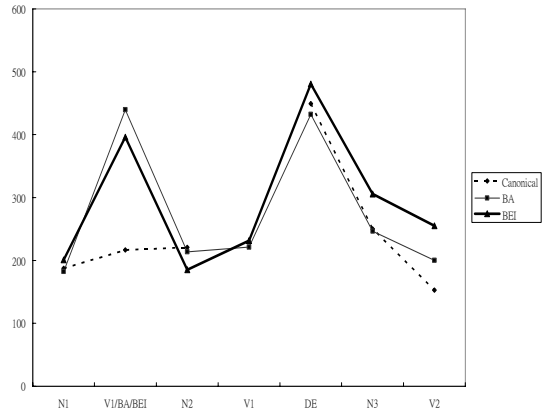


Figure 3: RTs (in ms) per Chinese character of adjunct RCs in three construction variants. (N3 = head noun, V2 = matrix verb)

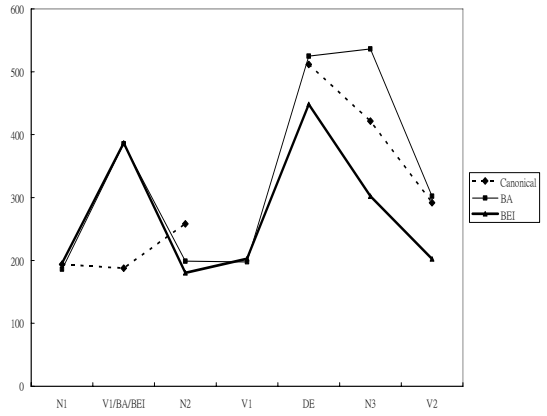


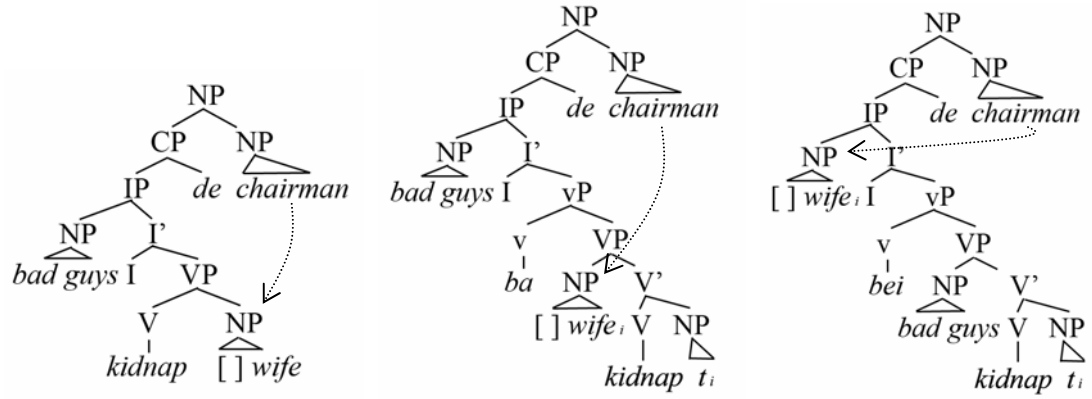
Figure 4: RTs (in ms) per Chinese character of PRCs in three construction variants. (N3 = head noun, V2 = matrix verb)

Discussion The on-line reading times showed that in adjunct RCs, where there is no filler-gap relation to be constructed, passives were read more slowly than both the canonical and *ba* variants. This result is consistent with the canonicity effect—passive constructions at the RC section made these sentences harder to comprehend at the post-RC regions. For PRCs, however, the parser started to look for filler-gap relations at the relativizor DE, producing longer reading times for PRCs than for adjunct RCs (e.g. with the *ba* construction). Such longer reading latencies consummated at the head noun (N3) and continued at the matrix verb (V2) for both canonical and *ba* PRCs, but not for the passive PRCs. The passive PRCs were read as fast as adjunct RCs at the relativizor and the head noun, and read faster at the matrix verb (V2). These results (*ba* > canonical > passives, in terms of RTs) contradicted the predictions of linear locality. Passive PRCs received an advantage, even though their gaps were located the farthest from the fillers. The effect of canonicity observed on adjunct RCs was not found on PRCs either. In summary, passive PRCs provided an advantage for gap-searching, which overrode both linear locality and canonicity. Such effects were not observed in adjunct RCs, which suggested that it is a process only related to gap-searching.

5. General discussion and conclusion

These three experiments consistently suggest that passive PRCs received advantage for filler-gap processing, which was not observed either in the canonical or in the *ba* variants. In Experiment 1, passive PRCs were considered more natural and more grammatical than both canonical and *ba* PRCs, countering the trend in all baseline conditions. Experiment 2 showed that even though sentences with PRCs were unusual, they were highly comprehensible. Passive PRCs were most frequently understood correctly, and most likely to incur the mention of possessive relations in the interpretations. In Experiment 3, we saw that the canonical and *ba* PRCs both required additional reading latency in the construction of filler-gap dependencies at the head noun and the matrix verb regions. However, passive PRCs were read as fast as (at the relativizor and the head noun) or faster (at the matrix verb) than their adjunct RC counterparts.

These results suggest that in parsing Chinese PRCs, the construction of filler-gap relations does not start until the head noun is reached. This is different from typical Chinese RCs with the gaps detected first and creating an expectation for the filler head noun. Though Chinese PRCs are prenominal, the parser is not aware of any gaps since without missing arguments, PRCs look just like thematically-complete sentences. The region prior to the relativizer *de* is therefore processed as a regular clause. When the parser reaches the relativizer, it incorporates the prior region as part of a relative clause. The head noun, as a filler with the expectation of a gap, then initiates the process of gap-searching. The parser probes at the subject position of the previous CP first, and then searches other NPs down the hierarchy. Figure 4 provides the syntactic diagrams of the PRCs in different constructions.



(a) PRC with canonical orders. (b) PRC with *ba* construction (c) Passive PRCs (with *bei*)

Figure 4: Syntactic diagrams for PRCs with different constructions.[4]

Such a goal-searching strategy requires knowledge of syntactic structure. The parser does not start by probing at the NP that appeared most recently. (That is, the NPs were not stacked into a push-down machine for gap-filling.) Instead, it probes at the topmost left-edge NP, thus producing an advantage for passive PRCs. In canonical PRCs, the parser has to search all the way to the object position to reach the gap. The same is true with the *ba* PRCs. This top-down searching mechanism overrides the effect of NP recency (i.e. linear locality), and passive complexity (i.e. canonicity). PRCs with *ba* are harder than canonical PRCs for two reasons. The parser has to search through one additional XP node in *ba* PRCs than in canonical PRCs to reach the target NP. The non-canonicity of *ba* constructions in terms of theta location and additional NP movements also make them more difficult.

The above evidence also suggests that even though RCs in Chinese are prenominal, when the gaps are not overt as in PRCs, the parser may gap-search in a similar way for English, where the fillers precede and search for the gaps. This produces a subject preference for Chinese PRCs, which is consistent with the same preference in regular English RCs. The results thus suggest that the top-down, structure-based parsing mechanism is universal, working for both prenominal and postnominal relative clauses.

This proposal also receives independent support from other languages with prenominal relative clauses like Japanese and Turkish. In Japanese, Fong and Hirose (2005) reported that object PRCs with objects scrambled to the surface subject position were rated no harder than non-scrambled canonical PRCs, while in their baseline sentences, object scrambling was rated more difficult. Similar preferences were obtained in Turkish, where an ambiguous bare noun phrase appearing in a PRC is predominantly taken as a subject than an object (see also discussion in Fong & Hirose, 2005). These suggest that similar gap-searching mechanisms

[4] We follow Tsai (2003) for the syntactic structures of *ba* and *bei* sentences.

operate in languages with prenominal and postnominal RCs. The parser searches the structure hierarchically in a top-down manner. Gaps at subject positions universally receive advantage when the filler is recognized prior to the gap.

Acknowledgements

We are thankful to the following people for suggestions on previous presentations: Kathleen Ahrens, San Duanmu, One-Soon Her, Yuki Hirose, Chu-Ren Huang, Michael Silverstein. Any errors are, nevertheless, our own responsibility.

References

Birch, S. L., J. E. Albrecht and J. L. Myers. 2005. Syntactic Focusing Structures Influence Discourse Processing. *Discourse Processes*, 30, pp. 285-304.

Chomsky, N. 2001. Derivation by Phase. In M. Kenstowicz (Ed.) *Ken Hale: A Life in Language* (pp. 1-52). Cambridge, MA: MIT Press.

Ferreira, F. 2003. The Misinterpretation of Noncanonical Sentences. *Cognitive Psychology*, 47, pp. 164-203.

Fong, S. and Y. Hirose. 2005. Minimalist Parsing and Japanese: On Theory and Experiential Factors. *The Proceedings of Tokyo Conference on Psycholinguistics* (TCP 2005).

Frazier, L. and C. Clifton Jr. 1989. Successive Cyclicity in the Grammar and the Parser. *Language and Cognitive Processes*, 4, pp. 93-126.

Frazier, L., C. Clifton Jr. and J. Randall. 1983. Filling Gaps: Decision Principles and Structure in Sentence Comprehension. *Cognition*, 13, pp. 187-222.

Frazier, L., C. Clifton, K. Rayner, P. Deevy, S. Koh and M. Bader. 2005. Interface Problems: Structural Constraints on Interpretation? *Journal of Psycholinguistic Research*, 34, pp. 201-231.

Frazier, L. and G. Flores d'Arcais. 1989. Filler-Driven Parsing: A Study of Gap-Filling in Dutch. *Journal of Memory and Language*, 28, pp. 331-344.

Gibson, E. 1998. Linguistic Complexity: The Locality of Syntactic Dependencies. *Cognition*, 68, pp. 1-76.

Gibson, E., T. Desmet, D. Grodner, D. Watson and K. Ko. 2005. Reading Relative Clauses in English. *Cognitive Linguistics*, 16, pp. 313-353.

Hsiao, F. and E. Gibson. 2003. Processing Relative Clauses in Chinese. *Cognition*, 90, pp. 3-27.

King, J. and M. A. Just. 1991. Individual Differences in Syntactic Processing: The Role of Working Memory. *Journal of Memory and Language*, 30, pp. 580-602.

Mazuka, R., K. Itoh and T. Kondo. 2002. Costs of Scrambling in Japanese Sentence Processing. In M. Nakayama (Ed.) *Sentence Processing in East Asian Languages* (pp. 131-166), Stanford, CA: CSLI Publications.

Stowe, L. A. 1986. Parsing *WH*-Constructions: Evidence for On-Line Gap Location. *Language and Cognition Processes*, 1, pp. 227-245.

Tamaoka, K., H. Sakai and J. Kawahara. 2003. The Effects of Phrase-Length Order and Scrambling in the Processing of Visually Presented Japanese Sentences. *Journal of Psycholinguistic Research*, 32, pp. 431-54.

Tamaoka, K., H. Sakai and J. Kawahara. 2005. Priority Information Used for the Processing of Japanese Sentences: Thematic Roles, Case Particles or Grammatical Functions? *Journal of Psycholinguistic Research*, 34, 281-332.

Traxler, M. J., R. K. Morris and R. E. Seely. 2002. Processing Subject and Object Relative Clauses: Evidence from Eye Movements. *Journal of Memory and Language*, 47, pp. 69-90.

Tsai, W-T. D. 2003. Three Types of Existential Quantification in Chinese. In A. Li and A. Simpson (Eds.) *Functional Structure(s), Form and Interpretation: Perspectives from Asian Languages*, London: Routledge.