

ORIGINAL RESEARCH REPORT

When Structure Competes with Semantics: Reading Chinese Relative Clauses

Peiyun Zhou*, Yun Yao† and Kiel Christianson‡§

An ongoing debate in Chinese psycholinguistics is whether subject-relative clauses or object-relative clauses are more difficult to process. The current study asks what happens when structure and plausibility are pitted against each other in Chinese relative clause processing. Chinese relative clause structures and semantic plausibility were manipulated to create both plausible and implausible versions of subject- and object-relative clauses. This method has been used in other languages (e.g., English) to elicit thematic role reversal comprehension errors. Importantly, these errors—as well as online processing difficulties—are especially frequent in implausible versions of dispreferred (noncanonical) structures. If one relative clause structure in Chinese is highly dispreferred, the structural factor and plausibility factor should interact additively. If, however, the structures are relatively equally difficult to process, then there should be only a main effect of plausibility. Sentence reading times as well as analyses on lexical interest areas revealed that Chinese readers used plausibility information almost exclusively when reading the sentences. Relative clause structure had no online effect and small but consistent offline effects. Taken together, the results support a slight preference in offline comprehension for Chinese subject-relative clauses, as well as a central role for semantic plausibility, which appears to be the dominant factor in online processing and a strong determinant of offline comprehension.

Keywords: Chinese relative clauses; semantic plausibility; syntactic parsing; eye movements; Good Enough Theory

Introduction

Chinese relative clauses processing

Chinese has a Subject-Verb-Object (SVO) structure, but the word order in Chinese relative clauses is different from most other languages with SVO structure. Unlike English or French, whose relative clause structure is head-initial, Chinese has head-final relative clauses, with the relative clause preceding the head noun (see (1)). Thus, in the Chinese subject-relative clause (SRC), the word order is “V-N(patient)-de-N(agent)” while in object-relative clause (ORC), the word order is “N(agent)-V-de-N(patient).” On one hand, the distance between the relative clause verb and the head noun is longer in SRCs, which may cause longer integration and processing time for SRCs. In addition, the canonical agent-first word order is reversed in SRCs, potentially also increasing processing difficulty. On the other hand, the *de* in the ORC is temporarily

ambiguous. The relativizer *de* also appears as a post-nominal genitive (possessive) marker. This ambiguity in ORCs is not resolved until the main clause verb. Thus, it may take longer for readers to process and reanalyze the *de* in ORCs to reach the correct interpretation. In both cases, the main clause verb is a critical point of integration and/or disambiguation (cf. Staub, Dillon, & Clifton, 2017).

- (1) a. 传唤(summon)证人(witness)的(that)法官(judge)喜欢(like)京剧(Beijing opera).
The judge that summoned the witness liked Beijing opera. (SRC, plausible)
- b. 传唤(summon)法官(judge)的(that)证人(witness)喜欢(like)京剧(Beijing opera).
The witness that summoned the judge liked Beijing opera. (SRC, implausible)
- c. 法官(judge)传唤(summon)的(that)证人(witness)喜欢(like)京剧(Beijing opera).
The witness that the judge summoned liked Beijing opera. (ORC, plausible)
- d. 证人(witness)传唤(summon)的(that)法官(judge)喜欢(like)京剧(Beijing opera).
The judge that the witness summoned liked Beijing opera. (ORC, implausible)
- e. The judge/witness summoned the judge/witness.
The witness/judge liked Beijing opera (True/False)

* Department of Psychology, University of Washington, Seattle, Washington, US

† Department of World Languages, Literature, and Cultures, University of Arkansas, Fayetteville, Arkansas, US

‡ Department of Educational Psychology, University of Illinois Urbana-Champaign, Illinois, US

§ Beckman Institute for Advanced Science and Technology, University of Illinois Urbana-Champaign, Illinois, US

Corresponding author: Peiyun Zhou (peiyun@uw.edu)

Given these characteristics, there is controversy about whether SRCs are easier to process in Chinese (as they are in English), or whether ORCs are easier. Researchers have found that Chinese SRCs are easier to process and comprehend than ORCs in a series of self-paced reading and eye-tracking experiments (e.g., Jäger, Chen, Li, Lin, & Vasishth, 2015; Lin & Bever, 2006; Vasishth, Chen, Li, & Guo, 2013). They argue that the findings support experience-based or frequentist theories: because SRCs are more frequent than ORCs in Chinese, readers have been exposed to SRCs more and thus process and comprehend Chinese SRCs more quickly and easily (Gennari & MacDonald, 2008; Hsiao & MacDonald, 2013; Levy, 2008).

Conversely, other researchers have observed that SRCs are processed more slowly than ORCs, as evidenced by inflated reading times on the first noun, relative clause verb, relativizer *de*, and second noun in several self-paced reading studies (Chen, Ning, & Dunlap, 2008; Gibson & Wu, 2011; Hsiao & Gibson, 2003). Based on these findings, these authors propose working-memory-based theories (e.g., Gibson, 1998) account for the difference: when reaching the head noun in a Chinese relative clause, readers have more difficulty retrieving the relative clause verb from their working memory in SRCs than ORCs, because the distance between the relative clause verb and the head noun is greater in SRCs than ORCs. Specifically, the working-memory-based theory predicts this difficulty in retrieving the relative clause verb should occur at the relativizer *de* and head noun.

Of course, a third possibility exists: Chinese SRCs and ORCs might be relatively equivalent in terms of parsing difficulty, for different reasons. And in fact, a third research camp has reported that no preference exists between relative clauses among Chinese readers. Chen and colleagues (2006) observed no preference for either SRCs or ORCs in a Chinese sentence complexity rating task. As described above, there are grounds to suspect that both structures present parsing difficulty for different reasons. To be more specific, compare Chinese RCs to English RCs. In English, ORCs are the dispreferred structure for multiple reasons: they are less frequent than SRCs (Wells, Christiansen, Race, Acheson, & MacDonald, 2009), they impose a greater working-memory burden (Gibson, 1998), and thematic role assignment is non-canonical (e.g., Christianson, Luke, & Ferreira, 2010; Ferreira, 2003; Lim & Christianson, 2013; Townsend & Bever, 2001; Zhou & Christianson, 2016a, b; Zhou, 2017). Given that these various factors leading to processing difficulty are distributed between SRCs and ORCs in Chinese, there is reason to question whether an *a priori* assumption that Chinese must have **one** highly preferred RC structure is warranted. Nevertheless, precisely this assumption as has largely been either explicitly or implicitly made in the previous, conflicted literature.

One potential factor contributing to these mixed results is the plausibility of the thematic role assignments (i.e., agent and patient roles) in the sentences that have been used in previous studies. Since Crain and Steedman (1985), it has been recognized that relative plausibility can modulate processing speed and accuracy of difficult

syntactic structures. Although some previous studies of Chinese RC processing have controlled for plausibility (e.g., Gibson & Wu, 2011; Jäger et al., 2015), there has been no attempt to pit plausibility against syntactic structure explicitly. The closest example to this sort of a design was Wu, Kaiser, & Anderson (2012), who manipulated the animacy of the RC heads. They found that processing preferences for the two RC structures could be reversed, depending on the animacy configurations of the two nouns (animate-inanimate, inanimate-animate). This finding strongly suggests that neither memory load nor the relative frequency of the structures can account for the previously observed processing asymmetries. If in fact neither relative clause structure is strongly preferred, i.e., if both structures contain characteristics that could complicate parsing and/or comprehension, then the semantic content might weigh heavily in guiding the online behaviors taken as indicators of parsing and/or the offline comprehension of the structures. In other words, it is possible that previous conflicting results are attributable to semantic confounds in the materials. This is the hypothesis we explore here. In order to do so, we employ an experimental paradigm that has been used over the past 15 years within the Good Enough Processing literature (cf. Christianson, 2016; Ferreira, Christianson, & Hollingworth, 2001; Ferreira & Patson, 2007), which pits asymmetrically plausible thematic role assignments against (purportedly) asymmetrically complex syntactic structures. We describe below how this manipulation results in strong additive interactions in structures such as English passives vs. actives and ORCs vs. SRCs. Specifically, when implausible thematic assignments occur in syntactic structures that are more difficult to process, online processing speed is slowed and offline comprehension accuracy is reduced. The working hypothesis that we test here, then, is that if one Chinese RC structure is in fact preferred over the other, we should find interactively additive effects of plausibility and structure on the dispreferred RC structure. If, however, the two RC structures are approximately equally difficult to process (for different reasons), then we should find only an effect of plausibility such that implausible sentences cause slower processing and worse comprehension accuracy.

Good-Enough (GE) Processing theory

Traditional psycholinguistic theories hold that readers comprehend sentences by computing syntactic representations relatively effortlessly and automatically, especially for arguments of main clause verbs (e.g., Frazier & Clifton, 1996; see Ferreira et al., 2001). From these syntactic structures, semantic representations are derived, based on syntactically licensed interpretations of the verbatim content of the sentences.

Misinterpretations do often happen though, and moreover, they can occur in systematic and predictable ways, not just haphazardly (e.g., due to inattention; see Christianson, 2016). Evidence suggests that misinterpretations are derived from competition between the syntactic structure of the verbatim input and other information sources, i.e., various processing heuristics

that can, under certain circumstances, overwhelm the syntactic structure (e.g., Christianson et al., 2001, 2010; Christianson & Luke, 2011; Christianson, Williams, Zacks, & Ferreira, 2006; Ferreira, 2003; den Ouden, Dickey, Anderson, & Christianson, 2016; Ferreira, Bailey, & Ferraro, 2002; Hussey, Ward, Christianson, & Kramer, 2015).

Good-Enough (GE) Processing theory focuses on accounting for how misinterpretations arise (Christianson, 2016; Ferreira et al., 2002; Ferreira & Patson, 2007). It suggests that readers often rely on a heuristic processing route to build representations that are judged to be good enough to handle the current language comprehension task, but often not good enough to reach an accurate interpretation of the actual language input. GE processing-related phenomena have been observed when readers (both native and non-native speakers) process garden-path sentences (Christianson et al., 2001, 2006; Lassotta, Omaki, & Franck, 2016), active/passive sentences (Christianson et al., 2010; Christianson & Lim, 2013a; Ferreira, 2003), and plausible/implausible relative clauses (RCs) (Lim & Christianson, 2013b; Zhou & Christianson, 2016a, b).

When native English speakers read garden-path sentences, such as (2a), they often misinterpret *deer* as the object of the subordinate verb *hunted* and answer “yes” to (2b).

- (2) a. While the man hunted the deer ran into the woods.
 b. Did the man hunt the deer?
 c. While the man hunted the deer paced in the zoo.

Although this interpretation is certainly available to inferential processes, it is not strictly licensed by the syntax. Furthermore, most readers also assign *deer* to the subject position of *ran*, yet still maintain the syntactically unlicensed interpretation that the man hunted the deer and the deer ran into the woods (Patson, Darowski, Moon, & Ferreira, 2009). Note also that readers are far less likely to answer “yes” to questions such as (2b) after reading sentences like (2c), despite the fact that (2c) presents the same structural difficulties as (2a). Thus, these readers derive an interpretation that is good enough to satisfy a general constraint on plausibility, but not good enough to derive an interpretation that is strictly faithful to the content of the input (i.e., that the man was hunting in general, but not necessarily the particular deer that was running away).

Lassotta and colleagues (2016) have observed the same good-enough processing among native French speaking adults and children when these readers (mis) interpreted garden-path sentences in French. In garden-path sentences like *Où est-ce qu'Aline a expliqué dans le salon qu'elle allai attraper des papillons?* (“Where did Aline explain in the living room that she was going to catch butterflies?”), the filled-gap prepositional phrase (*in the living room*) should help readers to disambiguate the interpretation and associate the *wh*-phrase to the embedded-clause verb. However, French adults and children still preferentially attached the *wh*-phrase to the first verb in the sentence (main clause verb) rather than the embedded-clause verb. The authors attributed the main clause misinterpretation to good-enough syntactic

representations. When both the *wh*-phrase and the filler-gap prepositional phrase are competing for attachment to the main clause verb, readers maintain the incomplete, and ultimately incorrect, syntactic representation that they first build and fail to revise it correctly to reach the true interpretation.

More evidence of good enough processing has been found in processing sentences containing thematic role assignment difficulties, such as active/passive sentences (Ferreira, 2003). When readers are asked to determine the thematic roles of plausible active sentences (e.g., “The dog bit the man”), implausible active sentences (e.g., “The man bit the dog”), plausible passive sentences (e.g., “The man was bitten by the dog”), and implausible passive sentences (e.g., “The dog was bitten by the man”), they make the most errors in response to the implausible passive sentences because the canonical noun–verb–noun syntactic order in English tends to map onto AGENT–verb–PATIENT thematic structure (Townsend & Bever, 2001), and this order generates a plausible interpretation. As a result, the interpretation that should be derived from the noncanonical (passive) syntactic structure is over-ridden by a plausible semantic interpretation, and thematic roles are reversed (cf. Christianson et al., 2010, in which implausible active sentences primed the production of passives, but implausible passives did not). The results suggest that readers rely on a word-order heuristic as well as a plausibility heuristic, which conspire to cause miscomprehension. The same processing pattern has been observed in second language (L2) English speakers’ comprehension as well (Jacob & Felser, 2016; Lim & Christianson, 2013a).

Recently, Zhou and Christianson (2016a) found that readers engage in good-enough processing when reading unambiguous RCs with similarly competing syntactic and semantic information. Computation of the more challenging ORC structure in (3a) is interfered by the implausible semantic information, causing longer fixation durations, longer whole sentence reading times, more regressions in sentence processing, and lower comprehension accuracy in a paraphrase verification task (3b).

- (3) a. The cat that the mouse chased was fast.
 b. The mouse chased the cat. The cat was fast. (Paraphrase verification. T/F)

Zhou and Christianson found that readers were more heavily influenced by the semantic plausibility information in computing the meanings of these sentences compared to when the same implausible thematic roles appeared in SRCs. This pattern suggests that when the algorithmic syntactic route runs into increased computational complexity, and its output conflicts with the output of the heuristic semantic route, readers tend to rely on the heuristic route to guide attentional allocation during parsing and in the offline interpretation. This “fast and frugal” processing strategy often yields an incorrect interpretation when sentences are purposely constructed in such a way so as to pit structure and semantics against

one another. It could be argued that the application of heuristics is not particularly “fast,” given that reading times are often reported to be elevated in more difficult semantic sentences. However, the “fast and frugal” characterization (Ferreira, 2003) should be evaluated in the full context of the processing routines hypothesized within the GE framework. Slattery, Sturt, Christianson, Yoshida, and Ferreira (2013) present evidence that, contra an earlier proposal by Christianson et al. (2001, 2006), syntactic structures might be fully revised. These revised structures and the semantic representations they generate enter into competition with previously built, or partially built, structures and their representations. Adjudication between the resulting semantic representations, including full inhibition of one of the representations if required by the task at hand, requires time and effort (cf. Hussey et al., 2015).

“Good-enough” processing is usually quite effective and accurate in normal linguistic environments; however, when materials are constructed to contain temporary ambiguities or to pit structure and semantics against one another (as is the case here), comprehension errors are frequently observed. Importantly for the present study, previous results within the good-enough framework predict the following interaction of structural difficulty and plausibility. When one structure is strongly dispreferred (i.e., when it is non-canonical due to less frequent thematic role assignment mappings), implausible thematic role assignments create even more difficulty in both online parsing and offline comprehension, resulting in an additive interaction between structure and plausibility such that implausible non-canonical structures are most difficult to process and comprehend.

In the case of Chinese relative clauses, however, the prediction is potentially different. If one relative clause structure is strongly preferred, the same interaction as in English should hold. But if there is not a strongly preferred structure, plausibility alone should affect online and offline measures. In other words, if *both* SRCs and ORCs present parsing difficulties (for different reasons), one might expect a relatively equal application of semantic plausibility information across both structures as a strategy for extracting the meaning in a hopefully “good-enough” fashion. Recall that the previous literature disagrees on which structure is preferred, and that the one previous manipulation of plausibility information pushed around the structural preference (Wu et al., 2012). As such, it might be that what has contributed to the conflicting results is plausibility, which sometimes has been controlled, but not manipulated. In this case, we would expect main effects of plausibility *only*, as structural cues are difficult to compute in both relative clause structures.

Current Study

This study examined structural preferences in Chinese relative clauses by examining how Chinese readers read and comprehend these structures when syntactic structure and semantic plausibility information point toward conflicting interpretations. Because previous studies have shown that animacy significantly affects relative clause processing in Chinese (consistent with our

suspicions about the importance of plausibility in Chinese relative clause processing) (Wu, 2009; Wu et al., 2010, 2012), we used only animate nouns as the objects and subjects in our study. Specifically, we used sentences that were similar to those in Zhou and Christianson (2016a), manipulating both syntactic (relative clause types) and semantic complexity (plausibility). According to Good-Enough Processing theory, the relative clause structure that is more difficult to process should show more evidence of plausibility heuristics influencing both the online processing and ultimate interpretations. If the two relative clauses are relatively equal in processing difficulty, plausibility should significantly affected interpretations, and plausibility should not interact with structure. Finally, if initial interpretations are derived solely from the licit, verbatim, syntactic structure, plausibility should have no effect on early first-pass eye movement measures, which are traditionally interpreted as indices of syntactic structure-building, independent of semantic content or sentence-level plausibility or pragmatics (e.g., Rayner, Carlson, & Frazier, 1983); later online measures and offline comprehension might still be influenced by plausibility, however.

In addition, we anticipate online processing effects, realized as inflated reading times on regions that require more syntactic processing effort. Specifically, we expect a significantly shorter processing time for ORC at the second noun, relativizer *de*, and main clause verb if working memory capacity affects the retrieval of the subject/object in relative clauses (Gibson, 1998; Hsiao & Gibson, 2003). On the other hand, we expect faster processing of SRCs if frequency and experience with relative clauses dominate readers' processing (Hsiao & MacDonald, 2013; Jäger et al., 2015; Lin & Bever, 2006; Vasishth et al., 2013). We will examine these effects at the relativizer *de*, second noun, and main clause verb regions. The rationale for analyzing these regions will be elaborated upon in the Results section. In the Discussion, we will discuss the results in connection to previous Chinese relative clause processing studies, as well as general theories of language processing, including Good Enough Processing theory.

Methods

Participants

Fifty-six native Chinese speakers (33 females) who were undergraduate and graduate students at a Midwestern American university participated in an eye-tracking experiment. Their ages ranged from 18 to 30 years old ($M = 24$; $SD = 4.38$). They received either \$10 or 1 research credit. Six participants' data were excluded due to either fatigue or difficulty in calibrating their eyes, leaving data from 50 participants in the final analyses. Before the experiment, an additional 22 native Chinese speakers from the same university completed a plausibility norming task for the target sentences. These participants did not participate in the eye-tracking experiment.

Materials

The Chinese sentence reading task included 55 target¹ sentences (see Appendix A), most of which were adapted from Ferreira (2003) and Lim and Christianson (2013b),

Table 1: Means and Standard Deviations (in parentheses) for Whole Sentence Reading Time and Comprehension Probe Accuracy.

	Whole Sentence Reading Time				Accuracy			
	Implausible		Plausible		Implausible		Plausible	
Object-RC	4184.97	(2686.73)	3350.27	(1972)	0.73	(0.44)	0.87	(0.42)
Subject-RC	3994.65	(2499.85)	3169.42	(1764.78)	0.78	(0.33)	0.89	(0.32)

Note: Values in whole sentence reading time are in milliseconds (msec); values for accuracy are proportion correct.

and 180 filler sentences. Structure and plausibility were manipulated in the target sentences, resulting in a 2 (SRC, ORC) × 2 (plausible, implausible) fully-crossed, within-participant and within-item design. Examples of the four sentence conditions are presented in (1). The plausibility of all target sentences (whether adapted or original) was normed for Chinese readers in plausible (e.g., 法官judge 传唤summon证人witness) and implausible (e.g., 证人 witness 传唤summon法官judge) forms on a 1-to-7 Likert scale in simple sentence structures (1= very implausible; 7 = very plausible; plausible: *M* = 6.86, *SD* = 0.25; implausible: *M* = 2.27, *SD* = 0.94; *p* < 0.001; see Appendix B). Materials were assigned pseudo-randomly for each participant across four lists in a Latin-square design, so that each participant saw each item only once.

Apparatus

Participants' eye movements and offline comprehension accuracy were collected using an SR Research EyeLink 1000 remote desktop eye tracker running SR Research Experiment Builder software and sampling at a rate of 1000Hz. We used a chin and forehead rest to stabilize participants' heads. Sentences were displayed in black SimSun font (15pt) on a white background on a computer monitor, which was approximately 70 cm away from the participants. All viewing was binocular, but data were recorded from only the right eye.

Procedure

After providing informed consent, as specified and approved by the university's Office for the Protection of Subjects Internal Review Board (IRB), participants sat in front of the eye tracker's computer monitor. After calibration, they were instructed to read the sentences at a normal reading speed while their eye movements were recorded. After each sentence, a paraphrase verification probe was used to measure comprehension. Readers were asked to decide whether the verification probe was true or false based on the sentence they had just read. The verification probe always contained statements about both the relative clause and the main clause, as illustrated in (3e). The answer was always "True" for the target items, but the number of True and False correct responses was equal across all items in the experiment.

Results

All reading time data points that were three standard deviations from the condition mean and shorter than 80ms were excluded (0.07%). Additional data trimming was performed for each numeric measure in the interest area analyses. **Table 1** provides the descriptive statistics for

the whole sentence reading time and response accuracy. Additionally, we analyzed the first fixation durations, gaze durations, go-past times, total dwell times, regressions-in, regressions-out, and skipping rates in three interest areas: the relativizer *de* (的), the second noun in the RC, and the main clause verb.² These interest areas in ORC (4a) and SRC (4b) are exemplified below.

- (4) a. 法官judge传唤summon的that 证人witness 喜欢 like京剧Beijing opera.
 Relativizer *de* **Second Noun**
 Main Clause Verb
- b. 传唤summon法官judge的that 证人witness 喜欢 like京剧Beijing opera.
 Relativizer *de* **Second Noun**
 Main Clause Verb

The relativizer *de* (roughly similar to 'that') was selected because it is the relative clause marker in Chinese. The second noun was examined because it is the end of the relative clause. Previous self-paced reading studies (e.g., Chen et al., 2008; Gibson & Wu, 2013) have found that the differences between SRC and ORC were salient at the relativizer *de* and the second noun region. With respect to the main clause verb, we should see wrap-up effects as well as how readers integrate the relative clauses into the whole sentence (Staub et al., 2017). In addition to these regions, previous studies also examined reading times on the first noun and the relative clause verb. We did not investigate these two regions because they showed up in the first positions of the sentences, prior to the relevant structure.

First fixation duration refers to the length of the initial fixation on an interest area (word). It is measured from the onset of fixation until the eyes move again in either direction, and is generally taken as a measure of lexical recognition (Rayner, 1998). Gaze duration is the sum of total fixations on the interest area before the eyes move out of it in either direction. It is generally taken as a measure of lexical access (Rayner, 1998). First fixation duration and gaze duration are considered to be "early" measures of eye movements, indexing readers' cognitive processes when they recognize the orthographic properties of a word and access its meaning, respectively. Go-past time, also called regression path duration, is measured from the time a word is first fixated until the eyes move past the word in a forward direction, including all fixations resulting from regressions back into the prior text. Go-past time is taken as a measure of the difficulty of integrating the currently fixated word into the unfolding global sentence representation (structural and semantic)

(Rayner, 1998). Total time is the sum of all fixations in the interest area until the offset of the stimulus, which is also a later measure, and which is generally taken to index total processing time on a given word (Rayner, 1998). Regressions-in is the proportion of trials in which readers regress into the interest area, while regressions-out is the proportion of trials they regress out of the region. Skipping shows when readers skip the interest area. Regressions-in and regressions-out reflect processing and integration difficulties. Skipping of a region suggests that the region was identified to some level of recognition prior to fixation. This is generally assumed to be accomplished through parafoveal processing, and is typical for short, frequent, and/or highly predictable words (cf. Drieghe, Rayner, & Pollatsek, 2005). Regressions-in, regressions-out, and skipping are binomial measures, whereas the rest are continuous measures.

Linear mixed-effects modeling (Bates, Mächler, Bolker, & Walker, 2014) with a maximal random effect structure was applied to analyze the continuous reading data, including whole sentence reading times, as well as first fixation durations, gaze durations, go-past times, and total reading times in each interest area. Plausibility, relative clause structure, and their interaction (reference coded) were added as fixed main effects. Logistic mixed-effects modeling with the same main effects and random effects structure was used to analyze the binomial accuracy data as well as regressions-in, regressions-out, and skipping in the interest areas.

Whole sentence reading time

Table 1 presents the descriptive data for whole sentence reading times and response accuracy. Model results for the whole sentence reading time are provided in **Table 2**. Analyses revealed that plausibility was the only significant

predictor of whole sentence reading times. Readers spent more time processing the implausible sentences than the plausible sentences. Structure did not significantly influence whole sentence reading times. The difference in both syntactic conditions between implausible and plausible conditions was approximately 800 msec.

Interest Area Analyses

Relativizer *de*

Trials in which readers skipped the relativizer *de* were excluded (36%) from all analyses except the skipping data analysis. **Table 3** contains the descriptive statistics for the seven interest area measures for the relativizer *de*. **Table 4** provides linear mixed-effects model outputs for the first fixation durations, gaze durations, go-past times, and total dwell times on the relativizer *de*. Plausibility significantly affected go-past times. Readers had longer go-past times on the relativizer *de* in implausible sentences than in plausible sentences. No variable significantly predicted first fixation durations, gaze durations, or total dwell times on the relativizer *de*.

Table 5 provides logit mixed-effects model outputs for the regressions-in, regressions-out, and skipping on the relativizer *de*. Results showed a marginal interaction of plausibility and structure for regressions-out of the relativizer *de* such that in the ORC condition, readers regressed out the *de* in implausible sentences more than plausible ones; in the SRC condition, readers regressed out of the *de* in the plausible sentences more than the implausible ones. No variable significantly predicted regressions-in or skipping on the relativizer *de*.

The high skipping rate of *de*, a short, frequent, morphosyntactic marker, suggests that it was identified in the parafovea on over one-third of the trials and skipped, irrespective of structure; however, the marginal

Table 2: Estimates, SEs, *t*-Value, and *p*-Value of Fixed Effects for Whole Sentence Reading Time.

Predictor	<i>b</i>	SE	<i>t</i>	<i>p</i>
Intercept	0.10	0.09	1.10	0.28
Plausibility	-0.30	0.05	-5.95	<0.001
Structure	-0.04	0.06	-0.80	0.43
Plausibility × Structure	-0.01	0.06	-0.11	0.92

Note: Significant effects are bolded.

Table 3: Means and Standard Deviations (in Parentheses) for First Fixation Durations, Gaze Durations, Go-past times, Total Dwell Times, Regressions-in, Regressions-out, and Skipping Rates on the Relativizer *de* Interest Area.

	ORC				SRC			
	Implausible		Plausible		Implausible		Plausible	
First Fixation Duration	216.41	(95.16)	229.88	(96.35)	222.94	(95.90)	217.4	(89.69)
Gaze duration	219.96	(98.39)	232.1	(109.01)	238.79	(139.93)	220.14	(91.78)
Go past	484.98	(430.53)	396.75	(109.01)	450.79	(385.02)	417.31	(306.19)
Total Time	300.13	(331.47)	238.6	(271.01)	277.07	(343.54)	205.87	(250.95)
Regressions in	0.44	(0.50)	0.46	(0.50)	0.41	(0.49)	0.33	(0.47)
Regressions out	0.25	(0.43)	0.20	(0.40)	0.23	(0.42)	0.29	(0.46)
Skip	0.64	(0.48)	0.64	(0.48)	0.63	(0.48)	0.62	(0.49)

Table 4: Estimates, SEs, *t*-Values, and *p*-Values of Fixed Effects for Continuous Reading Measures on the Relativizer *de* Interest Area.

Predictor	First fixation duration				Gaze duration				Go-past times				Total dwell time			
	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>
Intercept	-0.06	0.08	-0.8	0.43	-0.06	0.08	-0.84	0.41	0.09	0.08	1.19	0.24	0.04	0.09	0.47	0.64
Plausibility	0.15	0.1	1.53	0.13	0.11	0.1	1.19	0.24	-0.22	0.09	-2.36	<0.05	-0.11	0.11	-0.96	0.34
Structure	0.07	0.1	0.7	0.49	0.11	0.1	1.12	0.27	-0.07	0.1	-0.73	0.47	-0.02	0.1	-0.22	0.83
Plausibility × Structure	-0.22	0.16	-1.35	0.19	-0.23	0.14	-1.67	0.10	0.19	0.13	1.44	0.15	-0.1	0.15	-0.67	0.50

Table 5: Estimates, SEs, *z*-Values, and *p*-Values of Fixed Effects for Binomial Measures on the Relativizer *de* Interest Area.

Predictor	Regressions in				Regressions out				Skip			
	<i>b</i>	SE	<i>z</i>	<i>p</i>	<i>b</i>	SE	<i>z</i>	<i>p</i>	<i>b</i>	SE	<i>z</i>	<i>p</i>
Intercept	-0.32	0.18	-1.75	0.08	-1.25	0.23	-5.49	<-0.01	0.6	0.13	4.51	<-0.01
Plausibility	0.09	0.21	0.46	0.65	-0.35	0.26	-1.33	0.18	-0.01	0.18	-0.04	0.97
Structure	-0.16	0.21	-0.78	0.44	-0.12	0.26	-0.47	0.64	-0.04	0.18	-0.25	0.81
Plausibility × Structure	-0.41	0.3	-1.37	0.17	0.71	0.37	1.90	0.06	-0.04	0.25	-0.17	0.86

Table 6: Means in Msec and Standard Deviations (in Parentheses) for First Fixation Durations, Gaze Durations, Go-past times, Total Dwell Times, Regressions-in, Regressions-out, and Skipping Rates on the Second Noun.

	ORC				SRC			
	Implausible		Plausible		Implausible		Plausible	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
First Fixation Duration	234.89	(93.33)	231.76	(90.67)	237.38	(86.90)	240.12	(102.68)
Gaze duration	273.37	(140.20)	278.16	(166.87)	290.89	(165.82)	278.53	(154.78)
Go past	694.19	(715.61)	502.67	(427.23)	694.69	(691.98)	530.75	(427.56)
Total Time	709.99	(605.14)	538.16	(495.96)	656.11	(473.61)	499.44	(408.50)
Regressions in	0.32	(0.47)	0.33	(0.47)	0.29	(0.45)	0.25	(0.43)
Regressions out	0.37	(0.48)	0.28	(0.45)	0.37	(0.48)	0.33	(0.47)
Skip	0.15	(0.36)	0.23	(0.42)	0.17	(0.37)	0.20	(0.40)

interaction suggests that readers were aware of its importance to the ongoing parse, and its relationship to the plausibility of the thematic roles of the surrounding noun phrases. Because *de* is so short, it is almost certain that both the upcoming noun and the previous noun were in the parafovea while *de* was fixated (Rayner, 1998). This rather wide and complex collection of information sources would explain why all of the fixation measures on *de*, especially the measures associated with integration into the unfolding parse (total time, go-past time), were relatively large. A precise explanation of how the various complexities of both SRCs and ORCs drive integration of *de* (as indexed here by the marginally significant interaction) requires further investigation.

Second Noun

Trials in which readers skipped the second noun were removed (7.5%) from analyses except for the skipping data analyses. **Table 6** contains the descriptive statistics for the seven interest area measures for the second noun. **Table 7** presents the model results for continuous measures on the second noun. Results showed that plausibility was a

significant predictor for go-past times and total reading times on the second noun, with implausible sentences displaying inflated times in all measures. It appears that the level of disruption observed at the relativizer *de* had been resolved by the time the second noun was fully fixated, and plausibility alone dictated the relatively long later measures associated with integration into the unfolding parse (total time, go-past time).

Plausibility also significantly modulated readers' regressions-out and skipping on the second noun; they made fewer regressions out of the second noun and skipped it more in plausible sentences compared to implausible sentences. **Table 8** presents the model results for the binomial measures on the second noun.

Main Clause Verb

Trials in which readers skipped the *Main Clause Verb* were removed (4.1%) from analyses except for the skipping data analyses. **Table 9** contains the descriptive statistics for the seven interest area measures for the main clause verb. **Table 10** presents the model results for the continuous measures on the main clause verb. Analysis of go-past

Table 7: Estimates, SEs, *t*-Values, and *p*-Values of Fixed Effects for Continuous Reading Measures on the Second Noun Interest Area.

Predictor	First fixation duration				Gaze duration				Go-past times				Total dwell time			
	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>
Intercept	-0.01	0.06	-0.14	0.89	-0.03	0.06	-0.43	0.67	0.09	0.08	1.13	0.26	0.15	0.11	1.29	0.20
Plausibility	-0.04	0.08	-0.45	0.65	-0.02	0.09	-0.24	0.82	-0.27	0.10	-2.85	<0.05	-0.29	0.12	-2.44	<0.05
Structure	0.03	0.08	0.43	0.67	0.09	0.08	1.09	0.28	0.05	0.09	0.55	0.59	-0.01	0.1	-0.06	0.95
Plausibility × Structure	0.04	0.11	0.31	0.75	-0.07	0.12	-0.58	0.57	0.05	0.12	0.38	0.70	-0.11	0.13	-0.82	0.41

Table 8: Estimates, SEs, *z*-Values, and *p*-Values of Fixed Effects for Binomial Measures on the Second Noun Interest Area.

Predictor	Regressions in				Regressions out				Skip			
	<i>b</i>	SE	<i>z</i>	<i>p</i>	<i>b</i>	SE	<i>z</i>	<i>p</i>	<i>b</i>	SE	<i>z</i>	<i>p</i>
Intercept	0.89	0.2	4.38	<0.01	-0.59	0.16	-3.63	<0.01	-2.06	0.25	-8.12	<0.01
Plausibility	0.03	0.18	0.17	0.87	-0.44	0.19	-2.3	<0.05	0.58	0.24	2.4	<0.05
Structure	0.17	0.19	0.91	0.37	0.01	0.18	0.07	0.94	0.1	0.25	0.39	0.69
Plausibility × Structure	0.24	0.27	0.91	0.36	0.26	0.27	0.95	0.34	-0.34	0.34	-1	0.32

Table 9: Means and standard deviations (in parentheses) for first fixation durations, gaze durations, go-past times, total dwell times, regressions-in, regressions-out, and skipping rates on the main clause verb.

	ORC				SRC			
	Implausible		Plausible		Implausible		Plausible	
First Fixation Duration	249.18	(113.99)	234.48	(91.20)	248.07	(109.63)	243.79	(101.77)
Gaze duration	435.63	(335.60)	451.37	(337.85)	445.88	(345.42)	493.73	(424.12)
Go past	2368.4	(2110.7)	1881.19	(1604.63)	2264.37	(2055.8)	1648.78	(1372.46)
Total Time	952.54	(753.96)	875.54	(749.26)	911.89	(852.59)	814.18	(674.92)
Regressions in	0.01	(0.09)	0.02	(0.14)	0.01	(0.08)	0.01	(0.11)
Regressions out	0.83	(0.38)	0.82	(0.39)	0.85	(0.36)	0.80	(0.40)
Skip	0.02	(0.14)	0.05	(0.22)	0.05	(0.21)	0.03	(0.17)

times showed a significant plausibility effect: plausible sentences led to shorter go-past times. There was also a marginal effect of structure on total time, with SRC main verbs exhibiting approximately 50 msec faster reading total times than ORC main verbs.

Table 11 shows model results for the continuous measures on the Main Clause Verb. No variable significantly affected the early measures, regressions-in, or regressions-out of the main clause verb. With respect to skipping, plausible sentences displayed a marginally higher skipping rate over implausible sentences, and the interaction between plausibility and structure was significant. What seems to be driving the interaction is that readers tended to skip the main clause verb in implausible SRCs more than in implausible ORCs; however, all skipping rates were =/<5%, so it is not clear how robust this effect is.

Comprehension accuracy

Table 12 presents the model outputs for comprehension accuracy analyses. The results revealed that both plausibility and structure significantly affected the

response accuracy to the paraphrase verification probes. Readers comprehended plausible sentences more accurately than implausible sentences. They also showed higher comprehension accuracy for SRCs than ORCs. There was no interaction between plausibility and structure.

Discussion

This study examined the online processing and offline comprehension of relative clauses in Chinese (Mandarin). Previous research on Chinese relative clause processing has returned three basic results. As detailed in the introduction, some studies claim to show that subject-relative clauses (SRCs) are easier to process and comprehend, as predicted by experience-based processing theories (e.g., Jäger et al., 2015; Lin & Bever, 2006; Vasishth et al., 2013). Other studies claim to show that object-relative clauses (ORCs) are easier, as predicted by working-memory-based theories (e.g., Chen et al., 2008; Gibson & Wu, 2011; Hsiao & Gibson, 2003). A smaller third set of findings suggests that there is no clearly preferred relative clause structure in Chinese (Chen et al.,

Table 10: Estimates, SEs, *t*-Values, and *p*-Value of Fixed Effects for Continuous Reading Measures on the Main Clause Verb Interest Area.

Predictor	First fixation duration				Gaze duration				Go-past times				Total dwell time			
	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>	<i>b</i>	SE	<i>t</i>	<i>p</i>
Intercept	0.04	0.07	0.51	0.62	-0.05	0.08	-0.66	0.52	0.13	0.13	0.97	0.34	0.05	0.12	0.45	0.66
Plausibility	-0.1	0.09	-1.12	0.27	0.03	0.09	0.33	0.74	-0.25	0.12	-2.08	<0.05	-0.11	0.11	-0.98	0.33
Structure	-0.01	0.08	-0.06	0.95	0.02	0.09	0.27	0.79	-0.03	0.11	-0.32	0.75	-0.01	0.1	-0.06	0.95
Plausibility × Structure	0.08	0.12	0.67	0.51	0.05	0.13	0.43	0.67	-0.05	0.15	-0.32	0.75	-0.09	0.15	-0.58	0.56

Table 11: Estimates, SEs, *z*-Values, and *p*-Values of Fixed Effects for Binomial Measures on the Main Clause Verb Interest Area.

Predictor	Regressions in				Regressions out				Skip			
	<i>b</i>	SE	<i>z</i>	<i>p</i>	<i>b</i>	SE	<i>z</i>	<i>p</i>	<i>b</i>	SE	<i>z</i>	<i>p</i>
Intercept	-11.66	2.66	4.38	<0.01	2.06	0.29	7.20	<0.001	-5.16	0.63	-8.13	<0.01
Plausibility	0.58	2.64	0.22	0.83	0.16	0.28	0.57	0.57	1.07	0.57	1.89	0.06
Structure	-0.51	3.41	0.15	0.88	0.19	0.29	0.66	0.51	0.90	0.58	1.56	0.12
Plausibility × Structure	0.05	4.27	0.01	0.99	0.25	0.40	0.62	0.54	-1.58	0.79	-2.01	<0.05

Table 12: Estimates, SEs, *z*-Values, and *p*-Values of Fixed Effects for Response Accuracy.

Predictor	<i>b</i>	SE	<i>z</i>	<i>p</i>
Intercept	1.28	0.36	3.52	<0.001
Plausibility	2.46	0.42	5.79	<0.001
Structure	0.95	0.41	2.34	<0.05
Plausibility × Structure	0.10	0.63	0.17	0.87

Note: The significant effects are in bold.

2006). This lack of preference is explained by appealing to the competing theories of Chinese relative clause processing. Specifically, there is reason to believe that *both* structures present processing challenges of different sorts, and there is no *a priori* reason to assume that one or some combination of these challenges should produce a clearly “preferred” structure. Furthermore, a survey of the existent literature reveals that although several previous studies have controlled for plausibility, only two manipulated plausibility (operationalized as animacy) as an independent variable. In those studies (Chen et al., 2006; Wu et al., 2012), relative clause preferences were made to flip depending on the relative plausibility of the nouns. Thus, we applied a processing principle from Good-Enough Processing (cf. Christianson, 2016; Ferreira et al., 2003) to explore Chinese relative clause processing patterns. Specifically, Good-Enough Processing predicts that heuristic (i.e., non-syntactic) information, such as plausibility, should exert more influence on the online processing and offline interpretation of syntactic structures that are more difficult to parse, compared to those that are easier to parse.

In order to test the importance of plausibility in Chinese relative clause processing and comprehension, we crossed

SRC and ORC structures with plausible and implausible thematic role assignments in Chinese in a fully-crossed 2 × 2 design. We hypothesized that if Chinese does in fact have a strongly preferred relative clause structure, then the addition of implausible thematic role assignments to the more difficult (dispreferred) structure should produce an additive interaction such that the implausible, dispreferred structure becomes extremely difficult to process and comprehend. This prediction derives from several studies in English relative clause processing and passive sentence processing resulting in just such additive interactions (e.g., Christianson et al., 2010; Ferreira, 2003; Lim & Christianson, 2013b; Zhou & Christianson, 2016a). Alternatively, if Chinese relative clauses are relatively equivalent in their difficulty, perhaps for different reasons, we expected to observe a significant main effect for plausibility alone (Chen et al., 2006; cf. Stoops, Luke, & Christianson, 2014, on the importance of animacy information in Russian processing). This approach represents a novel contribution to the current debate in the literature as to whether SRCs or ORCs are easier to process in Chinese (Chen et al., 2006; Gibson & Wu, 2011; Jäger et al., 2015; Vasishth et al., 2013).

The results showed that plausibility alone significantly affected whole sentence reading times. Readers spent less time reading overall when processing plausible sentences compared to implausible sentences, regardless of relative clause structure. The results are consistent with those of Chen et al. (2006), who manipulated the animacy of the nouns in Chinese relative clauses (an indirect manipulation of plausibility). Chen also found no significant effects of structure, only a significant effect of animacy, and concluded that, “...semantic cues play a more significant role than syntactic cues in Chinese subjects’ sentence processing...” (Chen et al., 2006, p. 8). Importantly, the results are inconsistent with both predictions generated

by experience-based processing theories and working-memory-based theories, because Chinese readers did not show a reading time advantage for either of the relative clause structures.

Interest area analyses generally yielded data consistent with the whole sentence reading results. There were no effects of plausibility or structure in early reading measures, first fixation durations and gaze durations, on the relativizer (relative clause marker) *de*, the second noun, or the main verb. However, there were clear signs of processing disruption on *de*, as syntactic structure and semantic plausibility information appeared to influence eye movement measures in a not altogether clear pattern. Overall, though, readers seemed relatively insensitive to structural considerations in all IAs. There was no significant main effect of structure in early or late measures, in any IA analysis, contrary to syntax-first, two-stage models (e.g., Rayner, Carlson, & Frazier, 1983). Rather, readers seemed more sensitive to plausibility information largely with respect to integration of the contents of the various IAs, once they had been lexically identified and accessed, as suggested by significant plausibility effects on go-past times on *de*, the second noun, and the main clause verb. Readers also appeared to rely primarily on plausibility information when reanalyzing and integrating the words into the structure and deriving an interpretation. Thus, the IAs results also failed to support experience-based processing theories or working-memory-based theories of Chinese RC processing. Rather, the IA data suggest that readers employed a good-enough processing strategy to follow a semantic route to comprehension.

Both plausibility and structure modulated readers' offline comprehension; however, these effects were notably non-interactive. Plausible sentences were easier to comprehend than implausible ones. SRCs were easier to comprehend than ORCs. This result provides some evidence in support of experience-based processing accounts of Chinese relative clause processing (e.g., Jäger et al., 2015; Lin & Bever, 2006; Vasishth et al., 2013), as SRCs are more frequent. However, in the previous experience-based processing studies, none of the comprehension differences reached significance in the three self-paced reading experiments in Vasishth et al. (2013). Similarly, Jäger et al. (2015) failed to find any comprehension accuracy differences between ORCs and SRCs in either their self-paced reading experiment or their eye-tracking experiment. This offline preference for SRCs in the current study was notably slight, and not consistently observed in the online processing time measures. Taken together, these results suggest that any *a priori* assumptions about the strong canonicity of either of the relative clause structures in Chinese should be viewed with caution.

Finally, these results are consistent with observations in other languages regarding the way in which competing syntactic and semantic information is integrated. Within the Good Enough framework (Christianson, 2016; Ferreira et al., 2002), it is proposed that input is subjected to "algorithmic" syntactic parsing as well as "heuristic"

processing. The syntactic parse is fragile, however (Ferreira, 2003; Sachs, 1967), and in cases in which the output from the two processing routes conflict, it may be overwhelmed by the output of the heuristic route. Such a system predicts that structures that are non-canonical – less frequent or otherwise more costly to compute – should be more vulnerable to intrusions from heuristic processing. As a result, implausible sentences are more likely to end up misinterpreted with their thematic roles reversed, thereby generating a plausible but syntactically unfaithful interpretation. In English, ORCs are also more susceptible to these thematic role reversals (e.g., Zhou & Christianson, 2016a). Unlike the case of English relative clauses, however, plausibility and structure did not interact in the current Chinese relative clauses in online or offline measures. The significantly lower accuracy rates for offline comprehension were numerically in the expected direction, though: lower accuracy in implausible sentences, and lowest accuracy in implausible ORCs, in which both meaning and, according to some of the results here (mostly marginal effects in the online measures), structure conspire to flip the thematic roles. Of particular interest to future research is the broader pattern observed here, however, suggesting that semantic plausibility appears to have a much stronger effect on online processing measures than do structural considerations, and that these factors do not interact in Chinese relative clause structure. If this pattern were to be replicated, and expanded to other Chinese structures, it would suggest that Chinese language processing might rely on "good-enough" routines more so than English or, say, French, in which both structure and plausibility have been shown to have early significant and interactive online effects (Lassotta et al., 2016; Zhou & Christianson, 2016a, b). If this observation is borne out in future research, it is possible Chinese susceptibility to good-enough processing could be related to the extreme paucity of inflectional morphology in the language and occasional indeterminacy with respect to word boundaries (Lin, Anderson, Ku, Christianson, & Packard, 2011) in the unspaced orthography, which may induce further processing load in complex structures.

In sum, our reading time data demonstrated that readers showed no preference for either Chinese relative clause structure, which is inconsistent with the predictions from experience-based processing accounts and working memory-based accounts. Although there was a slightly advantage in comprehending SRCs, supporting experience-based accounts, this SRC comprehension advantage has not been consistently observed in previous studies and may not be reliable. Both online reading and offline comprehension data showed strong plausibility effects, suggesting that Chinese readers employ a heuristic, good-enough processing strategy when processing relative clauses manipulated to pit plausibly against structure. This study thus provides a novel explanation for the debate in the Chinese relative clause processing literature, and it is the first to test Good-Enough Processing theory in Chinese.

Data Accessibility Statement

All the stimuli can be found in the Appendix A. The plausibility norming results for all the stimuli can be found in Appendix B. Data available upon request.

Notes

¹ Fifty-six target sentences were created or adapted/translated from previous studies and normed for plausibility. The plausibility norming score was not significant for one of them (No. 15, see Appendix B), so we ended up with 55 target sentences in the final analysis.

² For ease of exposition, we use the term “main clause verb;” however, in some of the sentences, the word immediately following the RC – the disambiguating word – was not a verb, as copular verbs are omitted in Chinese. For example, the word “were” is omitted in Chinese in this example:

1. 嘉宾participant邀请invited的de组织者organizer非常very忙busy。
The organizers that the participants recruited were very busy.

Acknowledgements

Earlier versions of this work have been presented at the CUNY Human Sentence Processing Conference in 2015 and at the Psychonomic Society's 57th Annual Meeting in 2016. We would like to thank EdPsych Psycholinguistics Lab members, Aubrey Sumaydeng, Cassandra Phelps, Laeh Ragans, and Nayoung Kim for their valuable help in collecting data.

Funding Information

The project was funded by the research grant awarded to Kiel Christianson and Peiyun Zhou from the Confucius Institute at University of Illinois as well as the Beckman Institute Graduate Fellowship to Peiyun Zhou.

Competing Interests

The authors have no competing interests to declare.

References

- Bates, D., Mächler, M., Bolker, B., & Walker, S.** (2014). Fitting linear mixed-effects models using lme4. *arXiv preprint arXiv:1406.5823*.
- Chen, B., Ning, A., Bi, H., & Dunlap, S.** (2008). Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica, 129*(1), 61–65. DOI: <https://doi.org/10.1016/j.actpsy.2008.04.005>
- Chen, C. H.** (2006). Chinese L2 learners' use of structural and lexical information in processing English subject and object relative clauses. *Cahiers linguistiques d'Ottawa/Ottawa Papers in Linguistics, 34*, 11–23.
- Christianson, K.** (2016). When language comprehension goes wrong for the right reasons: Good-enough, underspecified, or shallow language processing. *The Quarterly Journal of Experimental Psychology, 69*(5), 817–828. DOI: <https://doi.org/10.1080/17470218.2015.1134603>
- Christianson, K., Hollingworth, A., Halliwell, J. F., & Ferreira, F.** (2001). Thematic roles assigned along the garden path linger. *Cognitive Psychology, 42*(4), 368–407. DOI: <https://doi.org/10.1006/cogp.2001.0752>
- Christianson, K., & Luke, S. G.** (2011). Context strengthens initial misinterpretations of text. *Scientific Studies of Reading, 15*(2), 136–166. DOI: <https://doi.org/10.1080/10888431003636787>
- Christianson, K., Williams, C. C., Zacks, R. T., & Ferreira, F.** (2006). Younger and older adults' “good-enough” interpretations of garden-path sentences. *Discourse Processes, 42*(2), 205–238. DOI: https://doi.org/10.1207/s15326950dp4202_6
- Crain, S., & Steedman, M.** (1985). On not being led up the garden path: The use of context by the psychological syntax processor. In: Dowty, D. R., Karttunen, L., & Zwicky, A. M. (eds.), *Natural language parsing*, 320–358. Cambridge University Press. DOI: <https://doi.org/10.1017/CBO9780511597855.011>
- Den Ouden, D. B., Dickey, M. W., Anderson, C., & Christianson, K.** (2016). Neural correlates of early-closure garden-path processing: Effects of prosody and plausibility. *Quarterly Journal of Experimental Psychology, 69*, 926–949. DOI: <https://doi.org/10.1080/17470218.2015.1028416>
- Drieghe, D., Rayner, K., & Pollatsek, A.** (2005). Eye movements and word skipping during reading revisited. *Journal of Experimental Psychology: Human Perception and Performance, 31*(5), 954. DOI: <https://doi.org/10.1037/0096-1523.31.5.954>
- Ferreira, F.** (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology, 47*(2), 164–203. DOI: [https://doi.org/10.1016/S0010-0285\(03\)00005-7](https://doi.org/10.1016/S0010-0285(03)00005-7)
- Ferreira, F., Bailey, K. G., & Ferraro, V.** (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science, 11*(1), 11–15. DOI: <https://doi.org/10.1111/1467-8721.00158>
- Ferreira, F., Christianson, K., & Hollingworth, A.** (2001). Misinterpretations of garden-path sentences: Implications for models of sentence processing and reanalysis. *Journal of Psycholinguistic Research, 30*(1), 3–20. DOI: <https://doi.org/10.1023/A:1005290706460>
- Ferreira, F., & Patson, N. D.** (2007). The ‘good enough’ approach to language comprehension. *Language and Linguistics Compass, 1*(1–2), 71–83. DOI: <https://doi.org/10.1111/j.1749-818X.2007.00007.x>
- Frazier, L., & Clifton, C., Jr.** (1996). *Construal*. MIT Press.
- Gennari, S. P., & MacDonald, M. C.** (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language, 58*(2), 161–187. DOI: <https://doi.org/10.1016/j.jml.2007.07.004>
- Gibson, E.** (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition, 68*(1), 1–76. DOI: [https://doi.org/10.1016/S0010-0277\(98\)00034-1](https://doi.org/10.1016/S0010-0277(98)00034-1)
- Gibson, E., & Wu, H. H. I.** (2013). Processing Chinese relative clauses in context. *Language and Cognitive*

- Processes*, 28(1–2), 125–155. DOI: <https://doi.org/10.1080/01690965.2010.536656>
- Hsiao, F., & Gibson, E.** (2003). Processing relative clauses in Chinese. *Cognition*, 90(1), 3–27. DOI: [https://doi.org/10.1016/S0010-0277\(03\)00124-0](https://doi.org/10.1016/S0010-0277(03)00124-0)
- Hsiao, Y., & MacDonald, M. C.** (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Frontiers in Psychology*, 4, 767. DOI: <https://doi.org/10.3389/fpsyg.2013.00767>
- Hussey, E. K., Ward, N., Christianson, K., & Kramer, A. F.** (2015). Language and memory improvements following tDCS of executive-control brain regions. *PLoS One*. Nov. 3, 2015. DOI: <https://doi.org/10.1371/journal.pone.0141417>
- Jacob, G., & Felser, C.** (2016). Reanalysis and semantic persistence in native and non-native garden-path recovery. *The Quarterly Journal of Experimental Psychology*, 69(5), 907–925. DOI: <https://doi.org/10.1080/17470218.2014.984231>
- Jäger, L., Chen, Z., Li, Q., Lin, C. J. C., & Vasishth, S.** (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language*, 79, 97–120. DOI: <https://doi.org/10.1016/j.jml.2014.10.005>
- Lassotta, R., Omaki, A., & Franck, J.** (2016). Developmental changes in misinterpretation of garden-path wh-questions in French. *The Quarterly Journal of Experimental Psychology*, 69(5), 829–854. DOI: <https://doi.org/10.1080/17470218.2015.1054845>
- Levy, R.** (2008). Expectation-based syntactic comprehension. *Cognition*, 106(3), 1126–1177. DOI: <https://doi.org/10.1016/j.cognition.2007.05.006>
- Lim, J.-H. & Christianson, K.** (2013a). Integrating meaning and structure in L1-L2 and L2-L1 translations. *Second Language Acquisition*, 29, 233–256. DOI: <https://doi.org/10.1177/0267658312462019>
- Lim, J. H., & Christianson, K.** (2013b). Second language sentence processing in reading for comprehension and translation. *Bilingualism: Language and Cognition*, 16(03), 518–537. DOI: <https://doi.org/10.1017/S1366728912000351>
- Lin, C. J. C., & Bever, T. G.** (2006). Subject preference in the processing of relative clauses in Chinese. In: *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 254–260. Somerville, MA: Cascadilla Proceedings Project.
- Lin, T.-J., Anderson, R. C., Ku, Y.-M., Christianson, K., & Packard, J.** (2011). Chinese children's concept of word. *Writing Systems Research*, 3, 41–57. DOI: <https://doi.org/10.1093/wsr/wsr007>
- Patson, N. D., Darowski, E. S., Moon, N., & Ferreira, F.** (2009). Lingering misinterpretations in garden-path sentences: Evidence from a paraphrasing task. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 35(1), 280–285. DOI: <https://doi.org/10.1037/a0014276>
- Rayner, K.** (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372. DOI: <https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K., Carlson, M., & Frazier, L.** (1983). The interaction of syntax and semantics during sentence processing: Eye movements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22(3), 358–374. DOI: [https://doi.org/10.1016/S0022-5371\(83\)90236-0](https://doi.org/10.1016/S0022-5371(83)90236-0)
- Sachs, J. S.** (1967). Recopition memory for syntactic and semantic aspects of connected discourse. *Attention, Perception, & Psychophysics*, 2(9), 437–442. DOI: <https://doi.org/10.3758/BF03208784>
- Staub, A., Dillon, B., & Clifton, C.** (2017). The Matrix Verb as a Source of Comprehension Difficulty in Object Relative Sentences. *Cognitive Science*, 41(S6), 1353–1376. DOI: <https://doi.org/10.1111/cogs.12448>
- Stoops, A., Luke, S. G., & Christianson, K.** (2014). Animacy information outweighs morphological cues in Russian. *Language, Cognition, & Neuroscience*, 29, 584–604. DOI: <https://doi.org/10.1080/01690965.2013.813560>
- Townsend, D. J., & Bever, T. G.** (2001). *Sentence comprehension: The integration of habits and rules*, 1950. Cambridge, MA: MIT Press.
- Vasishth, S., Chen, Z., Li, Q., & Guo, G.-L.** (2013). Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS One*, 8(10), e77006. DOI: <https://doi.org/10.1371/journal.pone.0077006>
- Wells, J. B., Christiansen, M. H., Race, D. S., Acheson, D. J., & MacDonald, M. C.** (2009). Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology*, 58(2), 250–271. DOI: <https://doi.org/10.1016/j.cogpsych.2008.08.002>
- Wu, F.** (2009). *Factors affecting relative clause processing in Mandarin*. University of Southern California.
- Wu, F., Kaiser, E., & Andersen, E.** (2010). Subject preference, head animacy and lexical cues: A corpus study of relative clauses in Chinese. In: *Processing and producing head-final structures*, 173–193. Springer Netherlands. DOI: https://doi.org/10.1007/978-90-481-9213-7_9
- Wu, F., Kaiser, E., & Andersen, E.** (2012). Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes*, 27(10), 1489–1524. DOI: <https://doi.org/10.1080/01690965.2011.614423>
- Zhou, P.** (2017). *Auditory perceptual simulation during silent reading: Effects on language processing and comprehension* (Doctoral dissertation, University of Illinois at Urbana-Champaign).
- Zhou, P., & Christianson, K.** (2016a). I “hear” what you’re “saying”: Auditory perceptual simulation, reading speed, and reading comprehension. *The Quarterly Journal of Experimental Psychology*, 69(5), 972–995. DOI: <https://doi.org/10.1080/17470218.2015.1018282>
- Zhou, P., & Christianson, K.** (2016b). Auditory perceptual simulation: Simulating speech rates or accents? *Acta Psychologica*, 168, 85–90. DOI: <https://doi.org/10.1016/j.actpsy.2016.04.005>

Appendix A

Examples of the target sentences in the implausible ORC condition.

1. 羊看护的狗看起来毛茸茸的。
The dog that the sheep herded was very furry.
2. 嘉宾邀请的组织者非常忙。
The organizers that the participants recruited were very busy.
3. 代言人雇用的企业家需要宣传。
The entrepreneur that the publicist employed needed exposure.
4. 模特打扮的设计师很有创新精神。
The designer that the model dressed was innovative.
5. 蝴蝶展出的标本师很漂亮。
The taxidermist that the butterfly exhibited was beautiful.
6. 嫌疑犯调查的侦探是个素食者。
The detective that the suspect investigated was a vegetarian.
7. 暴力受害者安慰的心理师有条蓝色的围巾。
The psychiatrist that the victim comforted had a blue scarf.
8. 双胞胎抚养的父母曾经住在纽约。
The parents that the twins raised lived in New York.
9. 花匠雇佣的业主很热爱园艺。
The homeowner that the gardener paid loved the garden.
10. 抢劫犯代理的律师很擅长跳舞。
The lawyer that the robber represented was good at dancing.
11. 运动员处罚的裁判是阿拉伯人。
The referee that the athlete penalized was from Arabia.
12. 士兵指挥的将军打赢了这场战役。
The general that the private commanded won the battle.
13. 师傅崇拜的学徒学得非常快。
The apprentice that the master worshipped learned very fast.
14. 顾客打劫的强盗扛着一个大袋子。
The mugger that the shopper robbed carried a big bag.
15. 模特画的画家是四川人。
The painter that the model sketched finished drawing quickly.
16. 孩子惩罚的父母跑下楼去了。
The parents that the child punished went to church.
17. 市民拘捕的警察长得很帅。
The policeman that the citizen arrested was handsome.
18. 员工解雇的老板在接电话。
The boss that the worker fired was answering a phone call.
19. 业余爱好者指导的专家很有耐心。
The expert that the amateur instructed was patient.
20. 证人传唤的法官爱好京剧。
The judge that the witness summoned liked Beijing Opera.
21. 那只老鼠追的黄猫跑得非常快。
The cat that the mouse chased was fast.
22. 明星崇拜的粉丝穿了条红裙子。
The fan that the star idolized wore a red shirt.
23. 少年训斥的父亲坐在客厅里。
The father that the teenager scolded was in the living room.
24. 猎物跟踪的老虎留下了脚印。
The tiger that the prey followed was hungry.
25. 球员训练的教练赢过两次冠军。
The coach that the player trained won the championship twice.
26. 苍蝇吃掉的青蛙很敏捷。
The frog that the fly ate was agile.
27. 人质劫持的恐怖分子藏在一栋楼里。
The terrorist that the hostage held was spotted in the building.
28. 歌手跟踪的粉丝很喜欢上网。
The fan that the singer stalked loved to wear big sunglasses.
29. 老师服从的小孩暑假玩得很开心。
The kids that the teacher obeyed enjoyed the summer break.
30. 客人保护的保镖接受过专业培训。
The guard that the guest secured received professional training.
31. 孤儿监护的修女住在小村庄里。
The nuns that the orphans supervised lived in a small village.
32. 大臣赐婚的皇帝有两个女儿。
The King that the peasant awarded has two daughters.
33. 观众娱乐的相声演员非常受欢迎。
The comedian that the spectator entertained was popular.
34. 初学者救起的救生员打电话叫了救护车。
The lifeguard that the beginner rescued called the ambulance.
35. 学生辅导的家教解决了很难的数学题。
The tutor that the student taught solved the math problem.
36. 猪宰了的屠夫满身是汗。
The butcher that the pig slaughtered was sweaty.
37. 游客欺骗的奸商看起来三十出头。
The greedy businessman that the tourist cheated is around 30.
38. 总理协助的秘书在会议室里。
The secretary that the President assisted was in a conference room.
39. 女孩绑架的罪犯上了央视头条。
The criminal that the girl kidnapped was on CNN news.
40. 学生测试的老师在教室里。
The teacher that the student quizzed was in the classroom.
41. 麋鹿射中的猎人藏在深山里。
The hunter that the moose shot hid in the mountains.

- 42. 动物治愈的兽医喝了很多水。
The veterinarian that the animal cured drank a lot of water.
- 43. 狗遛的男人在公园里。
The man that the pet walked was in the park.
- 44. 虫子吃的鸟儿非常小。
The bird that the worm ate was small.
- 45. 嫌犯监视的调查员非常细心。
The investigator that the suspect monitored was meticulous.
- 46. 选民欺骗的政客喜欢韩国食物。
The politician that the voter deceived liked Korean food.
- 47. 犯人看守的狱警在吃午饭。
The guard that the prisoner restrained was having lunch.
- 48. 狐狸逮住的猎户住在山里。
The trapper that the fox seized lived in the mountain.
- 49. 猫喂养的主人睡在沙发上。
The owner that the cat fed was sitting on a sofa.
- 50. 病人治疗的医生带着条粉晶手链。
The doctor that the patient treated wore a pink bracelet.
- 51. 人咬的狗正在院子里。
The dog that the man bit was in the yard.
- 52. 船员带领的船长非常有冒险精神。
The captain that the crew led was adventurous.
- 53. 小猴子欺负的大猩猩在吃东西。
The gorilla that the monkey intimidated was eating.
- 54. 演员采访的记者在咖啡店。
The reporter that the actress interviewed was at the coffee shop.
- 55. 盗贼追捕的警察很会开车。
The cop that the thief pursued was driving a car.

Appendix B

Averaged plausibility norming score for each item from 22 native Chinese speakers (1 = very implausible; 7 = very plausible).

There is no tense in Chinese, so the English translation does not have tense.

Item ID	Implausible	Norming Score	Plausible	Norming Score	Wilcoxon signed-rank test <i>p</i> value
1	羊看护狗 Sheep herd dog	1.50	狗看护羊 Dog herd sheep	6.55	<0.001
2	嘉宾邀请组织者 Organizer recruit participant	2.91	组织者邀请嘉宾 Participant recruit organizer	7.00	<0.001
3	代言人雇用企业家 Publicist employ entrepreneur	2.20	企业家雇用代言人 Entrepreneur employ publicist	6.91	<0.001
4	模特打扮设计师 Model dress designer	1.82	设计师打扮模特 Designer dress model	7.00	<0.001
5	蝴蝶展出标本师 Butterfly exhibit taxidermist	1.00	标本师展出蝴蝶 Taxidermist exhibit butterfly	6.73	<0.001
6	嫌疑犯调查侦探 Suspect investigate detective	2.73	侦探调查嫌疑犯 Detective investigate suspect	7.00	<0.001
7	暴力受害者安慰心理师 Victim comfort psychiatrist	1.80	心理师安慰暴力受害者 Psychiatrist comfort victim	6.91	<0.001
8	双胞胎抚养父母 Twins raise parents	4.36	父母抚养双胞胎 Parents raise Twins	7.00	<0.001
9	花匠雇佣业主 Gardener pay homeowner	2.40	业主雇佣花匠 Homeowner pay gardener	6.64	<0.001
10	抢劫犯代理律师 Robber represent Lawyer	1.73	律师代理抢劫犯 Lawyer represent robber	6.60	<0.001
11	运动员处罚裁判 Athlete penalize referee	1.40	裁判处罚运动员 Referee penalize athlete	7.00	<0.001
12	士兵指挥将军 Private command general	1.27	将军指挥士兵 General command private	7.00	<0.001
13	师傅崇拜学徒 Master worship apprentice	3.80	学徒崇拜师傅 Apprentice worship master	6.27	<0.001
14	顾客打劫强盗 Shopper rob mugger	1.64	强盗打劫顾客 Mugger rob shopper	7.00	<0.001
15	公司审查检验员 Company inspect inspector	5.60	检验员审查公司 Inspector inspect company	5.45	0.91
16	模特画画家 Model Sketch painter	1.82	画家画模特 Painter sketch model	7.00	<0.001

(contd.)

Item ID	Implausible	Norming Score	Plausible	Norming Score	Wilcoxon signed-rank test <i>p</i> value
17	孩子惩罚父母 Child punish parents	3.50	父母惩罚孩子 Parents punish child	6.45	<0.001
18	市民拘捕警察 Citizen arrest policeman	1.73	警察拘捕市民 Policeman arrest citizen	6.90	<0.001
19	员工解雇老板 Worker fire boss	4.00	老板解雇员工 Boss fire worker	7.00	<0.001
20	业余爱好者指导专家 Amateur instruct expert	2.18	专家指导业余爱好者 Expert instruct Amateur	7.00	<0.001
21	证人传唤法官 Witness summon judge	1.60	法官传唤证人 Judge summon witness	6.91	<0.001
22	老鼠追黄猫 Mouse chase cat	2.55	黄猫追老鼠 Cat chase mouse	7.00	<0.001
23	明星崇拜粉丝 Star idolize fan	2.60	粉丝崇拜明星 Fan idolize star	6.91	<0.001
24	少年训斥父亲 Teenager scold father	2.09	父亲训斥少年 Father scold teenager	6.90	<0.001
25	猎物跟踪老虎 Prey follow tiger	2.10	老虎跟踪猎物 Tiger follow prey	7.00	<0.001
26	球员训练教练 Player train coach	1.73	教练训练球员 Coach train player	7.00	<0.001
27	苍蝇吃掉青蛙 Fly eat frog	1.10	青蛙吃掉苍蝇 Frog eat fly	6.91	<0.001
28	人质劫持恐怖分子 Hostage hold terrorist	1.91	恐怖分子劫持人质 Terrorist hold hostage	7.00	<0.001
29	歌手跟踪粉丝 Singer stalk fan	2.60	粉丝跟踪歌手 Fan stalk singer	6.82	<0.001
30	老师服从小孩 Teacher obey kid	1.55	小孩服从老师 Kid obey teacher	6.80	<0.001
31	客人保护保镖 Guest secure guard	2.90	保镖保护客人 Guard secure guest	6.91	<0.001
32	孤儿监护修女 Orphans supervise nuns	1.27	修女监护孤儿 Nuns supervise orphans	6.90	<0.001
33	大臣赐婚皇帝 Peasant award King	1.60	皇帝赐婚大臣 King award peasant	6.91	<0.001
34	观众娱乐相声演员 Spectator entertain comedian	2.09	相声演员娱乐观众 Comedian entertain spectator	7.00	<0.001
35	初学者救起救生员 Beginner rescue lifeguard	2.70	救生员救起初学者 Lifeguard rescue beginner	6.73	<0.001
36	学生辅导家教 Student teach tutor	1.73	家教辅导学生 Tutor teach student	7.00	<0.001
37	猪宰了屠夫 Pig slaughter butcher	1.10	屠夫宰了猪 Butcher slaughter pig	6.73	<0.001
38	游客欺骗奸商 Tourist cheat greedy businessman	2.09	奸商欺骗游客 Greedy businessman cheat tourist	7.00	<0.001
39	总理协助秘书 President assist secretary	3.40	秘书协助总理 Secretary assist President	6.82	<0.001
40	女孩绑架罪犯 Girl kidnap criminal	1.73	罪犯绑架女孩 Criminal kidnap girl	7.00	<0.001
41	学生测试老师 Student quiz teacher	4.40	老师测试学生 Teacher quiz student	7.00	<0.001
42	麋鹿射中猎人 Moose shot hunter	1.00	猎人射中麋鹿 Hunter shot moose	7.00	<0.001
43	动物治愈兽医 Animal cure veterinarian	2.30	兽医治愈动物 Veterinarian cure animal	7.00	<0.001

(contd.)

Item ID	Implausible	Norming Score	Plausible	Norming Score	Wilcoxon signed-rank test <i>p</i> value
44	狗遛男人 Pet walk man	1.91	男人遛狗 Man walk pet	6.80	<0.001
45	虫子吃鸟儿 Worm eat bird	1.70	鸟儿吃虫子 Bird eat worm	6.91	<0.001
46	嫌犯监视调查员 Suspect monitor investigator	2.45	调查员监视嫌犯 Investigator monitor suspect	7.00	<0.001
47	选民欺骗政客 Voter deceive politician	2.20	政客欺骗选民 Politician deceive voter	6.82	<0.001
48	犯人看守狱警 Prisoner restrain guard	1.18	狱警看守犯人 Guard restrain prisoner	7.00	<0.001
49	狐狸逮住猎户 Fox seize trapper	2.50	猎户逮住狐狸 Trapper seize fox	6.64	<0.001
50	猫喂养主人 Cat feed owner	1.27	主人喂养猫 Owner feed cat	7.00	<0.001
51	病人治疗医生 Patient cure doctor	3.00	医生治疗病人 Doctor cure patient	6.55	<0.001
52	人咬狗 Man bite dog	2.36	狗咬人 Dog bite man	7.00	<0.001
53	船员带领船长 Crew lead captain	3.10	船长带领船员 Captain lead crew	6.91	<0.001
54	小猴子欺负大猩猩 Monkey intimidate gorilla	3.18	大猩猩欺负小猴子 Gorilla intimidate monkey	7.00	<0.001
55	演员采访记者 Actress interview reporter	3.20	记者采访演员 Reporter interview actress	7.00	<0.001
56	盗贼追捕警察 Thief pursue cop	1.82	警察追捕盗贼 Cop pursue thief	7.00	<0.001

Peer review comments

The author(s) of this paper chose the Open Review option, and the peer review comments are available at: <http://doi.org/10.1525/collabra.131.pr>

How to cite this article: Zhou, P., Yao, Y., & Christianson, K. (2018). When Structure Competes with Semantics: Reading Chinese Relative Clauses. *Collabra: Psychology*, 4(1): 22. DOI: <https://doi.org/10.1525/collabra.131>

Senior Editor: Rolf Zwaan

Editor: Max Coltheart

Submitted: 19 December 2017

Accepted: 09 June 2018

Published: 29 June 2018

Copyright: © 2018 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.