Limited syntactic parallelism in Chinese ambiguity resolution

Yufen Hsieh a, Julie E. Boland a, Yaxu Zhang b, Ming Yan c

a University of Michigan, Ann Arbor, MI, USA b Peking University, Peking, PR China c Beijing Normal University, Beijing, PR China

First Published: September 2009
Limited syntactic parallelism in Chinese ambiguity resolution

Yufen Hsieh and Julie E. Boland

University of Michigan, Ann Arbor, MI, USA

Yaxu Zhang

Peking University, Peking, PR China

Ming Yan

Beijing Normal University, Beijing, PR China

Using the stop-making-sense paradigm (Boland, Tanenhaus, Garnsey, & Carlsen, 1995) and eye-tracking during reading, we examined the processing of the Chinese Verb NP₁ de NP₂ construction, which is temporarily ambiguous between a complement clause (CC) analysis and a relative clause (RC) analysis. Resolving the ambiguity as the more complex, less preferred CC was costly under some conditions but not under others. We took this as evidence for a limited parallel processor, such as Tabor and Hutchins’ (2004) SOPARSE, that maintains multiple syntactic analyses across several words of a sentence when the structures are each supported by the available constraints.

A central issue in sentence processing is how the human parser resolves syntactic ambiguity while reading or listening to sentences. Syntactically ambiguous regions are common across human languages, as is the...
experience of a ‘garden path’ – i.e., processing difficulty associated with disambiguation toward a less preferred meaning (e.g., Bever, 1970). Garden path effects make clear that we often commit to an analysis while the structure is still ambiguous, leading to the widespread adoption of serial parsing models over parallel models. In the context of this paper, we define serial parsing models as those in which a single structure is selected as each word is recognised – even if multiple analyses were considered as candidates. Correspondingly, we define parallel parsing models as those in which two or more analyses are maintained to some degree across several words. Specific examples of each type of parsing model are described below, with respect to the syntactic ambiguity we are investigating here.

The two experiments reported in this paper utilise the Chinese construction of *Verb NP₁ de NP₂*, which is temporarily ambiguous between a complement clause (CC) analysis and a head-final relative clause (RC) analysis (see Figure 1). The ambiguity hinges upon the lexical ambiguity of the homograph *de*, as illustrated in (1). In the CC analysis (1a), *de* is a possessive marker: NP₂ *room* serves as the object of the sentence-initial verb and is modified by NP₁. In the RC analysis, (1b), *de* is a relative clause marker in a head-final relative clause construction: *worker* is the head noun that is modified by the preceding relative clause.

![Figure 1. The tree structures of CC (1a) and RC (1b).](image-url)
1. [fén shuā gōng yù de fāng jiān] zhī hou, xiǎo chén hé le shuǐ
   After [painting the apartment’s rooms], Chen drank water.

2. [fén shuā gōng yù de gōng rén] hěn lèi
   [The worker that painted the apartment] was very tired.

Parsing models can be distinguished by the number of representations they construct and maintain when confronted with a syntactic ambiguity. Most current parsing theories assume little, if any, parallelism. For instance, the garden-path model (Frazier, 1987) proposes that the parser constructs only the structurally simplest analysis (the RC in (1)). The unrestricted race model (Traxler, Pickering, & Clifton, 1998; Van Gompel & Pickering, 2001) claims that although both the RC and CC analyses are activated in a horse race, only the simplest structure (RC) would be completed. On the other hand, the constraint-based competition models (e.g., McRae, Spivey-Knowlton, & Tanenhaus, 1998; Spivey & Tanenhaus, 1998) propose that both syntactic alternatives, RC and CC, are activated in parallel and compete for selection at each word, by getting graded support from the available syntactic, lexical, and pragmatic constraints. In short, both the unrestricted race model and the constraint-based competition models are serial parsers, by our criteria, because a single structure is selected at each word position.

In a serial parsing framework, if new material appearing in a sentence cannot be included into the present structure, the processor must restructure its analysis to incorporate the new information. This reanalysis requires extra effort, which is usually accompanied by longer reading times and/or regressive eye movements in an eye-tracking paradigm (Frazier & Rayner, 1982). In fact, the core argument for serial parsing has been challenged by the observation that some structural ambiguities do not yield noticeable processing difficulties even if they are disambiguated as the more complex or dispreferred structure (Gibson, 1991). However, theories of reanalysis suggest that not all types of reanalysis must produce a measurable garden path effect. For example, Lewis (1998) distinguished between easy garden paths, for which his SNIP operator could initiate a local repair, and difficult garden paths, for which reanalysis failed because the necessary repair was out of SNIP’s reach. Another approach was suggested by Fodor and Inoue (1994), who maintained that reanalysis difficulty is largely dependent upon the informativeness of the disambiguation cue: the more directly a disambiguation cue signals the appropriate repair, the lower the processing cost. In other accounts of reanalysis, the costs are lower when some of the constituents from the initial parse can be reused for the new parse (e.g., Abney, 1989; Konieczny, 1996). In sum, serial models generally predict some processing cost associated with
reanalysis, but the severity of the processing cost is determined by the specific details of the reanalysis mechanisms.

Alternatively, a parallel parser would construct multiple structures at points of syntactic ambiguity: under a fully parallel model, both CC and RC would be maintained throughout the region of Verb $NP_1$ de $NP_2$ in (1). A ranked parallel model (e.g., Gibson, 1991; Gorrell, 1987) allows alternative syntactic representations to be ordered according to various constraints, such as syntactic complexity, lexical frequencies, semantic information, and context. Given that the ranking of the alternative structures causes them to be differentially available to the processor, a ranked parallel model is also compatible with garden-path effects. That is to say, if disambiguation forces the parser to adopt a dispreferred structural analysis, a garden path effect arises because the structural alternatives must be reranked, either by changing their activation levels or by some other mechanism.

Despite widespread adoption of serial parsing assumptions, there have been some empirical results suggesting that ranked parallelism provides a better account of garden path effects. For example, Hickok (1993) maintained that the parser computed both the preferred sentential-complement and the dispreferred relative clause representations in parallel when processing the ambiguous sentence ‘The psychologist told the wife that the man bumped that her car was stolen’. On the one hand, the parser was garden-pthed when the disambiguation required the assignment of a relative clause structure of the ambiguous region, suggesting that the sentential-complement reading was preferred. On the other hand, the NP the wife was reactivated following the presentation of the embedded verb bumped, suggesting that the relative clause reading was also computed.

Tabor and Hutchins’ (2004) computational self-organising model (SOPARSE) proposes that each new word of a sentence activates possible attachments in parallel and that these structural alternatives compete until one of them reaches stabilisation. The structural alternatives are largely determined on the basis of lexicalised syntactic knowledge. Under this account, attachments corresponding to both the RC and the CC would be activated as $de$ is perceived during the processing of the ambiguous string Verb $NP_1$ de $NP_2$. SOPARSE is a type of ranked parallel processor, because there are temporal intervals during which multiple analyses are partially active and no analysis has reached a stable state. Furthermore, SOPARSE predicts greater ‘digging-in’ costs the longer the ranking has been established because, even without additional supporting evidence, the initially preferred attachment continues to grow in activation strength via a ‘rich-get-richer’ feedback mechanism designed to elevate the activation of the selected structure to a stable state over the course of several words.

Prior research on the Verb $NP_1$ de $NP_2$ construction, using a self-paced word-by-word reading paradigm, demonstrated that a semantic constraint
led to a parsing commitment to a particular structure during the ambiguous region (Zhang, Zhang, & Hua, 2000). A plausibility cue was provided at NP₂ in sentences like (2), to bias the ambiguous phrases towards a reading of RC or CC, or remain neutral. Zhang et al. found garden-path effects one word after the disambiguation when RC-biased items were disambiguated as CC (2a) and vice versa (2b). More importantly, garden path effects appeared in the semantically balanced phrases when they were disambiguated as CC (2c), which suggests that the RC is the default analysis. There are several reasons why the RC might be preferred. First, the RC is structurally simpler by the principle of minimal attachment, and allows immediate thematic role assignment for NP₁, as the direct object of the verb. Second, the RC has an explicit subject (in the final position) and thus provides a complete propositional meaning, whereas the CC does not have an external argument. Third, Zhang et al. found that the syntactically contingent frequency of de as a relative clause marker (as in the RC) in this construction is considerably higher than de as a possessive marker (as in the CC). In the context of Verb NP₁ de NP₂, 70% of the 1000 syntactically ambiguous items that were randomly selected from a corpus¹ were RC.

2. Example stimuli from Zhang et al. (2000)

a. RC-biased disambiguated as CC
   [dai4man4 ke4ren2 de hai2zi] zhi1hou4, zhou1li4 xin1li3 you3xie1 ao4hui3
   [slight guest POSS child] after, Zhou Li in the mind somewhat regretful
   After [slighting the guest’s child], Zhou Li felt somewhat regretful.

b. CC-biased disambiguated as RC
   [zhi3ze2 bao4she4 de ji4zhe3] ren4wei2 xin1wen2 bao4dao3 bi4xu1 ke4guan1
   [censure newspaper-office RC reporter] think news report must objective
   [The reporter that censured the newspaper office] thought that news reports must be objective.

c. balanced disambiguated as CC
   [zhuang4dao3 xiao1ming2 de che1zi] zhi1hou4, liang3ge4 hai2zi fei1chang2 hai4pa4
   [run into Xiao Ming POSS bicycle] after, two children very scared
   After [running into Xiao Ming’s bicycle], the two children were very scared.

¹These items were selected from the Corpus for Studies of Modern Chinese (Beijing Language and Culture University, 1995), which has 1.24 million words collected from a broad range of genres.
In Zhang et al.'s (2000) stimuli, the semantic cue that biased the parser toward CC or RC was not available at NP₁ to guide the initial analysis of the ambiguous construction. Rather, the CC-bias was created with plausibility information provided at NP₂. If only the simpler RC is built at *de*, then a serial parser would have to reanalyse that commitment when it encountered the semantic cue that required the dispreferred CC at NP₂. The revision would elicit a garden-path as it did in the disambiguation region of an RC-biased item when it was disambiguated as a CC.

On the other hand, a ranked parallel model that computes both the RC and the CC at *de* would predict a graded effect of re-ranking from a preferred syntactic analysis (RC) to a dispreferred one (CC). Under a parallel account, the cost of re-ranking would be influenced by the amount and the relative timing of evidence to support each analysis. For example, re-ranking costs should be minimal at word N + 1 if the two structures had nearly equivalent support (and thus nearly equal levels of activation) at word N. On the other hand, much higher re-ranking costs would be expected if all the cues supported one structure across several word positions, but the item was subsequently disambiguated as the other structure.

The primary aim of the current study was to provide evidence that distinguishes parallel from serial parsing models. The ambiguity investigated here and in Zhang et al. (2000) is well-suited to the difficult empirical problem of distinguishing serial and parallel syntactic processing, because a revision from RC (Figure 1b) to CC (Figure 1a) requires a complete reanalysis of the first part of the sentence. Experiments 1 and 2 both use two ambiguous conditions, which differ with respect to the word position at which revision from RC to CC was required: either NP₂ (word four in (3a) below) or the conjunction (word five in (4a) below). As we discuss below, there have been no reanalysis mechanisms proposed that could accomplish such restructuring without considerable processing costs, regardless of the word position at which revision is necessary. A secondary aim was to investigate whether a semantic constraint resulted in parsing preferences for a particular structure, if the constraint occurred late during the ambiguous region. Based on Zhang et al., we expected further confirmation of the early use of semantic information, which distinguishes multi-constraint based approaches (e.g., MacDonald, Pearlmutter, & Seidenberg, 1994; McRae et al., 1998) from the construal/garden-path theories (Frazier & Clifton, 1996) as well as the majority of reanalysis hypotheses, which assume that semantic information does not influence the initial parse (e.g., Ferreira & Henderson, 1991a; Fodor & Inoue, 1994).
Sentence completion survey

In order to justify the claim that an RC is the default structure, an important assumption in our argument, we conducted a sentence completion survey (included in Appendix A) with the critical stimuli used in Experiments 1 and 2. Twenty-four native Mandarin Chinese speakers from Taiwan who were not involved in Experiments 1 and 2 participated in the study. The participants were presented with the sentences up to but not including NP2 (i.e., *Verb NP1 de*) and were asked to complete the sentence fragments using the first words that come to mind. The 40 critical items were pseudorandomly mixed with 60 filler sentence fragments containing three words of various structures, such that two critical trials did not occur consecutively. Two experimental lists with different item orders were then created. For all the critical items, all the participants began their completion with a noun phrase. This noun phrase was part of a RC completion 95% of the time (911/960) as anticipated, given the comprehension data from Zhang et al. (2000). The other 5% of responses were CC completions. As shown in Appendix A, all items had at least 50% RC completions, and only three items had fewer than 80% RC completions (two Inanimate items and one Animate item). Thus, the RC analysis is strongly preferred over the CC analysis for our stimuli.

EXPERIMENT 1

The experiment presented here focused on the construction of *Verb NP1 de NP2*, which is temporarily ambiguous between a complement clause (CC) structure and a relative clause (RC) structure. Consider the examples in (3) and (4): (3a) and (4a) contained the ambiguous construction in the first four words; (3b) and (4b) were unambiguous controls for (3a) and (4a), respectively, where NP1 was replaced with an adjective, forcing de to be an attributive marker. Thus, both (3b) and (4b) contained unambiguous attributive structures. They served as control structures for (3a) and (4a) because they were matched for lexical content, but did not contain the syntactically ambiguous sequence. Because the ambiguous conditions were always disambiguated as the less preferred CC structure, a processing cost for the ambiguous conditions compared with the unambiguous conditions is

\footnote{2 It is not entirely impossible for the first four words in (3b) and (4b) to have a relative clause continuation, such as [*fen3shua1 lao3jiu4 de fang2jian1 de gong1ren2 . . .]* [paint old ATT room RC worker] ‘The worker that painted the old rooms . . .’. However, such a sentence with two nearly adjacent des should be rare. In fact, the first de is usually dropped in an expression like this to avoid redundancy, as the de following an adjective is omissible and in fact omitted about 90% of the time based on our corpus analysis.}
likely to reflect costs associated with reanalysis under a serial account or re-ranking under a parallel account.

As described above, both syntax-first and multi-constraint theories predict that NP\textsubscript{1} would be taken as the direct object of the initial verb in both (3a) and (4a). Then, at the homograph \textit{de}, a serial parser would continue to construct an RC, whereas a parallel parser would compute both an RC and a CC with the former ranked higher. Examples (3a) and (4a) differed with regard to our animacy manipulation at NP\textsubscript{2}, which served to either (semantically) disambiguate the ambiguous construction as the CC (3a) or support the RC (4a). Finally, the structure for both the ambiguous phrases in (3a) and (4a) was disambiguated as a CC at the conjunction (before/after/while).

3. Inanimate

a. [fen3shua1 gong1yu4 de fang2jian1] zhi1hou4, xiao3wang2 hai2 da3sao3 le ke4ting1
[paint apartment POSS room] after, Wang also clean PERF living room
\textit{After [painting the apartment's rooms], Wang also cleaned the living room.}

b. [fen3shua1 lao3jiu4 de fang2jian1] zhi1hou4, xiao3wang2 hai2 da3sao3 ke4ting1
[paint old ATT room] after, Wang also clean PERF living room
\textit{After [painting the old rooms], Wang also cleaned the living room.}

4. Animate

a. [xun4lian4 shi4bing1 de jiang1jun1] zhi1hou4, zong3silling4 fa1biao3 le jian3duan3 yan3shuo1
[train soldier POSS general] after, commander give PERF short speech
\textit{After [training the soldiers' general], the commander gave a short speech.}

b. [xun4lian4 nian2qing1 de jiang1jun1] zhi1hou4, zong3silling4 fa1biao3 le jian3duan3 yan3shuo1
[train young ATT general] after, commander give PERF short speech
\textit{After [training the young general], the commander gave a short speech.}

The Inanimate Ambiguous condition (3a) was semantically disambiguated as a CC because NP\textsubscript{2}, \textit{room}, must be the direct object of the verb phrase \textit{paint} rather than the head noun that performs the action of painting an
apartment. In the Animate Ambiguous condition (4a), although both the interpretations of RC (the general that trained the soldiers...) and CC (training the soldiers’ general) were possible, the initially adopted RC is most plausible. That is, it is more plausible that a general trained soldiers than that the general was trained. So general is likely to be assigned the thematic role of Agent. Thus, in the Animate Ambiguous condition we expected the RC analysis to become deeply entrenched as semantic evidence increased through NP₂.

**Predictions**

The control conditions (3b) and (4b) provide unambiguous attributive baselines in which we expect no processing difficulty, under any type of processing theory.

A construal/garden path processor that does not make early use of semantic cues would behave the same in (3a) and (4a): it would first assign the simpler RC structure to the ambiguous construction *Verb NP₁ de NP₂*. Hence, the parser would be garden-pathed in both (3a) and (4a) compared with (3b) and (4b) when encountering syntactic evidence of the CC at the conjunction, *after*. It is possible that the semantic incongruity of the Inanimate condition would trigger reanalysis slightly earlier, after initially attaching NP₂ to the RC structure. Either way, a measurable garden path would be predicted in both conditions. Assuming a backtracking reanalysis mechanism such as that outlined in Frazier and Rayner (1982), in both conditions, the parser would have to reanalyse the structural assignments of the first four words because none of the RC tree can be recycled for the CC tree (see Figure 1).

A serial multi-constraint based approach also predicts a garden path in both the Inanimate Ambiguous (3a) and Animate Ambiguous (4a) conditions, because at the homograph *de*, the RC structure would be selected for both (3a) and (4a), based on structural simplicity, immediate interpretation, and contingent frequency. In (3a), the parser should reanalyse the RC as a CC at NP₂, because *room* must be the direct object of the verb phrase rather than the head noun. In (4a), however, the semantic cue at NP₂ supports the syntactic analysis of a RC, so the parser would not switch to a CC analysis until it was forced at the next word *after*. Thus, reanalysis costs should be observed at NP₂ in comparing (3a) and (3b), but at the conjunction in comparing (4a) and (4b). The reanalysis costs might be somewhat less in the Inanimate condition (3a), because only three words would have to be restructured, whereas four words would have to be restructured in the Animate (4a). We are not aware of a well-specified mechanism for reanalysis within a serial constraint-based theoretical framework, so it is difficult to make clear predictions as to the relative costs. However, all of the reanalysis
accounts that we have considered predict that both garden paths will cause
difficult reanalysis and therefore easily detectable garden path effects (Abney,

In contrast, a fully parallel account would predict no garden path in either
c Condition. We do not seriously consider such a framework here, in light of
considerable evidence against full-fledged parallelism (e.g., Ferreira &
Henderson, 1991a; Frazier & Rayner, 1982; Pritchett, 1988; Sturt, Pickering,
& Crocker, 1999).

We do consider a ranked parallel approach, along the lines of Tabor and
Hutchins’ SOPARSE account, in which the syntactic processor attempts
multiple attachments in parallel as each new word is recognised, and the
relative accessibility of the alternatives is governed by all relevant constraints.
As mentioned above, in Tabor and Hutchins’s SOPARSE, two viable
attachments can remain activated across several words, until one of the
syntactic alternatives reaches a sufficiently high level of activation. Under
such an account, attachments corresponding to both a RC and a CC would
be attempted at de, with the RC initially more activated, based upon
structural simplicity, immediate thematic role assignment, and contingent
frequency. At NP2, the less preferred CC structure would still be somewhat
activated, and the semantic cue in the Inanimate condition (3a) would
increase CC activation — possibly high enough to outrank the RC. In
contrast, the CC would receive no additional support in the Animate
c Condition (4a), so the RC would become more entrenched as the highest
ranked structure, based on rich-get-richer feedback. Some processing costs
associated with re-ranking the RC and the CC might be observable at NP2 in
(3a), but a much greater re-ranking cost should be apparent at the
conjunction in (4a), where full syntactic disambiguation occurs and the
CC would have to be resurrected from a low state of activation.

In brief, we hypothesised that semantic constraints would be used
immediately as a disambiguation cue late in the ambiguous region, resulting
in either reanalysis or re-ranking at NP2 in the Inanimate Ambiguous
condition. Another revision, caused by the syntactic disambiguation, would
take place at the conjunction in the Animate Ambiguous sentences.

Method

Participants. Twenty-four native speakers of Mandarin Chinese from
Taiwan were paid a nominal sum to participate in the experiment. All the
participants were recruited in the University of Michigan and had normal or
corrected-to-normal vision. Although they also spoke English, Mandarin
Chinese was the primary language that they used outside of their classes.
Materials. A total of 40 sets of critical items (see Appendix A) were typed in traditional Chinese characters with 20 sets in each of the Inanimate and the Animate versions (adapted from Zhang et al., 2000). All the sentences were 10 words long plus a full stop at the end, and were displayed in one line on the computer screen. As shown in (3) and (4) above, structural ambiguity was a within-item factor, whereas the animacy of NP$_2$ was a between-item factor. Our experiment thus included four types of sentences: Inanimate Ambiguous (3a), Inanimate Unambiguous (3b), Animate Ambiguous (4a), and Animate Unambiguous (4b).

For the structural manipulation, each set of the critical sentences comprised two types of underlying structures: an ambiguous and an unambiguous control condition, which differed lexically only at the second word. For the manipulation of semantic plausibility (triggered by animacy), NP$_1$ following the initial verb was a potential object as apartment in (3a) and soldier in (4a), yet NP$_2$ was either inanimate and therefore an implausible agent, as room in (3a), or a plausible one such as the animate general in (4a). The ambiguous construction, Verb NP$_1$ de NP$_2$, was syntactically disambiguated at the conjunction (i.e., before/after/while) followed by a comma. Plausibility and structural ambiguity were counterbalanced across two presentation lists. Each list contained ten ambiguous and ten unambiguous sentences in each of the Inanimate and the Animate groups. Two versions of an item were never presented to the same participant. The 40 critical items were pseudo-randomly embedded within 64 filler sentences of various types in order to prevent participants from being aware of the experimental design. Sixteen of the fillers consisted of an equal proportion of syntactic and semantic violations. Altogether, 26 of the 104 sentences (25%) became anomalous at various points across the sentences.

Procedure. The experiment used a self-paced stop-making-sense paradigm (Boland et al., 1995). Sentences appeared on a computer screen one word at a time in a moving window display. Participants were instructed to click the mouse with their dominant hand to receive the next word. Meanwhile, if the sentence stopped making sense, they should press the ‘No’ button (the space bar) on the keyboard with their non-dominant hand. The trial then terminated immediately if the ‘No’ button was pressed. In the written instructions, it was stressed that participants need not wait until the end of the sentence to respond negatively. All times between presentation of a word on the screen and a response were measured to millisecond accuracy using E-Prime software. Before the experiment began, participants were provided with 8 practice sentences to familiarise them with the task. In most cases, participants completed the experiment within 20 minutes.
**Plausibility survey.** A questionnaire study was conducted in order to ensure that the Animate Ambiguous items were semantically biased toward the RC interpretation. A separate group of 30 native Chinese speakers completed the survey. The survey contained 60 phrases of the structure of Verb \( NP_1 \) de \( NP_2 \), which were ambiguous between CC and RC interpretations. Participants rated the plausibility of the two readings of these phrases on a 7-point scale. For example, the ambiguous phrase in (4a) could be interpreted as either to train soldiers’ general or the general that trained soldiers. Participants were instructed to make the judgement based on their real-world knowledge, namely, which interpretation was more likely to occur. One endpoint of the scale (coded as 7) represented that the phrase was highly likely to mean the general that trained soldiers (i.e., RC). The other endpoint (coded as 1) indicated that to train soldiers’ general (i.e., CC) was much more plausible. The middle number (coded as 4) indicated that the phrase was balanced between the two readings.

Using the results of the survey, we selected 20 items as the Animate stimuli, all of which received a score of 5 or higher with the average of 5.71. The plausibility ratings for each Animate item are presented along with the items in Appendix A.

**Results**

Two dependent measures were analysed: the number of ‘No’ responses at each word position and word-by-word reading times for affirmative button presses.

A summary of the cumulative percentage of ‘No’ responses is presented in Figure 2 for each condition from the verb to CONJ + 2 of the sentences. From the beginning to \( NP_2 \) of the sentence, there were fewer than 2% ‘No’ responses across all four conditions. However, an interaction appears to begin at the conjunction (i.e., before/after/while), and continue over the next two words. These three words were identical for each item across the ambiguous and unambiguous conditions.

To eliminate dependencies among sequential responses, the per cent ‘No’ responses illustrated in Figure 2 were converted to ‘remaining possible ‘No’s’ using the procedure in Boland et al. (1990). The percentage of ‘remaining possible ‘No’s’ in each condition by both participants and items were

---

3. There were ten items per condition for each participant, so at the beginning of the sentence the maximum number of trials on which a participant could respond ‘No’ was ten. If a participant made four ‘No’ judgements at the first opportunity, the participant’s score at this word was 40%. At the next word, the number of remaining possible ‘No’ responses would be six. Thus, if two ‘No’ responses were made at this word position, the score was 33%. If all of the trials in a given condition for a particular participant or item received ‘No’s before the end of the sentence, the later word positions were assigned a value of 100% rather than 0%.
submitted to 2 (list) × 2 (structural ambiguity) × 2 (animacy) repeated measure ANOVAs. We did separate ANOVAs at NP2, the conjunction, and CONJ + 1, because the critical predictions were localised to these regions, especially the first two. A summary of the analyses appears in Table 1. Neither effects of ambiguity nor animacy, nor an interaction between them was observed at NP2. At the conjunction, there was an interaction between ambiguity and animacy, which appears to be driven by the relatively high proportion of ‘No’ responses in the Animate Ambiguous condition. Also, there were main effects of ambiguity and animacy. The same pattern of effects was observed at CONJ + 1.

Word-by-word reading times for affirmative button presses provided a secondary index of processing difficulty. The reading time data (shown in Figure 3) exhibited the same numerical pattern as the ‘No’ data, suggesting that, even when readers failed to press the ‘no’ button in the Animate Ambiguous condition, they nonetheless experienced processing difficulty. However, we interpret these data more cautiously, because once a trial was stopped by a ‘No’ judgement, no further reading times were recorded, and thus, the amount of missing data differs across conditions, beginning around the conjunction. The amount of missing data in each condition is equal to the cumulative per cent ‘No’ data in Figure 2. For example, by CONJ + 1 in the Animate Ambiguous condition, there had been ‘No’ responses in about 33% of the trials, thus the mean response time was computed over the remaining 67% of trials in that cell.

**Figure 2.** The cumulative percentages of ‘No’ responses for each condition at each word position in Experiment 1.
Mean reading times were computed both by participants and by items. One participant had no reading time data starting from the conjunction in the Animate Ambiguous condition, because he or she had responded ‘No’ to all trials in that condition by the conjunction. The missing cells were replaced using the formula \[
\text{participant mean for the other three conditions at that word position} \div \text{condition mean for the other participants at that word position}.
\]

### TABLE 1

<table>
<thead>
<tr>
<th>‘Remaining possible ‘No’s’ analysis</th>
<th>$F_1$ (df)</th>
<th>$F_2$ (df)</th>
<th>$p$</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP2 Structural Ambiguity</td>
<td>3.19 (1, 22)</td>
<td>2.31 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>NP2 Animacy</td>
<td>0.31 (1, 22)</td>
<td>2.31 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>NP2 Ambiguity $\times$ Animacy</td>
<td>2.30 (1, 22)</td>
<td>0.24 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>CONJ Structural Ambiguity</td>
<td>8.52 (1, 22)</td>
<td>37.34 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>CONJ Animacy</td>
<td>12.47 (1, 22)</td>
<td>57.30 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>CONJ Ambiguity $\times$ Animacy</td>
<td>8.37 (1, 22)</td>
<td>22.87 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>CONJ +1 Structural Ambiguity</td>
<td>16.94 (1, 22)</td>
<td>77.84 (1, 36)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CONJ +1 Animacy</td>
<td>23.54 (1, 22)</td>
<td>93.45 (1, 36)</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>CONJ +1 Ambiguity $\times$ Animacy</td>
<td>19.35 (1, 22)</td>
<td>60.36 (1, 36)</td>
<td>&lt;.001</td>
</tr>
</tbody>
</table>

**Figure 3.** Mean reading times in ms for each condition at each word position in Experiment 1.
position] – grand mean of all participants and conditions at that word position]. Participant and item means were submitted to 2 (list) × 2 (structural ambiguity) × 2 (animacy) repeated measure ANOVAs, conducted at NP2, the conjunction, and CONJ+1, respectively. A summary of the analyses is provided in Table 2. At NP2 (where all conditions had affirmative responses in over 95% of the trials), there were neither effects of ambiguity nor animacy, nor an interaction between them. At the conjunction, the interaction between structural ambiguity and animacy was significant by items but not by participants, while the main effects of ambiguity and animacy were fully reliable. The Animate Ambiguous sentences were 202 ms longer than their unambiguous counterparts (z = .05 by participants and by items), though there were also more missing data in this condition compared to the other three conditions. None of the other conditions differed from one another in pairwise comparisons. At CONJ+1, there was a robust interaction, as well as the main effects for ambiguity and animacy. The response times were 252 ms longer in the Animate Ambiguous condition than in the Animate Unambiguous condition (z = .05), though again, the Animate Ambiguous condition had more missing data than the other three conditions. The other three conditions did not differ.

### Discussion

Both the judgement data and the reading time data revealed a striking difference between processing of the Animate Ambiguous stimuli and the Inanimate Ambiguous stimuli. While there were robust garden path effects for the Animate Ambiguous condition, beginning at the conjunction, there was no comparable effect for the Inanimate Ambiguous condition. While

### Table 2

<table>
<thead>
<tr>
<th>Reading times analysis</th>
<th>F1 (df)</th>
<th>F2 (df)</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>NP2</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>2.00 (1, 22)</td>
<td>1.08 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Animacy</td>
<td>0.17 (1, 22)</td>
<td>0.35 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Ambiguity × Plausibility</td>
<td>0.00 (1, 22)</td>
<td>0.00 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>CONJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>8.36 (1, 22)</td>
<td>12.15 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>8.23 (1, 22)</td>
<td>11.03 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>1.49 (1, 22)</td>
<td>5.63 (1, 36)</td>
<td>&gt;.10, &lt;.05</td>
</tr>
<tr>
<td>CONJ+1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>26.61 (1, 22)</td>
<td>13.66 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>22.42 (1, 22)</td>
<td>11.46 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>13.65 (1, 22)</td>
<td>11.83 (1, 36)</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>
this might seem to suggest that a CC is the default analysis for this ambiguity, in fact there is considerable evidence against that possibility, based on our sentence completion data, the findings in Zhang et al. (2000), and processing factors known to guide syntactic ambiguity resolution, such as structural complexity, immediate thematic role assignment, and the contingent frequency of the alternative forms of *de*. Furthermore, there were no semantic cues to support the CC until NP2 in the Inanimate Ambiguous condition.

The serial parsing models predicted a garden path in both critical conditions, however, a garden path effect was observed only in the Animate Ambiguous condition. The Inanimate Ambiguous sentences showed no processing difficulty at either NP2 (i.e., the semantic cue) or the conjunction (the syntactic disambiguation). A serial parsing model could still accommodate this finding if the parsing model included a reanalysis mechanism such that reanalysis was very easy in the Inanimate Ambiguous condition, but much harder in the Animate Ambiguous condition. Therefore, we considered several reanalysis proposals to see if they predicted a difference in reanalysis difficulty across the Animate and Inanimate Ambiguous conditions.

First, we considered Fodor and Inoue’s (1994) Diagnosis Model of reanalysis in which most of the variability in reanalysis difficulty is linked to determination of what structural alterations are necessary to get from the current structure to the correct structure. This is an interesting proposal and has received some empirical support (Meng & Bader, 2000). However, it fails to predict a difference in processing difficulty between the Inanimate and Animate conditions in our study. Fodor and Inoue maintain that revision difficulty is a function of how easy it is for the parser to diagnose the source of processing error using ‘symptom’ provided by the breakdown of the first analysis. According to the diagnosis model, the sentence repairing procedures are only initiated by syntactic violations, presumably to make sure that revision does not take place unless it is absolutely necessary. Thus, for our stimuli, repairs would be triggered at the conjunction in both the Animate and Inanimate Ambiguous conditions, and no difference in revision costs would be predicted, because the same structural symptom would cue diagnosis in both cases. Regardless of whether the conjunction is considered an informative or an uninformative diagnostic cue, Fodor and Inoue predict equivalent garden path effects in the Animate and Inanimate conditions.

Second, Lewis’ (1998) SNIP account predicts successful reanalysis when a single, local structural dependency must be broken, as in ‘Thad knows Shaq is tall’. Using this example, Lewis suggested that the CP projected from *is* tries to attach as the complement of the VP projected from *knows*, a site that is already filled by *Shaq*. This structural incompatibility triggers SNIP to disconnect *Shaq*, allowing LINK operators to make the correct new
attachments. The SNIP operator is able to detach Shaq because it falls under the maximal projection containing the incompatibility (the VP). Lewis maintained that a single SNIP repair adds about 50–100 ms of processing time. Changing from an RC structure to a CC structure at NP₂ or five⁴ in our stimuli would require multiple instances of detaching and relinking constituents, so it presumably would take considerably longer under Lewis’ metric – if SNIP could reach into the CP and IP embedded within the complex NP (see Figure 1b) to detach NP₁ and de from the tree. But in fact, such operations are not local enough for SNIP, and as a result, reanalysis should fail for both the Animate and Inanimate conditions.

Pritchett’s (1988, 1991) Relicensing Constraint distinguishes easy and hard reanalyses in terms of syntactic government. Pritchett notes that reanalysis of water from direct object to subject is easy in John expected the water to taste bad, but difficult in After John drank the water tasted bad. According to Pritchett, the crucial difference is that water continues to be governed by expect, but not by drank. Sturt et al. (1999) tested Pritchett’s predictions by measuring self-paced reading times during the disambiguating region for these two types of sentences. In line with Lewis’ predictions, the easy reanalysis increased reading time by 87 ms – a modest, but statistically significant, processing cost. The difficult reanalysis increased reading times by 400 ms – a much greater processing cost. If we apply the Relicensing Constraint to the reanalysis of a RC (Figure 1b) as a CC (Figure 1a) NP₁, we find that the government relationships are altered by the global structural changes. Thus Pritchett’s account, as well as the similar account of Sturt et al., would predict a difficult garden path, even if reanalysis were initiated at NP₂ in the Inanimate condition.

In sum, while there have been a variety of reanalysis mechanisms proposed, there is general agreement concerning which garden paths should be difficult. Difficult garden paths are characterised by major structural changes, at least some of which are non-local. All of the reanalysis accounts that we have considered predict that difficult garden path effects should result from reanalysis of an RC as a CC in both critical conditions. None of the proposed reanalysis mechanisms are consistent with our finding that reanalysis costs are considerably less in the Inanimate condition compared with the Animate condition.

Our findings are most consistent with a limited parallel account along the lines of Tabor & Hutchins (2004). The absence of a garden path effect at the disambiguation of the Inanimate sentences suggests that the parser

⁴Lewis’ SNIP operator appears to be sensitive only to syntactic inconsistencies, so the most straightforward assumption is that reanalysis would be attempted at the conjunction in both the Animate and Inanimate Ambiguous conditions. In fact, the repair surpasses the locality constraints of SNIP, regardless of whether the repair is initiated at NP₂ or the conjunction.
already had the correct CC analysis at the previous region. Given that a relevant semantic cue was available only at NP₂, the parser probably reranked the structural alternatives, demoting the simpler RC and promoting the more complex CC at this point, with minimal processing costs. There was no evidence of a re-ranking cost in our dependent measures, but re-ranking at NP₂ could be very low cost if the lower ranked representation had been constructed and maintained up to this point – i.e., for one to two words.

The interaction of structural ambiguity and animacy demonstrated clear effects of semantic plausibility. This finding is inconsistent with syntax-first parsing models, which predicted that the RC should be computed in both (3a) and (4a) and then abandoned as the semantic and/or syntactic disambiguation required the assignment of CC. If this had been the case, processing difficulty should have been observed in the disambiguation regions of both Animate and Inanimate sentences.

EXPERIMENT 2

The primary goal of this experiment was to further explore possible reanalysis or re-ranking effects at NP₂ of the Inanimate sentences from Experiment 1. Although the stop-making-sense task is usually quite sensitive to processing difficulty, participants are asked to identify at which point (if any) a sentence becomes anomalous, by pressing a different button, and as a result reading speed is often slowed below normal rates. Thus, we also used the same Inanimate and Animate stimuli in an eye-tracking paradigm, which imposes minimal costs over normal reading and furthermore provides a number of dependent measures so that processing load can be measured in terms of early and late measures of processing.

Research on the eye movements during reading has shown that average fixation durations are similar (about 225–250 ms) for Chinese and English readers (Chen, Song, Lau, Wong, & Tang, 2003; Rayner, 1998). Nevertheless, Chinese and English eye movement patterns differ in other aspects due to orthographic differences. Because the orthography of Chinese is much more lexically dense than the orthography of English, the perceptual span is correspondingly smaller, but the number of words encompassed with the perceptual span is not very different across the two languages. The perceptual span of Chinese readers extends from 1 character left of fixation to 3 characters to the right (i.e., approximately two to three words) when they are reading from left to right (Inhoff & Liu, 1998), while English readers have a span extending from 3–4 letters left of fixation to about 14–15 letters to the right (i.e., about three words) (Rayner, 1998). A second orthographic factor is that Chinese characters are presented in a continuous string, without
spaces between words – unlike English. Because Chinese words can be one or more characters long, the word boundaries are often ambiguous, and the lack of spaces between words could conceivably increase parafoveal processing of upcoming words when reading Chinese, compared to English. Unfortunately, little is known about possible differences between Chinese and English with respect to the parafoveal processing of upcoming words, nor is it clear how much phonological and semantic processing occurs for characters in the periphery during Chinese reading (see Feng, 2006, for an overview). Third, Chinese readers exhibit a slightly higher regression rate (about 15%) than English readers (about 10%) (Chen et al., 2003; Rayner, 1998). Finally, average saccade length is shorter in Chinese (about 2–2.5 characters) than in English (about 7–9 letters or 1.5 words), due to the higher information density of the Chinese text (Chen et al., 2003).

Method

Participants. Thirty-two native Chinese-speaking students from Beijing Normal University participated in the experiment. All had normal or corrected-to-normal vision and were paid a nominal sum for their participation.

Materials. Two experimental lists contained the same 20 Inanimate and 20 Animate critical items from Experiment 1. Minor alterations were made to a subset of the items to accommodate differences between Mandarin as spoken in Taiwan vs. Mandarin as spoken in Beijing. The data for two items (item 14 in the Inanimate condition and item 20 in the Animate condition) were omitted from all analyses because of a mistake in the input of the sentences.

For each list, 40 critical items were pseudo-randomly mixed with 72 filler sentences of various structures in order to prevent participants from being aware of the purpose of the experiment. Half of the fillers were critical items of an unrelated experiment, and the other half were complex sentences without intentional syntactic ambiguity.

Procedure. Participants were seated 60 cm from a computer screen with their head stabilised on a chin rest. Their eye movements and fixations were binocularly recorded every 2 ms by an EyeLink head-mounted eye-tracker. Calibration and validation of the recording apparatus were performed before the experiment started. At the beginning of a trial, participants were asked to look at a dot, which appeared on the left of the screen at the position where the first character of a sentence would appear. Once fixation was stable and any necessary correction was made, the dot disappeared and a sentence appeared instead. Participants were instructed to read for comprehension in
a natural way and to press a button when they had finished reading. The button pressing caused the sentence to disappear with a verification sentence (a declarative sentence with a cue of ‘Please judge:) presented in about one-third (38/112) of the trials. Participants indicated whether this sentence correctly expressed the content of the preceding sentence by pressing the right or left button on a response box. Fifty percent of the verification sentences were correct. The verification sentences remained on the screen until the participant had responded. Subsequent to the button pressing, a fixation point, indicating the beginning of the next trial, appeared.

There were 28 practice sentences before the experiment began. The whole experiment lasted about 30 minutes, and participants took a short break in the middle of the experiment. Participants could ask for a break at any time; a recalibration was performed before continuing the experiment.

Results

Eye-movement data were collected for seven word positions in the critical items, beginning at the sentence-initial verb and continuing until CONJ + 2. Five measures were calculated for each region. First-fixation durations are the times of the first fixation inside a region. Gaze durations include all fixation times inside a region beginning with the first fixation inside until the gaze travels outside the region, either to the left or to the right. Regression-path durations sum all fixation durations from the time a region was initially fixated until the reader’s gaze crossed the right boundary of the region. This measure includes re-fixations of preceding regions and the target region itself. Probability of first-pass regressions is the percentage of leftward eye movements that cross a region’s left boundary immediately following a first-pass fixation in the region. Total times is the sum of all fixation times within a region.

First-pass measures. We first report the four measures that provide early indications of processing difficulty: first-fixation durations (Figure 4), gaze durations (Figure 5), regression-path durations (Figure 6), and the probability of first-pass regressions (Figure 7). We focus our analyses on NP2 and the conjunction, where the theoretically important predictions are localised. For the two regions, we submitted the condition means, both by participants and by items, for each dependent measure to a 2 (list) × 2 (structural ambiguity) × 2 (animacy) repeated measure ANOVA. A summary of the analyses appears in Table 3.

First-fixation and gaze duration measures are thought to reflect the earliest stages of processing. In fact, the data pattern was quite similar to that observed in Experiment 1. At NP2, we obtained neither main effects nor interactions in either measure. At the conjunction, Animate Ambiguous
sentences were 39 ms (first-fixation duration) and 76 ms (gaze duration) longer than their unambiguous counterparts. There was a robust interaction between ambiguity and animacy, as well as main effects of both ambiguity and animacy. In paired comparisons for both dependent variables at the

Figure 4. Means for first-fixation durations for each condition at each word position in Experiment 2.

Figure 5. Means for gaze durations for each condition at each word position in Experiment 2.
conjunction, the Animate Ambiguous condition differed from the other three conditions, which did not differ from one another ($\alpha = .05$ by both participants and items).

As for the regression-path duration data, both the Inanimate and Animate Ambiguous sentences produced slightly longer reading times at
### TABLE 3
Analyses of mean first-fixation durations, gaze durations, regression-path durations, probability of first-pass regressions, and total times at NP₂ and the conjunction in Experiment 2

<table>
<thead>
<tr>
<th>Reading times analysis</th>
<th></th>
<th></th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>First-fixation durations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>0.86 (1, 31)</td>
<td>0.31 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Animacy</td>
<td>0.08 (1, 31)</td>
<td>0.03 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>0.02 (1, 31)</td>
<td>0.07 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>CONJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>11.16 (1, 31)</td>
<td>13.04 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>4.83 (1, 31)</td>
<td>4.97 (1, 36)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>7.34 (1, 31)</td>
<td>5.34 (1, 36)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td><strong>Gaze durations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>0.01 (1, 31)</td>
<td>0.00 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Animacy</td>
<td>0.41 (1, 31)</td>
<td>0.15 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>1.65 (1, 31)</td>
<td>1.33 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>CONJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>19.35 (1, 31)</td>
<td>17.99 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>14.85 (1, 31)</td>
<td>8.28 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>10.59 (1, 31)</td>
<td>9.20 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td><strong>Regression-path durations</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>10.53 (1, 31)</td>
<td>4.29 (1, 36)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Animacy</td>
<td>0.23 (1, 31)</td>
<td>0.03 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>0.72 (1, 31)</td>
<td>0.73 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>CONJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>19.04 (1, 31)</td>
<td>15.44 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>14.00 (1, 31)</td>
<td>5.79 (1, 36)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>14.92 (1, 31)</td>
<td>10.82 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td><strong>Probability of first-pass regressions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>8.14 (1, 31)</td>
<td>3.87 (1, 36)</td>
<td>&lt;.01, =.05</td>
</tr>
<tr>
<td>Animacy</td>
<td>2.24 (1, 31)</td>
<td>0.01 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>0.76 (1, 31)</td>
<td>0.33 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>CONJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>14.75 (1, 31)</td>
<td>15.41 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>0.43 (1, 31)</td>
<td>0.15 (1, 36)</td>
<td>&gt;.10</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>3.26 (1, 31)</td>
<td>2.38 (1, 36)</td>
<td>.08, .13</td>
</tr>
<tr>
<td><strong>Total times</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NP₂</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>22.10 (1, 31)</td>
<td>7.05 (1, 36)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>Animacy</td>
<td>7.74 (1, 31)</td>
<td>2.83 (1, 36)</td>
<td>&lt;.01, =.10</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>15.53 (1, 31)</td>
<td>4.66 (1, 36)</td>
<td>&lt;.05</td>
</tr>
<tr>
<td>CONJ</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Structural Ambiguity</td>
<td>23.42 (1, 31)</td>
<td>15.98 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Animacy</td>
<td>20.70 (1, 31)</td>
<td>9.42 (1, 36)</td>
<td>&lt;.01</td>
</tr>
<tr>
<td>Ambiguity × Animacy</td>
<td>19.55 (1, 31)</td>
<td>9.66 (1, 36)</td>
<td>&lt;.01</td>
</tr>
</tbody>
</table>
NP₂ in comparison with their unambiguous counterparts (29 ms in the Inanimate condition and 66 ms in the Animate condition). The ANOVA analyses revealed a main effect of ambiguity, but neither an effect of animacy nor an interaction between the two factors, suggesting that the ambiguous sentences imposed a greater processing load. At the conjunction, the Animate Ambiguous sentences induced substantially longer reading times, which indicated a garden path. In the ANOVAs at the conjunction, the interaction between ambiguity and animacy reached significance, and so did the two main effects.

The ANOVAs for the probability of first-pass regressions revealed a pattern similar to that of the regression-path durations. At NP₂, only the main effect of ambiguity was found, and only by participants. At the conjunction, the main effect of ambiguity was fully significant, while neither animacy nor the potential interaction reached significance. The pairwise comparisons at both NP₂ and the conjunction show that readers were more likely to return to earlier regions in the two ambiguous conditions (α = .05). Thus, the regression data suggest an overall increased processing load for the two ambiguous conditions compared with the unambiguous conditions.

**Total times.** The total time data (Figure 8) reflect both initial and re-processing, as they included first- and second-pass reading. Correspondingly, the interaction and main effects observed at the conjunction in the first-pass

![Figure 8](image_url)  
*Figure 8.* Means for total times for each condition at each word position in Experiment 2.
reading time measures occurred at both NP₂ and the conjunction in this
global measure of processing difficulty. At both NP₂ and the conjunction,
the Animate Ambiguous condition was slower than each of the other three
conditions, which did not differ from each other (\(z = .05\)).

Discussion

As in Experiment 1, we found a clear interaction of ambiguity and NP₂
animacy. There were substantial processing costs (reflected in raised first-
pass reading times and total times) at the conjunction for the Animate
Ambiguous condition compared to the other three conditions. Crucially, at
NP₂ Inanimate Ambiguous sentences had similar first-fixation and gaze
durations to the Unambiguous controls. In other words, there was no early
evidence of processing difficulty due to reanalysis at NP₂ in the Inanimate
Ambiguous condition, suggesting that the more complex CC analysis was
already available. Although the regression data showed that the ambiguous
sentences induced more processing difficulty overall than the unambiguous
sentences, reanalysis/re-ranking effects cannot account for the fact that there
were more regressions in the Animate Ambiguous sentences than in the
Inanimate Ambiguous sentences, because reanalysis/re-ranking would only
be required in the Inanimate Ambiguous condition at NP₂. Rather, the
ambiguous sentences, which contained two conceptual entities (NP₁ and
NP₂), seem to have placed heavier cognitive load on the comprehension
processes than the unambiguous counterparts, which had only one entity.
This might be due to maintaining multiple syntactic analyses in parallel, but
it could also be due to the higher referential load at the discourse level of the
Ambiguous conditions compared with the Unambiguous conditions. A third
possibility is suggested by the fact that there were more regressions at NP₂ in
the Animate Ambiguous condition compared to the Inanimate Ambiguous
condition. Recall that we did see a substantial garden path at the conjunction
in the Ambiguous Animate condition only. Thus, it is plausible that the extra
regressions at NP₂ in the Animate Ambiguous condition actually reflect
parafoveal processing of the conjunction, foreshadowing the upcoming
garden path effect.

In sum, serial parsing models predicted strong reanalysis effects, while
ranked parallel parsing models predicted little or no processing cost
associated with re-ranking. The minimal costs for the Inanimate Ambiguous
condition in the eye-tracking paradigm, taken together with the absence of
any processing costs in the stop-making-sense paradigm (Experiment 1), are
in sharp contrast with the strong garden path effects observed in the
Animate Ambiguous condition.
GENERAL DISCUSSION

The most important finding from the current set of experiments is that the processing costs for revising a Chinese relative clause (RC) structure as a complement clause (CC) structure differed, depending on whether a semantic cue (an inanimate noun) supporting the CC structure was available during the ambiguous region. Getting such a cue prior to syntactic disambiguation eliminated the garden path effect. The assumption that readers initially construct the RC analysis is not in doubt: Zhang et al. (2000) demonstrated this empirically, and their conclusion that the RC analysis is preferred is consistent with the syntactic ambiguity resolution mechanisms of current parsing models. Therefore, the contrast in processing difficulty between our Inanimate and Animate Ambiguous conditions poses a challenge for any parsing model that assumes that only one analysis remains available after a syntactic choice point.

Most of the constraint-based models (e.g., MacDonald et al., 1994; Spivey & Tanenhaus, 1998) assume that although syntactic representations are initially activated in parallel, a single analysis is quickly selected, based on support from the various constraints. To some extent, such models are similar to parallel parsers, with the dispreferred alternative(s) receiving much less activation than the selected structure. In a ranked parallel version, the parser could maintain multiple alternatives at varying levels of activation throughout the ambiguous region of a sentence.

In fact, a limited parallel version of a multi-constraint based theory, such as Tabor and Hutchins’ (2004) computational SOPARSE model, provides the best account of our findings. When reading the Verb NP, de NP, we expect that the parsing system initially attached NP as the direct object of the sentence initial verb. However, the lexical ambiguity of de introduced two structural possibilities: de could be attached to the existing structure as a relative clause marker or it could head a possessive modifier, with NP as the possessor. At de, we expect the first (RC) analysis to be ranked highest. However, at the next word NP, animacy supporting the second (CC) analysis becomes available in the Inanimate Ambiguous condition, raising the activation level of the CC analysis, so that syntactic disambiguation at the next word is consistent with a highly activated analysis. In the Animate condition, however, the CC remains in low activation at NP, while the animacy constraint further elevates the activation level of the RC. Our findings are a good demonstration of the SOPARSE model’s ‘digging-in’ effects: the longer the parser is committed to a misanalysis (i.e., the RC is supported one word longer in the Animate than in the Inanimate condition) the more severe a garden path is, as the misanalysis continue to gain activation strength via the ‘rich-get-richer’ mechanism.
Could a serial parser also account for our findings? Perhaps, but we do not think it likely. As noted above, reanalysis of the RC structure as the CC structure requires a complete overhaul rather than a selective revision. Therefore, it is not the type of reanalysis that proposed revision mechanisms could accomplish cheaply or quickly, if at all (Abney, 1989; Fodor & Inoue, 1994; Frazier & Clifton, 1998; Frazier & Rayner, 1982; Konieczny, 1996; Lewis, 1998; Pritchett, 1991; Stevenson, 1998; Sturt et al., 1999). Even allowing for new proposals about how reanalysis proceeds, it is difficult to image one that could easily transform the structure in Figure 1b into the structure in Figure 1a. An alternative that we have not considered here is underspecified parsing, in which a single parse tree is constructed incrementally, but leaves some structural relationships unspecified (e.g., Weinberg, 1993). Because the structural configurations of the RC and CC analyses are so different, we do not think that such an approach can salvage serial parsing for this structural ambiguity.

Can our limited parallel account explain other well-known garden path phenomena? Yes, but the explanatory capacity rests as much upon the lexical projection of structure as on parallelism. Consider Pritchett’s contrast between the easy garden path in (6) and the difficult one in (7). In (6), the verb’s argument structure motives postulation of two structural alternatives, a direct object complement and a sentential complement. While the former may initially be ranked higher, the water is consistent with both alternatives, and the sentential complement alternative should be easily recoverable for verbs like expect that occur frequently with sentential complements. In (7), drink also allows for two alternatives, a transitive and an intransitive structure. However, the water only supports the transitive analysis, allowing the intransitive structure to decay in activation. Central to these predictions is the assumption that alternative analyses will not linger unless they receive support from at least some of the available constraints.

(6) John expected the water to taste bad.
(7) After John drank the water tasted bad.

Our limited parallel account may remind some readers of a reanalysis account proposed by Ferreira and Henderson (1991b), in which the success of reanalysis depended upon an unused thematic representation not having decayed beneath recoverable levels. Thus, although the syntactic system constructed a single analysis, the thematic system offered covert parallelism that could be exploited for reanalysis. Like our account, their decay hypothesis predicts that reanalysis is difficult when the ambiguous region is long and thematic relationships are assigned early in the ambiguous region and maintained over several words. However, their account cannot predict our central finding – the difference between the Animate and Inanimate conditions – because their account assumes that reanalysis cannot be initiated until a syntactic error signal is received. Thus, for our stimuli,
reanalysis would be initiated at the conjunction in both Animate and Inanimate conditions, and no difference between the two conditions would be predicted.

In conclusion, our experiments suggest that sentence parsing in Chinese is at least somewhat parallel: multiple syntactic analyses remain activated during the ambiguous region of a sentence, with the more supported structure ranked higher. The re-ranking of alternatives could be very low-cost if the dispreferred interpretation receives enough activation early within the ambiguous region. Our results also support a multi-constraint based processor. Namely, all available information is utilised on a word-by-word basis to determine the ranking of structural alternatives of an ambiguous item, although the parser probably does not commit to a single analysis until the disambiguation point.

REFERENCES


APPENDIX A

Experimental stimuli. Within both the Inanimate and Animate sets, the words that distinguish the ambiguous and unambiguous conditions are given in parentheses, with the ambiguous condition first. Although these stimuli were used in both experiments, minor modifications (not shown here) to some items were made for Experiment 2, to increase naturalness for speakers of mainland Mandarin. The percentage in parentheses represents the percentage of RC completions for each item.

Inanimate

1.

粉刷((公寓/髒亂))的|房間|之後，|小王|還|打掃|了|客廳|。（100%）

fenshua((gongyu/zangluan)de|fangjian|zhihou,|xiaowang|hai|dasaol|eketing).
paint(apartment/dirty)(POSS/ATT)room|after,|Wang|also|clean|PERF|living room].

After painting the apartment’s/dirty room, Wang also cleaned the living room.

2.

興建((醫院/廣大))的|停車場|之後，|民眾|不再|抱怨|停車|問題。 （100%）

xingjian((yiyinguangda)de|tingchechang|zhihou,|minzhong|buzai|baoyuan|tingche|wenti].
construct((hospital/huge)(POSS/ATT)parking lot|after,|people|no longer|complain|parking|problem].

After constructing the hospital’s/huge parking lot, people no longer complained about parking problems.

3.

侵占((公司/公有))的|土地|之後，|小陳|還|企圖|掩飾|罪行|。（54%）

qinzhan((gongsi/gongyou)de|tudi|zhihou,|Chen|hai|qi|yanshi|zuixing].
invade((company/public)(POSS/ATT)land|after,|Chen|also|attempt|conceal|crime].

After invading the company’s/public land, Chen also attempted to conceal the crime.

4.

參觀((畫廊/著名))的|展覽|之餘，|小明|還|喜歡|看|歌劇。 （100%）

canguan((hualang/zhuming)de|zhanlan|zhiyu,|xiaoming|hai|xihuan|kan|geju].
visit((gallery/famous)(POSS/ATT)exhibition|while,|Ming|also|like|watch|opera].

While visiting galleries’/famous exhibitions, Ming also likes to watch operas.
5.

推銷(汽車/昂貴)的(音響)之餘，(小李)四處打工(賺/學費)。 (100%)

sell(car/expensive)(POSS/ATT)(stereo)while,Li many place work earn tuition.

While selling cars '/expensive stereos, Li works in many places to earn the tuition.

6.

準備((口試/無聊))的(考題)之餘，(小王)藉著(運動)放鬆(心情)。 (100%)

prepare(oral examination/boring)(POSS/ATT)question while,Wang by exercise relax mood.

While preparing for the oral examination's/boring questions, Wang relaxed by exercising.

7.

參加(社團/課外)的(活動)之餘，(學生)也(應該)兼顧(課業)。 (100%)

attend(clubs/extracurricular)(POSS/ATT)activity while,student also should care schoolwork.

While attending club/extracurricular activities, students should also care about schoolwork.

8.

拆除((學校/老舊))的(宿舍)之前，(校長)先(徵求)(學生)同意)。 (67%)

dismantle(school/old)(POSS/ATT)dormitory before,principal first ask for student's agreement.

Before dismantling the school's/old dormitory, the principle first asked for students' agreements.

9.

偽造((支票/複雜))的(簽名)之後，(小王)趁機(盜領)(大筆)公款)。 (100%)

forge(check/complex)(POSS/ATT)(signature)after,Wang take the chance embezzle much public money.

After forging the check's/complex signature, Wang took the chance to embezzle much public money.
1258    HSIEH ET AL.

10. 
抄襲(期刊/匿名)的文章之後，小明一直怕被揭發。 (96%) 

plagiarize(periodical/anonymouse)(POSS/ATT)article|after,Ming all the time|afraid|PASSIVE|expose.

After plagiarism the periodical's/anonymouse article, Ming was afraid of being exposed all the time.

11. 
收集(雜誌/新奇)的照片之餘，小陳也很喜歡攝影。 (83%) 

collect(magazine/novelty)(POSS/ATT)picture|while,|Chen also very|like|photography

While collecting magazine/new pictures, Chen also likes photography very much.

12. 
服從(師長/合理)的規定之餘，學生也需要自我約束。 (96%) 

follow(teacher/reasonable)(POSS/ATT)regulation|while,|student also need|self|restraint.

While following teachers' reasonable regulations, students also need self-restraint.

13. 
閱讀(報紙/有趣)的文章之餘，小王喜歡與朋友分享。 (100%) 

read(newspaper/interesting)(POSS/ATT)article|after,|Wang like|with|friend|share.

After reading newspaper/interesting articles, Wang likes to share with friends.

14. 
吃(麥當勞/不新鮮)的食物之後，小明覺得身體有點不舒服。 (100%) 

eat(Macdonald/stale)(POSS/ATT)food|after,|Ming feel|body|a bit|uncomfortable.

After eating Macdonald's/stale food, Ming felt a bit uncomfortable.
15. 

維修((教室/老舊)的|冷氣)之後，|那位|友善|接著|清掃|校園。 (100%)

weixiu|(jiaoshi/laojiu)de|lengqi|zhihou.|nawei|gongyou|jiezhe|qingsao|xiaoyuan|.
repair|(classroom/old)|(POSS/ATT)|air-conditioner|after|that|worker|continue|clean|campus|.
After repairing the classroom's/old air-conditioner, that worker continued to clean the campus.

16. 

清理((住家/陰暗)的|地下室|之前，|李四|請|朋友|一起|幫忙。 (92%)

qingli|(zhujia/yinan)de|dixiashi|zhiquan.|liosi|qing|pengyou|yiqi|bangmang|.
clean|(house/gloomy)|(POSS/ATT)|basement|before.|Lisi|ask|friend|together|help|.
Before cleaning the house's/gloomy basement, Lisi asked friends to help together.

17. 

破壞((公園/美麗)的|草坪|之後，|那群|小孩|絲毫|沒有|悔意。 (100%)

pohuai|(gongyuan/meili)de|caoping|zhihou.|naqun|xiaohai|sihao|meiyou|huiyi|.
destroy|(park/beautiful)|(POSS/ATT)|lawn|after|those|children|a bit|no|regret|.
After destroying the park's/beautiful lawn, those children had no regrets at all.

18. 

重建((災區/損損)的|房子|之後，|災民|希望|不再|有|天災。 (100%)

chongjian|(zaigu/huisun)de|fangzi|zhihou.|zaimin|xiwang|xiaozai|you|tianzai|.
rebuild|(disaster area/damaged)|(POSS/ATT)|house|after|victims|wish|no more|have|natural|disaster|.
After rebuilding the disaster area's/damaged houses, the victims hoped not to have any more natural disasters.
Animate

The plausibility rating for each item is provided in the second parentheses.

1. 虐待[(小孩/善良)]的保姆之后，黄姓夫妻被邻居检举。 (100%) (5.83)

nuedai(xiaohai/shanliang)de[baomu/zhihou],huangxing/fuqi|bei|linju|jianju|.
abuse(child/kind)|(POSS/ATT)nanny|after|,Huang|couple|PASSIVE|neighbor|accuse|.
After abusing the child’s/kind nanny, Mr. and Ms. Huang were accused by the neighbors.

2. 探望[(病人/无助)]的医生之后，李四捐赠给医院很多钱。 (96%) (5.90)

tanwang((bingren/wuzhu)de[yisheng/zhihou],lisijuan|yi|yuan|henduo|qian|.
visit(patient/hopeless)|(POSS/ATT)doctor|after|,Li|donate|hospital|much|money|.
After visiting the patient’s/hopeless doctor, Li donated much money to the hospital.

3. 拜访[(教授/优秀)]的学生之前，小林买了贵重礼物。 (100%) (5.59)

baifang(jiaoshou/youxiu)de|xuesheng/zhihqian,|xiaolin|mai|le|guizhong|liwu|.
visit(professor/outstanding)|(POSS/ATT)student|before|,Lin|buy|PERF|expensive|present|.
Before visiting the professor’s/outstanding student, Lin bought expensive presents.
4.

協助(醫生/年輕)的|護士|之後，|陳姓|助理|感到|很|疲勞|。

(xiezhu|(yisheng/nianqing)|de|hushi|zhihou,|chenxing|zhulilaodao|hen|piloai).

assist|(doctor/young)|(POSS/ATT)|nurse|after,|Chen|assistant|feel|very|tired).

After assisting the doctor’s/young nurse, assistant Chen felt very tired.

5.

保護|(主人/名貴)|的|獵犬|之餘，|那個|僕人|還要|做|家事|。

(baohu|(zhuren/minggui)|de|liequan|zhiyu,|nage|puren|haiyao|zuo|jiashil).

protect|(master/precious)|(POSS/ATT)|hunting dog|while,|that|servant|also|need|do|housework|.

While protecting the master’s/precious hunting dog, the servant also needs to do housework.

6.

訓練|(士兵/年輕)|的|將軍|之後，|總司令|發表|了|簡短|演說|。

(xunlian|(shibing/nianqing)|de|jiangjun|zhihou,|zongsiling|fabiao|le|jianduan|yanshuo).

train|(soldier/young)|(POSS/ATT)|general|after,|commander|give|PERF|short|speech|.

After training the soldiers’/young general, the commander gave a short speech.

7.

侮辱|(演員/知名)|的|導演|之後，|陳姓|評審|受到|大|批評|。

(wuru|(yanyuan/miming)|de|daoyan|zhihou,|chenxing|pingshen|shoudaol|ajia|piping).

insult|(actor/famous)|(POSS/ATT)|director|after,|Chen|critic|PASSIVE|everybody|criticize|.

After insulting the actor’s/famous director, critic Chen was criticized by everybody.

8.

陷害|(市長/漂亮)|的情婦|之後，|那些|歹徒|試圖|湮滅|證據|。

(xianhai|(shizhang/piaoliang)|de|qingfu|zhihou,|naxie|daitu|shitu|yanmie|zhengjiu).

incriminate|(mayor/beautiful)|(POSS/ATT)|mistress|after,|those|ruffian|attempt|destroy|evidence|.

After incriminating the mayor’s/beautiful mistress, those ruffians attempted to destroy the evidence.
9.

抢劫(银行/富有)的职员之后，小张内心觉得很不安。(100%) (5.14)

After robbing the bank's/rich staff, Zhang felt very uneasy at heart.

10.

责骂(学生/无辜)的老师之后，黄校长被/迫(公开/道歉)。(100%) (6.34)

After scolding the students' /innocent teacher, headmaster Huang was forced to publicly apologize.

11.

测验(徒弟/资深)的师傅之前，评审/详细说明了规则。(58%) (5.72)

Before testing the apprentice's/senior master worker, the judge explained the rules in detail.

12.

毁谤(总统/正直)之后，洪姓记者被/停职/处分。(96%) (5.31)

After defaming the President's/upright bodyguard, reporter Hong was suspended as punishment.

13.

仰慕(老板/美丽)的秘密之后，小王/常/故/意/亲近她。(92%) (5.14)

While admiring the boss's/beautiful secretary, Wang often intentionally got close to her.
14.

怠慢[客人/年長]的[司機]之後，[那位/年輕/服務生]非常懊悔。 (96%) (5.79)
slight[guest/elder](POSS/ATT)driver[after],[that/young/waiter]very/regretful.
After slighting the guest's/elder driver, the young waiter was very regretful.

15.

praise[players/dutiful](POSS/ATT)coach[while],[team/boss]decide/issue/reward.
While praising the players'/dutiful coach, the team boss decided to issue a reward.

16.

想念[孩子/慈祥]的[母親]之餘，[小王]立刻[搭/飛機]回家。 (100%) (5.66)
miss[child/kind](POSS/ATT)mother[while],Wang/immediately[take/airplane]go home.
While missing his child's/kind mother, Wang went home by airplane immediately.

17.

討厭[雇主/懶惰]的[僕人]之餘，[小李]決定[自己]完成[工作]。 (88%) (5.31)
taoyan[employer/lazy](POSS/ATT)servant[while],Li[decide]self/finish/work.
While hating the employer's/lazy servant, Li decided to finish the work himself.

18.

欺骗[主管/親切]的[助理]之後，[小明]終於[決定]承認[錯誤]。 (100%) (5.17)
deceive[director/cordial](POSS/ATT)assistant[after],Ming/finally[decide]admit/mistake.
After deceiving the director's/cordial assistant, Ming finally decided to admit his mistake.
19.

照顧(老人/窮苦)的鄰居之餘，小張還幫助很多孤兒。

while, Zhang also helped many orphans.

20.

暗殺(總統/不忠)的特務之後，李四潛逃到國外藏身。

After assassinating the President’s disloyal secret agent, Lisi absconded abroad to hide.