The structure-sensitivity of memory access: evidence from Mandarin Chinese

Brian Dillon*
Department of Linguistics, University of Maryland, College Park, MD 20742, USA
Department of Linguistics, University of Massachusetts, Amherst, MA 01003, USA

Wing Yee Chow
Department of Linguistics, University of Maryland, College Park, MD 20742, USA

Matthew Wagers
Department of Linguistics, Linguistics Research Center, University of California Santa Cruz, CA 95064, USA

Taomei Guo
State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, PR China

Fengqin Liu
State Key Laboratory of Cognitive Neuroscience and Learning, Beijing Normal University, PR China

Colin Phillips
Department of Linguistics, Program in Neuroscience and Cognitive Science, University of Maryland, College Park, MD 20742, USA

*Corresponding Author:

Brian Dillon
Department of Linguistics
226 South College
University of Massachusetts
150 Hicks Way
Amherst, MA 01003
brian@linguist.umass.edu
Abstract

In an attempt to better understand how linguistic dependencies are constructed in online sentence processing, we examined the processing of the Mandarin Chinese long-distance reflexive *ziji* using both the multiple-response speed-accuracy tradeoff (MR-SAT) paradigm and event-related potentials (ERPs). We manipulated the hierarchical/linear distance between the reflexive and its antecedent, forcing either local or long-distance interpretations. In Experiment 1, MR-SAT results suggest that dependencies with local antecedents are constructed more rapidly than those with long-distance antecedents. Experiment 2 used ERPs to provide converging evidence for this locality advantage in antecedent-anaphor relationships, and results showed that long-distance *ziji* interpretations elicited anterior negativities relative to local *ziji* interpretations upon reaching the critical anaphor. This pattern of results is not consistent with parallel consideration of all possible antecedents for *ziji*, as might be expected in a content-addressable memory architecture (McElree, 2000). Instead, we hypothesize that these patterns implicate a role for structured search during online sentence comprehension, whereby certain structural positions are accessed without regard to their semantic content during the construction of a *ziji* antecedent-anaphor dependency.
Introduction

In online language comprehension, the information contained in a sentence unfolds over time. Successful parsing thus requires mechanisms for integrating information across time. This is especially clear when one considers long-distance dependencies such as antecedent-anaphor relations, which involve relations between two non-adjacent (and possibly quite distant) elements in a sentence. An important goal for models of dependency-building processes is an explicit characterization of the mechanisms involved in retrieval and manipulation of linguistic structure in working memory. A basic distinction that has been made in this area is the contrast between content-addressable and search processes during memory retrieval operations (e.g. McElree, 2000). Convergent evidence from a number of different sources appears to implicate a significant role for content-addressable retrieval mechanisms in sentence comprehension. Evidence from time course analysis (McElree, 2000; McElree, Foraker, & Dyer, 2003), measures of retrieval interference (Drenhaus, Frisch, & Saddy, 2005; Van Dyke & McElree, 2006; Van Dyke, 2007; Vashisih, Brüssow, Lewis, & Drenhaus, 2008; Wagers, Lau, & Phillips, 2009), and patterns of processing difficulty (Badecker & Straub, 2002; Lewis, 1996) have all been taken to provide evidence that the parser relies upon a content-addressable memory architecture.

A related issue, which is distinct from the choice of architecture, is the nature of the information used to generate the set of candidates for retrieval. In the accessing linguistic memory, do comprehenders deploy a wide range of morphological and semantic features in parallel, or is structural information used to narrow the range of memories that might be retrieved at the point of memory access? The present study attempts to address this question by examining the relation between feature-based content-addressable memory retrieval and syntactically-structured access patterns in the online comprehension of the Mandarin long-distance reflexive ziji. In what follows we argue that dependencies between ziji and its antecedent are subject to a number of linguistic constraints that make structured retrievals a useful strategy for accessing linguistic memory. We suggest that ziji provides a useful test case for probing the mechanisms for memory access in language comprehension. We then present a time-course analysis of the processing of such dependencies using the speed-accuracy tradeoff (SAT; Experiment 1) and event-related potentials (ERP; Experiment 2) techniques. The observed pattern of results suggests that the construction of the antecedent-ziji dependency is faster when its antecedent occupies a local syntactic position, and that recovering a long-distance ziji interpretation is associated with greater anterior negativities in the ERP signal. We argue that this provides evidence that, though the parser may rely on a content-addressable memory architecture, it organizes retrieval events so as to ensure structure-sensitive access to preceding constituents.

Accessing linguistic memory
The present study is concerned with the working memory processes that underlie the retrieval of an antecedent for anaphoric expressions. The key issue is whether or not retrieving the antecedent of an anaphor like ziji proceeds by considering all possible antecedents in parallel, or if certain structural positions are evaluated prior to others. For our purposes we will consider any retrieval procedure that privileges structural information in accessing linguistic memory a structured search. Examples of structured search include the use of positional information as a hard, initial constraint on the retrieval procedure (Sturt, 2003), such that constituents that are otherwise good candidates for an antecedent are nonetheless not considered if they do not occupy a licit position. This may be contrasted with a parallel-access mechanism, which constructs the dependency between ziji and its antecedent by considering all possible antecedent positions simultaneously, retrieving the antecedent that best fits a range of syntactic, semantic, and morphological cues. The crucial difference between these two mechanisms is the nature of the information used to retrieve elements from memory. Structured search mechanisms narrow their retrieval cues to target particular structural positions, whereas a wider range of information is recruited to access antecedents in a parallel-access mechanism. Figure 1 contrasts the two types of access mechanism. Parallel-access is easily implemented in content-addressable memory systems, which allow direct access to all memory representations that match the relevant search cues by recalling all stored information that contains (or is indexed by) the relevant cues. In this system, positional or order information is generally not used to directly guide memory access, and instead, structural constraints are implemented as retrieval cues. This effectively gives them the status of violable constraints on the memory access procedure. Direct access mechanisms have been suggested to be the primary manner of memory access employed during sentence processing (McElree, 2000; McElree et al., 2003; Lewis & Vasishth, 2005).

The evidence that most directly bears on the question of the memory access mechanisms deployed in sentence comprehension is findings about the speed with which certain memory representations are accessed. One way of drawing inferences about the time course of memory access comes from speed-accuracy tradeoff (SAT) measures, which provide a relatively direct measure of the time course of information accrual. By providing information about task accuracy at multiple time points, SAT allows researchers to build a complete picture of the dynamics of task completion that dissociates the speed of completing the task (measured in the dynamics parameters, see below) and the accuracy of task completion (referred to as the asymptotic accuracy parameter). Traditional RT paradigms are limited in how informative they are about the dynamics of memory processes. Because participants can trade speed and accuracy in many standard judgment tasks (Wickelgren, 1976), merely estimating reaction time per condition (or a single RT/accuracy pair) can obscure differences between the probability of successfully terminating a process and the speed with which that process reaches completion. By tracking the availability of information as a function of time, the resulting SAT functions allow the experimenter to quantify both the speed and the probability of successfully completing a given process, and provide dissociable measurements of both.

Consider the process of forming the relation between an anaphor and its linguistic antecedent, which arguably involves contacting a stored memory representation for the antecedent (Dell, McKoon, & Ratcliff 1983; Nicol & Osterhout 1988; Bever &
McElree 1989; Gernsbacher 1989; but cf. Chafe 1974, Karmiloff-Smith 1980 for alternative views about the role of retrieval processes in processing anaphoric elements. If one considers the anaphor to be the probe, and its antecedent the target for a memory retrieval, then the predictions of different memory access mechanisms are fairly straightforward. The prediction for a direct access retrieval mechanism that uses all possible cues is that all memory representations corresponding to potential antecedents should be accessed with similar temporal dynamics. The access procedure will retrieve the target antecedent with the same speed regardless of its position relative to the anaphor and thus it will lead to constant SAT dynamics across sentences that vary in the structural position of the target antecedent. A structured search, by contrast, opens up the possibility that some representations are accessed before others by virtue of their syntactic position, and these extra access operations should be reflected in differences in temporal dynamics of the SAT function in response to changes in the structure of the search space (e.g., more intervening material or greater hierarchical distance to the retrieval target).

Similar time-course predictions have been tested using the SAT paradigm across various grammatical phenomena, including filler-gap dependencies (McElree, 2000; McElree et al., 2003), subject-verb thematic dependencies (McElree et al., 2003), pronoun antecedent resolution (Foraker & McElree, 2007), and verb phrase ellipsis (Martin & McElree, 2008, 2009). In each of these studies, the results indicate that information accrual during sentence processing has a constant rate regardless of the structural or linear distance from the probe point to the target, confirming the predictions of a parallel-access model of memory access. However, one important consideration raised by Martin & McElree (2008) and Wagers (to appear) is that in many of these configurations, the need to retrieve a given memory can be reliably anticipated before the hypothesized point of memory retrieval. For example, the process of integrating a *wh*-element with its associated gapped argument position might result in reaccess at some point prior to the gap site or the use of specialized storage structures that maintain privileged access to the *wh*-element (Wanner & Maratsos 1979; Frazier 1987; Frazier & Flores d’Arcais 1989). If such prospective processes are used in the formation of these dependencies, then any differences in access time due to structural distance between the *wh*-element and the gap site might be neutralized by the time comprehenders reach the gap site. As these studies have measured access speed at the gap site, then such prospective strategies would make it difficult to observe any dynamics differences.

However, not all SAT evidence provided to date is amenable to this line of argumentation. The verb-phrase ellipsis (VPE) cases studied by Martin & McElree (2008, 2009) are examples of a dependency that cannot reliably be anticipated before the point of retrieval (a point noted by Martin & McElree, 2008). Thus, any effects of memory access should be observed at the point of the elided verb phrase (where the authors measured the SAT function) in examples such as (1) below (from Martin & McElree 2008):

(1) a. The editor [*vp admired the author’s writing*], but the critics did not.
b. The editor [*vp admired the author’s writing*], but everyone at the publishing house was surprised to hear that the critics did not.
By manipulating the distance between the target VP and the site of ellipsis, Martin & McElree found that the amount of hierarchically intervening material had no effect on the dynamics parameters of the SAT functions they measured, confirming the predictions of a parallel-access mechanism. Instead, the effect of interpolating extra material manifested itself in a decreased likelihood of successfully completing the dependency (i.e. a lower asymptotic accuracy) in (1b) relative to (1a). Similarly, Foraker & McElree (2007) investigated the processing of cross-sentential anaphora in the SAT paradigm, which like VPE is arguably not predisposed to prospective processing routines. Again, the SAT results showed no effect of structural or linear distance on the speed of completing the anaphor-antecedent dependency. However, there are few structural constraints on where the antecedent for these types of cross-sentential dependencies can be found (see, e.g., Wasow 1972; Johnson 2001). For this reason it may not be entirely surprising that positional information is not used to guide memory access for these dependencies.

In addition to arguments about the time course of dependency completion, one possible second source of evidence that bears on mechanisms of memory access is the case of ‘grammatical illusions’ (Phillips, Wagers, & Lau, 2011). Grammatical illusions refer to the spurious sense of acceptability that arises in certain ungrammatical sentences. Some of these illusions have been attributed to parallel-access retrieval mechanisms. The ability to limit retrieval operations to only grammatically licit structural positions is not obviously consistent with parallel-access information retrieval. Moreover, the fact that parallel-access mechanisms do not limit their search space in such a manner is sometimes seen as an empirical advantage, offering a natural explanation for cases of spurious dependency formation (Vasishth et al., 2008; Wagers et al., 2009), as well as for patterns of difficulty in sentence processing (Lewis & Vasishth, 2005; Van Dyke, 2007). For example, it has been argued that the so-called ‘agreement attraction’ effects are an example of this (Badecker & Kuminiak 2007; Wagers et al., 2009). In comprehension, agreement attraction refers to situations where comprehenders occasionally fail to notice when a verb's agreement features mismatch the subject of the clause, due to the presence of another noun phrase in the sentence that matches the verb's agreement features, as in the runners that the driver wave to every day smiled. Direct access mechanisms that probe for specific feature values give a natural explanation to this phenomenon: illusions result from the occasional retrieval of partially feature matched elements, such as a structurally inappropriate target with appropriate morphosyntactic features required for a given dependency.

By the same reasoning, however, results that indicate accurate dependency completion in the face of feature-matched but structurally inappropriate items suggest the possibility of structurally targeted retrieval mechanisms. It has been shown that the initial stages of anaphor-antecedent dependency building in English reflexives such as himself and herself are not impacted by inaccessible but feature-matched antecedents in the same way (Nicol & Swinney, 1989; Sturt, 2003; Xiang, Dillon, & Phillips, 2009; Clackson, Felser & Clahsen 2011). However, this evidence usually comes from English reflexive dependencies, which have such narrow locality restrictions that one might question the generality of this apparent structural fidelity. Nonetheless, these data suggest that comprehenders can selectively retrieve the local subject, which is a suggestive source of
evidence about the memory access procedures used to resolve long-distance dependencies.

The current study: Chinese long-distance anaphors

Although there is substantial evidence for the use of parallel-access mechanisms in sentence processing, the evidence to date does not lead us to conclude that memory access during sentence comprehension exclusively relies on feature-based parallel access mechanisms. In the current study, we focus on one example of an unbounded long-distance dependency that is structurally constrained, requiring successful retrievals to be limited to certain structural positions, and which cannot reliably be anticipated ahead of the retrieval site. We propose that the most stringent test of the memory access mechanisms used during sentence processing come from dependencies of this kind; they provide a context where using structural cues to narrow the search space could confer advantages retrieval accuracy during online parsing processes.

The dependency we consider here is the Mandarin long-distance reflexive ziji. Upon reaching the reflexive element ziji, comprehenders need to initiate a retrieval for an antecedent. This is an entirely retrospective process, as there are no cues prior to ziji that signal the presence of a dependency (with the exception of inherently reflexive verbs in Chinese, see Jin 2003; Li & Zhou 2010). There are a number of key properties of ziji’s interpretation that make structured search a potentially useful option. These are i) syntactic constraints on possible linguistic antecedents, specifically they must be syntactically commanding subjects; ii) blocking effects; and iii) long-distance binding possibilities. In order to satisfy these constraints, (i) requires that ziji retrieve antecedents that occupy particular syntactic positions; (ii) requires that they stand in a certain configuration relative to each other, and (iii) requires that it do so in a manner that can accommodate a potentially unbounded search space. We argue that structured search—selective retrieval of NPs in increasingly distant subject positions—would provide an effective way to satisfy these constraints during the construction of a dependency between ziji and its antecedent.

The search space for ziji is unbounded because, unlike the English reflexives himself/herself, it does not require that its antecedent be in the same clause. Ziji is an example of the cross-linguistically well-attested class of long-distance reflexives, reflexive pronouns that may be bound outside of their minimal clause. Like their local counterparts, long-distance reflexives have structural requirements on their linguistic antecedents (Büring 2005). The local and long-distance nature of ziji can be seen in (2) and (3), respectively. Importantly, in (3), ziji can refer to either the local subject Lisi, or the distant subject Zhangsan.

(2) \[ \text{Lisi} \text{ nongshang-le ziji,} \]
    \[ \text{Lisi harm-PERF self} \]
    “Lisi harmed herself”

(3) \[ \text{Zhangsan} \text{ shuo Lisi nongshang-le ziji \_ziji} \]
    \[ \text{Zhangsan says Lisi harm-PERF self} \]
    “Zhangsan says that Lisi harmed him / herself”
Like many long-distance reflexives, *ziji* imposes a number of constraints on potential linguistic antecedents (Büring 2005; Huang, Li & Li 2009). There are significant syntactic constraints placed on antecedents: they must be subjects and they must be contained in dominating clauses (Huang & Liu, 2001). In addition to these syntactic constraints, there are a number of discourse-pragmatic constraints on the use of *ziji*. Antecedents must be animate and sentient, and must be prominent in the current discourse (Xue, Pollard & Sag 1994; Huang & Liu 2001). In the absence of an appropriate antecedent in the immediate sentential context, *ziji* may refer to the speaker, presumably as a reflex of the prominent discourse status that is automatically afforded to the speaker (Kuno 1972; Huang & Liu 2001). Though there are ongoing debates about the exact nature of *ziji*’s licensing conditions (Huang, Cole & Hermon, 2006), it is uncontroversial that resolving the antecedent-anaphor dependency requires the comprehender to systematically exclude inaccessible referents from consideration.

In addition to limiting a potentially unbounded search space to a subset of syntactically (or pragmatically) accessible antecedents, correct resolution of the antecedent of *ziji* also requires verifying that potential antecedents stand in particular syntactic relations to each other. In particular, long-distance third person subjects are blocked if a closer subject is first or second person. Consider the following sentences (from Huang & Liu 2001: p.2-3; example (6) modified from their (11), p.6):

(4) **Zhangsan renwei Lisi hen ziji**
Zhangsan think Lisi hate ziji
“Zhangsan thinks that Lisi hates himself/him.”

(5) **Zhangsan renwei ni hen ziji**
Zhangsan think you hate ziji
“Zhangsan thinks that you hates yourself/*him.”

(6) **Ni renwei Zhangsan hen ziji**
You think Zhangsan hate ziji
“You think that Zhangsan hates himself/?you.”

Sentences (4) and (5) illustrate the blocking effect for *ziji*: in (5), assessing the feature content of the local subject is necessary to determine whether or not the matrix subject is an acceptable antecedent. Although the matrix subject *Zhangsan* is generally considered an acceptable antecedent for *ziji* as in (4), when the local subject is a second-person pronoun *ni*, then this binding possibility is no longer considered acceptable (5). These examples show an instance of a person blocking effect, where the presence of a first- or second-person antecedent blocks binding from higher subjects (Huang & Liu 2001). Similarly, it has been reported that singular subjects block binding from higher plural subject positions (Tang 1989; Huang & Tang 1991). Crucially, if you reverse the order of the subjects, as in (6), *Zhangsan* becomes an acceptable antecedent for *ziji*. Thus, it is not only the presence of a pronoun with first- or second-person features that impacts the availability of other subjects, but also its relative position to those subjects.
Experimental evidence to date suggests that there is a locality bias in processing *ziji*. For example, Li & Zhou (2010) provide ERP evidence that long-distance binding of *ziji* elicits a P300/600 component relative to local or ambiguous binding of *ziji*. We take up this finding in more detail in our discussion of Experiment 2. In addition to ERP results, a number of cross-modal priming studies have shown a bias towards reactivating the local antecedent upon reaching *ziji*. In one study, Gao, Liu & Huang (2005) showed that in sentences such as “the teacher asked the journalist to respect *ziji*”, reaction times for probes related to the object the journalist were faster than reaction times for probes related to the matrix subject the teacher. Liu (2009) replicated and extended this result by varying the latency of the probe word relative to *ziji*, probing for lexical decisions at 0ms, 160ms and 370ms after the offset of *ziji*. In a very suggestive finding, reaction times to probes related to the local subject were faster at the 0ms SOA, while the pattern reversed at 160ms, showing greater reactivation for the long-distance antecedent. For these two SOAs, the opposite pattern of activation was observed for the pronoun *ta*. At the longest SOA, both antecedents were equally activated. This result is suggestive of a structured search, with early consideration of the local antecedent and relatively delayed consideration of the long-distance antecedent only for *ziji*, and not for other pronominal elements. However, the reaction times for the lexical decision task do not provide a direct window on the time course of activating antecedents for *ziji*.

We argue that taken together, the structural constraints and blocking configuration constraints on potential antecedents provides a strong context to observed a structured search, understood as a process in which potential binding positions are sequentially considered as potential antecedents. Although it is true that fast memory access for online processing is an important factor that cuts in favor of parallel-access mechanisms (c.f. Lewis, Vasishth & Van Dyke 2007), presumably the parser’s ability to reliably recover the correct interpretation is a factor in determining an optimal online strategy for constructing a given linguistic dependency. By limiting consideration to certain structural positions before considering all semantic and pragmatic features in parallel, the task of finding an antecedent for *ziji* can arguably more straightforwardly respect structural constraints as well as blocking-type constraints. If only structural cues are initially deployed, then the blocking constraints can be naturally implemented by terminating the retrievals if the local subject contains a ‘blocking’ feature specification. If, under even these circumstances, it can be shown that parallel-access mechanisms are also employed for memory retrievals, then a strong case can be made for an entirely parallel-access retrieval mechanism across all levels of sentence processing. If, on the other hand, structured search is implicated in the processing of *ziji*, then it would suggest that the parser can engage in a structured search of linguistic memory for some linguistic long-distance dependencies.

**Experiment 1**

Experiment 1 employed the multiple-response speed-accuracy tradeoff (MR-SAT; Wickelgren, Corbett, & Dosher, 1980) procedure to provide time-course evidence about whether or not *ziji* accesses all potential antecedent positions in parallel, or whether or not certain positions are checked before others. Figure 1 shows both a schematized comparison between a serial, structured access to potential antecedents, an alternative
mechanism, direct-access using semantic features. I follow the same logic employed in prior SAT studies (McElree 2000; McElree et al 2003; Foraker & McElree 2007; Martin & McElree 2008, 2009). If all possible antecedent positions are considered in parallel for *ziji*, then local and non-local antecedents should show similar processing speed (identical dynamics parameters in the SAT function). If, on the other hand, certain positions are checked prior to other positions, then we should observe different processing speeds for local and long-distance bindings. We also predict that due to the nature of the blocking effects described above, any speed advantage should favor local antecedents. The availability of the long-distance binder depends on the nature of the local binder, and structured would convey an advantage only if the local position was checked first. Furthermore, the linear recency of the local subject may itself confer an advantage over the long-distance subject.

**Method**

**Participants**

Twenty college students from Beijing Normal University participated in the experiment. Data from 2 participants were excluded due to atypical (sigmoidal) SAT curves. The remaining 18 participants included 10 females, and had a mean age of 23.5 years. Each participant completed six 1-hour sessions spaced at least a day apart, in addition to a 1-hour practice session for familiarization with the multiple-response speed-accuracy tradeoff procedure (MR-SAT; see below). All participants were native Mandarin Chinese speakers and had normal or corrected-to-normal vision. They were paid 35 RMB per hour for their participation in the experiment.

**Materials**

We investigated the processing of Mandarin sentences that contained a matrix attitude verb and an embedded transitive complement clause. Two experimental factors were manipulated. One was the position of an animate subject; it was either the subject of the matrix clause, the subject of the local clause, or not present. The second factor was the identity of the object, which was either *ziji*, an acceptable definite NP object, or an unacceptable definite NP object. Since *ziji* forms a long-distance binding dependency with its antecedent, the predictions of a structured search are that the position of the target antecedent should impact processing time. Control conditions with an acceptable definite NP (thus with no binding dependency) were included to ensure that any processing speed differences that are observed are not due to baseline differences due to the sentential context. Because a definite NP object does not form a binding dependency with its animate subject, no processing speed differences should obtain when the position of the animate NP is manipulated for control conditions.

This design provided three critical reflexive conditions designed to investigate the processing of *ziji* (conditions 1-3 in Table 1 below). Based on the position of the animate subject, *ziji* either took a long-distance antecedent (*LD antecedent* condition; 1), a local antecedent (*local antecedent* condition; 2), or had no antecedent in the sentence (*no antecedent* condition; 3). In these conditions, the animacy of the subjects was the factor...
that controlled which antecedent *ziji* was forced to take. As noted above, an animate NP in either the main clause or embedded subject NP position can function as a grammatical antecedent for *ziji*. In addition to the critical *ziji* condition, two control sentences were constructed with definite descriptions instead of *ziji* in the embedded object position (*LD control, local control, no antecedent control*; conditions 4-6). The well-formed control conditions replaced *ziji* with a full NP that was a plausible object of the embedded verb (e.g., *the batsman* in conditions 4-6 below); whereas the unacceptable control conditions replaced *ziji* with a full NP that was an implausible object of the embedded verb (e.g., *glasses* in conditions 7-9 below), resulting in a semantic anomaly. The anomaly conditions (conditions 7-9) and the conditions with no animate subjects (conditions 3, 6 and 9) were included for purposes of *d’* scaling, as described below, and so they were not directly reported. The primary experimental contrast was the effect of subject animacy on the speed and accuracy of processing sentences that contained *ziji*, and sentences that contained acceptable definite NP objects.

Therefore, each set of items consisted of three critical (conditions 1-3) and six control sentences (conditions 4-9). The control sentences additional helped to prevent participants from forming strategies based on the pattern of animate and inanimate referents that preceded the critical object NP. Thus, because of the structure of the control conditions, neither the presence of *ziji* nor the acceptability of the continuation was predictable from the preamble. This was done to ensure that all processing measures reflect processes initiated at the object NP itself. All three conditions consisted of a main clause that contained a verb of reporting, and an embedded transitive clause. Additionally, to increase complexity and difficulty of the task, a temporal adverbial clause was interpolated between the embedded subject and the embedded verb. In all conditions, an animate NP was also used as the subject of the temporal adverbial phrase. However, since it occupied a structural position that does not c-command *ziji*, it was not a grammatical antecedent for *ziji*. In the local antecedent and control conditions, the main clause NP was always a ‘media’ noun (e.g. book, documentary, memo) to ensure compatibility with the meaning of the main clause verb. In the no antecedent condition, both structurally accessible NPs were inanimate, and therefore this condition did not contain a grammatical antecedent for *ziji*. None of the inanimates used in any position could be construed metonymically; metonymic interpretations of inanimates (i.e. *the newspaper* being used to refer to the employess of the newspaper) may be used as antecedents for *ziji*.

Forty sets of the 9 sentence types (5 acceptable and 4 unacceptable) were generated. The 360 sentences were equally distributed in 6 presentation lists, one for each of the 6 sessions, to minimize the repetition of content material within a session. Crucially, no two instances of *ziji* sentences (conditions 1-3) from the same set were included in a single presentation list. Within a session, each participant viewed 206 sentences, of which 60 were drawn from the current study. Since only one third of target sentences contained *ziji*, the critical *ziji* conditions comprised around 10% of all sentences within and across sessions. The order of presentation within a session was randomized.

Procedure
We employed the multiple-response SAT paradigm, following Wickelgren, Corbett & Dosher (1980). Stimulus presentation, timing, and response collection were all carried out on a personal computer using the Linger software by Doug Rohde (available at http://tedlab.mit.edu/~dr/Linger/). Each trial began with a 500 ms fixation cross presented in the center of the screen. Each word appeared in the center of the screen for 400 ms, followed by 200 ms of blank screen. All words were presented using simplified Chinese characters, and the last word of each sentence was marked with a period (。). At the onset of the final word, a series of 18 auditory response cues (50 ms, 1000 Hz tone) was initiated. The cues occurred every 350 ms, and the final word of the sentence remained on the screen. Participants were trained to initially respond by pressing both response keys simultaneously to indicate an undecided response. They were then trained to give a response after each tone, and to switch their response to either the ‘accept’ or ‘reject’ key as soon as they could. Importantly, they were also trained to modify their responses if their assessment changed. During the 1-hour practice session, participants were told that some of the sentences were complex, but nevertheless were meaningful sentences. Each participant performed six 1-hour sessions, and in each they saw one of the lists of materials. The order of lists was randomized across participants.

<table>
<thead>
<tr>
<th>#</th>
<th>Condition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>LD antecedent</td>
<td>张教练 表明 那篇报导 [在 团队 未能 发挥 水准 的时候] 低估了 自己 Coach Zhang say [that report [when team not perform well time] underestimate ziji]  “Coach Zhang says that that report underestimated self when the team was doing poorly.”</td>
</tr>
<tr>
<td>2</td>
<td>Local antecedent</td>
<td>回忆录 表明 张教练 [在 团队 未能 发挥 水准 的时候] 低估了 自己 Auto-biography say [coach Zhang [when team not perform well time] underestimate ziji]  “The auto-biography says that coach Zhang underestimated self when the team was doing poorly.”</td>
</tr>
<tr>
<td>3</td>
<td>No antecedent</td>
<td>*回忆录 表明 那篇报导 [在 团队 未能 发挥 水准 的时候] 低估了 自己 *Auto-biography say [that report [when team not perform well time] underestimate ziji]  <strong>“The auto-biography says that that report underestimated self when the team was doing poorly.”</strong></td>
</tr>
<tr>
<td>4</td>
<td>LD control</td>
<td>张教练 表明 那篇报导 [在 团队 未能 发挥 水准 的时候] 低估了 那位击球手 Coach Zhang say [that report [when team not perform well time] underestimate that batsman]  “Coach Zhang says that that report underestimated the batsman when the team was doing poorly.”</td>
</tr>
<tr>
<td>5</td>
<td>Local control</td>
<td>回忆录 表明 张教练 [在 团队 未能 发挥 水准 的时候] 低估了 那位击球手 Auto-biography say [coach Zhang [when team not perform well time] underestimate that batsman]  “The auto-biography says that coach Zhang underestimated the batsman when the team was doing poorly.”</td>
</tr>
<tr>
<td>6</td>
<td>No antecedent control</td>
<td>回忆录 表明 那篇报导 [在 团队 未能 发挥 水准 的时候] 低估了 那位击球手 Auto-biography say [that report [when team not perform well time] underestimate that batsman]  “The auto-biography says that that report underestimated the batsman when the team was doing poorly.”</td>
</tr>
<tr>
<td>7</td>
<td>LD control anomaly</td>
<td>*张教练 表明 那篇报导 [在 团队 未能 发挥 水准 的时候] 低估了 眼镜 *Coach Zhang say [that report [when team not perform well time] underestimate glasses]  <strong>“Coach Zhang says that that report underestimated the glasses when the team was doing poorly.”</strong></td>
</tr>
<tr>
<td>8</td>
<td>Local control anomaly</td>
<td>*回忆录 表明 张教练 [在 团队 未能 发挥 水准 的时候] 低估了 眼镜</td>
</tr>
</tbody>
</table>
**Auto-biography say [coach Zhang [when team not perform well time] underestimate glasses]**

**“The auto-biography says that coach Zhang underestimated the glasses when the team was doing poorly.”**

**Table 1:** Summary of conditions in experiment: Critical *ziji* conditions. Critical conditions are 1-2 and 4-5; conditions 3 and 6-9 were included for purposes of *d*' scaling (see text).

**Data Analysis**

To derive the full time-course information in SAT analysis, *d*’ scores are calculated by comparing an acceptable and an unacceptable condition at each of the response tones, both within and across subjects. The resultant series of *d*’ values at each time point *t* is fit using a shifted exponential function:

\[
d' = \lambda(1 - e^{-\beta(t-\delta)}) , \quad t > \delta, \\
d' = 0 , \quad \text{otherwise}
\]

Here, *d'* is the standard signal-detection measure of discrimination: *d'* = *z*(hits) - *z*(false alarms) (Wickens, 2001). The shifted exponential in (7) describes the growth of accuracy over time *t* (in ms) with three parameters: asymptote (*λ*), rate (*β*), and intercept (*δ*). The current experiment used common scaling in its design: in order to derive the *d*’ scores for LD and local antecedent conditions, their hit rates for the critical conditions (Conditions 1 and 2) were scaled against the pooled false alarm rate at the corresponding time lag from the three ungrammatical controls (Conditions 7-9), following McElree & Griffith (1998). Both critical *ziji* and control conditions were scaled using this pooled scaling, to allow a more straightforward comparison of any potential differences in the resulting time courses that may arise. All differences between *ziji* and control conditions stem from differences in the hit rates between the two. One reason for adopting this measure, instead of scaling against the no-antecedent *ziji* condition was the somewhat high acceptability of the no antecedent condition (see below). If the no antecedent condition were used for scaling purposes, the false alarm rate used in the *d'* calculations would have been high, leading to low *d'* scores (less than 1 in most cases). For such low scores, small variations in hit rate would case very large changes in the observed *d’*, leading to unreliable SAT function estimates. An additional benefit is that pooled scaling produced SAT curves for the *ziji* conditions and the control conditions with comparable asymptotic accuracy, ensuring that neither floor nor ceiling effects could be responsible for any observed differences. Thus, this scaling measure allows direct comparison between the critical *ziji* conditions and their non-*ziji* acceptable control counterparts (conditions 4 and 5), and we can straightforwardly estimate the dynamics of the acceptable judgment alone (i.e., the successful completion of the dependency).

In order to determine whether the SAT functions for these conditions differed in asymptote (*λ*), rate (*β*), or intercept (*δ*), the analysis proceeded in two steps: a model
selection analysis and a parameter estimation analysis (Li & Smith 2009). In the model
selection analysis, the best fit model was determined using the adjusted $R^2$-statistic using
a hierarchical model-testing scheme over the averaged and individual data, an approach
pursued in prior work on SAT in sentence comprehension (McElree et al 2000, 2003;
Foraker & McElree 2007; Martin & McElree 2008, 2009). In the parameter estimation
analysis, only the fully specified model was considered, and any differences between the
critical conditions on the parameters of interest were assessed using familiar hypothesis
testing measures. In all analyses the SAT function given above was fit to the measured $d'$
values at each time lag. In order to obtain parameter estimates, it is necessary to fit non-
linear regressions of the SAT function (7) against the observed $d'$ score. In the present
study, we primarily employ Gibbs sampling, a Monte Carlo method for approximating
the posterior of a distribution of two or more random variables. Gibbs sampling is often
used as a form of Bayesian statistical inference (Gelman & Hill 2005). For the current
analysis, Gibbs sampling was used to estimate the posterior distributions over the three
SAT parameters, and the best-fit model for a given set of data was chosen by selecting
the median of the resulting posterior distributions. Note that this method of fitting SAT
curves contrasts with the method often employed in the psycholinguistic SAT literature
(STEPIT fits, McElree & Griffith 1998; McElree et al. 2000, 2003; Martin & McElree). In
this approach, the SAT function is fit with an iterative hill-climbing algorithm (Reed,
1976; similar to STEPIT, Chandler, 1969), and fit quality was assessed by an adjusted
$R^2$-statistic, which measures the variance accounted for by the fit weighted by a penalty
for an increasing number of model parameters. In almost all cases the two methods
provide identical results. We present the Gibbs sampling method because in addition to
providing information about the estimate of a value, it provides an estimate of the
variance associated with that value. However, to ensure that the estimation technique did
not change the qualitative pattern of results, we also report STEPIT fits in the parameter
estimation section for purposes of comparison. A direct comparison of STEPIT and
Bayesian fits, as well as code for implementing the models here, is available at
http://people.umass.edu/bwdillon/.

For all Monte Carlo simulations, 4000 samples were used as a burn-in period,
followed by 15,000 iterations of the sampler. 3 parallel sampling chains were run in each
simulation, and convergence was checked by evaluating the potential scale reduction
factor (‘$\hat{r}$’-hat’) statistic (Gelman & Hill 2005). For all reported parameter values, the
potential scale reduction factor was effectively 1, indicating satisfactory convergence of
the MCMC chains.

In addition to model-fitting on the $d'$ data, analysis was performed on participant
and item mean final acceptance rates (the empirical accuracy), which was obtained by
taking the average rate of acceptance over the last response point.

Data from two participants were excluded due to unreliable dynamics estimates. The empirical $d'$
scores from these participants appeared to be better fit by a sigmoidal rather than an exponential function, leading to unrealistically large and unreliable differences in the critical conditions in the crucial intercept and rate parameters when fit
with the SAT function in (7).

Empirical Accuracy Analysis
The rate of acceptance of for the critical *ziji* conditions was 85% for the LD antecedent condition, 81% for the local antecedent condition, and 48% for the no antecedent condition. The rate of acceptance for the corresponding acceptable control conditions was 91%, 87% and 91%. A mixed-effects logistic linear regression was fit with orthogonal fixed-effect contrasts for *locality* (comparing long-distance versus local antecedent conditions) and *binding* (comparing both *ziji* conditions with antecedents to the no antecedent condition), as well as random slopes for subject, item, and session. For critical *ziji* conditions, there was a significant effect of *binding* ($\beta = -1.3, z = -17.5, p < .0001$) and a marginal effect of *locality* ($\beta = .28, z = 1.94, p < .06$). For control conditions, there was a significant effect of *locality* only ($\beta = .70, z = 3.97, p < .0001$).

For unacceptable controls, accuracy for long-distance, local, and no antecedent controls was 99%, with no differences between these conditions.

*Model Selection Analysis*

For the model selection analysis, individual participants’ data was fit with a series of nested models separately for control and *ziji* conditions. This analysis compared long-distance and local animate configurations. The model-fitting analysis pitted a series of nested models (including shared or separate parameters for the two conditions of interest for each of the intercept, rate and asymptote) against each other on adjusted $R^2$, following McElree et al (2003) and Liu & Smith (2009). This was done by participants to determine if extra parameters led to a significantly better model fit, as measured by adjusted $R^2$-statistic. Adjusted $R^2$, shown in (2), gives an estimate of the variance accounted for by the (non-linear) regression with the SAT function, weighted by the number of parameters $k$ used in constructing that curve. In (8), $d$ refers to the observed $d'$ value, $\hat{d}$ refers to the predicted $d'$ value, and $n$ refers to the number of data points.

$$ (8) \quad R^2 = 1 - \frac{\sum_{i=1}^{n} (d_i - \hat{d}_i)^2 / (n-k)}{\sum_{i=1}^{n} (d_i - \bar{d})^2 / (n-1)} $$

For the critical *ziji* conditions, adding separate asymptote parameters for LD and local antecedents led to a reliable increase in adjusted $R^2$ across participants ($\Delta \mu = 0.03 \pm 0.01; t(17) = 3.9, p < 0.01$). Compared to the two-asymptote baseline, the addition of an extra rate parameter led to a small but reliable increase in adjusted $R^2$ ($\Delta \mu = 0.002 \pm 0.0005; t(17) = 4.3, p < 0.001$). Any model that included an extra intercept parameter led to a significantly poorer fit on adjusted $R^2$. Thus the best fitting model across participants allotted separate asymptotes and separate rate parameters for local and long-distance antecedents for *ziji* ($2\lambda - 2\beta - 1\delta$).

For control conditions, the addition of an extra asymptote for long-distance and local controls led to a marginally significant increase in adjusted $R^2$ ($\Delta \mu = 0.032 \pm 0.016; t(17) = 2.1, p < 0.055$). Neither additional rate parameters nor additional intercept parameters for LD and local control environments led to a reliable increase in adjusted $R^2$. Across participants, the best-fitting model for control conditions had shared dynamics parameters for LD and local conditions, but separate asymptotes ($2\lambda - 1\beta - 1\delta$).
The model-fitting analysis suggests that the two critical *ziji* conditions are processed at different speeds, as allotting separate rate parameters $\beta$ to LD and local antecedent conditions led to a reliable increase in adjusted $R^2$. However, with the model-fitting approach one might be concerned by the possibility of over-fitting, a concern magnified by the small differences in adjusted $R^2$ among the best fit models for any given participant’s data (Liu & Smith 2009). For this reason it is important to ensure that the parameter estimates across participants display a consistent ordering (see, e.g., McElree et al 2003) by directly analyzing the estimated SAT parameters across participants.

*Parameter Estimation Analysis*

The focus of the parameter estimation analysis, following Liu & Smith (2009), was the magnitude of the difference between parameters for local and long-distance antecedent configurations. Here this was done by fitting fully saturated models ($2\lambda - 2\beta - 2\delta$) for each participant and testing for consistent ordering of parameters across participants. In addition to presenting absolute parameter values for the processing of local and LD conditions, we also present the differences between the two values of a given parameter across these conditions. This can be understood as the advantage enjoyed by one condition over the over in speed (for differences in dynamics parameters) or accuracy (for differences in the asymptote).

Parameter estimates for each of the participants included in the analysis were estimated using the Bayesian method described above, although we also present fits from the hill-climbing algorithm employed in other SAT work (McElree 2000; McElree et al 2003; Martin & McElree 2008). For Monte Carlo fits, the median of the resulting posterior distribution for each participant was chosen as the estimate and submitted to further analysis. In addition to rate and intercept estimates, we report the composite processing measures speed ($\beta^{-1} + \delta$), a measure that helps to guard against parameter tradeoffs in the dynamics estimates (Carrasco, Giordano & McElree, 2006).

The average advantage for local antecedents in the *ziji* and control conditions is summarized numerically in Table 2. Values above 0 indicate faster processing speeds or greater accuracy for local antecedent configurations. Fully saturated model fits to average data are presented in Figures 2 and 3, scaled to proportion of asymptote in order to highlight the differences in the growth portion of the SAT curve. Figures 4-7 show the mean parameter values for accuracy, rate, intercept and speed for both *ziji* and control conditions. Finally, Figure 8 summarizes the differences in all dynamics parameters for both critical *ziji* and control conditions, for each individual participant.

The results from both methods of parameter estimation are largely in agreement. However, the STEPIT fits appear in general to trade greater speed in the rate parameter for local configurations for later intercepts; no such trade-off occurs for Bayesian fits. However, for the compound speed measure, where these dynamics parameter tradeoffs are controlled, the estimates from both approaches are closely matched.

<table>
<thead>
<tr>
<th>Parameter</th>
<th><em>Ziji</em></th>
<th><em>Control</em></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bayesian fit</td>
<td>STEPIT fit</td>
</tr>
<tr>
<td>Asymptote ($d’$)</td>
<td>-0.14 (± .15)</td>
<td>-0.15 (± .14)</td>
</tr>
</tbody>
</table>
Table 2: Difference in parameter estimates between local and long-distance configurations on the critical *ziji* and control conditions. Values greater than 0 indicate a processing advantage for local antecedents. Standard errors by subject are in parentheses.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Local</th>
<th>Control</th>
<th>Local</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Rate ($s^{-1}$)</td>
<td>0.42 (± 0.20)</td>
<td>0.86 (± 0.35)</td>
<td>0.04 (± 0.15)</td>
<td>-0.04 (± 0.13)</td>
</tr>
<tr>
<td>Delta ($s$)</td>
<td>-0.01 (± 0.01)</td>
<td>-0.32 (± 0.15)</td>
<td>0.00 (± 0.02)</td>
<td>0.02 (± 0.02)</td>
</tr>
<tr>
<td>Speed ($s$)</td>
<td>0.09 (± 0.04)</td>
<td>0.09 (± 0.04)</td>
<td>-0.02 (± 0.05)</td>
<td>0.03 (± 0.04)</td>
</tr>
</tbody>
</table>

For the critical *ziji* comparison, parameter estimates from STEPIT fits indicated no reliable difference in estimated asymptote ($\lambda$) parameters between the two critical *ziji* conditions. There was a reliable difference between the *ziji* conditions for rate ($\beta$) parameters, as well as for the compound speed measure ($\beta^{-1} + \delta$) ($t(17) = 2.47, p < 0.05$ and $t(17) = 2.52, p < 0.05$, respectively). In addition, for the intercept (delta) parameter there was a marginal effect of condition ($t(17) = -2.09, p < 0.06$). For Bayesian fits, the pattern of findings was qualitatively similar. There were no reliable differences for local versus long-distance configurations in either the asymptote parameter or the intercept parameter. There were, however, reliable effects in both the rate parameter ($t(17) = 2.21, p < 0.05$) and the compound speed measure ($t(17) = 2.28, p < 0.05$).

On control conditions, STEPIT fits revealed no reliable differences between the local and long-distance configurations for any parameter. Likewise, Bayesian estimates of the parameters revealed only a marginal effect in the asymptote ($\lambda$) parameter ($t(17) = -1.75, p < 0.1$), and no other reliable differences between the control conditions in any of the dynamics parameters.

It is unclear whether the processing speeds of the control conditions and the *ziji* conditions are directly comparable, as the processes involved in rendering a judgment in the two cases are distinct. Nonetheless, planned pairwise comparisons were performed to directly compare parameter differences for the *ziji* conditions to those of the control conditions. For STEPIT fits, the locality advantage in *ziji* conditions over control conditions was significant for both rate and speed parameters ($\beta$: $t(17) = 2.98, p < 0.01$, $\beta^{-1} + \delta$: $t(17) = 2.29, p < 0.05$). The same pattern was found for the Bayesian fits, although this effect was only marginally significant for the compound speed measure ($\beta$: $t(17) = 2.69, p < 0.05$, $\beta^{-1} + \delta$: $t(17) = 1.74, p < 0.1$).

One possible concern about the analysis presented here is that the fully saturated models presented here posit separate asymptote parameters for the critical *ziji* conditions. If the true model does not contain separate asymptotes for each condition, a potential worry is that non-significant trends in accuracy drive the observed dynamics effects. However, the accuracy effect for the critical *ziji* conditions was nearly reliable in the empirical accuracy analysis, and so it is unlikely that modeling a given participant’s data with a single asymptote model is appropriate. Furthermore, since almost every participant presented data that was significantly better fit with a two-asymptote model, modeling the data with a single asymptote would have led to significant distortions of the estimates of the dynamics parameters. For participants who have greater asymptotic accuracy for LD antecedent conditions, a single-asymptote model inappropriately increases the processing speed estimates for long-distance configurations, and conversely for participants who show the opposite pattern of accuracy.
Discussion

Both analyses of the SAT data indicate that the critical ziji conditions differed in processing speed, and that the control conditions did not. Model-fitting analyses showed that for the critical ziji comparison, the best fitting model attributed separate rate ($\beta$) parameters and asymptote parameters ($\lambda$) to local and long-distance antecedent conditions. This suggests that the two ziji conditions differ in speed and accuracy. The control comparison, on the other hand, showed no such advantage for extra dynamics parameters; only separate asymptotic accuracy parameters for each control condition reliably improved model fit. This analysis was supplemented with an analysis of the resulting parameter estimates, in order to check for consistent ordering of parameters. This analysis showed that in both rate and speed measures of the ziji conditions, the local antecedent was processed reliably faster across participants than was the LD antecedent condition. No such difference was observed in the control conditions, although there was a marginal effect of asymptotic accuracy for these conditions.

Thus the two analyses of the SAT data show that long-distance and local ziji conditions differ reliably in processing speed. Furthermore, the analysis of the direction of the parameters indicates that when ziji’s antecedent is local, it is retrieved faster than when it is contained in a dominating clause. The most direct measure of this effect is the compound speed measure indicated in Figure 8. The speed measure combines both rate and intercept parameters, and provides a measure of average processing speed in seconds. It can be seen that on average, local ziji antecedent configurations have approximately a 90ms processing advantage over the long-distance configurations.

This finding supports the structured search hypothesis. Comprehenders appear to access the local antecedent before accessing the long-distance subject, a finding that is compatible with a number of architectural implementations. If ziji activates its antecedent with purely structural information then this finding is expected; the sort of iterated structural access mechanism that would derive this processing profile is presented in Figure 1. It can be seen that, on the assumption that the local antecedent is more active, then recovery of the long-distance antecedent requires multiple retrieval processes.

This finding stands in contrast to previous SAT work, which has generally found that memory access in sentence processing occurs with essentially constant time. The evidence presented here provides initial evidence that the parser does engage in structured search. Although they stand in contrast to prior SAT findings, this finding about the access profile of ziji converges nicely with prior findings on the processing of ziji, suggesting both serial access (Gao et al 2005; Liu 2009) and a locality preference (Liu 2009; Li & Zhou 2010).

One unexpected finding was the high acceptance rate of the no antecedent ziji condition, which participants accepted on 48% of trials. It is known that in the absence of linguistic binding, ziji can refer to the speaker (Huang & Liu 2001). In the context of this experimental task, it is unclear how participants established the point of view of the sentences they read. The perceived point of view (or source of communication) of the sentence is presumably the main factor that controls the recoverability of this indexical interpretation of ziji. Since this aspect was not controlled, however, it is unclear how participants perceived these unbound ziji sentences. The relationship of unbound ziji to
the linguistically bound examples presented here is an important question that we address in Experiment 2.

Experiment 2 investigated long-distance and local ziji bindings using event-related brain potentials (ERP) to attempt to provide converging evidence for structured access in comprehension. ERP is an attractive technique to complement the SAT evidence presented here. Although the SAT evidence provides crucial support to the contention that it is processing speed that differs between local and long-distance ziji binding, the nature of the processes that slow computation for long-distance antecedents remains unclear. The ERP signal can be broken down into a number of distinct waveform components, each of which is thought to index different aspects of linguistic processing. Furthermore, the multidimensional ERP measure may be able to illuminate the relationship between the linguistically bound interpretation of ziji and the unbound interpretation of ziji.

**Experiment 2**

Experiment 1 suggests that the process of retrieving a long-distance antecedent for ziji occurs more slowly than it does for local antecedents. The explanation for this effect on a structured search account is that the local subject is first selectively retrieved by virtue of its syntactic position, and then rejected due to a poor fit with the animacy requirements of ziji. The extra processing steps involved in rejecting the local antecedent and accessing the long-distance antecedent cause a slow-down in processing time when compared to the local antecedent. However, the SAT data do not indicate the exact processes that cause this slow-down. Experiment 2 investigates the processing of ziji using ERP in order to determine which stages of processing are impacted when a long-distance dependency needs to be constructed. An important feature of ERPs is that they provide a multidimensional view on language processing, with multiple components indexing arguably distinct cognitive processes that support linguistic processing. For this reason, in determining which ERP component indexes the locality advantage in the SAT materials, we may gain further insight into the nature of the processed speed advantage enjoyed by local antecedents relative to long-distance antecedents.

This feature of ERPs also allows us to ask a secondary question about the relationship of unbound ziji to linguistically bound ziji. Although we have described these interpretations as distinct, it may be the case that they are in fact processed in a uniform manner. For instance, if the primary licensing conditions on ziji are entirely discourse-based, then retrieving the speaker as an antecedent might not be qualitatively different than retrieving other prominent antecedents from the discourse. An alternative possibility is that qualitatively different processes are employed to construct these two interpretations. This is suggested by the structured search diagrammed in Figure 1, which limits retrieval to linguistically represented positions. On this view, processing the unbound interpretation requires a qualitatively different set of processing routines. If the difficulty related to long-distance and no antecedent conditions is distributed across distinct ERP components, then a case for separate processing of linguistically bound and unbound ziji dependencies can be made. On the other hand, if they are both reflected in similar components, then the case for distinct processing strategies for the two interpretations is weakened.
The exact functional significance of many ERP components is a matter of active research and debate, and there appears to be a many-to-many mapping between hypothesized linguistic processes and ERP components. To a first approximation, however, semantic and syntactic processing are characteristically associated with different ERP components. For instance, words that are anomalous with respect to morphological or syntactic features have long been recognized to generate the P600 response, a late posterior positivity that generally peaks around 600 ms post-stimulus (Friederici, Pfeifer, & Hahne, 1993; Hagoort, Brown, & Groothusen, 1993; Osterhout & Holcomb, 1992), as well as an earlier anterior negativity termed the (E)LAN (Kluender & Kutas 1993; Coulson, King, & Kutas, 1998; Friederici et al., 1993; Hagoort, Wassenaar, & Brown, 2003; Lau, Stroud, Plesch, & Phillips, 2006; Neville, Nicol, Barss, Forster, & Garrett, 1991). Although these components are sensitive to similar factors (morphosyntactic wellformedness), they are widely regarded as reflecting distinct processes in the computation of a syntactic representation (Friederici 1995; Hahne & Friederici 1999; Hagoort 2003; Bornkessel & Schlesewsky 2006). On the other hand, semantic anomalies in otherwise syntactically well-formed sentences typically elicit a central negativity around 400 ms known as the N400 (Kutas & Hillyard, 1980; Kutas & Federmeier, 2000; Lau, Phillips, & Poeppel, 2008). These characterizations are far from exceptionless, however. For example, there appear to be instances of “semantic” error that engender P600 responses (Kolk, Chwilla, van Herten, & Oor, 2003; Kim & Osterhout, 2005; Kuperberg, 2007), as well as N400 effects that have been linked to processing case anomalies (Hopf, Bayer, Bader, & Meng, 1998) or syntactic reanalysis effects (Bornkessel, Schlesewsky & McElree, 2004).

Previous ERP work on the processing of ziji has demonstrated that long-distance antecedents elicit a posterior positivity (P300/600) relative to local antecedents, supporting the locality bias for ziji (Li & Zhou 2010). However, this study differed in important ways from the materials in the SAT study. In Li and Zhou’s study, both the local and the long-distance antecedent positions contained animate NPs that were featurally compatible with ziji. In their manipulation, the semantics of the embedded verb served to disambiguate the antecedent of ziji, as the verb was either inherently reflexive or anti-reflexive (Li & Zhou 2010). Thus it is contextual or thematic knowledge that comprehenders need to draw upon to exclude the local subject from consideration, but the dependency between the local subject and the anaphor is well-formed. In contrast, the dependency between the inanimate local subject and the anaphor in the SAT study is not formally well-formed, which drives a further retrieval for an acceptable antecedent as in panel A of Figure 1. Because of the differences between the two studies, it is difficult to conclude from this that the difficulty associated with retrieving the long-distance antecedent is reflected in the P300/P600 effect that Li and Zhou found. In order to more directly compare the ERP reflex of ziji’s locality bias with the SAT results, we present an ERP study of the processing of long-distance and local ziji in environments that were more parallel to those tested in Experiment 1.

Method

Participants
Twenty-four college students from Beijing Normal University participated in the experiment, including 13 females (mean age 22). The 24 participants (13 females) had a mean age of 22, were all healthy, native speakers of Mandarin Chinese with no history of neurological disorder, and all were strongly right-handed based on the Edinburgh handedness inventory (Oldfield, 1971). All participants gave informed consent and were paid 50 RMB/hour for their participation, which lasted around 2½ hours, including set-up time.

Materials

The experimental materials consisted of three conditions designed to investigate the processing of *ziji* (conditions 1-3 in Table 3 below). These three conditions matched the three critical *ziji* conditions in Experiment 6 on a number of important dimensions. *Ziji* either took a long-distance antecedent (*LD antecedent* condition; 1), a local antecedent (*local antecedent* condition; 2), or had no antecedent in the sentence (*no antecedent* condition; 3). As noted above, an animate NP in either the main clause or embedded subject NP position can function as a grammatical antecedent for *ziji*. As in Experiment 6, all three conditions consisted of a main clause that contained a verb of reporting, and an embedded clause, of which *ziji* was the object. Likewise, the binding possibilities for *ziji* were manipulated by manipulating the animacy of the subject NPs: LD antecedent conditions contained an animate matrix subject, and local antecedent contained an animate embedded subject. All other pre-critical NPs were inanimate, and in the no antecedent condition, both subject NPs were inanimate, and therefore this condition did not contain a grammatical antecedent for *ziji*. In order to ensure that the critical *ziji* was not in a sentence final position, all conditions followed *ziji* with a conjunction (e.g. *ye* ‘also’, *que* ‘but’) and second clause continuation. Although the processing profile of conjoined anaphors is known to be different from that of single argument anaphors (see, e.g. Harris, Wexler & Holcomb 2000; Burkhardt 2005), none of the conjunctions in the experimental materials can be used as NP-level conjunctions, ensuring that participants did not interpret *ziji* as part of a conjoined NP.

120 sets of the 3 sentence types were generated. The 360 sentences were equally distributed in 3 presentation lists and subjects were assigned to each list in a Latin Square fashion. Within each list, the 120 targets were interleaved with 240 unrelated fillers of similar complexity, and the list was divided into six blocks of 40 sentences. The order of the 6 blocks was randomized across subjects. Of the 360 sentences in each session, half were considered to be acceptable, half were considered to be unacceptable.

<table>
<thead>
<tr>
<th>#</th>
<th>Condition</th>
<th>Example</th>
</tr>
</thead>
</table>
| 1 | LD binding | 厨师表示[油锅曾经烫伤了自己]，所以辞职了。
Chef say [deep-fryer scalded *ziji*], so (pro) resigned.
“The chef said the deep fryer scalded him, so he resigned.” |
| 2 | Local binding | 医疗报告表示[厨师曾经烫伤了自己]，所以辞职了。
Medical report say [chef scalded *ziji*], so (pro) resigned.
“The medical report said the chef scalded himself, so he resigned.” |
Table 3: Summary of conditions in Experiment 2.

**Procedure**

Participants were comfortably seated in a dimly lit testing room around 100 cm in front of a computer monitor. Sentences were presented one word at a time in white letters on a black background in 30 pt simplified Chinese characters. Each sentence was preceded by a fixation cross. Participants pressed a button to initiate presentation of the sentence, which began 1000 ms later. Each word appeared on the screen for 400 ms, followed by 200 ms of blank screen. The last word of each sentence was marked with a period, and 1000 ms later a question mark prompt appeared on the screen. Participants were instructed to read the sentences carefully without blinking and to indicate with a button press whether the sentence was an acceptable Mandarin sentence. Each experimental session was preceded by a 12-trial practice session that included both grammatical and ungrammatical sentences. Participants received feedback and were able to ask clarification questions about the task during the practice session. The experimental session was divided into six blocks of 75 sentences each. Breaks were permitted after each block as necessary.

**EEG Recording**

EEG was recorded from 32 Ag/AgCl electrodes, mounted in an electrode cap (Electrocap International): midline: Fz, FCz, Cz, CPz, Pz, Oz; lateral: FP1/2, F3/4, F7/8, FC3/4, FT7/8, C3/4, T7/8, CP3/4, TP7/8, P4/5, P7/8, O1/2. Recordings were referenced online to the right mastoid, and re-referenced offline using linked mastoids. An additional electrode was placed on the left and right outer canthus, and above and below the left eye to monitor eye movements and eye blinks. EEG and EOG recordings were amplified and sampled at 1000 Hz using a bandpass filter of 0.1-70 Hz. Impedances were kept below 5 kΩ.

**EEG Analysis**

All analyses were conducted over single trial epochs, consisting of the 100 ms preceding and the 1000 ms following the critical presentation of zijī, normalized using a 100 ms pre-stimulus baseline. In order to exclude motion and ocular artifacts, normalized epochs with activity greater than ± 50 µV were removed, as were trials that had a peak-to-peak voltage difference of greater than 100 µV in the EOG (Luck 2005). The total rejection rate with these criteria was 19%, ranging between 18%-20% across critical conditions. Averaged waveforms were filtered offline using a 20 Hz low-pass filter for presentation purposes; however, all analyses were performed on unfiltered data. The latency intervals that were analyzed statistically were chosen based on previous conventions in the ERP sentence processing literature: 0-200 ms, 200-400 ms, 400-600 ms, 600-800 ms, and 800-1000 ms. Regions of interest were defined as follows: left anterior (FT7, F3, FC3), midline anterior (FZ, FCZ, CZ), right anterior (F4, FC4, FT8),
left posterior (TP7, CP3, P3), midline posterior (CPZ, PZ, OZ), and right posterior (CP4, P4, TP8).

In order to assess the reliability of the effects elicited by the experimental manipulations, we employed linear mixed effects (LME) modeling (Pinheiro & Bates 2000). There are a number of advantages to this approach over traditional approaches. One important advantage in the current context is that this approach readily accommodates missing data (Gelman & Hill 2005; Baayen et al 2008). Since epoch rejection rates tend to vary a good deal across participants in ERP studies, LME models are an attractive analysis option.

Analysis proceeded separately for each time interval, with average values for each epoch within that time interval estimated for the electrodes included in the analysis. The best-fit model for each time interval was selected by hierarchical model comparison, by adding terms for fixed effects and checking for a significant increase in model likelihood using a \( \chi^2 \)-test. The logic and interpretation of this approach is qualitatively similar to the hierarchical interpretation of ANOVA results that is often reported in ERP papers: fixed effects (factors) and their interactions are evaluated to ensure that they explain a significant amount of variance, and the resulting best-fit model is then evaluated to determine the nature of these effects. For the present experiment, there were two fixed effects of electrode position, \textit{anteriority} (with the levels \textit{posterior} and \textit{anterior}) and \textit{laterality} (with the levels \textit{left}, \textit{midline}, and \textit{right}), as well as two orthogonal experimental contrasts for binding and locality. Main effects and interactions for electrode position and order were entered into the models prior to the planned orthogonal contrasts using simple difference coding. The \textit{binding} contrast assessed the effect of having a linguistic antecedent in the sentence, and compared the no antecedent condition to both LD and local antecedent conditions (coding: .5 for the antecedent conditions, -.25 for both LD and local antecedent conditions). The \textit{locality} contrast assessed the effect of having a local antecedent, and directly compared local and LD antecedent conditions (coding: .5 for LD antecedents, -.5 for local antecedents), 0 for no antecedent). Random intercepts for trial and participants were included; random intercepts for items and random slopes for the experimental fixed effects did not significantly increase the likelihood for any model, and so were not included in the final model fits described here. This exclusion did not change the pattern of results found. For each time interval, we describe the best-fitting model and the fixed effect coefficients. All \( p \)-values for linear model coefficients were estimated using MCMC methods as implemented in the \textit{LanguageR} R package (Baayen 2008), with \( n=10,000 \) samples.

\textit{Results—Behavioral Data}

The average rate of acceptance was 81\% for the long-distance binding condition, 73\% for the local binding condition, and 51\% for the no-binding condition. A logistic linear mixed effects model with crossed random intercepts for participants and items revealed significant effects of both the \textit{binding} contrast (\( \hat{\beta} = -.81, z = -14.0, p < .0001 \)) as well as the \textit{locality} contrast (\( \hat{\beta} = .47, z = 4.2, p < .0001 \)). An analysis of reaction times did not reveal any significant differences among the conditions.

\textit{Results—ERP Data}
The grand average ERPs are presented in Figure 9. A summary of the fixed effects for binding and locality are presented in Table 4. In the 0-200ms time interval, there were no significant fixed effects for experimental contrasts. In the 200-400ms time interval, the best-fitting model included an interaction of binding with anteriority and laterality. Resolving this interaction revealed a broadly distributed negativity for unbound ziji, relative to the two bound ziji conditions. Numerically, this effect was largest over posterior regions. In the 400-600ms time window, analysis revealed a negativity for the LD antecedent relative to the local antecedent binding conditions, which reached significance in the mid and left anterior ROIs. A negativity was also observed in the 600-800ms time window, with a focus over the mid and right anterior ROI. No experimental fixed effects were observed between 800-1000ms.

<table>
<thead>
<tr>
<th></th>
<th>0-200ms</th>
<th>200-400ms</th>
<th>400-600ms</th>
<th>600-800ms</th>
<th>800-1000ms</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Binding</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left Ant.</td>
<td>-</td>
<td>-.64 ± .44</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mid Ant.</td>
<td>-</td>
<td>-1.18 ± .44**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Right Ant.</td>
<td>-</td>
<td>-.91 ± .44*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Left Post.</td>
<td>-</td>
<td>-.79 ± .44†</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Mid Post.</td>
<td>-</td>
<td>-1.22 ± .44**</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Right Post.</td>
<td>-</td>
<td>-1.07 ± .44*</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td><strong>Locality</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Left Ant</td>
<td>-</td>
<td>-</td>
<td>-1.02 ± .43*</td>
<td>-.73 ± .47</td>
<td>-</td>
</tr>
<tr>
<td>Mid Ant.</td>
<td>-</td>
<td>-</td>
<td>-.91 ± .43*</td>
<td>-.90 ± .47†</td>
<td>-</td>
</tr>
<tr>
<td>Right Ant.</td>
<td>-</td>
<td>-</td>
<td>-.75 ± .43†</td>
<td>-.99 ± .47*</td>
<td>-</td>
</tr>
<tr>
<td>Left Post.</td>
<td>-</td>
<td>-</td>
<td>-.84 ± .43†</td>
<td>-.72 ± .47</td>
<td>-</td>
</tr>
<tr>
<td>Mid Post.</td>
<td>-</td>
<td>-</td>
<td>-.18 ± .43</td>
<td>-.13 ± .47</td>
<td>-</td>
</tr>
<tr>
<td>Right Post.</td>
<td>-</td>
<td>-</td>
<td>-.31 ± .43</td>
<td>-.48 ± .47</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4: Table of experimental fixed effects (coefficients in µV, with standard error). Experimental fixed effects only shown if the best-fit model included a significant interaction of experimental effect with anteriority and laterality. † = p < 0.1, * = p < 0.05, ** = p < 0.01.

Discussion

The ERP results revealed two distinct components associated with the processing of ziji in our materials. For the binding contrast, an early negativity associated with the processing of unbound ziji was observed in the 200-400ms time window. This negativity had a primarily central-posterior distribution. This is consistent with an N400 effect, though visual inspection of the ERPs suggests that the peak of this negativity is slightly earlier than the canonical N400 effect. For the locality contrast, a qualitatively different component was observed, differing both in time course and distribution. LD antecedents
caused an increased negativity over the anterior ROIs in the 400-600ms window, with some effects also being observed in the subsequent 600-800ms window. The distribution of this component is consistent with a (L)AN effect, although the time course is somewhat later than the canonical LAN effect.

An important insight from these results is that the processes involved in recovering a long-distance bound interpretation and revising to an indexical interpretation of ziji are at least partially distinct. This is consistent with structured search because it supports an assumption that is implicit in the structured search account; namely, that linguistic binding and extrasentential reference are processed differently. The exact interpretation of these components for the processes used to resolve unbound and linguistically bound ziji turns on assumptions about the identity of the component and its functional significance. For example, if the component that indexes the difficulty for unbound ziji is best characterized as an N400, then it might be understood as indexing the difficulty in accessing information in the lexicon (e.g. Lau et al 2010). On this interpretation, this effect suggests that ziji is less predicted in the no antecedent conditions. However, an important distinction between indexical and bound ziji is that in order to recover an indexical interpretation, participants must change the point-of-view of the sentences (Huang & Liu 2001; Anand 2006). If this is necessary when encountering ziji in the context of a discourse with no animate participants, then the negativity may reflect this process of context-shifting.

The difficulty associated with long-distance interpretations of ziji was observed in an anterior negativity at approximately 400ms post-stimulus. Anterior negativities of this sort have been linked to a variety of morphosyntactic violations as well as working memory difficulty associated with forming long-distance dependencies (Neville et al 1991; Friederici et al., 1993; Kluender & Kutas, 1993; King & Kutas, 1995; Coulson et al, 1998; Hagoort et al 2003). If structured search proceeds as in Figure 1, then the observed anterior negativity is compatible with either interpretation of the (L)AN component. On the assumption that the LAN indexes a morphological violation, then it might be caused by the mismatch between the animacy features of the local subject and ziji when the local subject is retrieved and temporarily considered. Alternatively, if the observed negativity reflects working memory difficulty (as argued by Kluender & Kutas 1993), then it may be indexing the fact that more retrieval operations (or more difficult retrievals) are necessary to recover the long-distance antecedent, leading to a greater anterior negativity. Both of these interpretations are compatible with structured access, but further work is necessary to distinguish between them.

It is important to note that these results stand in contrast to those reported by Li & Zhou (2010), who did report that long-distance interpretations of ziji caused both P300 and P600 effects. However, in their study, the local subjects were always animate and compatible with ziji; long-distance interpretations of ziji were forced by manipulating the verb semantics. In this case, if retrieval of the local subject yielded a feature-matched antecedent, it is unclear what processing would be necessary to reanalyze to a long-distance interpretation. If structured search terminates upon finding a compatible antecedent, then it is possible that entirely different reanalysis processes were required to recover the distant antecedent in that study, causing a P300/600 rather than an anterior negativity. If this reasoning is correct, then this account generates testable predictions for future work about when the different ERP components will be observed in processing
In particular, the difficulty associated with the structured search necessary to find an initial, acceptable antecedent for *ziji* should be reflected in a LAN. On the other hand, difficulty associated with selecting among multiple antecedents based on contextual information should be associated with P600 effects.

Overall, the ERP effects provide an alternative source of evidence in support of structured search by confirming that fully grammatical (and more acceptable) long-distance interpretations of *ziji* are associated with extra processing difficulty. The (L)AN associated with the long-distance interpretation of *ziji* may reflect either the feature mismatch with the local subject, or the extra memory retrievals necessary to reaccess the distant antecedent. Interestingly, the response that indexes the shift to an indexical interpretation of *ziji* preceded the difficulty associated with long-distance interpretations. Furthermore, since no LAN was seen in this condition, it is possible that the decision to pursue an unbound interpretation of *ziji* obviated the need for the parser to check the local subject position at all. This may indicate that participants quickly decide to pursue a shift in point-of-view rather than attempting to bind *ziji* to an antecedent within the sentence. The nature of this decision remains unclear, but one possibility is that the lack of any animate discourse entities may have biased readers in favor of a first-person interpretation of *ziji*. This would amount to a heuristic use of top-down discourse information to disambiguate between an unbound *ziji* and a bound *ziji* interpretation. Only for the latter parse does the parser need to engage a structured search. However, this possibility remains speculative, and further study is required to determine how comprehenders decide to pursue a bound interpretation of *ziji* prior to actually constructing an antecedent-anaphor dependency.

**Discussion**

**Summary of results**

The current study presented time-course and electrophysiological evidence on the processing of antecedent-anaphor dependencies involving the Mandarin Chinese long-distance reflexive *ziji*. Results demonstrate that in constructing a *ziji*-antecedent dependency, local antecedents are accessed more rapidly than long-distance antecedents, as measured in the SAT function’s rate and speed parameters. Control conditions without *ziji* did not show this difference in processing dynamics. In Experiment 2, the ERP record indicated that this locality preference was indexed by a left-lateralized anterior negativity, providing converging support for structured access in comprehension. The observed negativity may index an increase in the working memory processes needed to construct the long-distance interpretation, or it may reflect a mismatch in features between the local subject and *ziji*. Difficulty related to processing an unbound *ziji* was reflected in a central-posterior negativity that was distinct in topography and time course from the difficulty related to recovering a long-distance interpretation. This pattern suggests that the processes involved in extrasentential reference and linguistic binding are in part distinct. Interestingly, the negativity in response to unbound *ziji* appeared earlier in the ERP record than did the negativity related to distant antecedents. Lastly, in both experiments the long-distance binding conditions were judged as more acceptable than the local binding conditions (though this reached only marginal significance in Experiment 1),
showing a dissociation between offline acceptance and patterns of online processing difficulty. In both experiments, the preferred long-distance interpretations are associated with greater processing difficulty.

Structured search in retrospective dependencies

The time course evidence presented in Experiment 1 is compatible with theories of dependency construction that engage structured search under certain circumstances. The finding that local antecedent dependencies show an advantage in SAT dynamics suggests that at the point of initiating the antecedent-anaphor dependency, the information necessary to complete the local dependency is accessed earlier than the information required to complete the long-distance dependency. This finding is consistent with existing experimental work on *ziji* that has consistently demonstrated a locality bias (Gao et al 2005; Liu 2009; Li & Zhou 2010). In particular, the results presented here fit very closely with the findings in Liu (2009). In this study, the local antecedent was most active at 0ms post-*ziji*, and at 160ms post-*ziji*, it was instead the most distant antecedent that was active. Taken together, these findings are not consistent with the view that all *ziji*-antecedent dependencies are considered in parallel, which would predict constant SAT dynamics for all antecedent positions. These results stand in contrast to previous SAT studies, which suggested that processing advantages due to locality only impacted asymptotic accuracy (McElree 2000; McElree et al 2003; Martin & McElree 2008). There are several possibilities for why this might be so.

One possibility is that all licit antecedent locations are sampled in a serial manner using only structural cues, as suggested by panel A in Figure 1. In this view, upon reaching *ziji* the parser first attempts to find an antecedent by retrieving the local subject position without regard to its semantic content. If the local subject is unacceptable, the process is repeated in progressively more distant, structurally higher subject positions until a match is returned. Constructing the antecedent-anaphor dependency in this way provides a natural way of implementing the structural constraints on *ziji*, as well as respecting the blocking constraints, because the retrievals could be prematurely terminated if the local subject contains ‘blocking’ morphosyntactic features. Note that this mechanism is not at odds with the conclusion that the memory structure employed in sentence processing is a content-addressable architecture (e.g. McElree 2000; McElree et al 2003). In particular, the time-course data for the *ziji* dependency may reflect iterated retrievals from an underlyingly content-addressable architecture, with the assumption that only structural cues (but not semantic/pragmatic cues such as animacy or sentience) are used in the initial retrieval. In a content-addressable implementation of this sort of structured search, the parser might deploy iterated retrievals with varying sets of cues to target increasingly distant subject positions. For instance, initial retrievals might use cues for ‘subject NP’ and ‘same clause’, and subsequently eliminate the same clause cues if retrieval does not return a licit antecedent. The extra retrievals required for this type of access procedure might be engaged when the need to faithfully implement structural constraints outweighs the need for speed in dependency formation. The relevant difference between the *ziji* dependency and the VP-ellipsis dependency, on this view, is that the linguistic constraints on *ziji* cause the parser to engage a potentially slower access
in order to maintain grammatical fidelity. In the case of VPE, this may not be necessary due to a lack of substantial constraints on where the antecedent VP can be located.

Alternatively, the difference in dynamics might reflect an advantage for linguistic material contained within the local clause relative to material outside the local clause. In other words, the local subject may be available more quickly by virtue of its being contained within the local clause, which is still in the process of being parsed when *ziji* is encountered. In contrast to the account presented above, this account does not entail a fully serial sampling routine. Instead, it is serial in a somewhat more restricted manner: retrievals are initially limited to the local clause, and if failure results, then a more global retrieval is engaged. This account makes no specific reference to the constraints on *ziji*, and does not necessarily make predictions about accuracy in dependency completion beyond the local clause. This account is consistent with studies on sentence recall that suggest that the local clause has a privileged role in online sentence processing (Jarvella, 1971; Jarvella & Pisoni, 1970). This may also be equivalent to the claim that that some elements remain concurrently available in the focus of attention while others are displaced and must be later retrieved (McElree, 2006; Jonides, Lewis, Nee, Lustig, Berman & Moore, 2008). If the local subject remains in the focus of attention, while the long-distance subject requires retrieval, then the observed distinction in dynamics would be predicted. This interpretation seems less likely, in light of findings from list memory experiments that indicate that the focus of attention is extremely limited in size and scope, corresponding to just one task-relevant encoding (McElree & Dosher, 1989; McElree, 1998). If only one element occupies focal attention before *ziji* is processed, it is likely to be the verb. However, the available data on the contents of the focus of attention is somewhat sparser for connected linguistic representations, which have considerably richer structure than do lists. It is known that full clauses are sufficient to displace information about their embedding environment (McElree et al., 2003; Wagers & McElree, 2009). But it is presently unknown whether information about the subject is focally available during the initial processing of verb-phrase internal arguments.

Both of the retrieval-based accounts presented above share a common feature in that they require the use of positional or structural information in retrieval. In order to implement an access procedure that serially retrieves subject positions, or that preferentially accesses information in the local clause, the positional information inherent in those two specifications needs to be available to guide retrieval. It is possible, in principle, to empirically distinguish the two accounts, as they make distinct predictions for antecedents for *ziji* that differ in hierarchical distance beyond the local clause. A serial-sampling account predicts that the rate difference observed here should also be observed when comparing antecedents that are 2 versus 3 clauses distant. Alternatively, if the current dynamics differences reflect the status of the local clause as a privileged domain of processing, the rate difference should no longer obtain when hierarchical distance is manipulated beyond the local clause. Future work will examine the role of hierarchical distance beyond the local clause boundary in an attempt to tease apart these competing hypotheses.

Reanalysis
An alternative account of the current results is that the difference in processing dynamics between dependencies with local and long-distance antecedents reflects a reanalysis from a local interpretation to the long-distance interpretation. This is consistent with some linguistic accounts of ziji that have suggested that ziji as a local anaphor is distinct from ziji when its antecedent is distant, based on differences in meaning and pragmatics of usage in these different environments (e.g., Huang & Liu, 2001; Anand 2008). Reanalysis has been noted to cause delays in SAT dynamics parameters (McElree 1993; Bornkessel, Schlesewsky & McElree, 2004), and thus if the first option that comprehenders attempt upon recognizing ziji is the local-antecedent interpretation, they should fail and require reanalysis in the long-distance antecedent conditions.

Note, however, that the LAN we observed in the ERP record is not generally associated with reanalysis. One ERP component that is commonly associated with reanalysis processes is the P600 (Friederici et al., 2002; Hagoort, 2003; Hopf et al., 2003; Kaan & Swaab, 2003), though Bornkessel and colleagues also argue that certain subcases of reanalysis (specifically, reanalysis related to case marking) engender N400 effects (Bornkessel et al 2004). While it is possible that the LAN observed in the present experiment reflects another, distinct type of reanalysis process, such an account is not obviously supported by previous ERP studies that have elicited a LAN component.

Lastly, it is unclear whether an account that invokes reanalysis is an entirely distinct alternative to accounts that invoke structured search. In particular, the notion of structured search that we have suggested posits that in order to recognize that a local antecedent for ziji is inappropriate in a sentence with a long-distance antecedent for ziji, the local subject position must first be retrieved and rejected due to its unacceptability as an antecedent for ziji. In this sense, this intermediate step in a structured search procedure is a sort of reanalysis. However, in order to observe this pattern, it seems that only structural or positional information is deployed in retrieval. This is because the local subject is retrieved despite the fact that the long-distance subject has a number of semantic features (e.g., animacy, sentience, being a source of communication) that are known to invite anaphoric co-reference (Kaiser, Runner, Sussman, & Tanenhaus, 2009). In contrast, the local subject contained no semantic cues to support retrieval. None of the inanimate nouns used in this position supported metonymic interpretations, such as using Washington to refer to US government employees, that could license ziji.

It is possible that the local subject is first accessed purely because of a local activation advantage, perhaps due to greater decay of the matrix subject. However, the nature of this advantage would need to be substantial to outweigh the semantic compatibility of the distant noun on this account, as well as the semantic incompatibility of the local subject. In addition, this view is not consistent with the results obtained in cross-modal priming studies. Nicol & Swinney (1989) and Liu (2009) both found that the local subject was selectively accessed at the point of processing English and Chinese reflexives, respectively. Interestingly, this pattern did not occur for pronouns, which were able to initially exclude the local subject from consideration. If the local subject were retrieved due to a more recent encoding of the NP, even when it is a poor fit to retrieval cues, then this local advantage should be apparent independent of which anaphor initiates the retrieval. These results suggest that whatever locality advantage is present for the local subject, it is not so great as to force consideration of the local subject in cases where it is a poor fit for an anaphor’s requirements. If this line of reasoning is correct, then it
seems that the locality advantage for ziji cannot be entirely attributed to decay of the matrix subject. The nature of the locality advantage observed here and elsewhere during the resolution of ziji’s reference remains an important question for future research.

Linguistic and discourse antecedents

One consistent finding from both Experiments 1 and 2 is that sentences that do not contain an explicit linguistic antecedent for ziji are not consistently rejected as unacceptable by speakers. This is not entirely surprising, as it is possible to interpret ziji as referring to the speaker of a given utterance. In Experiment 2, we failed to find any consistent difference between sentences that contained a linguistic antecedent for ziji and those that did not. Interestingly, the sentences that lacked a linguistic antecedent appeared to pattern with local antecedent binding conditions in the ERP record. If this is correct, then it suggests that the egocentric interpretation is not an ‘elsewhere’ interpretation that is adopted only after an exhaustive search of the parse fails to return a licit antecedent. However, the offline judgment data appears to suggest that this interpretation is not as readily accepted as interpretations with explicit linguistic antecedents, which may suggest some difficulty in recovering this interpretation. This apparent difficulty may be an artifact of the tasks used in the present set of studies: it is difficult to know what perspective participants adopt when interpreting out-of-context sentences in a laboratory setting.

If it’s the case that the egocentric interpretation is adopted before an exhaustive search of the parse, then this raises the question of what information allows comprehenders to decide between linguistically binding ziji and adopting an egocentric interpretation. One possibility is that top-down information about the discourse model is applied to disambiguate early: in the case where there are no sentient entities in the discourse, the egocentric interpretation is the only licit option, and structured search is deployed only to check that discourse entities stand in an appropriate structural relation. At this point this suggestion remains speculative, as the present study does not provide evidence to support this notion. The relationship between linguistically and non-linguistically bound ziji in online processing routines remains an important unresolved question, and we leave this to future research.

Conclusion

The present study examined the time-course of antecedent-anaphor dependency construction using the Mandarin Chinese long-distance anaphor ziji. It was found that local antecedents are accessed more rapidly than long-distance antecedents, suggesting that the information necessary to complete local antecedent dependencies is present before the information necessary to complete long-distance antecedent dependencies. This finding was supported by converging evidence from ERPs, which showed that long-distance bindings of ziji showed a LAN component relative to local bindings. This pattern of ERP results suggests that syntactic reanalysis processes are not implicated in the slowdown observed in the SAT results. Although these findings are compatible with several implementations, the crucial features that all accounts share is that they invoke a serial consideration of candidate NPs that are retrieved primarily on structural cues,
without regard to the semantic fit of their content. These results are suggestive of structured search strategies as one option the parser may use to satisfy certain types of linguistic constraints, though further work is needed to examine whether or not such structured retrieval processes are invoked in a wider range of linguistic long-distance dependencies.

Acknowledgments

This work was supported in part by NSF EAPSI-091387 to BD, by NSF IGERT DGE-0801465 to the University of Maryland, by NSF BCS-0848554 to CP, and by a Natural Science Foundation of China grant (30600179) to TG. We would like thank Brian McElree for his help on both theoretical and technical aspects of the current work. We are grateful to Tim Hunter, Dave Kush, Roger Levy, Jeff Lidz, Shravan Vasishth and Ming Xiang for useful discussions of the material presented here, and to Mehmeti Abdullah, Peiyao Chen, and Yu Kai for assistance in collecting data and preparing experimental materials.

References


Figure 1: Panel A: an example of a structured search for finding ziji’s antecedent in the sentence *Lisi shuo fengbao hai-le ziji* “Lisi said the storm harmed him”. The hypothetical structural cues do not allow comprehenders to rule out consideration of the local subject *fengbao* “storm”. Thus comprehenders must evaluate multiple subject positions to find the correct antecedent *Lisi*. This structured access predicts that processing time should grow with the number of subject positions that need to be evaluated. Panel B: Example of feature-based parallel access for finding ziji’s antecedent in the sentence *Lisi shuo fengbao hai-le ziji* “Lisi said the storm harmed him”. The mixture of structural and semantic cues allow direct access to the correct antecedent *Lisi*. Feature-based access predicts that processing time should be constant with the number of subject positions that need to be evaluated.
Figure 2: SAT functions for LD and local antecedent *ziji* conditions with fully saturated models \((2\lambda - 2\beta - 2\delta)\), over average data (not averaged parameters). Accuracy is scaled to show proportion of asymptote; vertical bars indicate time point at which 50% accuracy is reached.
**Figure 3:** SAT functions for LD and local control conditions with fully saturated models \((2\lambda - 2\beta - 2\delta)\), over average data (not averaged parameters). Accuracy is scaled to show proportion of asymptote; vertical bars indicate time point at which 50% accuracy is reached.
Figure 4: Average asymptotic accuracy ($\lambda$) across individual participant SAT function fits with Bayesian parameter estimation. Error bars show ±1 SE, corrected for between-participant variance.
Figure 5: Average rate ($\beta$) across individual participant SAT function fits with Bayesian parameter estimation. Error bars show $\pm 1$ SE, corrected for between-participant variance.
Figure 6: Average intercept (δ) across individual participant SAT function fits with Bayesian parameter estimation. Error bars show ±1 SE, corrected for between-participant variance.
Figure 7: Average speed ($\delta + \beta^{-1}$) across individual participant SAT function fits with Bayesian parameter estimation. Error bars show ±1 SE, corrected for between-participant variance.
Figure 8: By-participants summary of dynamics advantage (in both rate and speed parameters) for local antecedent over LD antecedent conditions for *ziji* and control conditions. Participants are ranked by size of advantage; order is not identical across *ziji* and control conditions. Error bars represent 1 standard deviation of the posterior distribution of that parameter’s estimate.
Figure 9: Grand average ERPs for *ziji* conditions in Experiment 7, low-pass filtered for visual presentation. Only electrodes included in the ROI analyses are presented.