University of Potsdam

Doctoral Thesis

Working Memory and Prediction in Human Sentence Parsing

Cross-linguistic evidence from anaphoric dependencies and relative clauses

by
Lena Ann Jäger

Doctoral dissertation submitted to the Faculty of Human Sciences at the University of Potsdam in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Cognitive Science.

Submitted on April 23, 2015
Defended on October 15, 2015
Grade: summa cum laude

Supervisors:
Prof. Dr. Shravan Vasishth, Department of Linguistics, University of Potsdam
Prof. Dr. Reinhold Kliegl, Department of Psychology, University of Potsdam

Reviewers:
Prof. Dr. Shravan Vasishth, Department of Linguistics, University of Potsdam
Prof. Dr. Richard Lewis, School of Psychology, University of Michigan

Published online at the
Institutional Repository of the University of Potsdam:
URN urn:nbn:de:kobv:517-opus4-82517
http://nbn-resolving.de/urn:nbn:de:kobv:517-opus4-82517

Erklärung

Hiermit erkläre ich, dass ich bei der Abfassung der vorliegenden Arbeit alle Regelungen guter wissenschaftlicher Standards eingehalten habe. Weiter erkläre ich, dass ich die vorliegende Arbeit selbständig verfasst habe und über die Beiträge meiner Koautoren hinaus, welche in der beiliegenden *Erklärung über die Beiträge zu Gemeinschaftsveröffentlichungen* spezifiziert sind, keine Hilfe Dritter in Anspruch genommen habe.

Lena Jäger

Potsdam, den 9. April 2015

Abstract

This dissertation investigates the working memory mechanism subserving human sentence processing and its relative contribution to processing difficulty as compared to syntactic prediction. Within the last decades, evidence for a content-addressable memory system underlying human cognition in general has accumulated (e.g., Anderson et al., 2004). In sentence processing research, it has been proposed that this general content-addressable architecture is also used for language processing (e.g., McElree, 2000). Although there is a growing body of evidence from various kinds of linguistic dependencies that is consistent with a general content-addressable memory subserving sentence processing (e.g., McElree et al., 2003; Van Dyke and McElree, 2006), the case of reflexive-antecedent dependencies has challenged this view. It has been proposed that in the processing of reflexive-antecedent dependencies, a syntactic-structure based memory access is used rather than cue-based retrieval within a content-addressable framework (e.g., Sturt, 2003). Two eye-tracking experiments on Chinese reflexives were designed to tease apart accounts assuming a syntactic-structure based memory access mechanism from cue-based retrieval (implemented in ACT-R as proposed by Lewis and Vasishth, 2005). In both experiments, interference effects were observed from noun phrases which syntactically do not qualify as the reflexive's antecedent but match the animacy requirement the reflexive imposes on its antecedent. These results are interpreted as evidence against a purely syntactic-structure based memory access. However, the exact pattern of effects observed in the data is only partially compatible with the Lewis and Vasishth cue-based parsing model. Therefore, an extension of the Lewis and Vasishth model is proposed. Two principles are added to the original model, namely *cue confusion* and *distractor prominence*.

Although interference effects are generally interpreted in favor of a content-addressable memory architecture, an alternative explanation for interference effects in reflexive processing has been proposed which, crucially, might reconcile interference effects with a structure-based account. It has been argued that interference effects do not necessarily reflect cue-based retrieval interference in a content-addressable memory but might equally well be accounted for by interference effects which have already occurred at the moment of encoding the antecedent in memory (Dillon, 2011). Three experiments (eye-tracking and self-paced reading) on German reflexives and Swedish possessives were designed to tease apart cue-based retrieval interference from encoding interference. The results of all three experiments suggest that there is no evidence that encoding interference affects the retrieval of a reflexive's antecedent. Taken together, these findings suggest that the processing of reflexives can be explained with the same cue-based retrieval mechanism that has been invoked to explain syntactic dependency resolution in a range of other structures. This supports the view that the language processing system is located within a general cognitive architecture, with a general-purpose content-addressable working memory system operating on linguistic expressions.

Finally, two experiments (self-paced reading and eye-tracking) using Chinese relative clauses were conducted to determine the relative contribution to sentence processing difficulty of working-memory processes as compared to syntactic prediction during incremental parsing. Chinese has the cross-linguistically rare property of being a language with subject-verb-object word order *and* pre-nominal relative clauses. This property leads to opposing predictions of expectation-based accounts and memory-based accounts with respect to the relative processing difficulty of subject vs. object relatives. Previous studies showed contradictory results, which has been attributed to different kinds local ambiguities confounding the materials (Lin and Bever, 2011). The two experiments presented are the first to compare Chinese relatives clauses in syntactically unambiguous contexts. The results of both experiments were consistent with the predictions of the expectation-based account of sentence processing but not with the memory-based account. From these findings, I conclude that any theory of human sentence processing needs to take into account the power of predictive processes unfolding in the human mind.

Acknowledgements

First and foremost, I want to express my deepest gratitude to my supervisor Shravan Vasishth for his ongoing support and encouragement. I want to thank you, Shravan, for your wise guidance on the one hand and for the freedom to develop my own ideas and interests on the other hand. I also want to express my sincere gratitude to my second supervisor Reinhold Kliegl for his support and his overwhelming enthusiasm for science, which inspired and motivated me even in the moments when I was facing seemingly insurmountable difficulties. I am also very grateful to Richard Lewis, who kindly agreed to review my dissertation.

Moreover, I want to thank the Studienstiftung des deutschen Volkes who supported me throughout my studies and provided me with financial support in the form of a PhD scholarship allowing me to freely develop and pursue my personal scientific interests.

I want to express my thanks to Hua Shu and the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University for allowing me to conduct my experiments in their eye-tracking lab. I also want to thank Yan Ming, Pan Jing'er and Zhou Wei for their very kind assistance in preparing, running and evaluating the experiments in Beijing. Moreover, I want to thank Zhao Xiaoran, Wang Yuyan and Wang Bei for their warm welcome and friendship in Beijing. I thank the Department of English at National Taiwan Normal University, Taipei, for allowing me to conduct my experiment in their eye-tracking lab. In particular, I am very grateful to Chan Shiaohui and Liang Sun-Chieh for their hospitality and their assistance. I thank the School of Foreign Languages at Dalian University of Technology for conducting one of the reading experiments reported in this thesis.

I want to express my deep thanks to my collaborators and friends Lena Benz, Felix Engelmann, Chen Zhong, Charles Lin, Brian Dillon, Jens Roeser and Qiang Li, who were involved in the research presented in this thesis. It was a great pleasure to work with you and I hope we will carry on our fruitful and inspiring collaborations in the future.

I thank all present and past lab members for the stimulating research environment in our group and my colleagues and friends, whose comments, criticisms and suggestions were very valuable for my research. In particular, I want to name Bruno Nicenboim, Maja Stegenwallner, Anne Mucha, Anne Steinke, Tina Marusch, Pavel Logačev, Samar Husain, Titus von der Malsburg, Dario Paape, Umesh Patil, Thomas Kern, Carla Keßler, Ulla Behr and Paul Metzner.

I am grateful to Ted Gibson for sharing his experimental materials from a previous experiment on Chinese. I also want to thank Lu Yi, Chuang Chih-Chun, Jinxing Yue and Rui Qin for their assistance in the preparation of the Chinese experimental materials as well as Johanna Thieke together with all lab assistants for their practical help in the conduction of the German experiments.

Finally, I want to thank my mother, Brigitte Jäger, who never ceased to support and encourage me throughout my life, Sascha, who has been a constant source of strength and serenity, and my dear friends, who gave me happiness even in the difficult moments of life.

Contents

List of Figures

List of Tables

To my mother

Chapter 1

Introduction

Human sentence parsing refers to the psychological process of creating a syntactic representation from a linear stream of words. In order to assign a meaning to a sentence that goes beyond the mere lexical semantics of its individual words, we have to identify the syntactic relations between the words, for example, we need to identify the subject and the object of a verb. This finally allows us to assign the corresponding thematic roles, that is to decide *who* did *what* to *whom*. The aim of psycholinguistic research is to develop a general theory of human sentence processing which captures the general mechanisms underlying human language processing and thus goes beyond the specific characteristics of a particular language. Even very prominent theories of sentence processing which had been developed mainly based on evidence from English (e.g., the *Garden Path* theory proposed by Frazier, 1987) had to be revised when they were faced with broader cross-linguistic evidence (Pickering and Van Gompel, 2006). Thus, it is crucial for psycholinguistic research to develop and test theories focussing on evidence from different languages. Therefore, I adopt a cross-linguistic perspective in this thesis by presenting evidence from different languages and discussing the implications of the results in the light of previous literature from a broad range of languages.

One critical cognitive component involved in human sentence parsing is working memory. When creating the syntactic representation of a sentence, we often have to link together two elements which are not adjacent to each other, so-called unbounded dependencies. Indeed, in many syntactic dependencies, such as subject-verb or anaphor-antecedent dependencies, the amount of material linearly intervening between the first and the second element of the dependency is (theoretically) unlimited. For the incremental human parser (e.g., Marslen-Wilson, 1973, 1975), this means that when encountering the second part of a dependency, a memory process needs to be triggered to retrieve its previously parsed counterpart. Already early psycholinguistic research indicated

that working memory constraints pose central limitations to what kind of sentences humans are actually able to comprehend as compared to what sentences are theoretically grammatical in a language (Miller and Chomsky, 1963).

With respect to the working memory system subserving sentence processing, two major issues are to be addressed. The first issue concerns the architecture of the memory system. Here, the central questions are i) what the defining properties of this system are, and ii) whether this system is shared with other cognitive tasks, or whether it is a language-specific system with mechanisms that are qualitatively different from the ones used for other tasks. The second issue concerns the relative contribution of working memory constraints to sentence processing difficulty as compared to other factors.

The first question has resulted in a great deal of controversy in the history of psycholinguistic research. On the one hand, it has been proposed that the working memory system used for inherently automatic linguistic processes such as sentence parsing is detached from the working memory system used for other language related cognitive tasks such as more consciously controlled verbal tasks (Caplan and Waters, 1996, 1999). On the other hand, Just and Carpenter (1992) (see also King and Just, 1991) have proposed a shared memory system serving all kinds of linguistic and verbal tasks. More recent research has proposed that not only all kinds of verbal tasks make use of the same memory system, but that language draws on the same memory system as any other cognitive processes. Within the last decades, evidence for a content-addressable memory system underlying human cognition in general has accumulated (e.g., Watkins and Watkins, 1975; Anderson and Lebiere, 1998; Anderson et al., 2004; McElree, 2006). In sentence processing research, it has been proposed that this general content-addressable architecture also subserves language processing (e.g., McElree, 2000; Van Dyke and Lewis, 2003; Lewis and Vasishth, 2005; Van Dyke and McElree, 2011). In particular, Lewis and Vasishth (2005) developed a model of human sentence processing which is based on the general content-addressable cognitive architecture Adaptive Control of Thought–Rational (ACT-R) (Anderson and Lebiere, 1998; Anderson et al., 2004). Indeed, there is a growing body of evidence from various kinds of linguistic dependencies that is consistent with a general, i.e., independently motivated, content-addressable memory subserving sentence processing (e.g., Van Dyke and Lewis, 2003; Van Dyke and McElree, 2006; Van Dyke, 2007; Wagers et al., 2009; Van Dyke and McElree, 2011; Dillon et al., 2013; McElree et al., 2003; Martin and McElree, 2008; Vasishth et al., 2008). However, the case of reflexive-antecedent dependencies has challenged the view of one general-purpose content-addressable memory being used for syntactic parsing. One central prediction of a content-addressable memory architecture is that in the assumedly cue-based retrieval process interference effects from non-target items which match some of the retrieval relevant features (the so-called *retrieval cues*) should be observed. The

absence of such interference effects in the processing of reflexive-antecedent dependencies has been taken as evidence against cue-based retrieval as implemented in the Lewis and Vasishth (2005) model (e.g., Nicol and Swinney, 1989; Sturt, 2003; Phillips et al., 2011) and has been interpreted favoring a structure-based account in which syntax has some kind of priority over other information available in the input. Note that this view is tightly related to the question of modularity of language in general. Modular accounts of language processing assume that i) the language system as a whole is a cognitive subsystem which is largely detached from other cognitive subsystems and ii) within the language system, the different components (syntax, semantics etc.) are informationally encapsulated, i.e., operate independently from each other (usually in a serial fashion with syntax having temporal priority over semantics in sentence comprehension) (e.g., Frazier, 1979; Rayner et al., 1983; Ferreira and Clifton, 1986; Frazier, 1987; Abney, 1989; Crocker, 1995; Pritchett, 1992). Against this background, experimental evidence on the processing of reflexives has wide-ranging implications for theories of sentence processing. However, there are several concerns with the evidence which has been presented favoring a structure-based account of reflexive-antecedent processing. First, evidence for the structure-based account comes from the *absence* of interference effects, which statistically is a null result and hence inconclusive. In particular in the light of the low statistical power of psycholinguistic experiments in general, these findings should be interpreted with caution. Second, most of the evidence presented so far focusses on English reflexives and interference effects induced by a noun which matches the number or gender feature of the reflexive. As mentioned above, it is crucial for any general theory of sentence processing to be exposed to cross-linguistic tests from a wide range of typologically varying languages. In this thesis, the processing of reflexives is investigated from a more cross-linguistic perspective. Evidence from experiments with very large sample sizes to overcome the statistical power issue on the processing of reflexives in Mandarin Chinese, German and Swedish are presented. Based on the experimental results, an extension of the content-addressable ACT-R model of sentence processing (Lewis and Vasishth, 2005) is proposed and evaluated by comparing its predictions with previous experimental evidence on reflexives.

The second major question about working memory in sentence processing, namely the relative contribution of working memory constraints to processing difficulty, is important to answer because the predictions of any theory of sentence processing are formulated in terms of (some operational definition of) processing difficulty associated with one linguistic constellation as compared to another. Thus, any comprehensive theory of sentence processing needs to make explicit statements about the sources of processing difficulty. Hence, it is a central endeavor in psycholinguistic research to disentangle

the various factors which contribute to processing difficulty. Indeed, the different theories of human sentence processing attribute processing difficulty to different factors. The most prominent factors proposed in the literature are working memory constraints and top-down predictions about upcoming material. Although it is a well-established fact that the human parser heavily involves into predictions about upcoming material (e.g., Ehrlich and Rayner, 1981; Altmann and Kamide, 1999; Federmeier and Kutas, 1999; Kamide et al., 2003; McDonald and Shillcock, 2003; Kaiser and Trueswell, 2004; Van Berkum et al., 2005; DeLong et al., 2005; Staub and Clifton, 2006; for a recent review see Huettig, 2015), their relative contribution to sentence processing difficulty, in particular as compared to working memory limitations, is not entirely clear yet. Among the theories which assume that working memory processes are the main source of relative sentence processing difficulty, different concrete memory demands during sentence processing have been proposed to be responsible for the processing difficulty. On the one hand, it has been proposed that processing difficulty arises as a result of memory decay and/or interference at the moment of memory retrievals as a function of the distance between the currently parsed element and the target of the retrieval, that is, the first part of the dependency (e.g., Stevenson, 1994; Gibson, 1998, 2000; Gordon et al., 2001, 2004; McElree, 2000; Lewis and Vasishth, 2005; Van Dyke and McElree, 2011). I will refer to these accounts as *retrieval metrics*. However, note that there are substantial differences between these theories concerning the exact mechanisms underlying memory interference. On the other hand, it has been proposed that tracking (syntactic) predictions about upcoming material consumes working memory resources and hence leads to increased processing difficulty as a function of the number of the currently maintained predictions. The most well-known representative of this metric is the *storage cost* component of the Dependency Locality Theory (DLT) (Gibson, 1998, 2000). I will refer to this as *storage metrics*. Although storage metrics take into account prediction as an important source of processing difficulty, it is theoretically orthogonal to the proposal made by a wide range of theories which I will refer to as *expectation-based* accounts. This class of accounts attributes the main source of processing ease or difficulty to predictions about the incoming material which are either met or dashed (e.g., Mitchell and Holmes, 1985; Trueswell et al., 1993; Mitchell et al., 1995; Trueswell, 1996; Jurafsky, 1996; McRae et al., 1998; Tabor and Tanenhaus, 1999; Hale, 2001; Levy, 2008). In contrast to storage metrics, expectation-based theories assume that predictions about upcoming material *facilitate* processing. Note, however, that the locus of the predicted effect differs: storage metrics predict increased processing difficulty as long as a prediction has to be maintained, i.e., has not been confirmed or dashed yet, whereas expectation-based theories predict processing facilitation at the moment when the incoming material is in agreement with a prediction. This thesis presents reading experiments investigating the processing of Chinese relative clauses that were designed to tease apart the relative contribution

to processing difficulty of the different working memory metrics (retrieval and storage metrics) on the one hand and syntactic prediction on the other hand. Evidence for one very influential theory attributing sentence processing difficulty to working memory constraints largely came from English relative clauses (Gibson, 1998, 2000). However, due to the syntactic properties of English relative clauses, expectation-based accounts make similar predictions as Gibson (2000)'s Dependency Locality Theory. Chinese is an important test case for disentangling prediction and working memory constraints in the processing of relative clauses since, due to the syntactic properties peculiar to Chinese relative clauses, both accounts make opposing predictions.

This thesis is composed of three journal articles. The first and the second article concern the first question raised in the introduction, i.e., the architectural properties of the working memory system subserving human sentence parsing. The third article concerns the second question raised above, i.e., the relative contribution of working memory and prediction to sentence processing difficulty.

The first article, which is presented in Chapter 2, investigates the processing of the Chinese reflexive *ziji* in order to tease apart structure-based accounts of reflexive processing (Nicol and Swinney, 1989; Sturt, 2003; Phillips et al., 2011; Dillon et al., 2013, 2014) from content-addressable accounts of memory retrieval which, in addition to syntactic features, also use non-structural features as retrieval cues (Lewis and Vasishth, 2005). In contrast to English reflexives, Chinese *ziji* is neither marked for gender nor for number but requires its antecedent to be animate. Two eye-tracking-while-reading experiments show that retrieval ease of the reflexive's antecedent is affected by the presence of a syntactically illicit distractor noun that matches the animacy feature, which the reflexive arguable uses as a retrieval cue. The results are interpreted as evidence against a structure-based account of reflexive processing and are modeled in the ACT-R framework as proposed by Lewis and Vasishth (2005). As the current implementation of the Lewis and Vasishth (2005) model was not able to account for the interference pattern observed in the data, an extension of this model is proposed and is discussed against the background of previous literature on reflexive processing. In Chapter 2, only those properties of Chinese reflexives which are critical for the experimental design are presented. As the syntactic and discourse constraints of Chinese reflexives are highly complex, a chapter dedicated to the properties of Chinese reflexives has been added to the appendix of this dissertation (see Appendix D).

Although the interference effects presented in the first article are interpreted as evidence against structure-based memory access and in favor of a content-addressable memory architecture, an alternative explanation for interference effects in reflexives has been

proposed which, crucially, might reconcile interference effects with a structure-based account. Dillon (2011) has argued that interference effects observed when a reflexive triggers the retrieval of its antecedent do not necessarily reflect cue-based retrieval interference as predicted by a content-addressable memory but might equally well be accounted for by interference effects which have already occurred at the moment of encoding the antecedent in memory (Nairne, 1988, 1990; Lewandowsky et al., 2008; Oberauer and Kliegl, 2006). The idea is that the presence of a similar distractor leads to a decreased memory trace (e.g., in terms of distinctiveness or in terms of activation level; Oberauer and Kliegl, 2006) of the antecedent and, as a consequence, to increased latencies at later retrievals. Hence, both cue-based retrieval and structure-based memory access with the additional assumption of encoding interference predict increased retrieval latencies at the reflexive due to the presence of a feature-matching distractor. The second publication included in this thesis (see Chapter 3) presents three experiments on German reflexives and Swedish possessives which were designed to tease apart cue-based retrieval interference from encoding interference. Despite high statistical power, in neither of these three experiments, evidence for encoding interference affecting the retrieval of a reflexive's antecedent could be observed. Together with the results of the first article of this thesis and previous literature on reflexives, these results are interpreted in favor of a content-addressable memory architecture in sentence processing.

The third publication included in this thesis (see Chapter 4) investigates the processing of Chinese relative clauses. As mentioned above, Chinese relative clauses are an important test case to tease apart theories of sentence processing which assume that the major determinant of processing ease or difficulty is prediction from those theories which attribute it to memory processes. Chinese has the cross-linguistically rare property of being a language with subject-verb-object word order *and* pre-nominal relative clauses. This property leads to opposing predictions of expectation-based accounts and memory-based accounts. Expectation-based accounts such as the surprisal metric proposed by Hale (2001) and Levy (2008) predict subject-extracted relative clauses to be easier to process than object-extracted relative clauses. In contrast, working-memory based theories of sentence processing (e.g., Gibson, 1998, 2000; Lewis and Vasishth, 2005) predict object-extracted relative clauses to be easier to process. The experiments on Chinese relative clauses presented in this thesis, are designed to pit the predictions of expectation-based accounts against those of the different working memory-based processing metrics.

Chapter 2

Frontiers in Psychology
Volume 6, Issue 617, 2015

Retrieval Interference in Reflexive Processing: Experimental Evidence from Mandarin, and Computational Modeling

Lena Jäger, Felix Engelmann, and Shravan Vasishth
Dept. of Linguistics, University of Potsdam

Abstract

We conducted two eye-tracking experiments investigating the processing of the Mandarin reflexive *ziji* in order to tease apart structurally constrained accounts from standard cue-based accounts of memory retrieval. In both experiments, we tested whether structurally inaccessible distractors that fulfill the animacy requirement of *ziji* influence processing times at the reflexive. In Experiment 1, we manipulated animacy of the antecedent and a structurally inaccessible distractor intervening between the antecedent and the reflexive. In conditions where the accessible antecedent mismatched the animacy cue, we found inhibitory interference whereas in antecedent-match conditions, no effect of the distractor was observed. In Experiment 2, we tested only antecedent-match configurations and manipulated locality of the reflexive-antecedent binding (Mandarin allows non-local binding). Participants were asked to hold three distractors (animate vs. inanimate nouns) in memory while reading the target sentence. We found slower reading times when animate distractors were held in memory (inhibitory interference). Moreover, we replicated the locality effect reported in previous studies. These results

are incompatible with structure-based accounts. However, the cue-based ACT-R model of Lewis and Vasishth (2005) cannot explain the observed pattern either. We therefore extend the original ACT-R model and show how this model not only explains the data presented in this article, but is also able to account for previously unexplained patterns in the literature on reflexive processing.

Keywords: ACT-R; Reflexive processing; Cognitive Modeling; Eye-tracking; Interference; Cue-based retrieval; Chinese; Ziji; Content-addressable memory; Sentence processing; Working memory

2.1 Introduction

One major task the human parser has to accomplish is to syntactically link together two or more linguistic elements that are not adjacent to each other. For example, when a reflexive is being processed, it has to be somehow linked to its antecedent even if there is intervening material. Therefore, one central question in psycholinguistics is what mechanisms the human parser uses to identify and retrieve the previously processed part of a dependency. Theoretically, there are different options how this identification and retrieval of a linguistic constituent from working memory might be accomplished: different kinds of search mechanisms on the one hand (Sternberg, 1966, 1969) and cue-based, i.e., content-addressable, retrieval on the other hand (McElree and Dosher, 1989; Anderson et al., 2004)[1]. In general, a search mechanism checks certain items in memory based on their location in order to find the target. Cue-based retrieval, in contrast, assumes that retrieval targets are content-addressable and can be accessed directly by the use of certain features as retrieval cues. Over the last decade, evidence favoring a content-addressable memory underlying human sentence processing has accumulated (McElree, 2000, 2003; McElree et al., 2003; Van Dyke and McElree, 2006; Martin and McElree, 2008).

In the case of English reflexives, retrieval cues used in a content-addressable memory might be non-structural cues like gender or number along with structural cues like local c-command. Note that a reflexive's binding domain varies between languages (Büring, 2005; Reuland, 2011). Whereas in English it can be approximated by the local clause, in Chinese the reflexive *ziji* can be bound across clause boundaries (non-local binding; for a brief overview of the syntactic properties of Chinese *ziji* see below). For the sake of simplicity, we will refer to the structural feature of *c-commanding the reflexive and being contained in its binding domain* briefly as the *c-command* feature.

However, within the framework of cue-based retrieval, it is still an open question which features the parser uses as retrieval cues. On the one hand, it has been proposed that all available cues are used for retrieval with equal weights being applied to all cues (Lewis and Vasishth, 2005).

[1]Note that the different models of content-addressable memory differ with respect to their assumptions about the exact nature of similarity-based retrieval interference. While the model proposed by Anderson et al. (2004) predicts similarity-based retrieval interference to be observed in retrieval probabilities as well as in retrieval latencies, the model proposed by McElree (2000) predicts that similarity-based retrieval interference only affects retrieval probabilities and not retrieval latencies. In this article, we will focus on cue-based retrieval in the sense of Anderson et al. (2004).

We will refer to this account as the *standard cue-based retrieval account*. On the other hand, Van Dyke (2007); Van Dyke and McElree (2011) and others argue that syntactic cues (being in a certain tree-configurational position) have some kind of priority over non-syntactic cues. In particular, it has been proposed that for the processing of reflexive-antecedent dependencies, the set of features used for retrieving a reflexive's antecedent is limited to syntactic cues such as c-command within the reflexive's binding domain (Nicol and Swinney, 1989; Sturt, 2003; Xiang et al., 2009; Phillips et al., 2011; Dillon et al., 2013; Kush and Phillips, 2014). We will refer to this proposal as *structure-based account*.

If a structure-based retrieval is applied, a noun phrase that is in a structural position that disqualifies it from being the reflexive's antecedent should not have any effect on the processing of the reflexive-antecedent dependency, no matter whether it matches non-structural features of the reflexive such as gender or number. Thus, in a sentences like (2.1), the gender of *Jonathan* or *Jennifer* should not affect processing times of the reflexive since they do not c-command it and hence cannot syntactically bind the reflexive.

(2.1) a. *Antecedent-match; distractor-match*
The **surgeon** who treated **Jonathan** had pricked **himself** ...

b. *Antecedent-match; distractor-mismatch*
The **surgeon** who treated **Jennifer** had pricked **himself** ...

c. *Antecedent-mismatch; distractor-match*
The **surgeon** who treated **Jennifer** had pricked **herself** ...

d. *Antecedent-mismatch; distractor-mismatch*
The **surgeon** who treated **Jonathan** had pricked **herself** ...

The parsing architecture developed by Lewis and Vasishth (2005), which is based on Anderson and Lebiere (1998)'s cognitive architecture ACT-R (Adaptive Control of Thought–Rational) assumes a cue-based retrieval mechanism without syntactic constraints. This model has been used to explain interference effects in sentence processing and in reflexives in particular (e.g., Dillon et al., 2013; Parker and Phillips, 2014; Patil, Vasishth & Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript). According to the ACT-R model, both latency and probability of retrieving a certain target item are determined by i) the quality of the match between retrieval cues and target features and ii) similarity-based mutual inhibition between the target and other matching items. Retrieval speed and probability increase with the number of cues matching the target. If, however, certain cues match the features of multiple memory items, similarity-based interference leads to a higher retrieval latency, i.e., inhibitory interference effects. The latter is the case in (2.1a) as compared to (2.1b), because in (2.1a) both the target *surgeon* and the distractor *Jonathan* share the feature $+masculine$. In the antecedent-mismatch conditions (2.1c) vs. (2.1d), in contrast, the target *surgeon* and the cue-matching distractor *Jennifer* in (2.1c) do not share the feature $+feminine$, hence, no similarity-based interference arises. Consequently, no inhibition is predicted in (2.1c)

vs. (2.1d). On the contrary, because both target and distractor only partially match the retrieval cues in (2.1c), they are equally likely to be retrieved. Compared to (2.1d), this predicts a higher proportion of incorrect retrievals and a lower average retrieval latency, which is usually referred to as *facilitatory interference* or *intrusion*.

In sum, a major prediction that distinguishes standard cue-based retrieval from models assuming a limitation of the retrieval cues to structural features is that the former entails interference effects from non-target items that match (some of) the cues used for retrieval.[2]

In order to tease apart structure-based from standard cue-based retrieval, interference effects from feature-matching but syntactically illicit antecedents in the processing of reflexive-antecedent dependencies have drawn considerable attention in recent years. Several studies used a feature-match/mismatch design, where a non-syntactic feature (e.g., gender or number) was manipulated at the antecedent and at a structurally inaccessible distractor (see Example 2.1 for typical sentence material). In Table 2.1, we provide an overview of the studies examining interference effects in reflexives (including reflexives inside a prepositional phrase and possessive reflexives) and reciprocals using a feature-match/mismatch design. Studies on the processing of reflexives in so-called picture noun phrases have not been included in our review since their binding properties differ from other reflexives (Büring, 2005; Reuland, 2011). Moreover, experiments investigating specific populations such as children or L2 learners are not considered in the review. Table 2.1 summarizes whether or not inhibitory (i.e., a slowdown due to the presence of a cue-matching inaccessible distractor) or facilitatory (i.e., a speed-up due to the presence of a cue-matching inaccessible distractor) interference was observed in i) conditions with an accessible antecedent that matched the feature under examination and ii) conditions with an accessible antecedent that mismatched the feature under examination (i.e., sentences that are either ungrammatical or at least violating the stereotypical gender of the accessible antecedent). Some studies manipulated other factors in addition to the feature-match/mismatch manipulation. In these cases, we split the respective experiments into two entries in Table 2.1, with one entry for each level of the additional factor. In particular, for Felser et al. (2009), who manipulated feature type (gender vs. c-command) as additional within-participants factor and language proficiency (native speaker vs. L2 learner) as between-participants factor, one row in Table 2.1 refers to the manipulation of the c-command feature in native speakers and another row refers to the gender manipulation in native speakers. The results of the non-native group are not included in the table because this review concerns adult native speaker populations. For Chen et al. (2012b), who manipulated whether the Chinese reflexive *ziji* was locally or non-locally bound, one row in Table 2.1 refers to the interference effects observed in conditions with a local antecedent and a second row refers to the conditions with a non-local antecedent. Similarly, in the case of King et al. (2012), who

[2] It should be noted that cueing for a c-command feature is a simplification since it actually is a tree-configurational relation between items. There is no straightforward way of how to attribute a feature like that in an incremental parsing mechanism in content-addressable memory. In this paper, we do not provide a detailed account of how the attribution of a c-command feature could be implemented. As an example, Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English" (unpublished manuscript) in their ACT-R model for English reflexives approximated a c-command relation by cueing for a subject in the local clause. For a discussion of possible ways to encode tree-configurational information such as c-command in content-addressable memory see Alcocer and Phillips, "Using relational syntactic constraints in content-addressable memory architectures for sentence processing" (unpublished manuscript).

	Publication	Lang.	Method	Cue type	Depend. type	Interf. position	Distractor	Effect in antecedent-match		Effect in antecedent-mismatch	
								Interference	*AOI, Measure*	*Interference*	*AOI, Measure*
1	Xiang et al. '09	EN	EEG	gend	refl	retro	subj.	—		(inhib)	
2	Badecker&Straub '02 Exp5	EN	SPR	gend	refl	pro	genitive	*n.s.*		—	
3	Badecker&Straub '02 Exp6	EN	SPR	gend	refl	pro	prep.obj.	*n.s.*		—	
4	Chen et al. '12 local	CN	SPR	anim	poss.refl.	retro	subj.	*n.s.*		—	
5	Clackson et al. '11 Exp2 adults	EN	VW	gend	prep.refl.	pro	subj.(topic)	*n.s.*		—	
6	Clifton et al. '99 Exp1	EN	SPR	num	refl	retro	prep.obj.	*n.s.*		—	
7	Clifton et al. '99 Exp2	EN	SPR	gend,num	refl	retro	prep.obj.	*n.s.*		—	
8	Clifton et al. '99 Exp3	EN	SPR	gend	refl	pro	subj	*n.s.*		—	
9	Felser et al. '09 natives, gend	EN	ET	gend	refl	pro	subj.(topic)	*n.s.*		—	
10	Nicol&Swinney '89	EN	Primg	gend	refl	pro	subj.,obj.	*n.s.*		—	
11	Dillon et al. '13	EN	ET	num	refl	retro	obj.	*n.s.*		*n.s.*	
12	King et al. '12 adjacent	EN	ET	gend	refl	retro	obj.	*n.s.*		*n.s.*	
13	Sturt '03 Exp2	EN	ET	gend	refl	retro	obj.(topic)	*n.s.*		*n.s.*	
14	Cunnings&Felser '13 Exp1	EN	ET	gend	refl	pro	subj.(topic)	*n.s.*		(facil)	crit+2, FPRT
15	King et al. '12 non-adjacent	EN	ET	gend	prep.refl.	retro	prep.obj.	*n.s.*		facil	crit, FPRT
16	Parker&Phillips '14	EN	ET	nu/ge/an	refl	pro	subj.	*n.s.*		facil	crit, TFT
17	Cunnings&Sturt '14 Exp1	EN	ET	gend	refl	pro	subj.(topic)	*n.s.*		(inhib)	crit+1, FPRT
18	Kush&Phillips '14	HI	SPR	num	recipr	retro	prep.obj.	*n.s.*		(inhib)	crit+2
19	**this study Exp1**	CN	ET	anim	refl	retro	subj.	*n.s.*		inhib	crit, FFD/FPRT/RBRT/RPD
20	Cunnings&Felser '13 Exp2	EN	ET	gend	refl	retro	subj.(topic)	facil *lWM*	crit, FFD/FPRT	(inhib *lWM*)	crit, FFD
21	Sturt '03 Exp1	EN	ET	gend	refl	pro	subj.(topic)	facil	crit+2, RRT	*n.s.*	
22	**this study Exp2**	CN	ET	anim	refl	pro	memory	inhib	crit, FPRT/RBRT/RPD/TFT	—	
23	Badecker&Straub '02 Exp3	EN	SPR	gend	refl	pro	subj.	inhib	crit+1	—	
24	Badecker&Straub '02 Exp4	EN	SPR	num	recipr	pro	subj.	inhib	crit+1–crit+4	—	
25	Chen et al. '12 non-local	CN	SPR	anim	poss.refl.	retro	subj.	inhib	crit+1	—	
26	Clackson&Heyer '14	EN	VW	gend	prep.refl.	pro	subj.(topic)	inhib	gaze shift	—	
27	Felser et al. 09 natives, c-com	EN	ET	c-com	refl	pro	subj.(topic)	inhib	antecedent, Reg-in; crit, RPD	—	
28	Patil et al. unpublished	EN	ET	gend	refl	retro	subj.	inhib	crit, FPRP/(reg.-cont. FFD)	(facil)	crit, regr.-cont. FFD

TABLE 2.1: Overview of interference effects at reflexives categorized for feature-match/mismatch of the accessible antecedent and direction of the effect (inhibitory vs. facilitatory interference). Experiments are sorted by effect pattern (effect in antecedent-match first, effect in antecedent-mismatch second). Marginal effects are in parentheses. The experiments are classified by language (English EN, Mandarin Chinese CN, Hindi HI), experimental method (self-paced reading SPR, eye-tracking while reading ET, visual world eye-tracking VW, event-related potentials ERP, cross-modal priming Primg), retrieval cues that are examined (gender, number, animacy, c-command), dependency type (reflexive, prepositional reflexive, possessive reflexive, reciprocal), interference type (proactive vs. retroactive), and by syntactic position of the distractor (subject, object, genitive attribute, sentence external memory load and discourse topicality in parenthesis). For reading studies, the interest area (AOI) labeled *crit* refers to the region containing the reflexive/reciprocal. For eye-tracking experiments, Measure indicates the dependent variable in which the effect was observed (first-fixation duration FFD, first-pass reading time FPRT, total-fixation time TFT, regression-path duration RPD, right-bounded reading time RBRT, re-reading time RRT, first-pass regression probability FPRP, proportion of in-coming regressions Reg-in, first-fixation duration in regression-contingent trials reg.-cont. FFD). Effects in parentheses are only marginally significant. The entry '—' means that the respective conditions were not tested in the experiment while '*n.s.*' means that no significant effect was observed. *lWM* refers to participants with low working memory capacity.

manipulated whether the reflexive directly followed the verb or a preposition intervened, one table entry refers to the former configuration (labeled as *adjacent*) and another entry refers to the latter configuration (labeled as *non-adjacent*). In the review of Clackson et al. (2011), who primarily investigated the processing of reflexives in children, we only report the results of the adult control group. For the reviewed experiments, we report effects observed at the region containing the reflexive (labeled as *crit*) and the following regions (labeled as *crit+x*). Although the size of the interest areas in terms of number of words contained in one region differs between studies, which reduces the comparability of the time course of the observed effects to a certain extent, we keep the sectioning of the interest areas as in the respective publication.

In accessible antecedent-match conditions, previous studies found inhibitory interference in six cases (Badecker and Straub, 2002, Experiments 1 and 2; Felser et al., 2009, c-command manipulation in native speakers; Chen et al., 2012b, non-local reflexives; Clackson and Heyer, 2014; Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript). Statistically significant facilitatory interference in antecedent-match conditions was found in two experiments (Sturt, 2003, Experiment 1; Cunnings and Felser, 2013, Experiment 2). However, Sturt found the effect only in re-reading time two words after the reflexive and this effect could not be replicated by Cunnings and Sturt (2014), who used the same stimuli. Cunnings and Felser found the effect for readers with low working memory span (*lWM*), but not for high-span readers. In the majority of the experiments, in contrast, no interference effect was observed in antecedent-match conditions (Nicol and Swinney, 1989; Clifton et al., 1999; Badecker and Straub, 2002, Experiments 5 and 6; Sturt, 2003, Experiment 2; Felser et al., 2009, gender manipulation in native speakers; Clackson et al., 2011, adult control group of Experiment 2; King et al., 2012; Chen et al., 2012b, conditions with local reflexive binding; Cunnings and Felser, 2013, Experiment 1; Dillon et al., 2013; Kush and Phillips, 2014; Parker and Phillips, 2014; Cunnings and Sturt, 2014, Experiment 1).[3]

For conditions with a feature-mismatching accessible antecedent, two studies report significant effects of facilitatory interference (King et al., 2012; Parker and Phillips, 2014) and two studies report a marginal facilitatory effect (Cunnings and Felser, 2013, Experiment 1; Patil, Vasishth & Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript) — however, the latter effect was only found in a post-hoc analysis of regression-contingent first-fixation durations, and thus might be spurious. Marginal effects of inhibitory interference have been reported for participants with low working memory span (Cunnings and Felser, 2013, Experiment 2), in the processing of reciprocals (Kush and Phillips, 2014), and in Experiment 1 of Cunnings and Sturt (2014). The latter only report a marginal main effect of the distractor, but their reported means suggest that the effect was driven by the antecedent-mismatch conditions. This does, however, not seem very reliable because they used the same stimuli as Sturt (2003), Experiment 1, who, in contrast, had not found an effect in antecedent-mismatch conditions but a facilitation in antecedent match conditions. A general pattern is that interference effects in antecedent-match conditions are less frequently observed than effects in antecedent-mismatch conditions.

[3]King et al. (2012) report different results in their CUNY 2012 abstract and their final conference poster. We refer here to the results presented on the poster.

To summarize, the literature on reflexive interference contains a mixture of results, not favoring one particular of the retrieval models in question. Studies showing a general absence of interference support structure-based accounts (Nicol and Swinney, 1989; Sturt, 2003; Xiang et al., 2009; Phillips et al., 2011; Dillon, 2011; Dillon et al., 2013; Kush and Phillips, 2014). On the other hand, observations of significant interference effects have been interpreted as evidence against purely structure-based retrieval (Badecker and Straub, 2002; Chen et al., 2012b; Parker and Phillips, 2014; Clackson and Heyer, 2014). Crucially, however, taking into account the direction of the effects, there are patterns that cannot be explained by either account without employing additional assumptions: The cue-based retrieval account as implemented by Lewis and Vasishth (2005) and employed by Dillon (2011), Dillon et al. (2013), Kush and Phillips (2014), Parker and Phillips (2014) and Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English" (unpublished manuscript) is unable to explain facilitatory interference in antecedent-match conditions or inhibitory interference in antecedent-mismatch conditions.

The present article i) provides further experimental evidence relating to the current debate about the use of non-structural retrieval cues and ii) proposes two extensions to the standard cue-based retrieval architecture in order to account for the seemingly contradictory patterns of experimental results observed across studies.

We first present two eye-tracking experiments examining interference effects in the processing of the Mandarin Chinese reflexive *ziji*. There is a wide range of competing syntactic or pragmatic approaches of how to analyze *ziji* (for formal accounts see Yang, 1983; Manzini and Wexler, 1987; Pica, 1987; Kang, 1988; Tang, 1989; Huang and Tang, 1989, 1991; Cole et al., 1990, 1993; Cole and Sung, 1994; Cole and Wang, 1996; for pragmatic and non-uniform accounts see Huang et al., 1984; Yu, 1992; Xue et al., 1994; Yu, 1996; Pan, 1997; Pollard and Xue, 1998; Huang and Liu, 2001; Liu, 2010). We will restrict the following summary of the syntactic behavior of *ziji* to its properties that are relevant for the present experimental design. In contrast to English reflexives, *ziji* does not have any gender or number marking, but requires its antecedent to be animate (Tang, 1989).[4] Thus, animacy might be used as a non-structural cue to retrieve *ziji*'s antecedent. Similar to reflexives of many other languages including English, *ziji* needs to be c-commanded by its antecedent.[5] Moreover, the antecedent is required to be a subject (Huang, 1984). In contrast to English, the antecedent does not have to be contained in the local clause of the reflexive, but can also be contained in a superordinate clause (non-local binding). The processing of locally vs. non-locally bound *ziji* has been investigated by Gao et al. (2005), Liu (2009), Li and Zhou (2010), Dillon (2011), Chen et al. (2012b), and Dillon et al. (2014).

[4]There are some exceptions under which the animacy constraint can be violated, see Tang (1989) and Pan (1995) for a discussion. Crucially for our experimental design, in the syntactic literature, there is no example of non-emphatic, mono-morphemic *ziji* in argument position bound by a clearly inanimate NP.

[5]The c-command constraint might be violated in case of animate sub-commanding antecedents (Tang, 1989; Xue et al., 1994; Pollard and Xue, 1998), psychological verbs (Huang and Tang, 1991), passives and *ba*-constructions (Yu, 1992, but cf. Cole and Wang, 1996), and in case of cataphoric binding by the subject of a matrix clause that is preceded by an adjunct clause containing *ziji* (Huang and Liu, 2001). Moreover, *ziji* can refer to the speaker of the utterance (Li, 1991), the addressee, or even a third person salient in the discourse (Pan, 2000).

The present experiments examine whether animate nouns that are in a structurally inaccessible position (i.e., not c-commanding the reflexive) induce interference effects on the processing of *ziji*. So far, the literature on interference effects in reflexives has focused on morphologically marked phi-features (gender, number). Thus, the examination of animacy in the processing of Mandarin *ziji* does not only add cross-linguistic evidence to the debate that, so far, has been centered on English, but also extends the range of investigated retrieval cues to a purely semantic feature.

Both experiments have relatively large sample sizes in order to increase statistical power. Given that the prediction of the structure-based account is that no effect should be seen (i.e., a null result), it is particularly important to conduct high power studies.

2.2 Experiment 1

In Experiment 1, we tested whether locally bound *ziji* is subject to interference from a structurally inaccessible distractor that fulfills the animacy requirement of *ziji*. In a 2×2 design we manipulated animacy of the structurally accessible antecedent (henceforth labeled as *antecedent-match* vs. *antecedent-mismatch*) and of a structurally inaccessible distractor noun that intervened between the accessible antecedent and the reflexive (henceforth labeled as *distractor-match* vs. *distractor-mismatch*). This design extends the study reported by (Chen et al., 2012b), who were the first to test interference effects in Mandarin *ziji*, in several respects. In contrast to Chen and colleagues, in the present experiment, *ziji* was in object position rather than being a possessive modifier and we included antecedent-mismatch conditions which Chen et al. did not test. Moreover, we used the more time-sensitive eye-tracking method rather than self-paced reading.

The ACT-R model as implemented by Lewis and Vasishth (2005) predicts an inhibitory interference effect in antecedent-match conditions and a facilitatory interference effect in antecedent-mismatch conditions at the reflexive. The structure-based account (Nicol and Swinney, 1989; Sturt, 2003; Phillips et al., 2011; Dillon, 2011; Dillon et al., 2013; Kush and Phillips, 2014), in contrast, predicts the absence of an interference effect in both antecedent-match and antecedent-mismatch conditions. Moreover, the Lewis and Vasishth ACT-R model predicts incorrect retrievals of the animate distractor (misretrievals) in both antecedent-match and antecedent-mismatch conditions, but the proportion of misretrievals is predicted to be higher in antecedent-mismatch conditions. The structure-based account predicts no misretrievals of the animate inaccessible distractor.

2.2.1 Materials and method

2.2.1.1 Materials

We tested 48 experimental sentences which contained an either animate (antecedent-match) or inanimate (antecedent-mismatch) accessible antecedent in subject position (*yundongyuan* 'athlete' vs. *pihuating* 'kayak' in 2.2) and the reflexive as direct object. Due to the animacy requirement of *ziji*, the conditions with an inanimate accessible antecedent were ungrammatical.

Between the main clause subject and the main clause verb, an adverbial clause intervened that contained an either animate (distractor-match) or inanimate (distractor-mismatch) inaccessible distractor (*lingdui* 'team leader' vs. *meiti* 'media' in 2.2). This distractor was also a subject, but did not c-command the reflexive and was therefore not a legal antecedent. The reflexive was followed by a frequency phrase or a durational phrase which was analyzed as a spillover region.

(2.2) *Animate/Inanimate antecedent; Animate/Inanimate distractor*

运动员$_i$/*皮划艇$_i$				在	领队$_j$/媒体$_j$		施加	巨大
Yundongyuan$_i$/***Pihuating**$_i$			[PP	zai	**lingdui**$_j$/**meiti**$_j$		shijia	juda
athlete/kayak				when	team leader/media		exert	great
压力	的	情况	下	超越了	自己$_{i/*j}$		一共	三
yali	de	qingkuang	xia]	chaoyue-le	**ziji**$_{i/*j}$		yigong	san
pressure	MOD	circumstance	under	outperform-ASP	self		in total	three
次...								
ci...								
times...								

When the team leader/media exerted great pressure, the athlete/kayak outperformed himself/itself three times in total...

The experimental items were complemented with 72 filler sentences (48 grammatical, 24 ungrammatical) with varying syntactic structures including sentences containing the bare reflexive *ziji* as well as the bi-morphemic reflexive *ta-ziji* and pronouns in different syntactic positions.

Each sentence was followed by a multiple choice comprehension question that probed for the correct retrieval of the antecedent. Participants could choose between the antecedent, the distractor, an unrelated noun taken from a previous trial and the option "I am not sure". This design allowed us to examine not only whether the antecedent was retrieved correctly, but also to assess the proportion of misretrievals of the distractor. To ensure that participants also fully parsed the intervening adverbial clause containing the distractor, a second multiple-choice question targeted the adverbial clause. The same options were provided as in the first question. The questions following the filler sentences targeted various syntactic positions in the sentence.

Pretest. Since the exact binding properties of *ziji* are still subject to discussion in the syntactic literature, we conducted a paper-based questionnaire study to test our assumption that the main clause subject in the experimental items binds the reflexive. 40 native speakers of Mandarin recruited at Beijing Normal University participated in this study against payment of 25 RMB (approx. 3 EUR). None of them would participate in either of the eye-tracking experiments. Participants were presented with the antecedent-match conditions of the experimental items together with 90 filler sentences containing *ziji* in various syntactic positions and were instructed to circle the word in the given sentence *ziji* referred to or to explicitly write down the referent in case of an unbound interpretation of *ziji*.

Results. In 97.2% of all trials, participants selected the main clause subject as antecedent for the reflexive (97.0% and 97.3% when the distractor was animate or inanimate, respectively). This shows that in the experimental materials, Mandarin speakers indeed choose the main clause subject as antecedent for the reflexive.

2.2.1.2 Participants and procedure

The experiment was conducted in the eye-tracking lab of the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University. 150 students from different universities located in Beijing participated in the experiment against payment of 40 RMB (approx. 5 EUR). All participants were native speakers of Mandarin and had normal or corrected to normal vision.

Eye movements (right eye monocular) were recorded using an SR Research Eyelink 1000 eye-tracker at a sampling rate of 1000 Hz. Participants' head was stabilized using a forehead- and chin-rest. The screen-to-eye-distance was 82 cm, the camera-to-eye-distance 75 cm. Stimuli were presented in Simplified Chinese characters (font type SimSun, black font, font size 25) on a 22 inch monitor with light gray background using SR Research Experiment Builder software. Re-calibrations were performed between trials if necessary. Each experimental session began with 6 practice trials in which feedback to the comprehension questions was provided. In the experimental trials, no feedback was given. Short breaks were given according to the participants' individual needs. The sentences were presented according to a standard Latin Square. Items were pseudo-randomized such that at least one filler sentence intervened between two experimental sentences. Each sentence was followed by two multiple choice comprehension questions as described above.

2.2.2 Results

All statistical analyses were carried out in R using linear mixed effects models provided by the lme4 package version 1.0-6 (Bates et al., 2014). Binary dependent variables were analyzed using a logistic link function. For both, the analysis of response accuracies and eye movements, two sets of contrasts were applied. We first ran a model testing for a main effect of antecedent (animate antecedents coded as $+0.5$; inanimate antecedents coded as -0.5), a main effect of interference (animate distractors coded as $+0.5$; inanimate distractors coded as -0.5) and the interaction between the two main effects. Second, we applied nested contrasts testing for an interference effect within antecedent-match and antecedent-mismatch conditions separately. All models were fit with a full variance-covariance matrix for participants and items (Gelman and Hill, 2007); in case the model failed to converge or the variance-covariance matrix was degenerate, random slopes for items or participants were removed.

2.2.2.1 Comprehension questions

Comprehension questions targeting the reflexive-antecedent dependency were analyzed. We analyzed response accuracies and the proportion of incorrect selection of the inaccessible distractor. An overview of participants' answers is provided in Table 2.2. In the statistical analysis of response accuracies, only the main effect of antecedent reached marginal significance (estimate=0.34, SE=0.18, z=1.84, p=0.07). The antecedent (i.e., the correct option) was chosen more often in antecedent-match conditions. This effect was expected since in the antecedent-mismatch conditions, no fully grammatically correct answer to the comprehension question was available (the antecedent was coded as 'correct' answer, but the option 'not sure' was provided

as one response option in order to account for the ungrammaticality of the sentence). The analysis of the proportions of incorrect selection of the distractor revealed a main effect of antecedent: participants chose the distractor more often in antecedent-mismatch conditions than in antecedent-match conditions (estimate=-0.45, SE=0.18, z=-2.48, $p<0.05$). However, the size of this main effect was very small. We will therefore not base any conclusions on this effect. Moreover, the interaction between antecedent and distractor was significant (estimate=0.56, SE=0.15, z=3.61, $p<0.001$). Pairwise comparisons revealed that, within antecedent-match conditions, the distractor was chosen more often erroneously as answer to the comprehension question in case the distractor was animate (estimate=0.83, SE=0.31, z=2.70, $p<0.01$). But, as can be seen from Table 2.2, the animate distractor did not cause a decrease in selection probability of the antecedent but rather attracted selections from the unrelated noun. In antecedent-mismatch conditions, no interference effect was observed.

Antecedent	**Distractor**	**Chosen Answer**			
		antecedent	distractor	unrelated	'not sure'
match	*match*	82.3	5.1	0.9	11.7
	mismatch	81.6	3.6	2.4	12.4
mismatch	*match*	75.9	4.8	1.1	18.2
	mismatch	75.7	4.9	0.8	18.5

TABLE 2.2: Experiment 1. Chosen answer to the comprehension question by condition in percentages.

2.2.2.2 Eye movements

Eye movements were analyzed at the reflexive, the pre-critical region (verb) and the spillover material consisting of the frequency/durational phrase (post-critical). In order to provide a comprehensive picture of our data, and to make our results comparable to other studies we report the whole range of eye-tracking measures common in psycholinguistic research, although some of these measures are correlated by definition. As first-pass measures, we report first-fixation duration (FFD), i.e., the duration of the first fixation in first-pass reading, and first-pass reading time (FPRT, also called gaze duration), i.e., the sum of all first-pass fixations on a word before leaving it. As regression-related measures, we report regression-path duration (RPD, also called go-past time), i.e., the sum of all fixation durations starting from the first first-pass fixation on a word including regressive fixations to previous material until a region to the right of this word is fixated, right-bounded reading time (RBRT), i.e., the sum of all fixations on a word before another region to the right of this region is fixated, and first-pass regression probability (FPRP), i.e., the proportion first-pass regressions initiated from a word. As a later-pass measure, we analyzed re-reading time (RRT), i.e., the sum of all fixations on a word that are not contained in first-pass reading time. In addition, we analyzed total-fixation time (TFT), which is defined as the sum of FPRT and RRT. In order to achieve close to normally distributed model residuals, we log-transformed reading times (Box and Cox, 1964) and excluded all trials in which the respective continuous dependent variable was zero. First-fixation probability of the pre-critical region, the reflexive and the spillover region was 90%, 62%, and 87%, respectively. Re-readings occurred in 60%, 33%, and 45% of the trials at pre-critical region, the reflexive and the spillover

region, respectively. In all models, centered log-frequencies of the antecedent and the distractor taken from the SUBLETEX-CH database (Cai and Brysbaert, 2010) were included as covariates because items had not been matched for frequencies of the antecedents and distractors. Mean raw reading times with standard errors for the pre-critical, critical and post-critical regions are provided in Table 2.3. The results of the statistical analyses of participants' eye movements are summarized in Tables 2.4 and 2.5.

	Pre-critical				**Reflexive**				**Post-critical**			
Antecedent	*match*		*mismatch*		*match*		*mismatch*		*match*		*mismatch*	
Distractor	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*
FFD	279 (3)	277 (3)	285 (3)	279 (3)	258 (3)	259 (3)	264 (4)	251 (3)	270 (4)	274 (4)	274 (3)	268 (4)
FPRT	366 (6)	370 (6)	386 (6)	375 (6)	269 (4)	270 (4)	282 (5)	263 (4)	376 (8)	370 (8)	384 (7)	364 (7)
RBRT	397 (6)	407 (7)	425 (7)	413 (7)	286 (4)	284 (4)	302 (5)	284 (4)	436 (9)	430 (9)	448 (9)	432 (9)
RPD	484 (13)	508 (14)	537 (15)	533 (15)	430 (16)	410 (15)	484 (18)	494 (25)	688 (25)	662 (23)	759 (27)	755 (30)
FPRP	13 (1)	14 (1)	14 (1)	16 (1)	17 (1)	15 (1)	19 (1)	17 (1)	24 (1)	24 (1)	26 (1)	26 (1)
TFT	725 (14)	696 (14)	761 (15)	716 (14)	439 (10)	428 (9)	455 (10)	433 (9)	628 (14)	614 (15)	628 (15)	605 (13)
RRT	577 (17)	537 (16)	604 (17)	565 (16)	418 (15)	396 (14)	411 (14)	397 (13)	507 (18)	503 (20)	509 (21)	493 (17)

TABLE 2.3: Experiment 1. Means and standard errors of raw first-fixation duration, first-pass reading time, right-bounded reading time, regression-path duration, total fixation time, re-reading time in ms and first-pass regression probability in percentages at the pre-critical region, the reflexive and the post critical region. In the calculation of standard errors of continuous dependent variables, between-participants variance has been removed using the Cousineau (2005) normalization with Morey (2008)'s correction.

The main effect of antecedent (longer reading times or a higher proportion of regressions in antecedent-mismatch conditions) was significant across regression-related measures (RPD, RBRT, FPRP) and late measures (TFT, RRT). In RPD and RBRT, the effect of antecedent started already at the pre-critical region and remained significant at the reflexive and the post-critical region. In FPRP, the effect was significant at the reflexive only. In TFT, the effect also started at the pre-critical region and continued to be significant at the reflexive. In RRT, the effect reached significance only at the pre-critical region.

The main effect of interference (longer reading times or higher proportion of regressions in distractor-match conditions) reached significance across first-pass, regression-related and late measures. In RPD and FPRP, the effect reached significance at the reflexive itself, in FPRT and RBRT at the post-critical region and in TFT at the pre-critical region.

The interaction between antecedent and interference reached significance at the reflexive across first-pass and regression-related measures (FFD, FPRT, RBRT). In RBRT, this interaction was already present at the pre-critical region. Pairwise comparisons revealed that the interference effect was driven by the antecedent-mismatch conditions: Within antecedent-mismatch conditions, an inhibitory interference effect was observed across first-pass, regression-related and late measures (FFD, FPRT, RBRT, RPD, TFT).[6] In FFD, FPRT, RBRT and RPD, the effect

[6]In RPD, the effect predicted by the linear-mixed model is also an *inhibitory* one, although the opposite pattern is present in the raw means (cf. Table 2.3). This discrepancy is driven by a few very long (i.e., $>$ 6000 ms) regression-path durations in the antecedent-mismatch/distractor-mismatch condition of one particular item. Because of the concave nature of the log-function, the log-transformation of the data reduces the impact of these extremely high values. As all of these extreme values stem from the same experimental condition, the difference in means of the log-transformed RPDs even switches the sign in antecedent-mismatch conditions (log-transformed means in antecedent-mismatch conditions: distractor-match = 5.85 log-ms; distractor-mismatch = 5.80 log-ms). This explains why the linear-mixed model estimates an inhibitory rather than a facilitatory interference effect. Removing the item which caused the extreme values yields similar results as log-transforming the data, i.e., the sign of the interference

reached significance at the reflexive itself and, in FPRT, continued to be significant at the post-critical region. In TFT, the effect reached significance at the pre-critical region only. Within antecedent-match conditions, the interference effect did not reach significance in any measure or region.

Moreover, the models revealed that the higher frequency of the antecedent led to a significant slowdown at the reflexive in regression-based measures (RPD: estimate=0.03, SE=0.01, t=2.12; RBRT: estimate=0.02, SE=0.01, t=2.00) and RRT (estimate=0.05, SE=0.02, t=2.76). Frequency of the distractor, in contrast, did not affect reading times at the reflexive in any measure.

DV	**Predictor**	**Pre-critical**			**Reflexive**			**Post-critical**		
		coef	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*
FFD	Antecedent	-0.02	0.01	-1.72	0.00	0.01	-0.27	0.00	0.01	-0.10
	Interference	0.01	0.01	0.84	0.02	0.01	1.42	0.01	0.01	0.45
	Ant × Int	-0.01	0.01	-0.61	**-0.02**	**0.01**	**-2.06 ***	-0.02	0.01	-1.77
FPRT	Antecedent	-0.03	0.01	-1.93	-0.02	0.01	-1.02	-0.02	0.02	-0.96
	Interference	0.01	0.01	0.71	0.03	0.02	1.94	**0.04**	**0.01**	**2.69 ***
	Ant × Int	-0.02	0.01	-1.71	**-0.04**	**0.01**	**-2.92 ***	-0.02	0.01	-1.24
RBRT	Antecedent	**-0.04**	**0.01**	**-2.78 ***	**-0.04**	**0.02**	**-2.23 ***	**-0.04**	**0.02**	**-2.65 ***
	Interference	0.00	0.01	-0.20	0.02	0.02	1.59	**0.03**	**0.01**	**2.12 ***
	Ant × Int	**-0.03**	**0.01**	**-2.02 ***	**-0.03**	**0.01**	**-2.14 ***	-0.01	0.01	-0.56
RPD	Antecedent	**-0.06**	**0.02**	**-2.57 ***	**-0.09**	**0.02**	**-3.98 ***	**-0.09**	**0.03**	**-3.14 ***
	Interference	-0.02	0.02	-1.11	**0.04**	**0.02**	**2.25 ***	0.03	0.02	1.11
	Ant × Int	-0.02	0.02	-1.30	-0.01	0.02	-0.73	-0.01	0.02	-0.32
FPRP	Antecedent	-0.08	0.10	-0.79	**-0.18**	**0.08**	**-2.17 ***	-0.12	0.09	-1.41
	Interference	-0.14	0.10	-1.46	**0.16**	**0.07**	**2.14 ***	0.03	0.07	0.42
	Ant × Int	0.03	0.09	0.35	0.01	0.07	0.19	0.03	0.06	0.43
TFT	Antecedent	**-0.04**	**0.02**	**-2.20 ***	**-0.04**	**0.02**	**-2.29 ***	-0.01	0.02	-0.64
	Interference	**0.04**	**0.01**	**2.76 ***	0.02	0.02	1.48	0.03	0.02	1.93
	Ant × Int	-0.01	0.01	-0.83	-0.02	0.02	-1.12	0.01	0.02	0.57
RRT	Antecedent	**-0.06**	**0.03**	**-2.22 ***	-0.05	0.03	-1.47	-0.02	0.03	-0.68
	Interference	0.04	0.02	1.74	0.02	0.03	0.87	-0.01	0.03	-0.46
	Ant × Int	0.02	0.02	0.72	-0.01	0.03	-0.42	0.03	0.03	1.06

TABLE 2.4: Experiment 1. Main effects of antecedent and interference and their interaction at the pre-critical (*ziji*-1), critical (*ziji*) and post-critical (*ziji*+1) regions for the dependent variables (DVs) first-fixation duration, first-pass reading time, right-bounded reading time, regression-path duration, first-pass regression probability, total fixation time and re-reading time.

One potential issue with the data analysis reported here is the so-called multiple-testing problem, that is, testing more than one dependent variable but keeping the significance threshold α unchanged at 0.05. Although in the field of psycholinguistics it is uncommon to apply an α-level correction when multiple eye-tracking measures are analyzed, we applied a Bonferroni correction to the α-level (Bonferroni, 1936; Dunn, 1959, 1961) and checked whether the effects reported

effect also switches from negative to positive (raw means in antecedent-mismatch conditions with the item causing extremely long RPDs being removed: distractor-match = 476 ms; distractor-mismatch = 469 ms).

DV	Predictor	Pre-critical			Reflexive			Post-critical		
		coef	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*
FFD	Interference [ant. match]	0.00	0.02	0.26	0.00	0.02	-0.23	-0.01	0.02	-0.93
	Interference [ant. mismatch]	0.02	0.02	1.09	**0.04**	**0.02**	**2.3 ***	0.02	0.02	1.55
FPRT	Interference [ant. match]	-0.01	0.02	-0.71	-0.01	0.02	-0.40	0.02	0.02	1.05
	Interference [ant. mismatch]	0.03	0.02	1.68	**0.07**	**0.02**	**3.16 ***	**0.06**	**0.02**	**2.79 ***
RBRT	Interference [ant. match]	-0.03	0.02	-1.46	-0.01	0.02	-0.52	0.02	0.02	1.13
	Interference [ant. mismatch]	0.02	0.02	1.24	**0.05**	**0.02**	**2.07 ***	0.04	0.02	1.89
RPD	Interference [ant. match]	-0.04	0.03	-1.59	0.03	0.03	1.11	0.02	0.03	0.61
	Interference [ant. mismatch]	0.00	0.02	0.03	**0.06**	**0.03**	**2.07 ***	0.03	0.03	1.04
FPRP	Interference [ant. match]	-0.11	0.13	-0.84	0.17	0.10	1.64	0.03	0.09	0.33
	Interference [ant. mismatch]	-0.17	0.13	-1.36	0.14	0.10	1.40	0.00	0.09	0.00
TFT	Interference [ant. match]	0.03	0.02	1.40	0.01	0.02	0.26	0.04	0.02	1.72
	Interference [ant. mismatch]	**0.05**	**0.02**	**2.52 ***	0.04	0.02	1.80	0.02	0.02	1.04
RRT	Interference [ant. match]	0.06	0.03	1.84	0.01	0.04	0.33	0.02	0.04	0.41
	Interference [ant. mismatch]	0.03	0.03	0.77	0.04	0.04	0.90	-0.04	0.04	-1.06

TABLE 2.5: Experiment 1. Pairwise comparisons of animacy of the distractor (interference) nested within animate/inanimate antecedent (antecedent match/mismatch) at the pre-critical (*ziji*-1), critical (*ziji*) and post-critical (*ziji*+1) regions for the dependent variables (DVs) first-fixation duration, first-pass reading time, right-bounded reading time, regression-path duration, first-pass regression probability, total fixation time and re-reading time.

above remained significant under this more conservative analysis. This is important in order to reduce the Type I error probability because, as has been noted for example by Ioannidis (2005), false positives are serious issues in empirical science and in psychological science in particular (Simmons et al., 2011). With respect to reading studies, von der Malsburg and Angele, "The elephant in the room: False positive rates in standard analyses of eye movements in reading" (unpublished manuscript) recently showed by means of Monte Carlo simulations that testing multiple eye-tracking measures leads to a more dramatic increase of Type I errors as compared to what had been generally believed in the field. Von der Malsburg and Angele therefore recommend to apply a Bonferroni correction to the α-level. Given that we have analyzed seven dependent variables, the Bonferroni correction yields a corrected α-level of 0.007, which corresponds to an approximate t-value of ± 2.69.[7] With this adjusted α-level, the main effect of antecedent remained significant in RBRT at the pre-critical region and in RPD at the reflexive and at the post-critical region. The main effect of interference reached significance in FPRT at the post-critical region and in TFT at the pre-critical region. The interaction between antecedent and interference was significant in FPRT at the reflexive. In pairwise comparisons, the interference effect in antecedent-mismatch conditions in FPRT at the reflexive and at the post-critical region remained significant. The antecedent-frequency effect reached the Bonferroni-corrected significance threshold in RRT, but not in RPD and RBRT. In sum, although the Bonferroni correction and the considerable loss in statistical power that goes along with it makes some

[7] This t-value was approximated by using a normal distribution.

effects lose statistical significance, the overall pattern of results remains unchanged: An early interference effect at the reflexive present only within antecedent-mismatch conditions, an effect of antecedent in regression-related dependent variables starting already at the verb preceding the reflexive and an effect of antecedent-frequency at the reflexive.

2.2.3 Discussion

Comprehension questions required participants to correctly identify the reflexive's antecedent and to select it from four response options. Although participants could choose the option 'not sure', they were highly likely to choose the antecedent even if it was inanimate and hence a semantically illicit antecedent. This shows that in their final interpretation of the reflexive they gave structural information a higher priority than semantic information. In antecedent-match conditions only, the distractor was chosen more often in case it was animate. But, crucially, this higher proportion of distractor choices was at the cost of choices of the unrelated noun, not of the antecedent. From this pattern we conclude that the observed effect reflects *offline* interference, i.e., an effect driven by meta-linguistic considerations at the moment of answering the comprehension question. If, in contrast, the effect reflected retrieval interference during the actual sentence reading, i.e., online effects, it would be expected to manifest itself in a higher proportion of misretrievals of the distractor leading to a lower proportion of choosing the *antecedent*, not the unrelated noun, because the latter is only introduced in the question.

The analyses of eye movements firstly showed that the presence of an animate distractor led to a processing slowdown (i.e., inhibitory interference) in antecedent-mismatch conditions. This slowdown was observed across first-pass, regression-related and late measures. In the more conservative analysis with Bonferroni-corrected significance threshold, this slowdown remained reliable in first-pass reading time. In antecedent-match conditions, this interference effect did not reach significance. This pattern cannot be explained by either of the two accounts under discussion: The parser's sensitivity to the presence of an animate distractor cannot be accounted for by a structure-based retrieval mechanism. ACT-R cannot explain the results either since, in its current implementation, ACT-R predicts facilitatory rather than inhibitory interference in antecedent-mismatch conditions caused by a higher proportion of misretrievals of an animate distractor. Kush and Phillips (2014) also found inhibitory interference in antecedent-mismatch conditions in a self-paced reading experiment on Hindi reciprocals. They explain this effect in terms of interference that occurs during a later repair process of the ungrammatical sentence rather than at the moment of retrieval. Crucially, in Kush and Phillips (2014)'s experiment, the interference effect reached marginal significance only two words after the reciprocal. For the present experiment, their explanation seems implausible since the interfere effect reaches significance already in first-pass measures at the reflexive.

Second, we did *not* find any interference effects in the antecedent-match conditions. Although these results are statistically inconclusive, it is worth mentioning that this is consistent with the findings of Chen et al. (2012b), who found interference effects in non-locally bound *ziji* but failed to find effects in locally-bound *ziji*.

Third, we observed a slowdown due to an inanimate antecedent in regression-related and late measures. This grammaticality effect is in line with both structure-based retrieval and the ACT-R model. In contrast to the interference effect, this effect is most pronounced at the pre-critical region. We will discuss possible explanations for this early appearance of the effect in the Discussion of Experiment 2.

Fourth, we found that lower frequency of the antecedent led to faster reading times at the reflexive. This effect might be explained by a low-frequency encoding advantage. It has been shown that the lower frequency of a word leads to a better memory encoding which results in a faster retrieval at a later point in time (Diana and Reder, 2006). Thus, low frequency antecedents might be better encoded in memory leading to a facilitated retrieval when reaching the reflexive, which shows the more prominent role of the antecedent in the retrieval process. Indeed, this facilitation due to infrequent antecedents replicates findings from English pronouns. In an eye-tracking-while-reading experiment, Van Gompel and Majid (2004) found faster first-fixation durations and first-pass reading times at the region following the reflexive as a function of lower frequency of the antecedent.

One potential concern with the present results might be that task-related influences on interference cannot be ruled out. One of the two comprehension questions following the experimental sentences targeted the reflexive-antecedent dependency, which — in particular in the ungrammatical conditions — might have caused readers to spend some additional reading time to rule out the animate distractor. This would explain the observed inhibitory interference in the target-mismatch conditions. In the design of the experiment, we had addressed this potential issue by including ungrammatical fillers containing *ziji* with questions that did not target the reflexive-antecedent dependency. Moreover, participants had the option to answer "not sure", which allowed them not to assign any meaning to an ungrammatical sentence. Most importantly, the interference effect reached significance already in first-fixation duration and first-pass reading time. Based on a large-scale review of eye movements in reading, Clifton et al. (2007) have suggested that early measures like first-fixation duration or first-pass reading time are unlikely to reflect repair processes since across studies, repair or reanalysis effects are typically observed in regression-related or later-pass reading measures. To the extent that Clifton et al. (2007)'s claim is correct, we can conclude that repair processes caused by the task-demands are unlikely to explain the observed results.

2.3 Experiment 2

This experiment extended Experiment 1 in several aspects. First, it examined proactive rather than retroactive interference; second it examined the influence of distractor items that are not a syntactic part of the sentence itself but presented as memory load; third, we tested the influence of syntactic locality on the retrieval and its interaction with interference. Previous studies report a processing slowdown in case *ziji* is non-locally bound compared to locally bound *ziji* (Gao et al., 2005; Li and Zhou, 2010; Dillon, 2011; Chen et al., 2012b; Dillon et al., 2014). In the present experiment, we aimed at replicating this locality effect and investigating whether interference effects are modulated by locality of the reflexive binding.

In a dual-task paradigm, similar to Van Dyke and McElree (2006), participants were asked to remember three animate or three inanimate distractor nouns while reading a sentence containing an either locally or non-locally bound reflexive. This resulted in a 2×2 design, with locality (local vs. non-local) and the distractors' animacy (animate vs. inanimate) as factors. Conditions with animate distractors are labeled as *distractors-match* and conditions with inanimate distractors as *distractors-mismatch*.

The structure-based account predicts no effect of animacy of the distractor nouns held in memory. In contrast, the standard ACT-R cue-based retrieval model predicts an inhibitory interference effect due to animacy of the distractors: retrieval times at the reflexive are predicted to be longer in distractors-match conditions. Moreover, ACT-R predicts a main effect of locality with non-local conditions being read slower. This prediction does not follow from the cue-based nature of the retrieval mechanism but rather from the ACT-R assumption of decay: The more recent, i.e., the local, antecedent has a higher level of activation than the non-local antecedent when reaching the reflexive. This difference in activation is predicted to be reflected in both, retrieval times and comprehension accuracies. Since this predicted locality effect is unrelated to the set of cues used for retrieval, the structure-based cue-based retrieval account (i.e., the ACT-R model with only structural features used as retrieval cues) makes the same prediction. Moreover, a structure-based serial search mechanism that first checks the local subject position and subsequently the non-local subject as proposed by Dillon (2011) and Dillon et al. (2014) for the processing of Mandarin *ziji* also predicts a processing slowdown in non-local conditions.

2.3.1 Materials and method

2.3.1.1 Materials

We tested 36 experimental sentences[8] which consisted of a superordinate clause and an embedded clause containing the reflexive *ziji* as direct object. The locality factor of the antecedent-reflexive dependency was achieved by manipulating animacy of the local subject (i.e., the subject of the embedded clause) and the non-local subject (i.e., the subject of the superordinate clause): in the local conditions, the local subject was animate and the non-local subject was inanimate (see 2.3a) while in the non-local conditions, the local subject was inanimate and the non-local subject was animate (see 2.3b). Since *ziji* requires its antecedent to be animate, this design ensured that in the local conditions, *ziji* was bound by the local subject whereas in the non-local conditions it was bound by the subject of the superordinate clause. Similar to Experiment 1, the reflexive was followed by a spillover region consisting of four characters that formed a frequency phrase or a durational phrase. Each sentence was followed by a yes/no-comprehension question that probed for the correct binding of the reflexive. 72 filler sentences containing reflexives and pronouns in varying syntactic positions were presented with memory load words of varying part-of-speech.

[8]Originally, we had 48 items, but 12 of these were excluded based on low acceptability judgements of native speakers.

(2.3) a. *Local binding*

这些	数据$_i$	表明	这个	少年$_j$	耽误了	自己$_{*i/j}$
Zhe-xie	**shuju**$_i$	biaoming	[zhe-ge	**shaonian**$_j$	danwu-le	**ziji**$_{*i/j}$
this-CL	data	demonstrate	this-CL	youngster	hinder-ASP	self
整整	三	年...				
zhengzheng	san	nian]...				
wholly	three	years...				

These data demonstrate that this youngster hindered himself three whole years...

b. *Non-local binding*

这个	少年$_i$	表明	这些	数据$_j$	耽误了	自己$_{i/*j}$
Zhe-ge	**shaonian**$_i$	biaoming	[zhe-xie	**shuju**$_j$	danwu-le	**ziji**$_{i/*j}$
this-CL	youngster	demonstrate	this-CL	data	hinder-ASP	self
整整	三	年...				
zhengzheng	san	nian]...				
wholly	three	years...				

This youngster demonstrates that these data hindered him three whole years...

Pretest. In order to verify that speakers of Mandarin indeed bind the reflexive to the local subject/the superordinate subject in the local/non-local condition, respectively, we presented 40 native speakers of Mandarin recruited at Beijing Normal University with the experimental sentences in form of a paper-based questionnaire against payment of 25 RMB (approx. 3 EUR). 90 filler sentences containing *ziji* in various syntactic positions were included. Participants were instructed to circle the word in the sentence *ziji* referred to, or, in case they found that no antecedent was available in the sentence, to write down which entity *ziji* referred to.

Results. Overall, 90.4% of all trials were answered as we had expected: In the local conditions, the animate local subject was chosen as antecedent and in the non-local conditions the animate matrix subject was selected. In the local conditions, accuracy was lower (85.1%) than in the non-local conditions (95.6%). A syntactic classification of the incorrect answers is provided in the appendix.

2.3.1.2 Participants and procedure

This experiment was conducted in the same laboratory as Experiment 1. 130 native speakers of Mandarin with normal or corrected-to-normal vision participated in the experiment against payment of 60 RMB (approx. 7 EUR). The general experimental set-up was the same as in Experiment 1. The experiment was split into two experimental sessions (40-70 minutes per session) conducted on two subsequent days. At the beginning of each trial, the three distractors were shown on the screen one below another for three seconds. When the words disappeared, the test sentence was displayed. After having finished reading the sentence, the comprehension question was presented. After having answered the comprehension question, participants were asked to serially recall the distractors: The three distractors together with three unrelated items (similarly animate or inanimate nouns) were displayed simultaneously on the screen as a numbered list in randomized order. Participants were asked to choose the distractors in their correct order from this list.

2.3.2 Results

For all dependent variables, we fit two sets of contrasts; the first tested for main effects of locality (local conditions coded as −0.5; non-local conditions coded as +0.5) and interference (animate distractors coded as +0.5; inanimate distractors coded as −0.5) and their interaction; in the second model pairwise comparisons of memory load nested within each level of locality were applied. In addition, experimental session (first vs. second session) was coded with sum-contrasts and its interaction with the other effects were included as predictors. All models were fit with random intercepts for items and participants, no random slopes were fit since they led to convergence failure in most of the models.

2.3.2.1 Comprehension questions

Mean accuracy scores by experimental condition are shown in Table 2.6. None of the comparisons reached statistical significance.[9]

Locality	**Distractors**	**Accuracy**
local	*match*	67.1
	mismatch	68.7
non-local	*match*	71.8
	mismatch	71.9

TABLE 2.6: Experiment 2. Comprehension question response accuracy in percentage by experimental condition.

2.3.2.2 Memory recall

Mean serial and non-serial recall accuracies for each of the three distractors and total serial and non-serial recall accuracy (i.e., all distractors recalled correctly) are presented in Table 2.7. In the statistical analyses of total serial recall accuracy none of the comparisons reached significance. In the analyses of total non-serial accuracies, the interaction between animacy of the distractors and locality was significant (estimate=-0.22, SE=0.10, z=-2.21, $p<0.05$). Pairwise comparisons revealed that this interaction was driven by a significant effect of distractors (lower recall accuracy

[9] In response accuracies the proportion of correctly answered yes-questions was strikingly higher than the proportion of correctly answered no-questions. We can exclude the possibility that this pattern can be explained by a general tendency of the participants to answer 'yes' since no such difference was observed in filler sentences. We also excluded the hypothesis that this pattern might be related to the difficult nature of the dual-task paradigm by running a follow-up eye-tracking experiment (N=14) with the same experimental set-up but without memory load that yielded a similar response pattern. As the pre-test on the materials had shown that native speakers indeed do the correct binding of the reflexive, we hypothesized that the response pattern was intrinsically related to the nature of the comprehension questions rather than to the experimental sentences themselves. We therefore ran another experiment (N=52) in which the experimental and filler sentences appeared on the computer screen together with the respective comprehension question. Again, we observed a similar response pattern as in the online experiments. We thus conclude that the observed tendency to answer 'yes' on the experimental comprehension questions reflects an offline effect, i.e., an effect which occurs at the moment when participants meta-linguistically think about how to answer the question, rather than an effect of online reflexive binding.

of animate distractors) that was present only in local conditions (estimate=-0.30, SE=0.14, z=-2.25, p<0.05).

	Serial Accuracy				**Non-serial Accuracy**			
Locality	*local*		*non-local*		*local*		*non-local*	
Distractors	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*
1st word correct	85	83	85	85	92	94	94	94
2nd word correct	79	75	81	79	93	94	94	93
3rd word correct	82	78	83	82	90	91	92	91
total correct	68	67	71	69	77	81	82	80

TABLE 2.7: Experiment 2. Mean serial and non-serial recall accuracy in percentage of the three memory load words separately and total accuracy in percentage presented by experimental condition.

2.3.2.3 Eye movements

The same log-transformed dependent variables as in Experiment 1 were analyzed at the reflexive, the verb preceding it (pre-critical), and the spillover material (post-critical). As in the analysis of Experiment 1, trials were excluded when the continuous variable on which the analysis was carried out was zero. First-pass fixations occurred at the pre-critical region, the reflexive, and the spillover region in 86%, 50%, and 85% of the trials, respectively. Re-readings were recorded in 55%, 25%, and 36% of the trials at pre-critical region, the reflexive, and the spillover region, respectively. Mean reading times with standard errors for each dependent variable are provided in Table 2.8.

	Pre-critical				**Reflexive**				**Post-critical**			
Locality	*local*		*non-local*		*local*		*non-local*		*local*		*non-local*	
Distractors	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*	*match*	*mism.*
FFD	267 (5)	268 (5)	267 (5)	270 (5)	251 (6)	239 (5)	240 (5)	244 (6)	257 (6)	255 (5)	253 (5)	258 (6)
FPRT	342 (9)	341 (9)	351 (9)	343 (9)	260 (7)	245 (6)	250 (6)	254 (7)	325 (10)	320 (10)	322 (9)	320 (9)
RBRT	398 (11)	409 (11)	433 (12)	447 (13)	278 (7)	259 (6)	263 (7)	277 (8)	378 (11)	375 (12)	383 (12)	375 (11)
RPD	575 (25)	573 (26)	596 (24)	638 (28)	486 (28)	419 (27)	448 (29)	484 (32)	636 (35)	628 (36)	667 (41)	710 (47)
FPRP	23 (2)	20 (2)	25 (2)	28 (2)	18 (1)	14 (1)	16 (1)	16 (1)	24 (2)	24 (2)	25 (2)	25 (2)
TFT	683 (25)	666 (22)	737 (26)	763 (28)	396 (14)	354 (13)	377 (14)	379 (14)	501 (19)	491 (18)	508 (19)	479 (17)
RRT	626 (33)	592 (28)	645 (33)	704 (34)	352 (22)	365 (26)	360 (24)	345 (22)	432 (29)	414 (30)	468 (32)	413 (26)

TABLE 2.8: Experiment 2. Means and standard errors of raw first-fixation duration, first-pass reading time, right-bounded reading time, regression-path duration, total fixation time, re-reading time in ms and first-pass regression probability in percentages at the pre-critical region, the reflexive and the post critical region. In the calculation of standard errors of continuous dependent variables, between-participants variance has been removed using the Cousineau (2005) normalization with Morey (2008)'s correction.

The output of the linear-mixed models is summarized in Tables 2.9 and 2.10. The effect of experimental session was significant across regions and measures: Participants read faster in their second experimental session.[10] The main effect of locality reached significance across regression-based and later-pass measures (RBRT, RPD, FPRP, RRT, TFT) at the pre-critical region only. The main effect of interference was significant only in RRT at the post-critical region (longer RRTs when distractors were animate, i.e., inhibitory interference). The interaction between

[10]The effect of experimental session is not of theoretical interest to our research question, therefore it is not presented in the results tables and will not be discussed further.

locality and interference was significant across first-pass, regression-based, and later-pass measures (FFD, FPRT, RBRT, RPD, TFT) at the reflexive. The pairwise comparisons revealed that the interaction was driven by a slowdown for animate distractors at the reflexive that was present only in local conditions. This inhibitory interference reached significance across first-pass, regression-based, and later-pass measures (FPRT, RBRT, RPD, TFT). For non-local conditions, a similar slowdown was observed only in RRT at the post-critical region.

DV	**Predictor**	**Pre-critical**			**Reflexive**			**Post-critical**		
		coef	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*
FFD	Locality	0.00	0.01	-0.06	-0.01	0.01	-0.92	0.00	0.01	-0.19
	Interference	-0.01	0.01	-0.98	0.01	0.02	0.35	-0.01	0.02	-0.58
	Locality × Interference	0.00	0.02	0.00	**0.04**	**0.02**	**2.20 ***	0.02	0.02	1.21
FPRT	Locality	0.01	0.02	0.52	-0.01	0.02	-0.83	0.00	0.02	0.24
	Interference	0.00	0.02	0.26	0.01	0.02	0.63	0.00	0.02	-0.23
	Locality × Interference	0.00	0.02	-0.23	**0.04**	**0.02**	**2.33 ***	0.02	0.02	0.85
RBRT	Locality	**0.08**	**0.02**	**5.37 ***	-0.01	0.02	-0.34	0.01	0.02	0.36
	Interference	-0.01	0.02	-0.49	0.01	0.02	0.61	0.00	0.02	0.01
	Locality × Interference	0.00	0.02	0.03	**0.06**	**0.02**	**3.21 ***	0.02	0.02	0.79
RPD	Locality	**0.11**	**0.02**	**5.15 ***	0.03	0.03	1.10	0.02	0.03	0.78
	Interference	0.00	0.03	-0.08	0.04	0.03	1.36	-0.01	0.03	-0.30
	Locality × Interference	0.02	0.03	0.71	**0.07**	**0.03**	**2.16 ***	0.03	0.03	0.95
FPRP	Locality	**0.46**	**0.08**	**5.80 ***	0.11	0.09	1.19	0.05	0.08	0.62
	Interference	0.04	0.10	0.43	0.15	0.11	1.34	0.00	0.09	0.02
	Locality × Interference	0.14	0.10	1.41	0.09	0.11	0.83	-0.01	0.09	-0.12
TFT	Locality	**0.10**	**0.02**	**5.55 ***	-0.01	0.02	-0.31	0.00	0.02	-0.15
	Interference	0.01	0.02	0.57	0.05	0.02	1.92	0.02	0.02	1.01
	Locality × Interference	0.01	0.02	0.34	**0.08**	**0.02**	**3.10 ***	0.00	0.02	-0.16
RRT	Locality	**0.10**	**0.03**	**3.71 ***	-0.01	0.04	-0.23	0.06	0.04	1.73
	Interference	-0.02	0.03	-0.55	0.00	0.05	0.04	**0.09**	**0.04**	**2.10 ***
	Locality × Interference	0.07	0.03	1.92	-0.03	0.05	-0.64	-0.04	0.04	-0.89

TABLE 2.9: Experiment 2. Main effects of locality and interference and their interaction at the pre-critical (*ziji*-1), critical (*ziji*), and post-critical (*ziji*+1) regions for the dependent variables (DVs) first-fixation duration, first-pass reading time, right-bounded reading time, regression-path duration, first-pass regression probability, total fixation time and re-reading time.

As we did for Experiment 1, we checked which of the observed effects remained significant with a Bonferroni-corrected significance threshold. Given seven dependent variables, the corrected α-level is 0.007, which corresponds to an approximate t-value of ± 2.69.[11] The significance of the main effect of locality was not affected by this correction in any dependent variable, it remained significant at the pre-critical region in RBRT, RPD, FPRP, TFT, and RRT. The main effect of interference at the post-critical region in RRT did not reach the adjusted significance threshold. The interaction between locality and interference remained significant at the reflexive in RBRT and TFT, but did not reach significance anymore in FFD, FPRT, and RPD. In pairwise comparisons, the interference effect in local conditions at the reflexive remained significant in RBRT and TFT, but did not reach the significance threshold anymore in FPRT and RPD. The

[11] This t-value was approximated by using a normal distribution.

DV	Predictor	Pre-critical			Reflexive			Post-critical		
		coef	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*
FFD	Interference [local]	-0.01	0.02	-0.69	0.04	0.03	1.79	0.01	0.02	0.45
	Interference [non-local]	-0.01	0.02	-0.69	-0.03	0.02	-1.31	-0.03	0.02	-1.26
FPRT	Interference [local]	0.00	0.03	0.02	**0.06**	**0.03**	**2.09 ***	0.01	0.03	0.44
	Interference [non-local]	0.01	0.03	0.35	-0.03	0.03	-1.21	-0.02	0.03	-0.76
RBRT	Interference [local]	-0.01	0.03	-0.33	**0.07**	**0.03**	**2.70 ***	0.02	0.03	0.57
	Interference [non-local]	-0.01	0.03	-0.37	-0.05	0.03	-1.83	-0.02	0.03	-0.55
RPD	Interference [local]	0.02	0.04	0.44	**0.11**	**0.05**	**2.49 ***	0.02	0.04	0.46
	Interference [non-local]	-0.02	0.04	-0.56	-0.03	0.05	-0.57	-0.04	0.04	-0.89
FPRP	Interference [local]	0.18	0.14	1.23	0.24	0.15	1.52	-0.01	0.13	-0.07
	Interference [non-local]	-0.09	0.13	-0.74	0.06	0.15	0.36	0.01	0.13	0.10
TFT	Interference [local]	0.02	0.03	0.64	**0.12**	**0.03**	**3.56 ***	0.02	0.03	0.60
	Interference [non-local]	0.00	0.03	0.16	-0.03	0.03	-0.83	0.03	0.03	0.82
RRT	Interference [local]	0.05	0.05	0.94	-0.03	0.07	-0.42	0.05	0.06	0.87
	Interference [non-local]	-0.08	0.05	-1.79	0.03	0.06	0.49	**0.13**	**0.06**	**2.09 ***

TABLE 2.10: Experiment 2. Interference effect nested within each level of locality (local vs. non-local) at the pre-critical (*ziji*-1), critical (*ziji*) and post-critical (*ziji*+1) regions for the dependent variables (DVs) first-fixation duration, first-pass reading time, right-bounded reading time, regression-path duration, first-pass regression probability, total fixation time and re-reading time.

interference effect in non-local conditions that was observed at the post-critical region did not reach the adjusted significance threshold. In sum, the main effect of locality as well as the interference effect in locally bound *ziji* remained significant in various dependent variables even with an adjusted α-level. The interference effect in non-local conditions, in contrast, was not reliable under the corrected α-level.

2.3.3 Discussion

In the comprehension questions, no evidence for an interference effect was found. In the memory recall task, in contrast, we found that, in local conditions only, animate words were more difficult to recall than inanimate words.

First, we found evidence for a processing slowdown associated with the non-local binding of the reflexive. This locality effect replicates findings from SAT (Dillon, 2011; Dillon et al., 2014), ERP (Li and Zhou, 2010; Dillon, 2011), cross-modal priming (Liu, 2009), and self-paced reading (Chen et al., 2012b), and is accounted for by the ACT-R model, no matter whether the set of retrieval cues is unconstrained or limited to structural cues. The structure-based serial search as proposed by Dillon (2011) and Dillon et al. (2014) is also in line with the observed locality effect. However, it is not fully clear why this locality effect appears at the verb preceding the reflexive rather than at the reflexive itself. One explanation would be a preview effect. Alternatively, it

might be the case that the observed effect does not reflect locality of the reflexive binding but rather the verb's preference for an animate subject since the locality manipulation is achieved by having the local subject either animate or inanimate. Along the same lines, one could explain why in Experiment 1, the effect of animacy of the antecedent becomes significant already at the verb preceding the reflexive. A strong indication that the observed effect at the verb indeed reflects the verb's preference for an animate subject comes from a re-analysis of the self-paced reading data reported by Chen et al. (2012b), where the locality manipulation was also achieved by varying the animacy of the local and non-local subjects, and the main clause verb also directly preceded the reflexive *ziji*. Chen et al. (2012b) analyzed only the region containing the reflexive and the regions *following* the reflexive, but not the verb *preceding* the reflexive. Re-analyzing their data at the verb region revealed that the locality effect in their data was already significant at the verb (t=2.5). As preview effects are ruled out as an explanation in self-paced reading, and given the high structural similarity of our experimental materials to the ones used by Chen et al. (2012b), we conclude that the effect observed at the verb in Experiment 2 is most likely due to an animacy preference of the verb. Given this — admittedly unforeseen — confounding animacy preference of the verb, we cannot draw any conclusions about the actual locality manipulation. A potential locality effect might have been masked by the stronger effect of animacy preference: when reaching the verb in the non-local conditions, readers are highly likely to re-read the previous material to overcome the difficulty associated with the verb's inanimate subject, as indicated by the highly significant effects in first-pass regression probability, regression-path duration, and right-bounded reading time. This leads to activation of the preceding materials in the non-local conditions *directly before* reaching the reflexive, which, in turn, might have canceled out a locality effect at the reflexive. Therefore, we conclude that our data is inconclusive with respect to the locality manipulation.

Second, we found clear evidence for inhibitory interference, but the time-course of this effect was different for local and non-local conditions. In local conditions, animate distractors led to a slowdown across first-pass, regression-based, and late eye-tracking measures at the reflexive itself. Even with a Bonferroni corrected significance threshold of α=0.007, this effect remained significant in right-bounded reading time and total-fixation duration. In first-pass reading time and regression-path duration, the inhibitory interference effect did not survive Bonferroni correction. However, since these measures numerically pattern with other measures — especially with right-bounded reading time, which is closely related — it could reflect a real effect. In non-local conditions, the interference effect appeared only later in processing (in re-reading times at the post-critical region). However, with Bonferroni adjusted significance threshold, this effect was not reliable. In sum, the observed interference pattern extends the findings of Experiment 1 in two respects. First, Experiment 2 shows that locally bound *ziji* is subject to early interference even in case a fully cue-matching antecedent is available. The difference to Experiment 1, where the interference effect did not reach significance in antecedent-match conditions, might be explained by the different experimental paradigms: rehearsal of the distractors during reading might cause stronger interference than the sentence-internal manipulation of Experiment 1. Second, the interference profile in non-locally bound *ziji* differs from the one in locally bound *ziji* in the sense that in non-local conditions no early effect was found, but there is weak evidence for a late effect. Although the late effect in non-local conditions was not significant under Bonferroni

correction, there is reason to believe in this effect when viewed against the background of previous findings by Chen et al. (2012b), who found an inhibitory interference effect in non-local *ziji*.

The observed interference effects are not compatible with a structure-based retrieval mechanism since no effect of the distractors is predicted. The ACT-R model, in contrast, can account for the inhibitory interference effect. However, ACT-R is unable to explain the delayed appearance of the effect in non-local conditions.

A possible explanation for the different interference patterns in local vs. non-local conditions could be that qualitatively different mechanisms are involved in the processing of locally and non-locally bound *ziji*. In the syntactic literature, it has been proposed that only the locally bound *ziji* should be regarded as a reflexive pronoun whereas non-locally bound *ziji* should be regarded as a logophoric pronoun which is subject to pragmatic and discourse constraints rather than to purely syntactic binding principles (Huang and Liu, 2001; Huang, 2002). One prominent argument favoring this idea of two lexically different instances of *ziji* are blocking effects observed in long-distance *ziji* but not in local *ziji* (Huang, 1984; Tang, 1989; Huang and Tang, 1991; Xue et al., 1994; Pan, 2000; Huang, 2002). A qualitative distinction between locally bound *ziji* and non-local *ziji* has also been proposed in the psycholinguistic literature. Based on previous work by Gao et al. (2005), Liu (2009) conducted a cross-modal priming experiment using sentences in which both a local and a non-local animate antecedent were present (i.e., globally ambiguous sentences in terms of binding) and manipulated stimulus-onset asynchrony (0 ms, 160 ms, 370 ms). When the probe was presented directly after the offset of the reflexive (SOA = 0 ms), a semantic priming effect for probes related to the local antecedent but not for probes related to the non-local antecedent was observed. At an SOA of 160 ms, in contrast, the pattern was reversed: There was a priming effect for probes that were semantically related to the non-local antecedent, but no priming effect for probes related to the local antecedent. At an SOA of 370 ms, both the local and non-local antecedent elicited a semantic priming effect. Liu (2009) interpreted these results as evidence for *ziji* being bound by the local subject in a first stage of processing and by the non-local subject in a second stage of processing, whereas in the final stage, both bindings are possible. Along the same lines, Dillon (2011) and Dillon et al. (2014) suggested that the parser tries to first access the local subject and only at a later stage accesses non-local antecedent positions. Such a temporal delay for the triggering of the retrieval of a non-local antecedent would indeed predict the pattern observed in Experiment 2: In the local conditions, the retrieval is triggered immediately at the moment when the reflexive is first encountered. The interference effects associated with this retrieval therefore appear already in early measures at the reflexive. In non-local conditions, in contrast, the retrieval of the non-local antecedent is triggered only after a certain delay, which causes the interference effects to occur only in re-reading times at the spillover region.

2.4 An extended cue-based retrieval model

As has been pointed out in the experimental discussions, the interference effects observed in the experiments presented here are not compatible with structure-based accounts. The current

implementation of the standard cue-based retrieval model in ACT-R (Lewis and Vasishth, 2005) cannot explain the observed patterns either. In particular, standard cue-based retrieval is unable to explain i) why there is an effect in antecedent-match conditions in Experiment 2 but not in Experiment 1, and ii) why there is inhibitory interference observed in antecedent-mismatch conditions in Experiment 1. We propose an explanation of the observed patterns by adding two independently motivated assumptions to standard cue-based retrieval: that i) similarity-based interference is modulated by *distractor prominence* and that ii) *cue confusion* can lead to similarity-based interference between non-similar items. As discussed earlier, the difference in the interference profiles of local and non-local *ziji* might be due to a qualitative difference in processing mechanisms and was therefore not included in our modeling.

2.4.1 Principle 1: Prominence

In Experiment 1, we found an interference effect in antecedent-mismatch conditions but not in antecedent-match conditions. According to Wagers et al. (2009), this is an expected prediction of cue-based retrieval and, in the context of subject-verb number attraction phenomena, the authors named it "grammatical asymmetry". Their intuitively plausible explanation was that a perfectly matching antecedent (as is the case in antecedent-match conditions) must clearly outcompete a partially matching distractor, while more interference is caused when both antecedent and distractor are only partially matching candidates.

Simulations with the current ACT-R implementation (Lewis and Vasishth, 2005) revealed that the latter does not predict such asymmetry (for details, see Engelmann et al., 2015, and our forthcoming paper Engelmann, Jäger & Vasishth, "Confusability of retrieval cues in dependency resolution: A computational model", manuscript in preparation) — at least not in a principled way: It is possible to adjust ACT-R's parameters to permanently reduce similarity-based interference. However, this would leave unexplained why in some cases effects in antecedent-match conditions do appear (see the General Discussion for details). Standardly, ACT-R predicts interference effects in match and mismatch conditions. We therefore extended the ACT-R model with a *prominence principle*, that scales similarity-based interference in relation to the difference in activation between antecedent and distractor.

In standard ACT-R, a memory item i receives an amount of spreading activation S_{ji} for each retrieval cue j it matches. This activation is reduced relative to the number of distractors that match the same retrieval cue j (this number is called the fan_{ji}):

$$S_{ji} = S - \ln(fan_{ji}) \tag{2.4}$$

where S is the *maximum associative strength* parameter (MAS), which defaults to 1.

In our model, the fan_{ji} is transformed into fan'_{ji} by a *prominence correction*, that takes into account the distractors' relative activation:

$$fan'_{ji} = \begin{cases} \frac{1}{1+e^{-C(x_0 - Diff)}} \times fan_{ji}, & \text{if } C > 0 \\ fan_{ji}, & \text{otherwise} \end{cases} \tag{2.5}$$

where *Diff* is the difference $A_i - \bar{A}_{Competitors}$ between the target activation A_i and the mean activation of all competitor items associated with cue j. The *prominence correction factor* C scales the steepness of the logistic *prominence correction* function and should not vary within the same model. In our simulations, we set it to 5. The function's *offset* x_0 is fixed at 1.3, which means that fan'_{ji} is $0.5 \times fan_{ji}$ at an activation difference between target and distractor of 1.3.

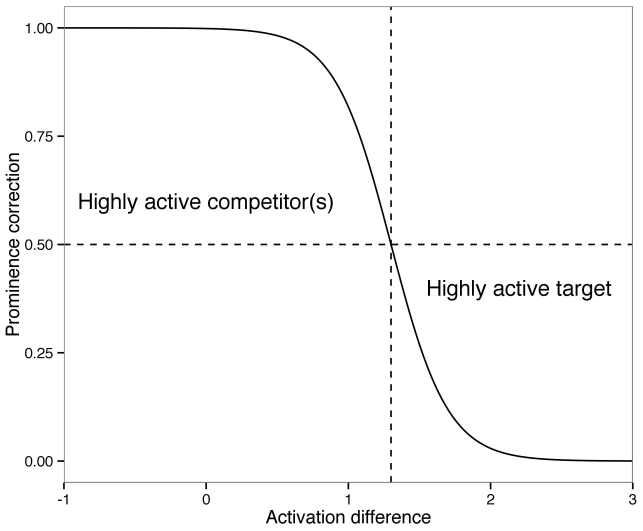

FIGURE 2.1: Prominence correction by activation difference *Diff* (target − distractors) with $C = 5$ and $x_0 = 1.3$.

Figure 2.1 shows the change in the multiplicative term (the *prominence correction*), that determines the relation between *fan* and its transformation fan'. When the target has lower activation than the mean activation of its competitors, *Diff* is negative and the prominence correction approaches 1, which implies that the fan will correspond to the standard calculation in ACT-R, and the activation of the target will be reduced by some amount. This is the case when there are highly activated distractors present: similarity-based interference occurs in this case. *Diff* will be positive when the mean activation of the competitors is relatively low. In this case, the prominence correction will be a value less than 1, and as a consequence the second term in Equation (2.4) will approach 0, leading to a relatively larger amount of spreading activation to the target. In other words, there will be less interference.

This implementation of a prominence principle adds two predictions to the standard cue-based retrieval model: First, there is generally less interference in antecedent-match conditions due to the presence of a highly activated fully matching antecedent; Second, similarity-based (inhibitory) interference in antecedent-match conditions is *increased* for distractors that are highly activated or when there are multiple distractors as in our Experiment 2.[12] Distractor base-level activation could be influenced by its grammatical role (subjects are more salient or accessible than objects, Chafe, 1976; Keenan and Comrie, 1977; Grosz et al., 1995; Brennan, 1995) and by

[12]Note that, for the case of multiple distractors, the original model, too, predicts increased interference. This, however, only explains the difference in effect size between Experiment 1 and 2, but neither the discrepancy between antecedent-match and antecedent-mismatch conditions in Experiment 1 nor the differences between other experiments that did not use multiple distractors.

its discourse topicality (Chafe, 1976; Givón, 1983; Du Bois, 1987; Grosz et al., 1995; Ariel, 1990; Gundel et al., 1993; Du Bois, 2003). Other factors contributing to the salience of the distractor and hence to its base-level activation might be first mention (Gernsbacher and Hargreaves, 1988), thematic role (Arnold, 2001), contrastive focus (Cowles et al., 2007) or animacy (Fukumura and van Gompel, 2011). In effect, the prominence principle accounts for both the absence of an effect in antecedent-match conditions of Experiment 1 and the presence of an inhibitory effect in Experiment 2. Furthermore, the prominence principle predicts greater interference effects in antecedent-match conditions for distractors in more salient positions. We will relate this prediction to the literature in the General Discussion.

2.4.2 Principle 2: Cue confusion

As explained in the introduction and resulting from Equation (2.4), similarity-based (inhibitory) interference (or the fan effect) in ACT-R only arises when multiple memory items match the same retrieval cues. Since this is not the case in the antecedent-mismatch conditions of Experiment 1, the observed inhibitory interference is incompatible with ACT-R theory. At least this seems to be the case. We argue that this assumption of incompatibility might not be justified.

In the application of cue-based retrieval to sentence comprehension, it is generally assumed that retrieval cues perfectly distinguish matching features from non-matching ones. For instance, a *+plural* cue always activates plural items and not singular items. For our first experiment, this means that *+animate* is perfectly different from *+ c-com* and no similarity-based interference is predicted in antecedent-mismatch conditions where the antecedent only matches *+ c-com* and the distractor only matches *+animate*. However, the language processor might not differentiate between features categorically but rather on a continuous scale of similarity. In fact, in the general ACT-R framework, features are memory items just like the items they belong to and, therefore, could be confused with each other if they have a sufficient degree of similarity. If we assume that cue-feature associations have to be learned from language experience, it follows that these associations would somehow reflect co-occurrence statistics in the language input. Consequently, cues in a retrieval specification could, depending on the retrieval-relevant context, be associated with several features to different degrees.

A co-occurrence-based account would predict differences between English reflexives and Mandarin *ziji* in the following way: *Ziji* invariably requires its antecedent to match $\{+\,c\text{-}com, +animate\}$, meaning that these two features frequently co-occur in the specific task of processing the Mandarin reflexive. English reflexives, on the other hand, have several alternative forms like *himself*, *herself*, *itself*, and *themselves*. All of these forms have the same structural requirement towards their antecedent but their non-structural retrieval cues vary in gender and number. The benefit of distinguishing features for number, gender, and structural relation in English reflexives results in a stronger one-to-one association between a cue and the corresponding feature. In the case of Mandarin *ziji*, however, there is no benefit from distinguishing *+ c-com* and *+animate* for the task of finding the appropriate antecedent. In consequence, retrieval cues might in this case be associated with both features to some degree in a kind of *crossed association*. In relation to the retrieval specification, antecedent and distractor would appear similar in this case, although they

theoretically do not share any features. This confusion-induced similarity can cause similarity-based interference as of Equation (2.4), predicting inhibitory effects in conditions where they would not be expected in terms of standard cue-based retrieval assumptions.

We implemented cue confusion by further adjusting the measure of similarity-based interference (the *fan*) from Equation (2.4) to take into account all features and their strength of association with a certain cue:

$$fan_{ji} = 1 + \sum_{k}(1 + Q_{jk}) \tag{2.6}$$

where Q_{jk} is the *associative strength* between cue value j and feature value k on a scale of $[-1, 0]$, with -1 meaning no association and 0 representing maximum association. We assume that this association is dynamically adaptive to individual dependency environments. Equation (2.6) predicts that the stronger a cue-feature association the more this feature will contribute to similarity-based interference related to that cue. For example, if $Q_{c\text{-}com;anim}$ for *ziji* is -0.5, the resulting fan for the $+$ *c-com* cue would be 1.5 instead of 1 as original ACT-R would predict. This increases similarity-based interference in comparison to English reflexives, where, say, $Q_{c\text{-}com;gend}$ would be standardly assumed -1, hence having a fan of 1 for each cue.

Another example of increased feature-co-occurrence are reciprocals like *each other*. In this case, the feature combination $\{+$ *c-com*, $+$*plural*$\}$ is invariably required. Hence, our account predicts an increased cue-confusion level in the case of English reciprocals just like in Mandarin reflexives, possibly leading to inhibitory interference in antecedent-mismatch conditions.

With the cue confusion account, we propose that task requirements (frequent co-occurrence of certain features in similar retrieval contexts) dynamically influence how cues are treated during a retrieval request. Cue confusion therefore predicts that inhibitory interference effects in antecedent-mismatch conditions should preferably be observed in constructions where cues frequently co-occur. An evaluation of these predictions beyond our own experimental results will be provided in the General Discussion.

2.4.3 Simulation results

We report model predictions for the full range of cue confusion values. ACT-R parameters were fixed to their defaults or to values used in previous simulations (Lewis and Vasishth, 2005): latency factor $LF = 1.5$, activation noise value $ANS = 1.5$, mismatch penalty $MP = 1.5$. We compare the model predictions with empirical first-pass reading times on *ziji* of Experiments 1 and 2. We refer to first-pass reading time in Experiment 2 although it was not significant under Bonferroni correction. It however patterned with an effect in right-bounded reading time, which had a similar magnitude. Figure 2.2 plots the prediction space of a cue-based retrieval model that implements cue confusion and prominence (values represent the means of 2000 simulations each). For comparison, the predictions of a model without prominence are plotted in gray. The cue confusion level is plotted on a percentage scale, with 100% confusion meaning that both features, $+$ *c-com* and $+$*animate*, are maximally associated with both the *c-com* and *animate* cues ($Q_{c\text{-}com;anim} = 0$ and $Q_{anim;c\text{-}com} = 0$). With *prominence correction factor* at 0 and *cue confusion level* at -1, the current model is equivalent to the original ACT-R model. The original

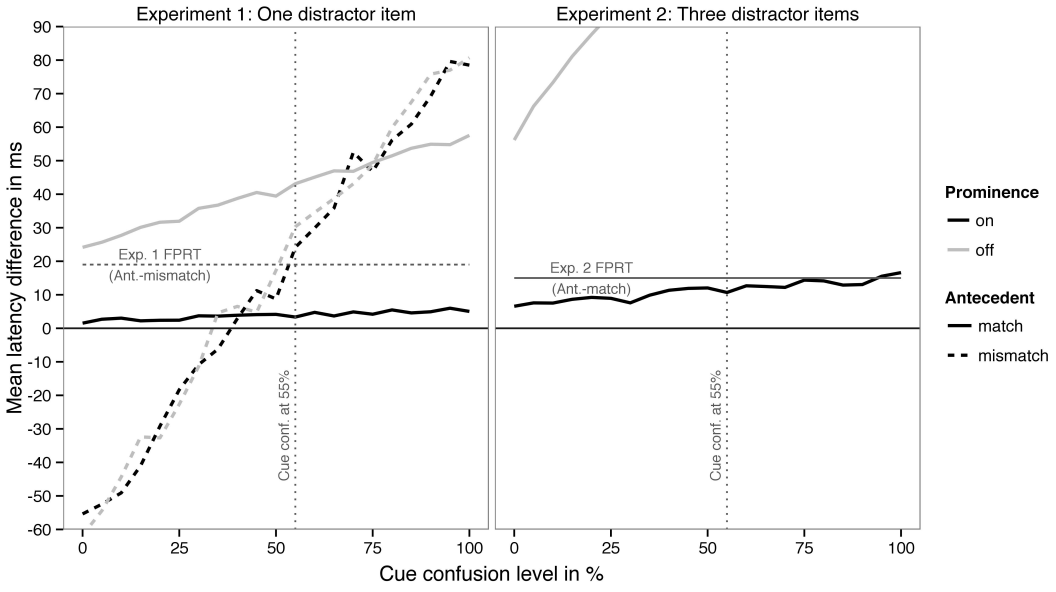

FIGURE 2.2: Predicted interference effect (distractor-match − distractor-mismatch) by cue confusion level for the default model (gray lines) and the prominence model (black lines). The left panel shows the predicted interference for a single distractor (Exp. 1); the right panel for three distractors (Exp. 2). Solid lines represent the conditions where the antecedent matches the semantic cue, mismatch conditions are represented by dashed lines. The gray horizontal lines indicate the observed effect size in antecedent-mismatch conditions in Exp. 1 (left panel) and local antecedent-match conditions in Exp. 2 (right panel) — both in first pass reading time FPRT. The gray dotted vertical line intersects the x-axis at the estimated cue confusion value (55%) in both panels.

model's predictions are therefore represented by the left-most points of the gray lines. The left panel shows the predictions for Experiment 1. With increasing cue confusion, the interference effect for the antecedent-mismatch conditions increases. At a confusion level of about 55% (indicated by the dotted vertical line), the model predicts an effect of the observed size in local conditions (19 ms in first-pass reading time, indicated by the dashed horizontal line). In contrast to the original model, the prominence model predicts an interference effect close to zero for antecedent-match conditions in Experiment 1 for all cue confusion levels. This is in line with the absence of an effect in the data.

The right panel of Figure 2.2 shows the predictions for a similar model as the left panel, but with three distractors instead of one, simulating the conditions of Experiment 2. The inhibitory effect for antecedent-match conditions increases with cue confusion in this scenario. An effect of about the observed size (15 ms in first-pass reading time) is predicted at the same cue-confusion level as for Experiment 1.

To summarize, the extended model with cue confusion and prominence predicts the observed data of both experiments with fixed parameters at a cue-confusion level of about 55%. More

specifically, the model predicts two patterns that the original ACT-R model does not predict: i) the absence (or near absence) of an inhibitory interference effect in the antecedent-match conditions of Experiment 1 in spite of an effect present in Experiment 2 and ii) an *inhibitory* interference effect in antecedent-mismatch conditions in Experiment 1.

2.5 General discussion

We conducted two eye-tracking experiments in which we investigated whether the reflexive *ziji* is subject to interference effects from structurally inaccessible distractor nouns that fulfill the animacy requirement of *ziji*. In Experiment 1, where only a single distractor was present in the sentence, we found inhibitory interference in antecedent-mismatch conditions but no effect in antecedent-match conditions. In Experiment 2, where three distractors were presented as memory load, we found interference effects also in antecedent-match configurations.

These results are clear evidence against a structure-based mechanism underlying memory retrieval in human sentence parsing. The interference effects observed in Experiments 1 and 2 are incompatible with a purely structure-based retrieval mechanism. However, Sturt (2003) and Kush and Phillips (2014) have proposed a potential explanation for interference effects within the structure-based account. These authors hypothesize that, in the case of retrieval failure, a later repair process might employ a retrieval with relaxed structural restrictions, giving rise to late interference effects. This late-interference account is a plausible explanation for the effect observed in the non-local conditions of Experiment 2, where the effect occurred only in re-reading times at the post-critical region. However, for the effects observed in locally bound *ziji* (Experiments 1 and 2), the late-interference account appears implausible given that the effects occur already in first-pass eye-tracking measures and at the critical region. Also note that the effect reported in Kush and Phillips (2014) does not necessarily reflect late processes, since in self-paced reading experiments, it is very common that effects triggered at the critical region appear several words downstream.

The standard ACT-R model of cue-based retrieval (Lewis and Vasishth, 2005) does predict immediate interference effects but is not fully compatible with our results either. First, it predicts facilitatory rather than inhibitory interference in antecedent-mismatch conditions and, second, it cannot explain the absence of an effect in the antecedent-match conditions of Experiment 1. In fact, in the literature on reflexive processing, hardly any study can be found that reports the exact pattern predicted by the standard ACT-R model, namely inhibitory interference in antecedent-match conditions and facilitatory interference in antecedent-mismatch conditions.[13] An approach of extending the ACT-R model in favor of a structure-based mechanism has been taken by Parker and Phillips (2014). They have proposed that structural cues are weighted higher than semantic

[13]It should be noted that the (marginal) facilitatory interference in antecedent-match conditions reported by three studies presented in Table 2.1 (Sturt, 2003; Cunnings and Felser, 2013) is compatible with the ACT-R model although this may not be intuitively obvious. An exceptionally highly activated distractor (in all three of these experiments, the distractor is a discourse prominent subject) can lead to facilitatory interference (see Engelmann et al., 2015, and our forthcoming publication Engelmann, Jäger & Vasishth, "Confusability of retrieval cues in dependency resolution: A computational model", manuscript in preparation).

or morphological cues, so that interference effects occur only in case of an abnormally poor match of the accessible antecedent. This is a plausible explanation for their data and offers an account for the fact that interference is hard to find in reflexives. However, with respect to our results, it neither explains the inhibitory interference in antecedent-match conditions nor the difference in effect sizes in antecedent-match vs. antecedent-mismatch conditions.

In order to account for our results and the diverse patterns in the literature, we have introduced two concepts as an extension of the standard cue-based retrieval model. The *prominence principle* implements the idea that a perfectly matching or otherwise highly activated antecedent is only marginally affected by similarity-based interference from comparably poorly matching distractors. This explains the discrepancy between Experiment 1 and 2 (absence of an effect in antecedent-match conditions in Experiment 1 vs. an inhibitory interference effect in Experiment 2). With the concept of *cue confusion*, we proposed that the retrieval cues can be associated with several features of memory items and that the strength of these associations depends on experience with a specific linguistic context. For special cases, this can cause similarity-based interference between items that do not match the same retrieval cues. We argued that *ziji* is such a special case, which would explain the observed inhibitory interference in antecedent-mismatch conditions of Experiment 1.

In the following, we compare the predictions of the extended ACT-R model with the literature on reflexives. Prominence predicts that interference in antecedent-match conditions is generally low compared to antecedent-mismatch conditions but increases as a function of distractor activation. If we assume that distractor position (grammatical role and discourse topicality) affects its base-level activation in memory, the literature summary in Table 2.1 seems to conform with these predictions: Among the studies which tested both antecedent-match and antecedent-mismatch conditions, about 75% report an interference effect (including marginal effects) in antecedent-mismatch conditions while only 50% of the studies found an effect in antecedent-match conditions. All studies that did report an effect in antecedent-match conditions had the distractor either in subject position (Badecker and Straub, 2002; Chen et al., 2012b; Patil, Vasishth & Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript), in topicalized subject position (Felser et al., 2009; Cunnings and Felser, 2013; Clackson and Heyer, 2014), or had multiple distractors (Experiment 2 reported here). On the other hand, only half of the studies reporting no interference effect in antecedent-match conditions had the distractor in subject position. Obviously, not all studies that have the distractor in subject position report an effect, but the literature review suggests that subject position increases the probability of finding one. For the absence of an antecedent-match interference effect in our Experiment 1, there might be a specific reason: Dillon et al. (2015) have shown that items within restrictive relative clauses cause more interference as compared to items in appositive relative clauses. They attribute this difference to the idea that in contrast to restrictive relative clauses, appositive relative clauses constitute a speech act separate from the one of the main utterance (Potts, 2005; Arnold, 2007). More generally, their results suggest that the embedding environment containing a distractor influences the strength of interference caused by this distractor. In terms of ACT-R, one might think of this as different base-level activations as a function of the type of embedding environment. It might be possible that the interposed adverbial structures which contain the distractor in our materials belong to those

embedding environments which cause a relatively low degree of interference. This seems a plausible assumption since in our materials, the adverbial clause can simply be ignored by the parser without affecting the grammaticality or plausibility of the whole sentence.

For antecedent-mismatch conditions, cue confusion predicts stronger inhibition the higher the crossed association between cues and features is assumed to be, that is, in contexts with frequently co-occurring cue combinations. However, note that cue confusion is compatible with both facilitatory and inhibitory effects, and even with the absence of an effect, as all this is part of the effect continuum that is illustrated in Figure 2.2. This raises the concern of how to determine a sensible confusion level in each case, since a model allowing arbitrary predictions is not useful. Currently, the model prediction can only be treated as a predicted difference between two conditions in one or the other direction along the effect continuum. In other words, a prediction should be stated in terms of whether the antecedent-mismatch interference effect of one dependency tends more towards inhibition or towards facilitation in comparison to another dependency like, e.g., English reflexives. In the reasoning we apply here, we refer to English reflexives as a baseline with zero cue confusion and spot special cases where a different feature-co-occurrence rate can be assumed that would motivate a higher confusion level. We have argued that inhibitory interference was observed in antecedent-mismatch conditions in our Experiment 1 because *ziji* is a special case in the sense that the feature combination $\{+\textit{c-com}, +\textit{animate}\}$ is constant compared to the variable combinations in the different forms of English reflexives. The same logic with respect to $\{+\textit{c-com}, +\textit{plural}\}$ would apply to reciprocals. In the literature there is one study by Kush and Phillips (2014) that tested the Hindi equivalent of the reciprocal *each other* and indeed found the predicted inhibitory interference in antecedent-mismatch conditions.

Although the post-hoc nature of our proposals here is an important limitation that needs to be addressed with new empirical tests, theory development necessarily is data-driven, and the existing data suggest that our proposal constitutes one possible explanation. Indeed, currently it is the only computational account of the patterns of findings discussed here. In order to empirically test the predictions of cue confusion, it is necessary to experimentally manipulate feature-co-occurrence within a minimal pair. A potential experiment could use stimuli like in Example (2.7) to compare the interference effect in antecedent-mismatch conditions for *themselves* and *each other*. Cue confusion predicts a smaller facilitation or even an inhibition for *each other*. Furthermore, it should be possible to derive a numerical metric of cue confusion for a range of dependencies by computing co-occurrence frequencies in a treebank that contains dependency information as well as information about retrieval relevant features such as gender, number, and animacy.

(2.7) a. *Reflexive; distractor-match*
The **nurse** who cared for the **children** had pricked **themselves**...

b. *Reflexive; distractor-mismatch*
The **nurse** who cared for the **child** had pricked **themselves**...

c. *Reciprocal; distractor-match*
The **nurse** who cared for the **children** had pricked **each other**...

d. *Reciprocal; distractor-mismatch*
The **nurse** who cared for the **child** had pricked **each other**...

A more thorough test of the extended model's predictions will be presented in a forthcoming publication (Engelmann et al., "Confusability of retrieval cues in dependency resolution: A computational model", manuscript in preparation) that includes quantitative simulations of a range of previous studies on reflexive processing and subject-verb dependencies.

As a rather speculative point we want to add that the cue confusion level of a certain dependency might not only be influenced by feature-co-occurrence but also by task demands and individual differences. If cue-feature associations are subject to an adaptive learning process, they might also be affected by resource-preserving strategies. An example where strategic adaptation of comprehension processes has been found are relative clause attachment ambiguities. Swets et al. (2008) and Logačev and Vasishth (2015) have found that processing effort in ambiguity resolution was adapted to the type of comprehension questions. Also, effects of individual differences in working memory span have been found by Traxler (2007) and von der Malsburg and Vasishth (2012) for the processing of attachment ambiguities. If analogously to task- and resource-related underspecification in attachment ambiguities, cue-feature associations are affected by resource-preserving strategies in the sense of *good-enough processing* (Ferreira et al., 2002), we would expect that low-span readers tend to have greater cue confusion and, thus, exhibit interference effects further towards inhibition in the continuum than high-span readers. The marginal inhibitory effect for low-span readers in antecedent-mismatch conditions of Experiment 2 by Cunnings and Felser (2013) would fit with this expectation. However, more experimental data is needed in order to evaluate effects of individual differences and task-demands on cue-feature associations.

2.6 Conclusion

We have presented experimental evidence that is incompatible with structure-based accounts of reflexive processing and also inconsistent with the original cue-based ACT-R model of sentence processing. In order to account for the observed pattern, we have proposed to add two new principles, prominence and cue confusion, to the ACT-R model. This extension to the ACT-R model is not only able to explain the pattern observed in the data presented in this article, but can also account for a range of previously unexplained patterns reported in the literature on reflexive processing. Naturally, this proposal needs to be evaluated with novel experimental data.

Acknowledgements

We thank Prof. Hua Shu and the State Key Laboratory of Cognitive Neuroscience and Learning at Beijing Normal University, China, for allowing us to conduct our experiments in their eye-tracking lab. We are grateful to Dr. Yan Ming, Dr. Jing'er Pan and Zhou Wei for their kind

assistance and their very helpful comments on the experimental materials. We thank Prof. Reinhold Kliegl for his valuable advice on the statistical analyses and Dr. Zhong Chen and Maja Stegenwallner-Schütz for constructive discussions. We thank the audience at AMLaP 2012 and 2014 for comments.

Chapter 3

Frontiers in Psychology
Volume 6, Issue 506, 2015

Teasing apart Retrieval and Encoding Interference in the Processing of Anaphors

Lena A. Jäger[1], Lena Benz[1], Jens Roeser[2], Brian W. Dillon[3], and Shravan Vasishth[1]

[1]Dept. of Linguistics, University of Potsdam [2]Dept. of Psychology, Nottingham Trent University [3]Dept. of Linguistics, University of Massachusetts, Amherst

Abstract

Two classes of account have been proposed to explain the memory processes subserving the processing of reflexive-antecedent dependencies. Structure-based accounts assume that the retrieval of the antecedent is guided by syntactic tree-configurational information without considering other kinds of information such as gender marking in the case of English reflexives. By contrast, unconstrained cue-based retrieval assumes that all available information is used for retrieving the antecedent. Similarity-based interference effects from structurally illicit distractors which match a non-structural retrieval cue have been interpreted as evidence favoring the unconstrained cue-based retrieval account since cue-based retrieval interference from structurally illicit distractors is incompatible with the structure-based account. However, it has been argued that the observed effects do not necessarily reflect interference occurring at the moment of retrieval but might equally well be accounted for by interference occurring already at the stage of encoding or maintaining the antecedent in memory, in which case they cannot be taken as evidence against the structure-based account. We present three experiments (self-paced reading and eye-tracking) on

German reflexives and Swedish reflexive and pronominal possessives in which we pit the predictions of encoding interference and cue-based retrieval interference against each other. We could not find any indication that encoding interference affects the processing ease of the reflexive-antecedent dependency formation. Thus, there is no evidence that encoding interference might be the explanation for the interference effects observed in previous work. We therefore conclude that invoking encoding interference may not be a plausible way to reconcile interference effects with a structure-based account of reflexive processing.

Keywords: Anaphors; Reflexives; Possessives; Eye-tracking; German; Swedish; Working-memory; Interference

3.1 Introduction

A central task the human sentence processing mechanism has to accomplish is to link two parts of a syntactic dependency, irrespective of how much linguistic material separates the two dependents. Many theories of sentence processing therefore assume that upon encountering the second dependent, the parser triggers a memory retrieval to access the first dependent in order to integrate it with the current node (Gibson, 2000; Lewis and Vasishth, 2005). Interference effects have recently come into focus in sentence processing research because they are taken to be informative about the more precise nature of the retrieval mechanisms that subserve sentence processing. However, the relationship between empirically observed similarity-based interference effects and theories of retrieval is somewhat indirect, because there are multiple distinct mechanisms that could give rise to similarity-based interference effects in online processing. Indeed, whether or not the observation of interference effects can be interpreted as evidence favoring one or another account of sentence processing depends on the exact mechanisms causing the interference effects. In this article, we will present different mechanisms that have been proposed to account for interference effects in sentence comprehension and present three experiments with different methodologies and languages to tease them apart. We will first give an overview of two kinds of mechanisms, *cue-based retrieval interference* and *encoding interference*, which in the working memory literature have been proposed to underly similarity-based interference. Subsequently, we will turn to the implications for sentence processing and antecedent-retrieval in the processing of reflexives in particular.

Similarity-based interference has long been known to be a major cause of forgetting (Anderson and Neely, 1996). In memory models which represent items as bundles or vectors of features, similarity-based interference is assumed to arise as a function of the degree of overlap between an item's features with the features of other items in memory (Nairne, 1988, 1990; Anderson and Neely, 1996; Anderson and Lebiere, 1998; Anderson et al., 2004; Oberauer and Kliegl, 2006; Lewandowsky et al., 2008). However, the various memory models differ with respect to the mechanisms which they assume to underlie similarity-based interference. Generally speaking, one can distinguish between two kinds of similarity-based interference. On the one hand, similarity-based interference is assumed to affect the encoding or maintenance of an item (Nairne, 1988, 1990; Lewandowsky et al., 2008; Oberauer and Kliegl, 2006). We will refer to this proposal as *encoding interference*. On the other hand, similarity-based interference is assumed to arise during

the retrieval of an item (Anderson and Neely, 1996; Anderson and Lebiere, 1998; Anderson et al., 1998, 2004; McElree, 2006; Oberauer and Kliegl, 2006). We will refer to this second proposal as *cue-based* retrieval interference.

Encoding interference is assumed to arise from the competition between the features of similar items that occurs at the moment of encoding or maintaining items in memory. Nairne (1990), for instance, proposed that whenever two items share a feature, they compete for this feature. In a certain proportion of cases, the memory representation of one of these items therefore loses this feature.[1] Hence, this item's memory representation becomes less distinct from other items and, as a result, retrieval probability decreases. An important, but subtle, point here is that even though encoding interference arises at the stage of encoding or maintaining an item in memory, it has an impact on the ease of this item's later *retrieval*. Oberauer and Kliegl (2006), who adopted Nairne (1990) 's concept of feature-overwriting, implemented the idea of an item's memory representation being degraded by decreasing this item's activation level. At the moment of later retrieval, this lower activation level leads to lower retrieval probability and a slow-down in processing times. In their model, the retrieval of an item from working memory is implemented as its gradual activation into the focus layer of the memory system. The processing speed of this gradual activation is defined as a function of this item's activation level prior to retrieval. Thus, if an item's activation level is decreased due to encoding interference from competitor items, a slow-down in the retrieval process is predicted. Note that Oberauer and Kliegl (2006) do not make any predictions about retrieval latencies. Their model is designed to explain data collected in speed-accuracy tradeoff experiments, where they experimentally controlled the time point when retrieval was supposed to happen. In their model, the slow-down in the retrieval process therefore is reflected in a higher proportion of retrieval failures rather than in increased retrieval latencies because participants are forced to interrupt the retrieval process after an experimentally defined time lag. Translating the Oberauer and Kliegl (2006) model to sentence processing, where the participant has more time to carry out retrieval, leads us to the assumption that the slow-down in the retrieval process is reflected in longer retrieval latencies. For the predictions of the experiments reported in this article, we will refer to encoding interference as implemented in the Oberauer and Kliegl (2006) model, with the additional assumption that a slow-down in the retrieval process leads to increased retrieval latencies. In sum, although encoding interference acts at the moment of encoding and maintenance rather than at retrieval, it indirectly affects the success and the speed of the retrieval process because it results in a representation that is more difficult to access.

Cue-based retrieval interference, in contrast, is assumed to arise due to cue-overload at the moment of retrieval. In a content-addressable memory architecture, cue-overload refers to a scenario when the cues used for retrieval do not point to a unique target, but rather match multiple items (Watkins and Watkins, 1975). This is assumed to lead to misretrievals of partially matching distractor items (Anderson and Lebiere, 1998; Anderson et al., 2004; McElree, 2006) and mutual inhibition between the distractors and the target resulting in a higher retrieval latency in case the target and the distractor have one or more retrieval relevant features in common (Anderson

[1] Nairne (1990) did not use the term *encoding interference* but rather *feature-overwriting* to refer to his conception of interference.

and Lebiere, 1998; Anderson et al., 2004).[2] To summarize, encoding interference is predicted to occur whenever items share features, no matter whether these features are used for retrieval or not. Cue-based retrieval interference, in contrast, is predicted to occur when more than one item matches the retrieval features. Inhibition between these items occurs only when they match the same retrieval features, otherwise cue-based retrieval interference is reflected only in misretrievals (Anderson et al., 2004). Note that encoding interference and cue-based retrieval interference are not mutually exclusive concepts. Indeed, in Oberauer and Kliegl (2006)'s working memory model, both retrieval and encoding interference are assumed and the authors show that their interference model is indeed able to account for a large range of data.

In sentence processing research, early studies investigating interference effects point rather towards encoding than cue-based retrieval interference, but they were not designed to disentangle the two. For example, Gordon et al. (2002) conducted a self-paced reading experiment where participants held a set of nouns in memory while reading the target sentence. The authors report a slow-down in reading times when the noun type (common noun /proper name) of the memory load words matched the nouns in the sentence compared to when the memory load nouns and the nouns in the sentence were of different types. These results are further supported by Fedorenko et al. (2006), who also observed similarity-based interference in a memory-load paradigm. Gordon and colleagues report similar results for studies that manipulated similarity between sentence internal nouns rather than memory load (Gordon et al., 2001, 2004, 2006). An example item taken from Gordon et al. (2006) is shown in (3.1).

(3.1) *Interference/No interference*
The banker that **the barber/Sophie** praised climbed the mountain ...

Since in all of these studies, similarity of the nouns was manipulated while the efficiency of the retrieval cues (i.e., the degree to which the retrieval cues uniquely identify the target) remained constant across experimental conditions, the data reported by Gordon and colleagues favor rather encoding than cue-based retrieval interference as an explanation. However, as Van Dyke and McElree (2006) noted, the above cited studies found interference effects only in the region where the critical noun phrase was retrieved (i.e., at the region containing the verb). This might indicate that the observed effect should rather be attributed to cue-based retrieval interference since encoding interference should also affect processing ease at the moment of encoding, i.e., at the moment when the second of the similar nouns is first being encountered. Van Dyke and McElree (2006) conducted a memory load experiment where, in contrast to the memory load experiments reported by Gordon et al. (2002) and Fedorenko et al. (2006), the memory load words were held constant across experimental conditions, but the retrieval cues at the verb were manipulated. The experimental items consisted of object-cleft sentences in which the main clause object preceded the main clause verb (the critical region where retrieval was triggered); for an example taken from Van Dyke and McElree (2006) see (3.2).

[2]Note that the model proposed by McElree (2006) predicts that cue-based retrieval interference is reflected only in retrieval probability, not in retrieval latency. In contrast, the ACT-R architecture developed by Anderson and Lebiere (1998); Anderson et al. (2004), on which the Lewis and Vasishth (2005) model of sentence processing is based, predicts retrieval interference to be reflected in both, retrieval probability and retrieval latency.

(3.2) *Interference/No interference*
It was the boat that the guy who lived by the sea **sailed/fixed** in two sunny days.
Memory load: table, sink, truck

When the memory load words fit the semantic constraints of the verb, a slow-down in self-paced reading times was observed. These results cannot be attributed to encoding interference since the degree of similarity between the memory load words and the verb's object NP is constant across conditions. Van Dyke and McElree (2006)'s data are thus clear evidence for cue-based retrieval interference playing a role in sentence processing. However, note that the possibility that both retrieval and encoding interference affect sentence processing ease cannot be excluded by Van Dyke and McElree (2006)'s study since their data is clear evidence for cue-based retrieval interference but no evidence *against* encoding interference affecting sentence processing in general.

In recent years, interference effects in the processing of reflexive-antecedent dependencies have drawn considerable attention. The underlying research question was whether unconstrained cue-based retrieval, as proposed by Badecker and Straub (2002) and Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English" (unpublished manuscript), or a structure-based access mechanism, as proposed by Nicol and Swinney (1989) and Sturt (2003), subserves the processing of reflexive-antecedent dependencies. Unconstrained cue-based retrieval assumes that all available cues are used to retrieve a reflexive's antecedent. Structure-based accounts, in contrast, assume that structural, i.e., syntactic tree-configurational, information guides the retrieval process. Interference effects in reflexive processing have been generally interpreted in terms of cue-based retrieval interference and taken as evidence for a cue-based retrieval mechanism since *retrieval* interference from syntactically inaccessible constituents is incompatible with the structure-based account. However, as pointed out by Dillon (2011) and Dillon et al. (2013), many of the observed effects — which we will describe more in detail below — can equally well be accounted for by *encoding* interference and hence are not necessarily incompatible with the structure-based account. Indeed, for the kind of materials commonly used to investigate the processing of reflexives (see 3.3; example taken from Sturt, 2003), encoding interference makes the same predictions for all experimental conditions as the unconstrained cue-based retrieval account implemented in the Lewis and Vasishth (2005) sentence processing model, which is based on the general cue-based architecture Adaptive Control of Thought-Rational (ACT-R) (Anderson and Lebiere, 1998; Anderson et al., 2004) and has been widely used for modeling the processing of reflexives (Dillon, 2011; Dillon et al., 2013; Parker and Phillips, 2014; Kush and Phillips, 2014; Jäger et al., 2015; Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript).[3] Thus, for the question of structure-based vs. unconstrained cue-based retrieval in reflexives, it is crucial to disentangle encoding from cue-based retrieval interference. If evidence can be found showing that encoding interference plays a role in the type of materials generally used to investigate the processing of reflexives, this implies that the interference effects

[3]The Lewis and Vasishth (2005) model per se does not make any commitments with respect to the question which features are used as retrieval cues. Hence it is also possible to implement the structure-based account in this framework by restricting the set of retrieval cues to structural features.

that have been interpreted as evidence favoring unconstrained cue-based retrieval are equally well compatible with a structure-based account.

(3.3) a. *Antecedent-match; distractor-match*
The **surgeon**$_i$ who treated **Jonathan**$_j$ had pricked **himself**$_{i/*j}$...

b. *Antecedent-match; distractor-mismatch*
The **surgeon**$_i$ who treated **Jennifer**$_j$ had pricked **himself**$_{i/*j}$...

c. *Antecedent-mismatch; distractor-match*
The **surgeon**$_i$ who treated **Jennifer**$_j$ had pricked **herself**$_{i/*j}$...

d. *Antecedent-mismatch; distractor-mismatch*
The **surgeon**$_i$ who treated **Jonathan**$_j$ had pricked **herself**$_{i/*j}$...

Studies investigating interference effects in the processing of reflexives mostly tested sentences in which the reflexive was bound by the local subject which c-commanded the reflexive (*surgeon* in example 3.3; henceforth referred to as *antecedent*). We will express the antecedent's conformance to the structural requirements for binding the reflexive by attributing the feature $c-com:+$ to it.[4] The interference manipulation was achieved by inserting another noun phrase in a structurally inaccessible position, i.e., not c-commanding the reflexive ($c-com:-$) and hence not qualifying as a binder for the reflexive (*Jonathan/Jennifer* in example 3.3; henceforth referred to as *distractor*). A non-structural feature (e.g., gender or number in English reflexives) of this distractor was manipulated. Crucially, the feature which was manipulated might theoretically be used as a retrieval cue. For example, in the processing of English reflexives, the gender feature ($gender:masc/fem$) marked at the reflexive *himself* or *herself* might be used as a cue to retrieve the antecedent. Thus, if gender is used as a retrieval cue, a gender-matching distractor is predicted to cause cue-based retrieval interference as compared to a distractor which does not match the gender of the reflexive. Therefore, interference effects caused by a cue-matching distractor can be interpreted as evidence favoring an unconstrained cue-based retrieval account. If, in contrast, no effect of a cue-matching distractor is observed, this can be taken as evidence for a structure-based account. This experimental design (or a variation thereof) was used by a large number of studies which aimed to decide whether an unconstrained cue-based retrieval or a structure-based access underlies the processing of reflexive antecedent-dependencies (Nicol and Swinney, 1989; Badecker and Straub, 2002; Sturt, 2003; Xiang et al., 2009; Chen et al., 2012b; King et al., 2012; Cunnings and Felser, 2013; Dillon et al., 2013; Kush and Phillips, 2014; Parker and Phillips, 2014; Clackson and Heyer, 2014; Jäger et al., 2015; Patil, Vasishth and Lewis,"Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript). Some of the cited studies also manipulated feature-match

[4]It should be noted that using $c-com:+$ as a feature is a simplification since a tree-configurational relation is not as straightforward to code as a feature of an item as, e.g., gender or number. For a discussion of how tree-configurational information such as c-command could be encoded as an item's feature see Alcocer and Phillips, "Using relational syntactic constraints in content-addressable memory architectures for sentence processing" (unpublished manuscript). On a theoretical basis, Kush (2013) argues against the representation of c-command as a feature and discusses how, in online sentence processing, the human parser might distinguish between c-commanding and non-c-commanding antecedents.

of the structurally accessible antecedent (*surgeon* in example 3.3).[5] An effect of antecedent match/mismatch can be accounted for by both unconstrained cue-based retrieval and structure-based accounts.

The results of the above cited studies are mixed. In antecedent-match conditions, increased processing difficulty due to the presence of a cue-matching distractor has been reported by Badecker and Straub (2002, Exp. 3, 4), Chen et al. (2012b), Clackson and Heyer (2014), Jäger et al. (2015, Exp. 2) and Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English" (unpublished manuscript). By contrast, Sturt (2003, Exp. 1) and Cunnings and Felser (2013, Exp. 2) found a facilitation due to a cue-matching distractor. It should be noted that in Sturt (2003)'s experiment, the effect appeared only delayed and in Cunnings and Felser (2013)'s study, the interference effect was only observed in participants with low working-memory span. Importantly, in a large number of studies, no interference effect in antecedent-match conditions was observed (Nicol and Swinney, 1989; Badecker and Straub, 2002, Exp. 5, 6; Sturt, 2003, Exp. 2; King et al., 2012; Dillon et al., 2013; Kush and Phillips, 2014; Parker and Phillips, 2014; Jäger et al., 2015, Exp. 1). In antecedent-mismatch conditions, a significant processing speed-up due to a cue-matching distractor is reported by King et al. (2012) and Parker and Phillips (2014). The opposite direction of the effect was only observed in Jäger et al. (2015, Exp. 1). The absence of an effect in antecedent-mismatch conditions is reported by Sturt (2003), Xiang et al. (2009) and Dillon et al. (2013). For an overview and a discussion of the literature on interference effects in reflexives, see Jäger et al. (2015).

As mentioned above, unconstrained cue-based retrieval as implemented in the Lewis and Vasishth (2005) ACT-R model of sentence processing makes precisely the same predictions as encoding interference for sentences like the ones shown in (3.3). For conditions with a feature-matching antecedent (see 3.3a, b), the Lewis and Vasishth (2005) model predicts cue-based retrieval interference when the distractor matches the gender of the reflexive (3.3a in the example sentence). This retrieval interference is predicted to be reflected in inhibition between the antecedent and the distractor because in (3.3a), but not in (3.3b), the antecedent (*surgeon*) and the distractor (*Jonathan*) share the gender cue $gender : masc$. Thus, longer retrieval latencies (and hence longer reading times at the reflexive) are predicted in (3.3a) compared to (3.3b). Moreover, misretrievals of the partially feature-matching distractor (*Jonathan* in 3.3a) are predicted. These misretrievals are predicted to be reflected in response-accuracies if the comprehension questions target the reflexive-antecedent dependency. For conditions with a mismatching antecedent (see 3.3c, d), the unconstrained cue-based retrieval model (Lewis and Vasishth, 2005) also predicts cue-based retrieval interference due to a matching distractor (3.3c). As in antecedent-match conditions, this retrieval interference is predicted to be reflected in a higher proportion of misretrievals of the matching distractor. But, in contrast to antecedent-match conditions, no inhibition between the antecedent and the distractor is predicted because they do not share any of the experimentally manipulated retrieval relevant features (in 3.3c and d, the antecedent and the distractor neither share the gender cue $gender : fem$ nor the structural cue $c - com : +$). Since ACT-R predicts faster retrieval latencies in the case of misretrievals as a result of a race-like

[5] In some experiments, only the stereotypical gender of the accessible antecedent was violated (as in 3.3c, d), whereas in other studies, real feature violations were used resulting in ungrammatical sentences in the antecedent-mismatch conditions.

configuration, the trials with misretrievals are predicted to lead to a decreased mean retrieval latency. Therefore, in the absence of inhibition between the distractor and the antecedent in antecedent-mismatch conditions, faster processing times are predicted when a feature-matching distractor is present.

Encoding interference predicts increased retrieval latencies and a higher proportion of misretrievals as a function of the number of features the target (here the antecedent) shares with other items in memory (Oberauer and Kliegl, 2006).[6] Thus, in conditions with a feature-matching antecedent (see 3.3a, b), a slow-down and a higher proportion of misretrievals due to a feature-matching distractor (3.3a) is expected. By contrast, in conditions with a mismatching antecedent (see 3.3c, d), a slow-down and a higher proportion of misretrievals due to a *feature-mismatching* distractor (3.3d) is predicted since the mismatching antecedent and the mismatching distractor have the same gender feature *gender* : *masc*.[7]

To summarize, for materials as the ones presented in (3.3), both encoding interference and cue-based retrieval interference predict that a matching distractor leads to a processing slow-down in antecedent match conditions and to a speed-up in antecedent mismatch conditions. For online reading time measures, both accounts thus make precisely the same predictions and can account for the inhibitory effects in antecedent-match conditions reported by Badecker and Straub (2002), Chen et al. (2012b), Clackson and Heyer (2014), Jäger et al. (2015) and Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English" (unpublished manuscript) as well as for the facilitatory effects in antecedent-mismatch conditions reported by King et al. (2012) and Parker and Phillips (2014). For retrieval probabilities (to be reflected in response accuracies of adequate comprehension questions), both accounts also make the same predictions for antecedent-match conditions but differ in their predictions for antecedent-mismatch conditions. Hence, if online evidence for encoding interference in reflexives can be found, we need to reconsider the theoretical implications of interference effects in reflexives with respect to the debate about structurally-guided vs. unconstrained cue-based retrieval. In the following, we present two experiments on German and one experiment on Swedish designed to disentangle encoding from cue-based retrieval interference.

[6] To be precise, the number of distractors sharing a certain feature with the target also affects retrieval latencies and retrieval probability because the more distractors share this feature with the target, the higher the probability that one of these distractors "robs" this feature from the memory representation of the target.

[7] Because we set out to determine whether invoking encoding interference is a way to reconcile interference effects with structure-based retrieval, for the predictions of encoding interference we are assuming that only structural retrieval cues are used. If, by contrast, one assumes that gender is used as a retrieval cue, the feature matching distractor (3.3c) is predicted to be misretrieved more often than the feature mismatching distractor (3.3d). This prediction is orthogonal to the question of encoding interference, but follows from the basic assumption that an item's retrieval probability depends on its features' match with the retrieval cues. This basic assumption is shared by models of encoding interference (Nairne, 1990; Oberauer and Kliegl, 2006). (Note that this point is unrelated to the cue-based retrieval interference component in the Oberauer and Kliegl, 2006 model which is assumed to cause inhibition between items sharing the same retrieval cues.)

3.2 Experiment 1: German reflexives (self-paced reading)

The German reflexive *sich* 'himself'/'herself' is an interesting test case for teasing apart encoding from cue-based retrieval interference. The third-person singular reflexive *sich* is gender neutral and, roughly speaking, requires its antecedent to be a c-commanding noun phrase contained in the reflexive's local clause. For more details about the syntactic properties of German reflexives see Everaert (1986), Reinhart and Reuland (1993), Reuland and Reinhart (1995), Reuland (2001), Gast and Haas (2008) and Reuland (2011). Since *sich* is gender neutral and thus gender can be assumed to not be used as a retrieval cue, we do not expect any cue-based retrieval interference from a structurally inaccessible distractor that shares its gender with the antecedent. Encoding interference, in contrast, predicts that a distractor of the same gender as the antecedent leads to a degradation of the antecedent's memory representation resulting in longer processing times when retrieving the antecedent upon encountering the reflexive. Moreover, encoding interference predicts a lower retrieval probability of the antecedent when a gender-sharing distractor is present. We will use the term *gender-overlap* to refer to the situation where the antecedent and the distractor share their gender in order to reserve the term *gender-match* for the match of an item's feature with a retrieval cue as in example (3.3) discussed above.

3.2.1 Materials and method

3.2.1.1 Materials

The experimental items consist of a matrix clause whose subject is the antecedent of the third person singular reflexive *sich* (see 3.4 for an example). The reflexive is the first constituent of a conjoint determiner phrase (*sich und die Kollegen* in 3.4) which as a whole is the direct object of the matrix verb. The antecedent (*der Dieb/die Diebin* in 3.4) is modified by an object-extracted relative clause that intervenes between the antecedent and the reflexive. The subject of this relative clause (*der Hehler/die Hehlerin* in 3.4) does not c-command the reflexive and hence syntactically disqualifies as antecedent. We will refer to this noun phrase as *distractor*. Both the antecedent and the distractor were always animate common nouns with a definite article. King et al. (2012) have shown that interference effects in reflexives are more likely to be detected when the verb, which triggers the retrieval of its subject — which, in turn, is also the reflexive's antecedent — does not directly precede the reflexive. In order to increase the chances of detecting an effect, we chose perfective tense for our materials, because, as opposed to present tense or simple past, the reflexive precedes the main verb in perfective sentences (for another study on interference effects using pre-verbal reflexives see Kush and Phillips, 2014). Moreover, we inserted a relatively long adverb between the perfective auxiliary *hat* and the reflexive. As in the classical gender-match/mismatch design, we manipulated the antecedent's and the distractor's gender. This resulted in a fully crossed 2×2 design with gender of the antecedent (masculine vs. feminine) and interference (gender-overlap vs. no gender-overlap between the distractor and the antecedent) as factors. For our research question, the gender manipulation of the antecedent was not of interest per se. It was included in order to experimentally control for lexical properties such

as word length or frequency which, due to the nature of the German language, are inseparable from the gender manipulation. We will discuss this issue more in detail in the Results section.

(3.4) *Masc/Fem antecedent; Masc/Fem distractor*

Der Dieb$_i$/Die Diebin$_i$,				dem/der	**der Hehler$_j$/die Hehlerin$_j$**				
the thief-MASC/the thief-FEM				whom	the dealer-MASC/the dealer-FEM				
befohlen	hat	zu	stehlen,	hat	überraschenderweise	**sich$_{i/*j}$**	und	die	
obliged	has	to	steal	has	surprisingly	**self**	and	the	
Kollegen	angezeigt,	berichtete	das	Hochglanzmagazin.					
colleagues	denounced	reported	the	magazine					

The thief whom the dealer obliged to steal surprisingly denounced himself/herself and the colleagues, reported the magazine.

Each sentence was followed by a *yes/no* comprehension question targeting the reflexive-antecedent dependency. One half of the comprehension questions tested whether the antecedent was retrieved successfully (to be answered with *yes*) and the other half tested whether the distractor was misretrieved instead (to be answered with *no*). Question types were balanced across items and held constant within the four conditions of each item.

3.2.1.2 Participants and procedure

144 undergraduate students from the University of Potsdam who were all native speakers of German participated in the study for credit or payment of 5 EUR. We chose a relatively large sample size in order to increase statistical power, i.e., reduce Type II error probability. For our research question, high statistical power is particularly important since if encoding interference in the processing of reflexives is absent, a null result is predicted. The number of participants was determined based on a statistical power test assuming an effect of 20 ms and a standard deviation of 75 ms. In order to achieve power of 90%, 149 participants would be needed. Due to the restricted nature of our participant pool, the actual sample size was slightly smaller, which yielded a statistical power of 0.89%. 16 test items and 32 filler sentences were presented in a moving-window self-paced reading paradigm (Just et al., 1982). Items were arranged according to a Latin Square with a different randomization for each participant. Each trial was followed by a *yes/no* comprehension question.

3.2.2 Results

Statistical analyses were carried out in GNU-R (R Development Core Team, 2009) using linear mixed effects models provided by the lme4 package version 1.0-6 (Bates et al., 2014). Binary dependent variables were modeled using generalized linear mixed models with a logistic link function. For the analyses of comprehension questions and reading times, we fit models testing for a main effect of gender of the antecedent, a main effect of interference (i.e., effect of whether

or not the distractor overlapped in gender with the antecedent) and an interaction between the two. All models were fit with random intercepts and slopes for participants and items (Baayen et al., 2008). No correlations between random effects were estimated since in many of the models the correlation matrix of random effects was degenerate.

	Length antecedent	**Frequency antecedent**	**Length distractor**	**Frequency distractor**
Main effect interference	0	0	0	0
Main effect gender antecedent	-0.44	0.63	0	0
Interference×gender antecedent	0	0	-0.24	0.39

TABLE 3.1: Experiments 1 and 2. Correlations of word length and log lemma frequencies of the antecedent and distractor with the experimental manipulations (Pearson correlation coefficient). Word length and log lemma frequencies were centered (z-scores).

In German, the feminine form of a noun is usually generated by adding the suffix *-in* and in many nouns, the masculine form is more frequent than the feminine one. Therefore, a correlation between gender and word length and word frequency could not be avoided in the stimuli. More precisely, correlations between the main effect of gender and frequency/length of the antecedent as well as correlations between the interaction antecedent gender×interference and frequency/length of the distractor are expected. Crucially, including the gender manipulation of the antecedent as a fully crossed within-items factor in our design ensured a zero correlation between frequency/length of the antecedent or the distractor with the critical main effect of interference. Along the same lines, correlations between frequency/length of the antecedent and the interaction antecedent gender×interference as well as correlations between length/frequency of the distractor and the main effect of gender of the antecedent cancel out due to the fully-crossed factorial design. To test these assumptions and to obtain estimates for the expected correlations, we computed Pearson-correlations of each of the contrasts to be tested in the linear-mixed model with centered word lengths measured in number of characters and centered log-transformed lemma frequencies taken from dlexDB[8] (Heister et al., 2011) of the antecedent and the distractor (see Table 3.1). As expected, there was a positive correlation (r=0.63) between the main effect of gender of the antecedent and frequency of the antecedent and a negative correlation (r=-0.44) between the main effect of gender of the antecedent and word length of the antecedent. Similarly, there was a positive correlation (r=0.39) between the frequency of the distractor and the interaction interference×gender of antecedent and a small negative correlation between word length of the distractor and the interaction interference×antecedent gender. Thus, a main effect of gender of the antecedent and the interaction between the two main effects should not be interpreted since they might be confounded with the effects of antecedent/distractor length and frequency.

3.2.2.1 Comprehension questions

Comprehension question response accuracies were analyzed using a linear mixed model with a logistic link function. Mean accuracy scores of question responses are provided in Table 3.2. Statistical analyses revealed a main effect of interference: accuracy was lower in conditions with a gender-sharing distractor (estimate=-0.25, SE=0.12, z=-2.02, $p<0.05$). Neither the main effect of gender nor the interaction were significant.

[8] www.dlexdb.de

Condition	**Accuracy**
gender-overlap; masculine antecedent	71
gender-overlap; feminine antecedent	73
no gender-overlap; masculine antecedent	75
no gender-overlap; feminine antecedent	77

TABLE 3.2: Experiment 1. Mean accuracy scores of question responses in percentage by experimental condition.

3.2.2.2 Reading times

An overview of raw reading times for each region of the sentence is provided in Table B.1 in the appendix. Reading times were analyzed at the reflexive, the following NP together with the preceding conjunction *und* 'and' (n+1), the main clause verb (n+2) as well as at the word preceding the reflexive as a sanity test of the baseline reading times. In order to achieve a close to normal distribution of the model residuals, we analyzed negative reciprocal reading times (Box and Cox, 1964). None of the comparisons reached significance at any region. Modeling log-transformed RTs instead of reciprocal RTs yielded similar results. The output of the linear-mixed models is provided in Table 3.3.

Predictor	**n-1**			**REFL**			**n+1**			**n+2**		
	coef	*SE*	*t*	*coef*	*SE*	*t*	*coef*	*SE*	*t*	*coef*	*SE*	*t*
Interference	4e-05	3e-05	1.39	2e-05	2e-05	0.67	0.00	1e-05	0.18	-1e-05	3e-05	-0.58
Gender antecedent	0.00	2e-05	-0.08	1e-05	3e-05	0.29	0.00	1e-05	-0.42	-4e-05	3e-05	-1.25
Gender ant.×Interf.	-4e-05	3e-05	-1.49	-3e-05	2e-05	-1.49	-1e-05	1e-05	-0.96	-1e-05	3e-05	-0.45

TABLE 3.3: Experiment 1. Main effects of interference and gender of the antecedent and their interaction on negative reciprocal RTs as dependent variable measured at the adverb preceding the reflexive (n-1), the reflexive (REFL), the coordinate NP following the reflexive (n+1) and the main clause verb (n+2).

3.2.3 Discussion

In reading times, we did not find any effect of gender-overlap between the antecedent and the distractor. However, in comprehension questions, we observed lower response accuracies when the distractor overlapped in gender with the antecedent. This effect might be explained by misretrievals due to encoding interference during online processing which, critically, did not affect processing times. Alternatively, the lower response accuracies in the gender-overlap conditions might reflect an offline effect that arises at the moment of answering the comprehension question. Crucially for our research question, we could not find any evidence supporting the idea that encoding interference affects online processing times at the reflexive. With respect to previous studies on reflexives, we can therefore conclude that there is no indication that the interference effects observed in previous studies reflect encoding rather than cue-based retrieval interference.

3.3 Experiment 2: German reflexives (eye-tracking)

Experiment 2 is a cross-methodological replication of Experiment 1. Already Ronald Fisher, the father of frequentist statistics, emphasized the importance of replication (Fisher, 1937, page 16). Indeed, non-replicable findings are a major problem in experimental psychology and psycholinguistics (Asendorpf et al., 2013; Simmons et al., 2011). Moreover, a potential concern about Experiment 1 is that our conclusions are based on a null result. Although we have addressed this issue by testing a large sample and thus gaining high statistical power, one could still argue that the self-paced reading method is not sensitive enough to detect a potential effect. We therefore tested the same materials as in Experiment 1 in an eye-tracking while reading paradigm, which presumably is a more sensitive method compared to self-paced reading (Staub and Rayner, 2007).

3.3.1 Materials and method

3.3.1.1 Materials

The same stimuli (including fillers) were used as in Experiment 1.

3.3.1.2 Participants and procedure

151 undergraduate students from the University of Potsdam with normal or corrected-to-normal vision who were all native speakers of German participated in the experiment against credit or payment of 7 EUR. None of the participants had participated in Experiment 1.

Participants' eye movements (right eye monocular tracking) were recorded with an SR Research Eyelink 1000 eyetracker at a sampling rate of 1000 Hz using a Desktop mount camera system with a 35 mm lens. The participant was seated at a height-adjustable table with his/her head stabilized using a forehead/chin-rest. Stimuli were presented on a 22 inch monitor (resolution of 1680×1050 pixels) with an eye-to-screen distance of 62 cm and an eye-to-camera distance of 60 cm. As a response pad, a Microsoft Button Box was used. Stimuli were presented using Experiment Builder software provided by SR Research. The experimental items were presented on a light gray background in black font, font type Times New Roman, font size 14. They were arranged according to a Latin Square and were pseudo-randomized for each participant separately such that every experimental trial was preceded by at least one filler sentence. A nine-point calibration was carried out at the beginning of the experiment and repeated during the experiment, if needed. Each experimental session started with 6 practice trials. At the beginning of each trial, participants had to fixate a drift correction point at the left center of the screen where the first word of the sentence was to appear.

3.3.2 Results

Linear mixed-effects models were fit with same predictors as for Experiment 1. As in the analysis of Experiment 1, all models were fit with varying intercepts and slopes for participants and items. No correlations between random effects were estimated since, as in the data of Experiment 1, the correlation matrix of random effects was degenerate in many of the models.

3.3.2.1 Comprehension questions

Mean accuracy scores by experimental condition are provided in Table 3.4. We observed a marginal main effect of interference with lower accuracies in conditions where antecedent and distractor had the same gender (estimate=-0.20, SE=0.10, z=-1.95, p=0.05). This replicates the pattern found in Experiment 1. None of the other effects was significant.

Condition	**Accuracy**
gender-overlap; masculine antecedent	74
gender-overlap; feminine antecedent	69
no gender-overlap; masculine antecedent	75
no gender-overlap; feminine antecedent	75

Table 3.4: Experiment 2. Mean accuracy scores of question responses in percentage by experimental condition.

3.3.2.2 Eye movements

An overview of raw reading times at each word of the sentence is provided in Table B.2 in the appendix. The same regions were analyzed as in Experiment 1. Raw fixation durations shorter than 20 ms or longer than 1000 ms (0.25% of the data) were excluded from all analyses. In eye-tracking data, the dependent measures can be partitioned into first-pass, regression-related (proportions of regressions and duration of regressive events) and later-pass measures. Since the exact mapping between syntactic effects and eye-tracking measures is still unclear (Clifton et al., 2007), we analyzed one representative measure from each group. As a first-pass measure, we analyzed first-pass reading time (FPRT, also referred to as gaze duration), which is defined as the sum of all first-pass fixations on a region. As regression related measures, we analyzed first-pass regression-probability (FPRP), i.e., a binary variable coded as 1 if a first-pass regression was initiated from a region, and regression-path duration (RPD), i.e., the sum of all fixation durations starting from the first fixation on a region until leaving this region to the right including all regressive fixations that fall into this time window. As a later-pass measure, we analyzed total-fixation time (TFT), i.e., the sum of all fixations on a region. Strictly speaking, TFT is not a pure late measure but rather the sum of FPRT and re-reading time. However, we chose to report TFT as a representative late measure since TFT is one of the most commonly reported measures in psycholinguistics; we do not analyze re-reading time because the critical region was re-read in only about 20% of the trials leading to very low statistical power. In order to achieve approximately normally distributed residuals, the continuous dependent variables were log-transformed (Box and Cox, 1964).

DV	Predictor	n-1			REFL			n+1			n+2		
		coef	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*
FPRT	Interference	-0.01	0.02	-0.25	-0.01	0.02	-0.31	-0.01	0.03	-0.38	-0.02	0.01	-1.14
	Gender antecedent	0.01	0.02	0.61	-0.03	0.02	-1.35	-0.03	0.03	-1.03	0.01	0.01	0.64
	Gender ant.×Interf.	-0.02	0.02	-1.06	-0.01	0.02	-0.64	-0.04	0.02	-1.91	0.00	0.02	0.25
RPD	Interference	0.04	0.03	1.59	0.01	0.02	0.44	-0.02	0.03	-0.67	0.00	0.03	-0.06
	Gender antecedent	0.02	0.02	0.86	-0.02	0.02	-0.76	-0.02	0.02	-1.01	0.01	0.02	0.55
	Gender ant.×Interf.	-0.03	0.02	-1.14	0.03	0.02	1.10	-0.02	0.02	-1.01	0.00	0.02	-0.22
FPRP	Interference	0.19	0.11	1.64	0.12	0.14	0.83	-0.05	0.11	-0.44	0.13	0.17	0.77
	Gender antecedent	0.09	0.12	0.72	-0.10	0.14	-0.7	-0.05	0.14	-0.34	-0.01	0.13	-0.05
	Gender ant.×Interf.	0.03	0.11	0.22	0.18	0.18	0.98	0.10	0.11	0.85	-0.01	0.14	-0.09
TFT	Interference	0.03	0.02	1.4	-0.01	0.02	-0.36	0.01	0.02	0.37	0	0.02	0.12
	Gender antecedent	0.03	0.02	1.54	0.02	0.02	0.68	0	0.02	0.13	0.04	0.02	2.22 *
	Gender ant.×Interf.	0.01	0.02	0.29	0.01	0.02	0.46	-0.04	0.02	-1.93	-0.01	0.02	-0.72

TABLE 3.5: Experiment 2. Main effects of interference and gender of the antecedent and their interaction on the dependent variables log-first-pass reading time, log-regression-path duration, first-pass regression probability and log-total fixation time measured at the adverb preceding the reflexive (n-1), the reflexive (REFL), the coordinate NP following the reflexive (n+1) and the main clause verb (n+2).

An overview of the output of the linear mixed-effects models is provided in Table 3.5. At the reflexive, the word preceding it and the region following the reflexive none of the comparisons reached significance in any of the dependent variables. At region n+2 (i.e., the main clause verb), a significant effect of gender of the antecedent was observed in TFT (longer fixation times in conditions with a feminine antecedent). However, as we have argued in the Results section of Experiment 1, this effect should not be interpreted since it correlates with frequency and length of the antecedent. For our research question, only the main effect of interference is relevant.

Moreover, a post-hoc analysis of the region containing the relative clause verb (*zu stehlen* in example 3.4) revealed a significant main effect of interference in TFT with longer fixation durations when the antecedent and the distractor overlapped in gender (estimate=0.047, SE=0.02, t=2.28).

3.3.3 Discussion

Experiment 2 replicated the findings of Experiment 1. As in Experiment 1, no evidence for encoding interference due to gender-overlap between the reflexive's antecedent and a structurally inaccessible distractor was observed neither at the reflexive, nor at the pre- or post-critical regions.

At the relative clause verb, however, gender-overlap between the main clause subject, i.e., the antecedent, and the relative clause subject, i.e., the distractor, led to significantly longer total-fixation times. At this region, the relative clause subject needs to be retrieved. Hence, the observed effect, which appears in a similar region as the effects reported by Gordon et al. (2001), might reflect encoding interference. However, it is disconcerting that this effect was observed

only in total-fixation time and was not present in Experiment 1, as a post-hoc analysis of the self-paced reading data showed. Thus, one might discount this effect as a possible Type I error. If one does not discount the effect, it raises the question why encoding interference affects argument-head dependency completion, but not reflexive-antecedent dependency formation. A possible explanation might be that the encoding interference effect (to the extent that it is not a Type I error) dies out by the time the reflexive is processed.[9] In any case, further replication attempts of this configuration are needed. In sum, it is possible that we are seeing encoding interference at the distractor, but, which is crucial for our research question, this encoding interference does not seem to have any effect at the reflexive.

Taken together, the results of Experiments 1 and 2 are a strong indication that in reflexive-antecedent dependency formation, the sharing of a non-structural feature such as gender does not lead to encoding interference reflected in a processing slow-down. More precisely, it indicates that in materials of the type used in this experiment, encoding interference does not affect retrieval latencies of the antecedent when processing the reflexive. However, the marginal interference effect in offline comprehension accuracies, which had been significant in Experiment 1, indicates that the antecedent was retrieved less often correctly when it shared its gender with the distractor. This can be interpreted as evidence for encoding interference affecting retrieval probability of the antecedent. In sum, neither experiment provides any evidence for the claim that encoding interference affects reading time at the reflexive. However, our offline results suggest that encoding interference might affect retrieval probability of the antecedent. Crucially, even if encoding interference affected retrieval probability of the antecedent or the offline interpretation of the sentence, there is no evidence that it affects the participants' online behavior at the reflexive measured in self-paced reading times or eye-movements. Hence, encoding interference is not a plausible explanation for the *online* effects previous studies have observed in eye-tracking or self-paced reading measures.

3.4 Experiment 3: Swedish possessives (eye-tracking)

Experiments 1 and 2 yielded converging results: we found no evidence for encoding interference affecting the online processing speed of German reflexives. However, there are still two potential concerns with these results: i) Our conclusion is based on two null-results, and ii) we need to cross-linguistically validate our conclusions. In Experiment 3, we addressed these issues by examining the processing of Swedish possessives in an eye-tracking experiment. In Swedish, there are two kinds of possessives: reflexive possessives that are not gender marked and pronominal possessives that need to agree in gender with their antecedent. The reflexive possessive *sin* 'his'/'her' can only be bound by a c-commanding antecedent inside its local clause. In contrast, the pronominal possessive *hans* 'his' must not be bound within its local clause, but requires an

[9] A reviewer noticed that the effect at the relative clause verb occurs in total-fixation time, a measure which can reflect processing difficulty encountered further downstream in the sentence, and therefore might actually reflect processing difficulties at the reflexive which triggers re-readings of the previous materials. However, if this were the case, one would expect an increase in the proportion of regressions or increased regression-path durations at the reflexive. As this is not the case, it is difficult to conclude that the effect observed at the verb reflects processing difficulty associated with the reflexive.

antecedent outside its clause domain (see Holmes and Hinchliffe, 1994 and Kaiser, 2003, p. 209). In a 2×2 factorial design, we manipulated anaphor type (pronominal possessive vs. reflexive possessive) and interference, i.e., whether or not a structurally inaccessible distractor shared the gender of the antecedent. For this design, encoding interference predicts increased processing difficulty in the interference conditions compared to the non-interference conditions, regardless of anaphor type. Cue-based retrieval interference, in contrast, predicts an interaction between anaphor type and interference: increased processing difficulty due to a gender-sharing distractor is predicted for the gender-marked pronominal possessives but not for the gender-unmarked reflexive possessives. This is because only in pronominal possessives, the gender-marked anaphor can trigger a retrieval process where gender is used as a retrieval cue. When both the antecedent and the distractor match the gender cue, cue-based retrieval interference predicts inhibition between the antecedent and distractor and a higher proportion of misretrievals of the distractor (Lewis and Vasishth, 2005). Thus, the present experiment allows us to directly pit encoding and cue-based retrieval interference against each other. In contrast to Experiments 1 and 2, cue-based retrieval interference predicts an interaction rather than a null-result.

3.4.1 Materials and method

3.4.1.1 Materials

The conditions with pronominal possessives (see 3.5a for an example item) consist of a superordinate clause whose subject is the antecedent (*Åke* in 3.5a) and a subordinate clause containing the distractor (*Alf* or *Ann* in 3.5a) which either matches or mismatches the gender of the antecedent and the gender-marked pronominal possessive (*hans* 'his' in 3.5a). The conditions with reflexive possessives (see 3.5b for an example item) consist of a main clause containing the antecedent (*Åke* in 3.5b) and the gender-unmarked reflexive possessive (*sina* 'his'/'her' in 3.5b). The distractor (*Alf* or *Ann* in 3.5b) is the subject of an appositive relative clause intervening between the antecedent and the reflexive possessive. As Swedish does not code masculine and feminine as grammatical gender, and the number of nouns with inherent gender such as *boy* or *girl* is very limited, both the antecedent and the distractor were proper names in all experimental sentences. Indeed, it is crucial for our research question to extend the findings of Experiments 1 and 2 to proper names, which differ from common nouns with respect to their referential properties (Longobardi, 1994; Elbourne, 2005), since several of the studies reporting interference effects in reflexives actually employed proper names (e.g., Badecker and Straub, 2002).

(3.5) a. *Pronominal possessives; gender-overlap/no gender-overlap*

Åke$_i$	säger	att	**Alf**$_j$**/Ann**$_j$	jobbade	med	**hans**$_{i/*j}$
Åke-masc	says	that	**Alf-masc/Ann-fem**	worked	with	**his**

sysslingar	på	helgerna.
cousins	at	the weekend

Åke says that Alf/Ann worked with his cousins at the weekend.

b. *Reflexive possessives; gender-overlap/no gender-overlap*

Åke$_i$	som	**Alf**$_j$**/Ann**$_j$	tackade	ringer	**sina**$_{i/*j}$	sysslingar
Åke-masc	whom	**Alf-masc/Ann-fem**	thanked	calls	**his**	cousins

på kvällen.
in the evening

Åke, whom Alf/Ann thanked, calls his cousins in the evening.

The nouns used as antecedents and distractors are all highly frequent, gender unambiguous Swedish first names taken from Statistics Sweden, a database which contains the 100 most frequently given and used male and female first names in Sweden.[10] Antecedents and distractors are all matched for word length (numbers of characters) within each item. Half of the items have a feminine antecedent and the other half a masculine antecedent. The possessed noun phrase (*sysslingar* in 3.5) is always a plural noun.

Two types of comprehension questions were designed. The first type probes for the correct interpretation of the anaphor-antecedent dependency, 50% of which were to be answered with *yes*. The second question type targeted various parts of the sentence, but not the interpretation of the anaphor. Again, 50% of these questions were to be answered with *yes*.

3.4.1.2 Participants and procedure

35 native speakers of Swedish currently living in Berlin or Potsdam with normal or corrected-to-normal vision participated in the experiment against payment of 5 EUR (plus 6.20 EUR to cover travel expenses). The sample size was smaller compared to Experiments 1 and 2 due to logistic limitations, but we tested a larger number of experimental items compared to Experiments 1 and 2. Participants' eye movements (right eye monocular tracking) were recorded while reading 48 experimental sentences and 70 filler sentences. The general technical set-up was the same as in Experiment 2. Stimuli were arranged in a Latin Square and pseudo-randomized such that each experimental trial was preceded by at least one filler sentence. Each trial was followed by a comprehension question. Two thirds of the comprehension questions targeted the correct interpretation of the anaphor and one third targeted other parts of the sentence. The experiment started with 5 practice trials to familiarize participants with the procedure.

3.4.2 Results

On all dependent variables, we fit linear mixed-effects models with main effects of anaphor type (pronominal vs. reflexive possessive), interference (whether or not the distractor had the same gender as the antecedent) and their interaction as predictors. When the interaction reached significance, nested contrasts testing for an interference effect within each anaphor type were fit. All models were fit with varying intercepts for participants and items. No varying slopes

[10] http://www.scb.se/BE0001-EN; we used the data of 2012.

were fit because the generalized likelihood-ratio test showed that they did not improve the model fit. The pattern of results was not affected by whether or not varying slopes were fit. For the interpretation of results, it should be kept in mind that the effect of anaphor type is not of theoretical relevance to our research question. As the two levels of anaphor type differ lexically at the pre-critical and the critical region, a main effect of anaphor type does not have any useful interpretation.

3.4.2.1 Comprehension questions

Mean accuracies by experimental condition and question type (i.e., whether or not the comprehension question targeted the anaphor) are provided in Table 3.6. We ran a linear-mixed effects model with a logistic link function with main effects of anaphor type, interference and question type and their interactions including the three-way interaction between all main effects as predictors. The model output is summarized in Table 3.7. The main effect of interference and the interaction between interference and question type reached significance. Moreover, a marginal three-way interaction between interference, anaphor type and question type was observed. A second model in which we applied nested contrasts testing for an interference effect within each level of anaphor type and question type[11] showed that the interactions were caused by a highly significant interference effect that was present only in questions targeting the anaphor in pronominal possessives (estimate=-1.16, SE=0.25, z=-4.62, p<0.0001). In sum, in questions targeting the anaphor-antecedent dependency, the presence of a gender matching distractor led to lower response accuracies in sentences with pronominal possessives but not in sentences with reflexive possessives. In questions not targeting the anaphor-antecedent dependency, no effects were observed.

Condition	**Accuracy**	
	anaphor	*other*
pronominal; gender-overlap	75	82
pronominal; no gender-overlap	90	82
reflexive; gender-overlap	85	80
reflexive; no gender-overlap	86	81

TABLE 3.6: Experiment 3. Mean accuracy scores of comprehension questions in percentage by experimental condition and question type, i.e., whether the question targeted the anaphor-antecedent dependency or another element of the sentence.

3.4.2.2 Eye movements

An overview of raw reading times at each region of the sentence is provided in Table B.3 in the appendix. We analyzed the pre-critical region containing the verb (plus postposition), the critical region containing the pronominal or reflexive possessive and the post-critical region containing the possessed noun. The same dependent variables were analyzed as in Experiment 2. Continuous

[11]The model predictors were main effects of anaphor and question type, interaction between anaphor type and question type and the four pairwise comparisons (interference effects in pronominal and reflexive possessives in question targeting the anaphor and questions not targeting the anaphor.)

Predictor	***coef***	***SE***	***z***
Interference	-0.17	0.07	-2.42*
Anaphor type	-0.01	0.07	-0.17
Question type	0.12	0.10	1.11
Interference×Anaphor type	-0.12	0.07	-1.61
Interference×Question type	-0.15	0.07	-2.04*
Anaphor type×Question type	-0.06	0.07	-0.86
Interference×Anaphor type×Question type	-0.14	0.07	-1.95

TABLE 3.7: Experiment 3. Analysis of comprehension questions: Main effects of interference, anaphor type and question type together with their interactions.

DV	**Predictor**	**n-1**			**REFL/PRON**			**n+1**		
		coef	*SE*	*t or z*	*coef*	*SE*	*t or z*	*coef*	*SE*	*t or z*
FPRT	Interference	0.02	0.02	0.75	0.01	0.02	0.57	-0.03	0.02	-1.29
	Anaphor type	0.03	0.02	1.43	-0.04	0.02	-1.97	0.05	0.02	1.99
	Anaph. type×Interf.	0.03	0.02	1.28	-0.01	0.02	-0.41	0.02	0.02	0.71
RPD	Interference	-0.01	0.03	-0.18	-0.02	0.03	-0.67	0.00	0.04	-0.05
	Anaphor type	-0.25	0.03	-8.11*	-0.13	0.03	-4.04*	-0.07	0.04	-1.94
	Anaph. type×Interf.	0.01	0.03	0.29	0.01	0.03	0.17	-0.02	0.04	-0.53
FPRP	Interference	-0.08	0.13	-0.62	-0.11	0.13	-0.85	-0.19	0.12	-1.56
	Anaphor type	-1.26	0.13	-9.4*	-0.32	0.13	-2.45*	-0.4	0.12	-3.36*
	Anaph. type×Interf.	-0.08	0.13	-0.63	0.09	0.13	0.70	**-0.25**	**0.12**	**-2.1***
TFT	Interference	0.04	0.03	1.46	0.02	0.03	0.64	0.01	0.03	0.50
	Anaphor type	-0.07	0.03	-2.44*	0.09	0.03	3.15*	0.03	0.03	0.89
	Anaph. type×Interf.	0.05	0.03	1.70	-0.01	0.03	-0.40	0.02	0.03	0.65

TABLE 3.8: Experiment 3. Main effects of interference and anaphor type and their interaction at the pre-critcal region n-1, the reflexive/pronominal possessive (REFL/PRON) and the post-critical region n+1. The dependent measures are log-first-pass reading time, log-regression-path duration, first-pass regression probability and log-total fixation time.

dependent variables were log-transformed in order to achieve approximately normally distributed residuals.

The effect of anaphor type reached significance across regions and dependent variables. However, as mentioned above, this effect was not of interest to our research question: conditions with pronominal and reflexive possessive differ from each other in syntactic structure, distractor position, lexicon, word length and number of words contained in the pre-critical region. At the pre-critical and the critical region, no other effect reached significance in any dependent variable. At the post-critical region, a significant interaction between anaphor type and interference was observed in first-pass regression probability. Pairwise comparisons revealed that this interaction was driven by a significant interference effect in pronominal possessives. When the distractor shared the gender of the antecedent and hence matched the gender-cue, less first-pass regressions were observed (estimate=-0.44, SE=0.18, z=-2.47, $p<0.05$). In order to test whether this facilitation due to a gender-matching distractor reflected misretrievals of the latter, we re-ran the

models on comprehension question response accuracies for trials with and without a first-pass regression from the post-critical region separately.

In trials without a first-pass regression from n+1, the interference effect in pronominal possessives in questions targeting the critical dependency (i.e., the effect observed in the overall data) was highly significant (estimate=-1.19, SE=0.28, z=-4.21, p< 0.0001). By contrast, in trials with a first-pass regression initiated at n+1, this effect did not reach significance (estimate=-0.94, SE=0.57, z=-1.66, p=0.09). This post-hoc analysis clearly shows that the interference effect in response accuracies in pronominal possessives was driven by trials in which no first-pass regression was initiated, i.e., by the trials responsible for the facilitation observed in FPRP.

3.4.3 Discussion

We did not find any evidence for encoding interference affecting processing times of Swedish anaphor-antecedent dependencies. Together with the results of Experiments 1 and 2, this suggests that in materials with a classical gender-match/mismatch manipulation, encoding interference does not affect retrieval latencies of the antecedent. In comprehension questions, we did not see evidence for encoding interference affecting retrieval probability of the antecedent either. This is in contrast to the pattern observed in response accuracies of Experiments 1 and 2.

Evidence for interference occurring at the moment of retrieval was observed in online and offline measures. The lower proportion of first-pass regressions initiated at the region directly after the gender-marked pronominal possessive in conditions with a gender-matching distractor indicates a processing facilitation due to a cue-matching distractor. Such a facilitation can be explained in terms of misretrievals of the gender-matching distractor under the assumption that misretrievals go along with shorter retrieval latencies. The lower response accuracies in comprehension questions targeting the retrieval of the antecedent support this explanation. Indeed, the post-hoc analysis of response accuracies for trials with and without a first-pass regression from the post-critical region clearly shows that the facilitation observed in first-pass regressions is directly connected to misretrievals of the gender-matching distractor.

The cue-based ACT-R model of sentence processing (Lewis and Vasishth, 2005) predicts misretrievals of the gender-matching distractor. These misretrievals are predicted to lead to shorter retrieval latencies, i.e., a processing facilitation, in the respective trials. However, the ACT-R model also predicts inhibition between the gender-matching distractor and the antecedent leading to longer retrieval latencies of the antecedent. Overall, the predicted direction of the interference effect therefore depends on the concrete parameter setting of the model. With the default parameter setting, inhibitory interference (i.e., the opposite effect than the one in the data) is predicted. If one assumes a particularly high activation of the distractor, ACT-R predicts the observed pattern because the highly activated distractor is misretrieved in a considerable proportion of the trials, which leads to a speed-up in the observed mean retrieval latencies (Jäger et al., 2015). Indeed, facilitation in a configuration similar to our materials has been observed in previous studies (Sturt, 2003; Cunnings and Felser, 2013; Laurinavichyute et al., 2015; Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding

in English", unpublished manuscript). An argument favoring the assumption that the distractor is highly activated in our materials is that, similar to the other experiments reporting facilitation, the distractor is in subject position. Moreover, the distractor has a recency advantage over the antecedent as it is linearly closer to the retrieval site. Indeed, ACT-R predicts a recency advantage which follows from the assumption that an item's activation level decreases as a function of the passage of time (decay) and intervening material (interference). In sum, under the plausible assumption that the distractor is highly activated in our materials, cue-based retrieval interference as implemented in the ACT-R model can account for the observed pattern. Hence, the interference effect in pronominal possessives can be interpreted as evidence favoring a cue-based retrieval mechanism. However, it should be kept in mind that pronominal possessives are not subject to Binding Principle A. Hence, the observed effects cannot be interpreted as evidence against theories of sentence processing claiming that Principle A is immune to interference from structurally illicit antecedents (Nicol and Swinney, 1989; Dillon et al., 2013; Phillips et al., 2011).

An alternative explanation that can account for the facilitation leading to misretrievals of the gender-matching distractor in pronominal possessives but not in reflexive possessives builds on the fact that we are comparing reflexive possessives which are subject to Binding Principle A with pronominal possessives which are subject to Binding Principle B. As mentioned above, pronominal possessives must not be bound in their local domain (Binding Principle B, see Chomsky, 1981). In the syntax-semantic literature about the interpretation of pronouns, it has been proposed that in the presence of a local c-commanding noun phrase which matches the gender feature of the anaphor (as the gender-matching distractor in the pronominal possessives conditions of Experiment 3), local binding is preferred over a non-local antecedent (Fox, 1998; Heim and Kratzer, 1998). This leads to a temporary violation of Binding Principle B. Only after the local binder has successfully been inhibited, the actual search for the structurally licit antecedent is initiated (Grodzinsky and Reinhart, 1993; Reinhart, 2000; Reuland, 2011). If in our materials, the syntactically local binder of the pronominal possessive (i.e., the distractor) is accessed in a first stage of the retrieval process, in a certain proportion of the trials, this local binder might be misretrieved in case it matches the gender of the pronominal possessive and the search for the antecedent terminates already after this first stage. Such a scenario would explain the misretrievals reflected in response accuracies and also the speed-up in trials where misretrievals occurred. This model correctly predicts that facilitatory interference should be observed only with Principle B pronouns, not with Principle A reflexives since in reflexives, the local binder is the licit antecedent. Crucially, the absence of an effect in our reflexive possessive conditions is not explained by them being unmarked for gender but rather by their syntactic binding properties.

To summarize, we found no evidence for encoding interference affecting the processing of Swedish possessives. We did observe evidence for retrieval interference in gender-marked pronominal possessives. The presence of a gender-matching distractor led to facilitated processing, presumably as a consequence of misretrievals of the latter in a certain proportion of trials. Although this pattern can be explained in terms of unconstrained cue-based retrieval, it is also consistent with the view that comprehending a pronoun constrained by Principle B requires comprehenders to temporarily consider and inhibit coreference with the local subject (the distractor in our materials). However, it should be noted that recent evidence from English pronouns reported by Chow et al. (2014) is inconsistent with the idea of first accessing and subsequently inhibiting a

local antecedent. In none of their five reading experiments did they observe a facilitatory effect on pronoun resolution from a feature-matching local antecedent.

3.5 General discussion

We set out to find evidence for encoding interference in the processing of reflexives. With respect to the current debate about structure-based vs. unconstrained cue-based retrieval subserving the processing of reflexives, the question whether encoding interference can be observed in reflexives is crucial because, as has been argued by Dillon (2011), encoding interference provides an alternative explanation for interference effects in reflexives which originally have been attributed to cue-based retrieval interference and hence taken as evidence for unconstrained cue-based retrieval (Badecker and Straub, 2002; Chen et al., 2012b; Jäger et al., 2015; Patil, Vasishth and Lewis, "Retrieval interference in syntactic processing: The case of reflexive binding in English", unpublished manuscript).

In order to decide whether encoding interference is present in the processing of reflexives, we conducted two experiments on the German reflexive *sich*. In contrast to previous studies, where encoding and cue-based retrieval interference made the same predictions, the gender-unmarked *sich* allowed us to pit against each other the predictions of retrieval and encoding interference. Cue-based retrieval interference predicts no effect of gender of a structurally inaccessible distractor whereas encoding interference predicts a slow-down when the gender of the distractor matches the gender of the antecedent. Neither with self-paced reading nor with eye-tracking did we find any indication for an online interference effect caused by a gender-sharing distractor, although the statistical power of our experiments was considerably higher than the one of previous experiments reporting interference effects in reflexives. We conducted a third experiment on Swedish possessives to cross-linguistically validate our finding. The interaction between interference and anaphor type provided further support for the conclusion that sharing the gender feature with a distractor does not lead to encoding interference in the processing of reflexives. Although we did not find any evidence that encoding interference affected online processing ease, response accuracies in the comprehension questions of Experiment 1 indicate that encoding interference might have caused misretrievals of the gender-sharing distractor. However, this effect was only marginal in Experiment 2 and could not be replicated in Experiment 3. Critically, these supposed misretrievals observed in Experiment 1 are not reflected in online processing measures. In sum, there is no evidence for encoding interference affecting online processing measures. Therefore, there is no evidence for the proposal that interference effects reported in previous studies on reflexives arise from encoding interference. This finding therefore provides support for the assumption that interference effects observed in reflexive processing arise at the moment of retrieval rather than at the encoding stage. In other words, encoding interference is not a plausible explanation for reconciling interference effects with a structure-based account of reflexive processing. Thus, taken together with the interference effects reported in previous studies on reflexive processing, our findings favor an unconstrained cue-based retrieval architecture.

Lastly, we want to emphasize that our results should not be interpreted as evidence for the absence of encoding interference in sentence processing per se. Indeed, the effect at the relative

clause verb in Experiment 2 might reflect encoding interference. The presence of encoding interference *as such* is in principle not incompatible with a content-addressable architecture since content-addressability is an architectural mechanism concerning the *retrieval*, but not the *encoding* or the *maintenance* of an item in working memory.

More generally, our findings provide support for a content-addressable memory architecture subserving language comprehension. This adds to a growing body of evidence from various kinds of syntactic dependencies such as filler-gap (McElree et al., 2003) and subject-verb dependencies (Van Dyke and Lewis, 2003; Van Dyke and McElree, 2006; Van Dyke, 2007; Wagers et al., 2009; Van Dyke and McElree, 2011; Dillon et al., 2013), the licensing of negative-polarity items (Vasishth et al., 2008) and verb-phrase ellipsis (Martin and McElree, 2008), suggesting that the parser uses a cue-based retrieval mechanism to process these dependencies. One fundamental question in sentence processing research is whether the human parser uses qualitatively different retrieval mechanisms in the processing of different kinds of dependencies. Indeed, proponents of the structure-based account of reflexive processing have argued that the retrieval mechanisms mediating the processing of reflexives differ qualitatively from the ones used, e.g., in the processing of subject-verb dependencies (Phillips et al., 2011; Dillon et al., 2013). Hence, evidence for cue-based retrieval subserving the processing of reflexives is one important piece of evidence towards a content-addressable model of working memory underlying sentence processing in general, which not only invokes qualitatively similar working memory mechanisms to explain the processing of different kinds of linguistic dependencies, but, even beyond that, locates the language processing system within a general cognitive architecture where independently motivated working memory mechanisms operate on linguistic representations.

Chapter 4

Journal of Memory and Language
Volume 79-80, 2015, pages 97-120

The Subject-Relative Advantage in Chinese: Evidence for Expectation-Based Processing

Lena Jäger[1], Zhong Chen[2], Qiang Li[3,4], Chien-Jer Charles Lin[5], and Shravan Vasishth[1,6]

[1]Dept. of Linguistics, University of Potsdam [2]Dept. of Modern Languages and Cultures, Rochester Institute of Technology [3]Dalian University of Technology [4]Dept. of Communicative Disorders, University of Louisiana at Lafayette [5]Dept. of East Asian Languages and Cultures, Indiana University Bloomington [6]School of Mathematics and Statistics, University of Sheffield

Abstract

Chinese relative clauses are an important test case for pitting the predictions of expectation-based accounts against those of memory-based theories. The memory-based accounts predict that object relatives should be easier to process than subject relatives because, in object relatives, less linguistic material intervenes between the head noun and the gap (or verb) that it associates with. By contrast, expectation-based accounts such as surprisal predict that the less frequently occurring object relative should be harder to process than the subject relative, because building a rarer structure is computationally more expensive. Previous studies on Chinese relative clauses have the problem that local ambiguities in subject and object relatives could be confounding the comparison. We compared reading difficulty in subject and object relatives (in both subject- and object-modifications) in which the left context leads the reader to predict

a relative clause structure as the most likely continuation; we validate this assumption about what is predicted using production data (a sentence completion study and a corpus analysis). Two reading studies (self-paced reading and eye-tracking) show that the Chinese relative clause evidence is consistent with the predictions of expectation-based accounts but not with those of memory-based theories. We present new evidence that the *prediction of upcoming structure*, generated through the probabilistic syntactic knowledge of the comprehender, is an important determiner of processing cost.

Keywords: Sentence comprehension; Relative clauses; Structural expectation; Working-memory; Eye-tracking; Surprisal; Chinese

4.1 Introduction

Tracking expectations is of central importance in language comprehension: as we read or hear a sentence, we constantly generate predictions about upcoming material at every linguistic level. According to one view, these predictions are a result of our production system being deployed as an "emulator" during language comprehension (Pickering and Garrod, 2007); a related assumption is that the comprehender maintains and uses linguistic knowledge probabilistically to incrementally parse a sentence (Hale, 2001; Levy, 2008). Common to these views is the idea that the frequency of the structures we produce in a language at least partly determines what we incrementally predict while engaged in a comprehension task.

Once an expectation has been generated, it is either met or not met. When an expectation is met, generally no disruptions occur; but when the expectation is dashed, we often experience processing difficulty. Although the role of predictive processing is well-established in the psycholinguistic literature, in sentence comprehension research, the work of Hale (2001) and Levy (2008) has considerably sharpened our understanding of what it means to have a syntactic expectation, and, more importantly, how to quantify the effect of a dashed expectation. The key idea here is that building a rarer syntactic structure, with which comprehenders have less experience, is more difficult than building a relatively more frequent structure. As a simple example, in English, reading the relative pronoun in a sentence like *The man who ...* raises an expectation for a subject relative because it is more frequent; if the sentence continues with an object relative (e.g., *The man who the woman ...*), a slowdown is predicted (relative to an appropriate baseline sentence). Staub (2010) has shown that this prediction is correct. The surprisal metric introduced by Hale (2001) thus formalizes the well-established idea of reanalysis that dates back to the early days of psycholinguistics (Frazier, 1979). Surprisal assumes a ranked-parallel parser that ranks the available parses by their conditional probabilities. When the input favors a parse that has not been ranked highest, a re-ranking occurs. This re-ranking due to dashed expectations is computationally costly and formalizes the idea of reanalysis.

Over the last few years, evidence has started to accumulate in favor of such an expectation-based account of sentence comprehension. For example, Levy et al. (2013) have recently shown, using Russian relative clauses, that dashed expectations lead to slowdowns in reading; in other words, building a rarer structure is more difficult than building a more frequent structure. However, it is

currently unknown how generally applicable the expectation-based account is cross-linguistically; as Pickering and Van Gompel (2006) have documented, even robust theories of sentence processing have floundered in the face of a cross-linguistic investigation. It is particularly important to stress-test the expectation-based account given that its predictions sometimes go directly against another, well-established class of explanation, the memory-based accounts of sentence comprehension. Such models attribute processing difficulty to limitations of memory resources (Miller and Chomsky, 1963; Frazier and Fodor, 1978; Frazier, 1979; Clifton and Frazier, 1989; Just and Carpenter, 1992; Gibson, 1998, 2000; Lewis and Vasishth, 2005; Lewis et al., 2006). A prominent example for this class of account is the Dependency Locality Theory (DLT) (Gibson, 1998, 2000). DLT assumes that processing difficulty depends on so-called integration cost and storage cost. Structural integration cost is defined as a linearly increasing function of the number of new discourse referents that intervene between the constituent that is currently being processed and the constituent(s) with which a syntactic dependency has to be built. Storage cost assumes that processing difficulty linearly increases as a function of the number of predicted heads.

Chinese relative clauses (RCs) are a critical test of the opposing predictions of the expectation-based account and memory-based accounts like the DLT. Due to the syntactic properties of Chinese, memory-based accounts and expectation-based accounts make diametrically opposed predictions about the processing difficulty associated with subject relative clauses (SRs) compared to object relative clauses (ORs). Because of the great importance of Chinese relative clauses in unpacking the relative contributions of expectations and memory cost (among other theoretical explanations), they have drawn considerable attention in recent psycholinguistic research. In contrast to other languages such as Japanese and Korean with pre-nominal relative clauses, which display a consistent SR advantage in all published studies (Japanese: Miyamoto and Nakamura, 2003; Ueno and Garnsey, 2008; Korean: Kwon et al., 2006, 2010; Yun et al., 2010; Kwon et al., 2013), the evidence from Chinese relatives is not conclusive. As discussed below, the conflicting results are likely due to local ambiguities in the experimental materials.

(4.1) a. *Subject-modifying SR*

		邀请	男孩	的		女孩	认识	老师。
[$_{RC}$	t_i	Yaoqing	nanhai	de	]	nühai$_i$	renshi	laoshi.
		invite	boy	REL		girl	know	teacher

The girl who invites the boy knows the teacher.

b. *Subject-modifying OR*

	男孩	邀请		的		女孩	认识	老师。
[$_{RC}$	Nanhai	yaoqing	t_i	de	]	nühai$_i$	renshi	laoshi.
	boy	invite		REL		girl	know	teacher

The girl who the boy invites knows the teacher.

c. *Object-modifying SR*

老师	认识			邀请	男孩	的		女孩。
Laoshi	renshi	[$_{RC}$	t_i	yaoqing	nanhai	de	]	nühai$_i$.
teacher	know			invite	boy	REL		girl

The teacher knows the girl who invites the boy.

d. *Object-modifying OR*

老师	认识		男孩	邀请		的		女孩。
Laoshi	renshi	[$_{RC}$	nanhai	yaoqing	t_i	de	]	nühai$_i$.
teacher	know		boy	invite		REL		girl

The teacher knows the girl who the boy invites.

Chinese relative clauses (4.1) are pre-nominal, i.e., the head noun (here, *nühai*, 'girl') appears after the relative clause. This has the interesting consequence that the distance between head noun and the gap inside the relative clause (or, equivalently, the relative clause verb) is greater in SRs than in ORs. By contrast, in English, the distance between the gap (or verb) and head noun is greater in ORs than SRs. As a result of this reversal in gap-head distance in Chinese, memory-based accounts like the DLT, which index processing cost in terms of the distance between co-dependent elements, predict longer reading times at the head noun in SRs than ORs. In other words, memory-based accounts predict an object relative advantage.[1]

By contrast, the expectation-based account outlined above, which assumes that rare structures are harder to process, predicts that in ORs, *longer* reading times should be seen compared to SRs (i.e., a subject relative advantage); this is because, similar to English, SRs are more frequent than object relatives in Chinese (Hsiao and Gibson, 2003; Wu, 2009; Wu et al., 2010; Wu, 2011; Chen et al., 2012a; Vasishth et al., 2013). The expectation-based account would predict an SR advantage as soon as the relative clause is built; and this SR advantage could plausibly spill over to the head noun and beyond as well. But, crucially, the predicted onset of the SR advantage depends on the point at which the structural properties of the experimental materials allow the comprehender to predict a relative clause. In the materials used in previous experiments, several local ambiguities might have confounded the results (for a discussion of local ambiguities in Chinese RCs also see Lin and Bever, 2006, 2011; Qiao et al., 2012; Vasishth et al., 2013). Indeed, Hsiao and MacDonald (2013) and Hsiao et al. (2014) have argued that the results of previous studies on Chinese RCs can be largely explained by the local ambiguities in the stimuli. Thus, although the theoretically interesting aspect of Chinese relative clauses lies in the diametrically opposed predictions of the expectation-based account versus the memory-based accounts, it is vital to bring the local ambiguities under experimental control before we can investigate these opposing predictions. In the present paper, we investigate the predictions of these two classes of account using an experiment design where we strongly constrain the local ambiguities that have confounded previous studies.

We will first give an overview of the various ambiguities present in Chinese RCs and then describe how they might have confounded previous studies and how we deal with these ambiguities in our experimental materials.

[1]Note that these predictions do not depend on whether we assume the psychological reality of empty categories (the gap inside the relative clause) or not (Pickering and Barry, 1991; Traxler and Pickering, 1996). For Chinese relative clauses the same predictions hold if we assume that the head noun is directly associated with the verb — not via a two-step mechanism involving a gap — because the distance between the head noun and the relative clause verb is shorter in object relatives compared to subject relatives.

4.1.1 Local ambiguities in Chinese relative clauses

Figure 4.1 provides an overview of how the various local ambiguities differently affect subject- and object-modifying SRs and ORs.[2]

Local ambiguity 1 in Figure 4.1 shows an alternative parse that is due to the lexical ambiguity of the relativizer *de*: when *de* directly follows a noun phrase, it can also be read as a possessive marker. In this case, the NP preceding *de* is interpreted as a possessor of a following NP. Therefore, in SRs, when reading the RC object followed by the RC head noun, the reader can interpret them as a complex NP (resulting in 'the boy's girl' in Example 4.1a and 4.1c). This parse becomes impossible at the main clause verb in subject-modifying SRs and at the end of the clause in object-modifying SRs.

Local ambiguity 2 in Figure 4.1 shows alternative parses that involve null subjects (for a discussion of null subjects in Chinese see Huang, 1989). In SRs, a dropped subject (*pro*) whose reference is determined by the context of the utterance is postulated before the RC verb. Thus, the RC verb together with the RC object can be read as a main clause in subject-modifications (resulting in '*pro* invites the boy' in Example 4.1a) and as sentential object in object-modifications (resulting in 'The teacher knows that *pro* invites the boy' in Example 4.1c).[3] These parses become impossible upon reaching the relativizer *de*. (If the relativizer is interpreted as a possessive marker, the parse becomes impossible at the main clause verb in subject-modifications and the end of the clause in object-modifications.) Note that in object-modifications, *Local ambiguity 2* is only present if the main clause verb allows a sentential object.

Local ambiguity 3 in Figure 4.1 shows an alternative possible parse in ORs. The RC subject is interpreted as the main clause subject (in subject-modifications) or the main clause object (in object-modifications). In object-modifying ORs, this main clause reading becomes impossible already at the following word, the RC verb. In subject-modifying ORs, the RC verb can still be integrated into this parse as main clause verb. This parse becomes impossible at the relativizer.

Local ambiguity 4 in Figure 4.1 shows how, in object-modifying ORs, the RC subject together with the RC verb can be interpreted as a sentential object (in case the main clause verb allows a sentential object). This parse becomes impossible at the relativizer.

Thus, it is clear that in order to fairly compare SRs and ORs, we need a syntactic configuration that allows us to eliminate the local ambiguities described above such that they cannot confound the comparison.

[2]Note that the list of ambiguities in Chinese RCs can never be exhaustive due to the recursive nature of syntax. E.g., all sentential nodes could be embedded below another sentential node as sentential subject or sentential object. The local ambiguities which result from the recursive application of a production rule on a sub-tree where the same local ambiguity is already present are therefore not discussed here. Crucially for the experimental design described below, the elimination of an alternative parse automatically eliminates all alternative parses that result from this parse via recursion.

[3]In the concrete Example given in 4.1c, this parse is not possible since the verb *renshi* ('know' or 'recognize') does not take sentential complements. If *renshi* is replaced by a verb like *zhidao* ('know' or 'acknowledge'), which allows sentential complements, this local ambiguity is present.

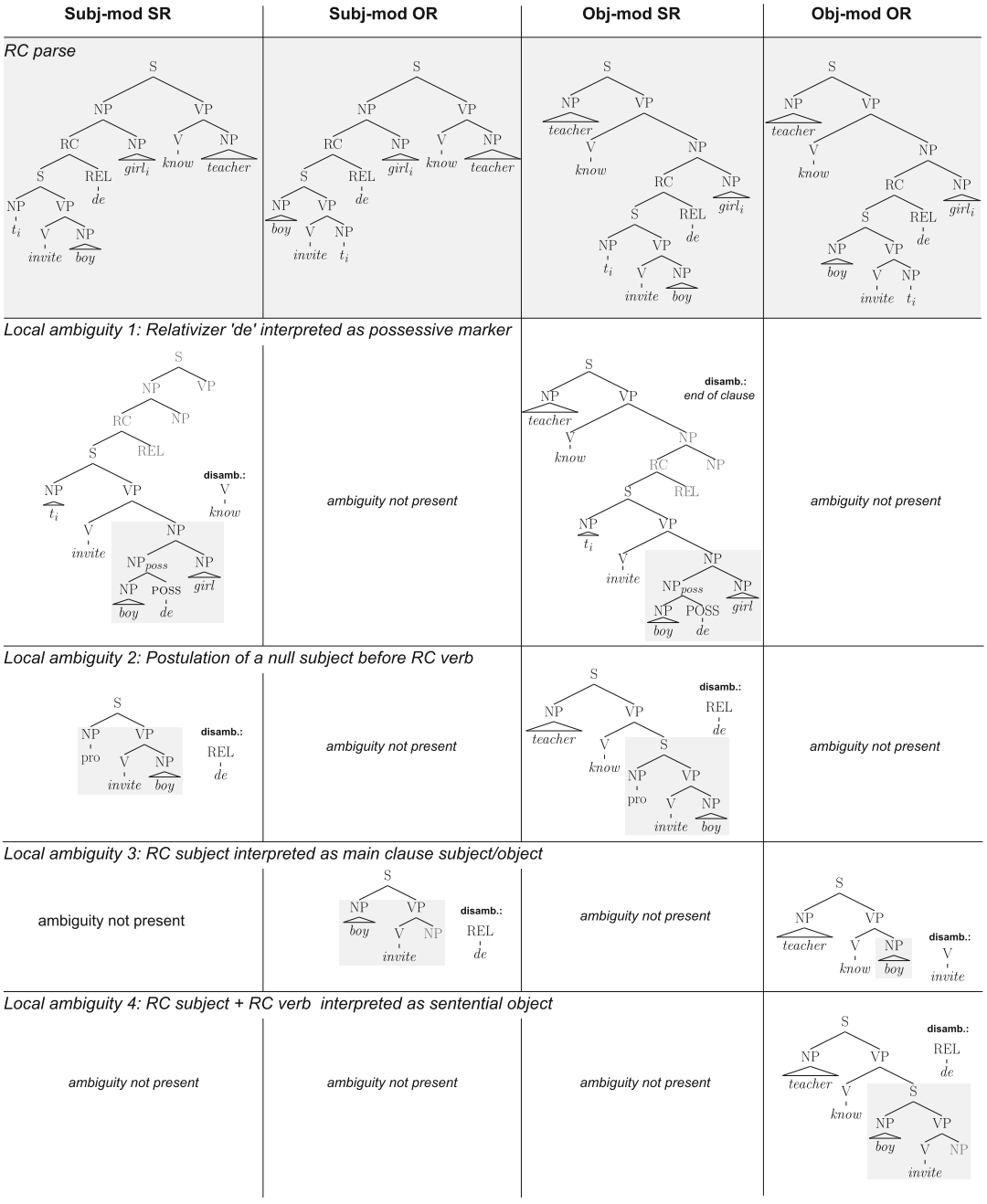

FIGURE 4.1: Local syntactic ambiguities in Chinese relative clauses. The first line shows the relative clause parses for the subject- and object-modifying SRs/ORs presented in Example 4.1. Lines 2-5 show the alternative parses up to the disambiguating word labeled as 'disamb'. The mis-analysed part of the tree that differs from the correct RC parse is highlighted with gray shading. Predicted nodes whose head has not been encountered yet are depicted in gray color. For the Chinese wording see Example (4.1).

4.1.2 Previous work on Chinese relative clauses

Research on Chinese relative clause processing has dealt with the local ambiguities discussed in the last section in different ways. Earlier studies of Chinese relative clause processing (e.g., Hsiao and Gibson, 2003; Lin and Bever, 2006) directly presented Chinese relative clauses such as those in (4.1) in isolated sentences without taking confounding ambiguities into account. Hsiao and Gibson (2003) conducted a 2×2 self-paced reading study in which they manipulated RC type (SR vs. OR) and number of embeddings (single vs. double). In single embeddings, the authors found faster reading times in ORs compared to SRs at the relative clause region, namely the region containing the RC verb and the RC object in SRs, and the region containing the RC subject and the RC verb in ORs. At the head noun, no effect was found. In double embeddings, the summed up reading time of the region containing the two relative clauses up to the second relativizer *de* was shorter in ORs than in SRs. The authors interpreted this OR advantage as evidence for a memory-based account and explained their results in terms of the Dependency Locality Theory (Gibson, 1998, 2000). Note that, as pointed out by Lin and Bever (2006), the differences between SRs and ORs found in the pre-relativizer regions can be explained by the fact that the NP+V sequence of an OR, but not the V+NP sequence of an SR, is actually the canonical order of a main clause (Chinese is an SVO language). In other words, *Local ambiguity 3* presented in column 2 of Figure 4.1 might have confounded Hsiao and Gibson (2003)'s results.

Lin and Bever (2006) carried out two studies which suggest that SRs are processed faster than ORs; however, the SR advantage was only found in cases where the object of the main clause was modified by a relative clause. Subject-modifying relative clauses did not show any difference between RC types. Gibson and Wu (2013) pointed out that the Lin and Bever (2006) finding can be explained by another local ambiguity confounding the materials, namely that the initial NP of an OR may have been mistaken as the object of the matrix clause, as can be seen in (4.1d) (*Local ambiguity 3*, column 4 of Figure 4.1).

Subsequent studies report processing facilitation for ORs whereas others report an SR advantage. An SR advantage has been reported by Wu (2009), Vasishth et al. (2013) and Liu et al. (2011). By contrast, several self-paced reading studies (e.g., Chen and Ning, 2008; Chen et al., 2008; Lin and Garnsey, 2011) as well as experiments using Maze tasks (e.g., Qiao et al., 2012) report an OR advantage. Wu et al. (2012) show that the subject relative advantage is further modulated by the animacy configuration of the head noun and the embedded noun. They found an SR advantage only in case the subject was inanimate and the object was animate. In the reverse animacy configuration, the SR advantage disappeared. Contrary to these results, Zhang and Jiang (2010) report an ERP study in which they found that SRs were harder to process than ORs in subject-modifying relatives while in object-modifying relative clauses, a processing facilitation was observed in SRs. In none of the studies mentioned so far has the issue of local ambiguity been taken into account.

One attempt to overcome the local ambiguity issue was undertaken by Gibson and Wu (2013), who examined SRs vs. ORs with disambiguating preceding context; they investigated only subject-modifying RCs. In their design, sentences similar to the materials of Hsiao and Gibson (2003) were preceded by a context that introduced the action described in the relative clause.

The target sentence itself was the answer to a question that required the reader to identify either the agent or the patient of the RC action. The assumption was that the presence of such a context ensures that the reader expects either an SR or an OR and therefore will not be garden pathed. At the relativizer *de*, Gibson and Wu (2013) found slower self-paced reading times in SRs than in ORs. This comparison reached significance in the by-participant analysis but not in the by-item analysis. At the head noun, SRs were read significantly slower than ORs. No other comparisons reached significance. In contrast to Hsiao and Gibson (2003), who did not find an effect at the head noun in single embeddings, Gibson and Wu (2013)'s results are consistent with the predictions of DLT's structural integration cost. Moreover, at the relative clause region preceding the relativizer, Gibson and Wu (2013) did not find any statistically significant effects. This absence of an effect is consistent with the DLT storage cost metric under the assumption that the reader is aware of an upcoming relative clause from the very outset due to the preceding context, so that the number of predicted heads is identical in SRs vs. ORs.

Vasishth et al. (2013) also replicated the OR advantage at the head noun that Gibson and Wu found. However, the effect was already significant at the relativizer preceding the head noun, which is not predicted by either DLT integration or storage cost metrics. In recent work, Lin (2014) has argued that the OR advantage observed in Gibson and Wu (2013)'s materials might reflect a syntactic priming from the context rather than a lower DLT integration cost in ORs. Using materials similar to Gibson and Wu (2013)'s, Lin (2014) additionally manipulated the different thematic orders in the context and found a stronger OR advantage when the thematic order in the preceding context was similar to the one of an OR.[4] Thus, the Gibson and Wu (2013) context manipulation may have introduced a new confound, thematic order of the context and target sentences, that rendered ORs easier to process than SRs.

As the above summary shows, the evidence about the processing of Chinese RCs is inconsistent across studies, and this is due at least in part to various confounds introduced by local syntactic ambiguities in the stimuli. Importantly, these ambiguities differ between SRs and ORs, as well as between subject-modifications and object-modifications.

[4]There is also an inconsistency in the predictions of the DLT as presented in Gibson (2000) and the DLT predictions derived for the experimental materials used in Gibson and Wu (2013). The structural integration cost metrics of DLT are defined as a function of the number of *new* discourse referents intervening between the two elements of the dependency that is being built. The following is the definition in (Gibson, 2000, p. 125) (the emphasis is ours):

> The structural integration cost associated with connecting the syntactic structure for a newly input head h_2 to a projection of a head h_1 that is part of the current structure for the input is dependent on the complexity of the computations that took place between h_2 and h_1. For simplicity, it is assumed that one EU is consumed for each *new* discourse referent in the intervening region.

However, in Gibson and Wu (2013)'s materials, the discourse referents in question are already introduced in the context sentence and hence are no new discourse referents anymore when processing the relative clause. Thus, in fact, for the Gibson and Wu materials the DLT should predict no processing difference at the head noun in SRs vs. ORs.

4.1.3 Using syntactic context to eliminate local ambiguities in Chinese relative clauses

We present an experimental design which leads the reader to strongly predict that a relative clause is coming up. Unlike Gibson and Wu (2013), we eliminate the local ambiguities not by providing a biasing context but rather by creating a syntactic configuration in which the comprehender is highly likely to predict a relative clause. Several previous studies have used structural cues to generate a prediction for an upcoming RC parse in Chinese. For instance, Hsu et al. (2005) showed that the presence of the OR marker *suo* led to a higher proportion of OR continuations in a sentence completion task and to shorter reading times at the relativizer and the head noun in self-paced reading. Other studies have used the semantic clash caused by local classifier-noun mismatches as an indicator of embedded clauses (Yoshida et al., 2004; Hsu et al., 2005, 2006; Wu et al., 2014). Note, however, that while a classifier-noun mismatch indicates RC boundaries in Japanese and Korean (Yoshida and Yoon, 2014), it alone may not be strong enough to predict an RC in Mandarin (Hsu et al., 2005; cf. Wu et al., 2014). Wu and colleagues also used the passive marker *bei* at the onset of the relative clause to lead comprehenders to expect an upcoming passive SR. Participants indeed showed faster self-paced reading times in the presence of this passive marker (Wu, 2013; Wu et al., 2014). Crucially, the above-cited studies differ from the present study in that they did not examine the processing asymmetry between SRs and ORs but the predictive strength of certain syntactic markers. Note also that the use of *suo* and the classifier-noun mismatch only indicated the existence of an OR, not an SR (see Lin and Bever, 2011 for a comprehensive review); by contrast, the passive marker *bei* can only mark a passivized SR. Thus, the syntactic markers *suo* and *bei* allow us to predict either an SR or an OR but not an RC in general and hence are not sufficient for our purposes.

We created a 2×2 factorial design that, across experimental conditions, leads readers to strongly predict a relative clause from the first word of the relative clause on. In addition to the manipulation of RC type, we included the manipulation of modification type for two reasons. First, the conflicting pattern reported in the literature might reflect different RC type preferences depending on modification type. For example, Hsiao and Gibson (2003) and Gibson and Wu (2013), who both report an OR advantage, only tested subject-modifications. Lin and Bever (2006) report an SR advantage, but this effect reached significance only in object-modifications. Second, for our materials, the different accounts not only differ in their predictions with respect to RC type but also with respect to modification type, as will be discussed below.

The experimental sentences consisted of a main clause whose subject or object was modified by an SR or an OR. This main clause was followed by another coordinate main clause as the spillover material; see (4.2) for example items and Figure 4.2 for the underlying syntactic structures.

(4.2) a. *Subject-modifying SR*

那个			上个月	邀请了	男孩	几次	的		女孩
Na-ge	[$_{RC}$	t_i	shanggeyue	yaoqing-le	nanhai	ji-ci	de	]	nühai$_i$
DET-CL			last month	invite-ASP	boy	several-CL	REL		girl
认识	王老师		因为	上过	她的	课。			
renshi	Wang laoshi		yinwei	shang-guo	ta-de	ke.			
know	teacher Wang		because	attend-ASP	her	class			

The girl who invited the boy several times last month knows teacher Wang because [she] has attended her class.

b. *Subject-modifying OR*

那个		上个月	男孩	邀请了		几次	的		女孩
Na-ge	[$_{RC}$	shanggeyue	nanhai	yaoqing-le	t_i	ji-ci	de	]	nühai$_i$
DET-CL		last month	boy	invite-ASP		several-CL	REL		girl
认识	王老师	因为	上过	她的	课。				
renshi	Wang laoshi	yinwei	shang-guo	ta-de	ke.				
know	teacher Wang	because	attend-ASP	her	class				

The girl who the boy invited several times last month knows teacher Wang because [she] has attended her class.

c. *Object-modifying SR*

王老师		认识	那个			上个月	邀请了	男孩
Wang laoshi		renshi	na-ge	[$_{RC}$	t_i	shanggeyue	yaoqing-le	nanhai
teacher Wang		know	DET-CL			last month	invite-ASP	boy
几次	的		女孩	因为	教过	她的	课。	
ji-ci	de	]	nühai$_i$	yinwei	jiao-guo	ta-de	ke.	
several-CL	REL		girl	because	teach-ASP	her	class	

Teacher Wang knows the girl who invited the boy several times last month because [she] has taught her class.

d. *Object-modifying OR*

王老师		认识	那个		上个月	男孩	邀请了	
Wang laoshi		renshi	na-ge	[$_{RC}$	shanggeyue	nanhai	yaoqing-le	t_i
teacher Wang		know	DET-CL		last month	boy	invite-ASP	
几次	的		女孩	因为	教过	她的	课。	
ji-ci	de	]	nühai$_i$	yinwei	jiao-guo	ta-de	ke.	
several-CL	REL		girl	because	teach-ASP	her	class	

Teacher Wang knows the girl who the boy invited several times last month because [she] has taught her class.

First, we ensured that the relativizer *de* cannot be interpreted as a genitive marker (*Local ambiguity 1* in Figure 4.1); this was done by inserting a frequency phrase (FreqP, *ji-ci* 'several times' in 4.2) consisting of a numeral adjective and a verbal classifier (Cl) before the relativizer. This FreqP attaches to the relative clause VP (see Figure 4.2).

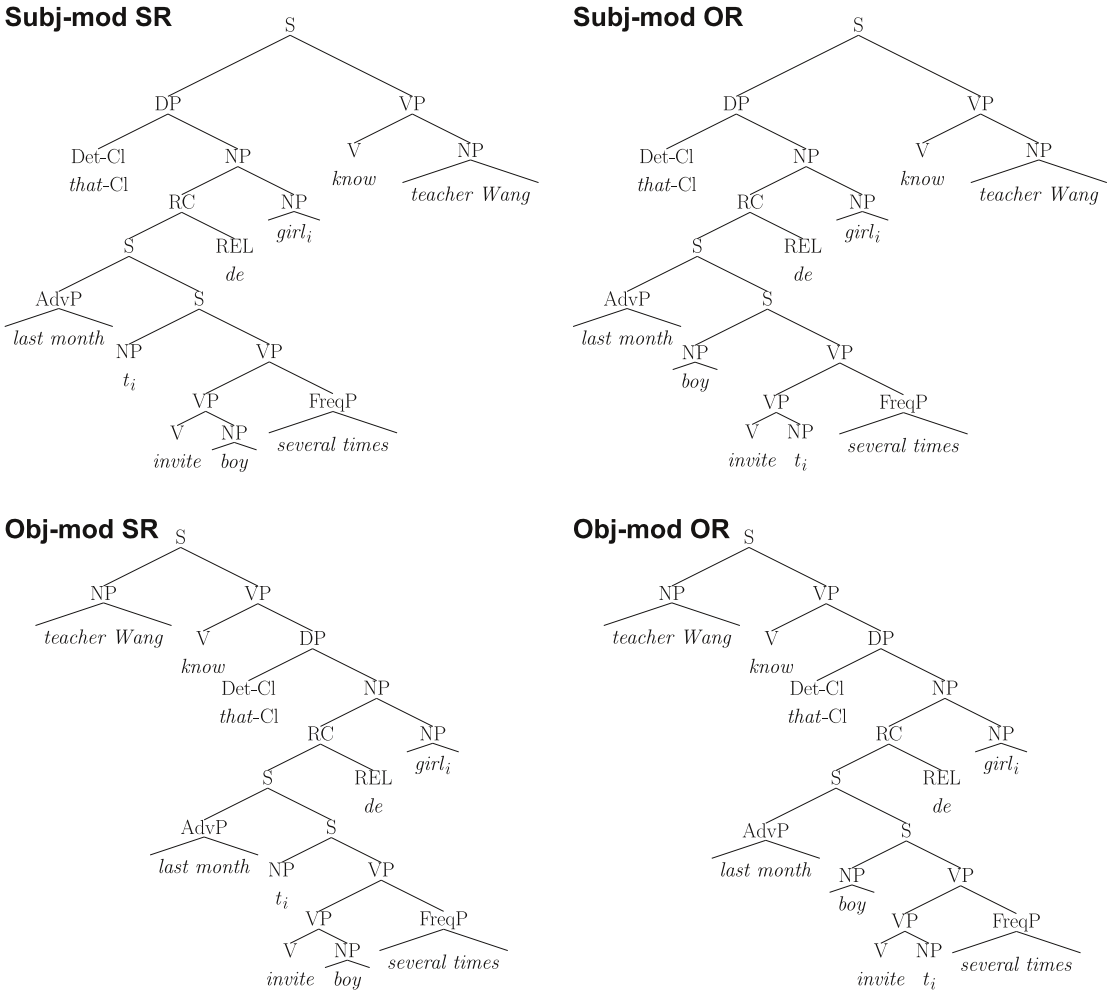

FIGURE 4.2: Syntactic structure of the experimental materials. For the Chinese wording see Example 4.2.

Second, we inserted a determiner (Det) (*zhe* 'this' or *na* 'that') followed by a nominal classifier (Cl) at the onset of the relative clause.[5] In Chinese, a Det+Cl sequence predicts an NP. [6] Therefore, the head noun is already predicted before the onset of the RC. This predicted NP, which will become the RC head, eliminates *Local ambiguity 2* (see Figure 4.1): it ensures that in subject- and object-modifying SRs, a *pro* subject is impossible since the subject position of the clause is already occupied by Det+Cl and the predicted NP.

Third, we inserted an adverbial phrase at the onset of the relative clause, i.e., between the Det+Cl sequence and the RC verb (in SRs) or the RC subject (in ORs). Together with the

[5]Note that the disambiguating character of the Det+Cl sequence in our materials only makes use of the general syntactic prediction that a classifier attaches to an NP; unlike Hsu et al. (2006) or Wu et al. (2014), our reasoning does not depend on any semantic match/mismatch of the classifier and a following noun.

[6]In case this NP is modified by a relative clause, the Det+Cl can be either located between the relative clause and the NP that is modified (i.e., the head of the RC), or it can precede the relative clause. In our materials, the Det+Cl sequence always precedes the relative clause.

inserted Det+Cl, the AdvP ensures that in ORs, the RC subject cannot be interpreted as a main clause subject/object (*Local ambiguity 3* in Figure 4.1). Moreover, it ensures that in object-modifying ORs, the RC subject and RC verb cannot be interpreted as a sentential object (*Local ambiguity 4* in Figure 4.1). This is because the AdvP predicts a clausal node to which it can attach (see Figure 4.2). It cannot directly attach to the main clause because it appears between the Det+Cl and the NP predicted by the Det+Cl combination. Hence, the only possible phrase for the adverb to attach to is an RC that modifies the predicted NP.

Although the local ambiguities presented in Figure 4.1 are eliminated in the present design, one new ambiguity has been introduced by the insertion of Det+Cl. The NP predicted by Det+Cl could be covertly realized as an elided NP, given a suitable context. In SRs, a parse structurally similar to *Local ambiguity 2* where the [$_{\text{NP}}$ pro] constituent is replaced by [$_{\text{DP}}$ Det Cl [$_{\text{NP}}$ ~~NP~~]] is still possible (*Na-ge ~~NP~~ shanggeyue yaoqing-le nanhai ji-ci* 'That one invited the boy several times last month'). This parse becomes impossible at the relativizer. In contrast to *pro*, Det+Cl+~~NP~~ is also allowed in object positions. Therefore, in object-modifications, a parse like *Wang laoshi renshi na-ge ~~NP~~* 'Teacher Wang knows that one.' is possible until encountering the AdvP. However, the interpretation of Det+Cl as Det+Cl+~~NP~~ needs a highly constraining discourse context. Indeed, such a reading is highly infrequent, as we will show with a corpus search and a sentence completion task reported below.

In sum, the present design ensures that, in both object and subject-modifications, the most frequent structure that comprehenders expect is a relative clause. This avoids the potential garden path effects that might have confounded the results of earlier studies on Chinese RCs.

In order to empirically validate our assumptions about the experimental materials and derive quantitative predictions of the expectation-based account, we conducted two pre-tests: a corpus analysis and a sentence completion study. These are described next.

4.2 Pre-tests

4.2.1 Corpus analyses

We conducted corpus analyses in order to (i) empirically validate the assumption that the local ambiguities presented in Figure 4.1 are eliminated in the experimental materials and that the most frequent structure that occurs after a determiner+classifier+adverb sequence is a relative clause; and (ii) derive empirically grounded region-by-region predictions of the expectation-based account for the present materials. Predictions of the modification type and RC type factors can be derived from the probabilities of the respective structure at the point in the sentence, where this structure is built given the left context of the sentence. We carried out four corpus searches on the Chinese Treebank 7.0 (Xue et al., 2005), which consists of 51,447 fully parsed sentences (1,196,329 words) using the pattern matching tool Tregex (Levy and Andrew, 2006).

In order to verify that the local ambiguities presented in Figure 4.1 are eliminated, we extracted all tokens containing Det+Cl+Adv sequences to obtain counts for the different structural types

that follow this structure. Since the experimental items only used the determiner 'zhe' or 'na' (*this* or *that*, respectively), we restricted the determiner to these lexical items. The raw corpus counts are provided in Table C.1 in the appendix. The results suggest that the zhe/na+Cl+Adv combination in Chinese indeed predicts a relative clause: 98.6% of all tokens of zhe/na+Cl+Adv are followed by a relative clause. Only a single token of a Det+Cl+Adv chunk (1.4% of all tokens) is not followed by a relative clause. In this sentence, an elided NP is postulated between the Det+Cl and the adverb (i.e., the newly introduced ambiguity in the present materials). These results confirm our syntactic analysis that the insertion of the Det+Cl+Adv sequence indeed eliminates all the local ambiguities presented in Figure 4.1 and that the possibility of a configuration involving an elided NP is highly infrequent and thus unlikely to lead to a garden path later in the sentence. Here, we make the assumption, uncontroversial in psycholinguistics, that incremental parsing consists of a ranking of alternative parses using their relative frequencies, with a preference to predict the structure that has the highest frequency (Gibson, 2000; Pickering and Garrod, 2007). Thus, no non-RC parses leading to garden paths confound the comparison between SRs and ORs.

Next, in order to assess the region-by-region conditional probabilities associated with each experimental condition of the present materials, we conducted three more corpus analyses.

First, to estimate the conditional probability of a subject-modifying vs. an object-modifying relative clause given that the subject/object NP (i.e., the RC head) has been predicted by a preceding Det+Cl, we carried out a corpus search on structures following a Det+Cl chunk that either attached to a subject NP or to an object NP. When the Det+Cl modified a subject, it was followed by an RC (i.e., a subject-modifying RC) in 13.9% of the tokens. When Det+Cl modified an object, it was succeeded by an RC (i.e., an object-modifying RC) in 12.3% of the tokens (see Table C.3 in the appendix). Thus, after having read a Det+Cl, an RC parse is similarly likely to follow in subject- and object-modifying contexts.

Second, in order to estimate conditional probabilities of an SR or OR appearing after Det+Cl+Adv for subject and object-modifications separately, we further categorized all tokens of the very first corpus search according to whether Det+Cl modified the subject NP or the object NP (or another constituent) of the sentence (see Table C.2 in the appendix). In subject-modifications, 77.5% of the Det+Cl+Adv-tokens are followed by an SR and 20% by an OR.[7] In object-modifications, the Det+Cl+Adv chunk is followed by an SR in 85% of the tokens and by an OR in 15% of the tokens.[8] This analysis shows that SRs are predicted overwhelmingly more frequently than ORs in the present materials.

Third, we conducted a corpus analysis to derive the conditional probability of an overt RC head noun given the Det+Cl+Adv+OR/SR prefix. We therefore checked all tokens of SRs/ORs preceded by a Det+Cl+Adv sequence (i.e., the tokens obtained in the first corpus search) for whether they are headed or headless RCs. In all tokens, the RC head was overtly realized (headed RC).

In sum, the corpus analyses show that:

[7] In all of these ORs, the RC subject is covertly realized as a *pro*.

[8] 5% of these ORs have an overt RC subject and 10% are ORs with a *pro* as RC subject.

(i) The Det+Cl+Adv sequence in the experimental items leads readers to strongly predict an RC parse. This ensures that readers do not follow a non-RC parse leading them into garden path, i.e., the present design rules out that non-RC parses confound the comparison between SRs and ORs.

(ii) The conditional probability of an RC appearing in subject-modifying position is similar to the conditional probability of an RC appearing in object-modifying position.

(iii) The conditional probability of an SR appearing after Det+Cl+Adv is higher than the conditional probability of an OR.

(iv) The conditional probability of an overt RC head noun appearing after an RC that is preceded by a Det+Cl sequence and starts with an adverbial phrase is very high (1.0) across all conditions.

Mod. Type	**Prefix**	**Target**	**Prob(Target | Prefix)**
subj-mod	Det+Cl	RC	0.14
	Det+Cl+Adv	SR	0.78
	Det+Cl+Adv	OR	0.20
	Det+Cl+Adv+SR+*de*	overt RC head	1.00
	Det+Cl+Adv+OR+*de*	overt RC head	1.00
obj-mod	NP+V+Det+Cl	RC	0.12
	NP+V+Det+Cl+Adv	SR	0.85
	NP+V+Det+Cl+Adv	OR	0.15
	NP+V+Det+Cl+Adv+SR+*de*	overt RC head	1.00
	NP+V+Det+Cl+Adv+OR+*de*	overt RC head	1.00

TABLE 4.1: Summary of the corpus analyses based on the Chinese Treebank 7.0 (see Tables C.2-C.3 for the underlying corpus counts). Conditional probabilities (Target|Prefix) associated with the upcoming structure (Target) at the various regions (Prefix) of the experimental materials grouped by modification type.

4.2.2 Sentence completion task

Despite the clear evidence in the corpus investigation that the local ambiguities discussed earlier have been eliminated, it is nevertheless possible that the relatively small size of the corpus may have yielded incomplete information. We carried out a direct test of the possible completions at the critical regions in the sentence by conducting a sentence completion study. The main aim of the sentence completion study was to cross-methodologically validate the corpus findings with the exact materials to be used in the reading experiments. In particular, we wanted to further validate the experimental design by testing whether participants indeed predict a relative clause when reading the Det+Cl+Adv sequence in the experimental materials and are not garden pathed by other non-RC-parses. Moreover, we wanted to obtain sentence continuations from participants so that we could derive region-by-region predictions of the expectation-based account for the factors RC type and modification type.

In contrast to the corpus counts, these predictions will be based on the exact materials to be used in the reading experiments. We therefore conducted a sentence completion task with all experimental items. The 32 experimental items to be used in Experiment 2 (i.e., a superset of the items

used in Experiment 1) were presented with a truncation either after Det+Cl or after the adverb. This resulted in sentence fragments consisting of Det+Cl or Det+Cl+Adv in subject-modifying conditions and NP+VP+Det+Cl or NP+V+Det+Cl+Adv in object-modifying conditions. We included 78 filler sentences (the same sentences as in Experiment 2) that were pseudo-randomly truncated at different points. Twenty Mandarin speakers from Taiwan currently living in the US (12 participants) or in Germany (8 participants) participated in this experiment against payment of 18 USD or 13 EUR respectively. The sentence completion task was implemented as an online questionnaire, with the items being presented in Traditional Chinese characters and arranged according to a Latin Square in pseudo-randomized order, such that each experimental item was preceded by at least one filler sentence.

The sentence completion data provide us with a quantitative estimate of the conditional probability of an RC parse being adopted given the left context of the experimental items up to Det+Cl (i.e., the conditional probability of an RC parse modulated by modification type) and the conditional probability of an SR or OR parse being adopted given the left context of the experimental items up to Det+Cl+Adv. Moreover, participants' productions were analysed for whether the RC head noun was overtly or covertly produced. This provided us with an estimate of the conditional probability of an overt head noun appearing given a Det+Cl+Adv+SR/OR sequence.

We can compute the conditional probability at each region by taking all alternative parse completions into account. Thus, in our estimation procedure, we make the independently motivated assumption (Pickering and Garrod, 2007, p. 108) that the parser predicts a range of syntactic alternatives, each associated with a probability conditional on its prefix. For example, the conditional probability of an SR given left context Det+Cl+Adv is then the number of SRs produced divided by the total count of all constructions produced given the same left context. The sentence completion task therefore allows us to empirically estimate surprisal (Hale, 2001). An overview of the results of the sentence completion study is provided in Table 4.2.

4.2.2.1 Sentence completions after Det+Cl and NP+V+Det+Cl

A classification of all sentence completions produced when the sentence was truncated after the sequence Det+Cl (subject-modification) or NP+V+Det+Cl (object-modification) is provided in the appendix (see Table C.4). Participants continued the Det+Cl sequence with an RC modifying an upcoming subject in 23% of the trials. In object-modifying contexts, the NP+V+Det+Cl fragment was continued with an RC modifying an upcoming object in 43% of the trials. Thus, the conditional probability of an RC appearing is higher in an object-modifying syntactic environment compared to a subject-modifying environment.

4.2.2.2 Sentence completions after Det+Cl+Adv and NP+V+Det+Cl+Adv

A classification of all sentence completions produced when the sentence was truncated after the sequence Det+Cl+Adv (subject-modification) or NP+V+Det+Cl+Adv (object-modification) is provided in the appendix (see Table C.5). Participants completed the sentence fragment with an

RC that started with the Adv and that modified an upcoming NP attaching to the Det+Cl in 88% of all trials in subject-modifications, and in 98% of the trials in object-modifying conditions. In the trials in which no RC continuation was produced, participants either postulated an elided NP between the Det+Cl and the Adv (i.e., the newly introduced ambiguity in the present design) (10% in subject-modifications, none in object-modifications)[9] or assigned a different part-of-speech to the adverb: In three trials of each modification type, they interpreted the adverb as an NP. (In Chinese, the mapping between lexical items and the part-of-speech categories is not always one-to-one; Kwong and Tsou, 2003. The temporal adverbs in our materials can function as genitive NPs in case they modify a second NP to their right, similar to *today's*+NP in English). In sum, the sentence completion study replicated the corpus finding that all the local ambiguities presented in Figure 4.1 are indeed eliminated in the present materials. Low-probability continuations are a parse with an elided NP interposed between Det+Cl and Adv (the newly introduced ambiguity) or a parse with the adverb being read as a genitive NP. Thus, the most probable structure predicted upon encountering the Det+Cl+Adv sequence is a relative clause and the local ambiguities that might have confounded previous studies have been brought under experimental control.

In order to derive conditional probabilities associated with each RC type modulated by modification type, the RCs produced after the Det+Cl+Adv and the NP+V+Det+Cl+Adv sequence were classified manually into SRs (72% of all trials in subject-modifications, and 86% in object-modifications), ORs (14% in subject-modifications and 11% in object-modifications), possessive RCs (1% in subject-modifications and 1% in object-modifications) and adjunct RCs[10] (1% in subject-modifications and 0 tokens in object-modifications). Since the SR/OR structures used in the experimental items are canonical SRs/ORs, i.e., RCs with the canonical SVO word order and, in the case of ORs, with an overt RC subject, the SRs and ORs observed in participants' sentence completions were further categorized into canonical SRs/ORs, passive RCs, topicalized RCs, RCs with an adjectival predicate and RCs with a dropped RC subject.

Canonical SRs were produced in 68% and 82% of the trials in subject- and object-modifications, respectively. Canonical ORs were produced in 2% and 5% of the trials in subject- and object-modifications, respectively. This pattern confirms the main finding of the corpus analyses: The probability of an SR continuation after Det+Cl+Adv is higher than the probability of an OR continuation. Note that in this highly constraining context, the difference in conditional probabilities of SRs vs. ORs is even bigger than the general frequency difference between SRs and ORs reported in other studies.

In order to derive estimates for the probability of an overt RC head, as opposed to a covert NP, noun being produced after a subject/object-modifying Det+Cl+Adv+SR or Det+Cl+Adv+OR,

[9]The higher proportion of productions with an elided NP compared to the corpus counts (one token) might be explained by the fact that truncating the experimental sentences after Det+Cl+Adv (three open clause boundaries) might have introduced a bias towards an elided NP reading. The reason for this is that the insertion of a covert NP before the adverb leaves the reader with only one open clause (only the main clause VP to be produced in order to complete the sentence) in contrast to three predicted heads in case no elided NP is postulated (the RC, the main clause subject and the main clause VP). This bias towards a reduction of open clauses by postulating an elided NP might have been even strengthened by the fact that the filler items in the sentence completion task were all cut between clause boundaries.

[10]Adjunct RCs are also called gap-less RCs.

we further categorized the SR and OR productions by whether their head noun was overtly produced or realized as a covert NP (headless RC). This estimation procedure is necessary to compute the conditional probability associated with the head noun in the experimental items, because the headless NP constitutes an alternative parse to the one in the experimental materials. In subject-modifications, one headless SR (canonical SR) (1% of all subject-modifying SRs) and one headless OR (OR with dropped subject) (5% of all subject-modifying ORs) were produced. In object-modifying contexts, no headless RCs were produced.

Mod. Type	**Prefix**	**Target**	**Prob(Target | Prefix)**
subj-mod	Det+Cl	RC	0.23
	Det+Cl+Adv	SR	0.72
	Det+Cl+Adv	canonical SR	0.68
	Det+Cl+Adv	OR	0.14
	Det+Cl+Adv	canonical OR	0.02
	Det+Cl+Adv+SR+*de*	overt RC head	0.99
	Det+Cl+Adv+ canonical SR+*de*	overt RC head	0.99
	Det+Cl+Adv+OR+*de*	overt RC head	0.95
	Det+Cl+Adv+ canonical OR+*de*	overt RC head	1.00
obj-mod	NP+V+Det+Cl	RC	0.43
	NP+V+Det+Cl+Adv	SR	0.86
	NP+V+Det+Cl+Adv	canonical SR	0.82
	NP+V+Det+Cl+Adv	OR	0.11
	NP+V+Det+Cl+Adv	canonical OR	0.05
	NP+V+Det+Cl+Adv+SR+*de*	overt RC head	1.00
	NP+V+Det+Cl+Adv+ canonical SR+*de*	overt RC head	1.00
	NP+V+Det+Cl+Adv+OR+*de*	overt RC head	1.00
	NP+V+Det+Cl+Adv+canonical OR+*de*	overt RC head	1.00

TABLE 4.2: Summary of the results of the sentence completion experiment (see Tables C.5-C.4 in the appendix for the underlying raw counts). Conditional probabilities (Target|Prefix) associated with the upcoming structure (Target) at the various regions (Prefix) of the experimental materials grouped by modification type. Canonical RCs are RCs with the canonical SVO word order and an overt RC subject in ORs. The experimental items are canonical SRs/ORs.

In sum, the sentence completion data show that:

(i) Consistent with the corpus findings, the Det+Cl+Adv sequence strongly predicts an RC. This effectively eliminates the possibility that non-RC continuations are predicted. In other words, the present design eliminates the confound present in previous experimental comparisons between ORs and SRs.

(ii) The conditional probability of an SR continuation after Det+Cl+Adv is higher than the conditional probability of an OR. This is also consistent with the corpus analyses.

(iii) The conditional probability of an RC appearing in object-modifying position is higher than the conditional probability of an RC appearing in subject-modifying position. This finding differs from the results of the corpus analyses where subject- and object-modifications did not differ in their conditional probability.

(iv) The conditional probability of an RC-head being overtly produced was close to 1.0 across conditions. This replicates the pattern observed in the corpus.

In summary, the corpus analyses and the sentence completion data allow us to derive predictions of the expectation-based account for the relative processing ease associated with subject- and object-modifying SRs and ORs for each regions of the sentence. Corpus counts and sentence completion data provide similar estimates of the conditional probability associated with RC type for all regions but differ with respect to the conditional probability of modification type and its interaction with RC type.

4.3 Experiment 1 (self-paced reading)

Before discussing the experiment, we present the predictions of the two competing accounts by region of interest: the adverb at the onset of the relative clause, the relative clause region containing the V+N/N+V sequence, and the head noun.

4.3.1 Predictions

4.3.1.1 The expectation-based account

The expectation-based account assumes that the relative frequency of a certain structure influences processing time. Specifically, surprisal (Hale, 2001; Levy, 2008) predicts that the difficulty associated with the integration of a certain word is proportional to its negative log-probability given the preceding context. Higher surprisal values are thus predicted to be reflected in increased processing difficulty, i.e., longer self-paced reading times.

Table 4.3 gives an overview of the surprisal values associated with each experimental condition at the adverb, the RC region (the V+N sequence *yaoqing-le nanhai*, 'invited the boy' and N+V sequence *nanhai yaoqing-le*, 'the boy invited' in SRs and ORs respectively) and the RC head noun. These surprisal values are calculated from the conditional probabilities based on the sentence completion data presented in Table 4.2. We used the counts reflecting the probabilities of the subtypes of SRs and ORs that exactly match the structure of our experimental materials (canonical SRs and ORs) as a basis for the calculation of the surprisal values.

	Adv	**V+N/N+V**	**RC head**
$SR_{Subj\text{-}mod}$	1.47	0.39	0.01
$OR_{Subj\text{-}mod}$	1.47	3.91	0.00
$SR_{Obj\text{-}mod}$	0.84	0.20	0.00
$OR_{Obj\text{-}mod}$	0.84	3.00	0.00

Table 4.3: Predicted surprisal values derived from sentence completion data associated with each experimental condition by region of interest. Surprisal is calculated as the negative natural logarithm of the conditional probabilities presented in Table 4.2. The estimates are based on those numbers that take into account only the exact syntactic structure of the experimental materials (canonical SRs/ORs).

From these surprisal values associated with each experimental condition, we derived predictions for the experimental factors (main effects of modification type and RC type, their interaction and

pairwise effects of RC type nested within modification type). The surprisal value associated with the main effect of modification type was calculated by subtracting the mean surprisal value of the subject-modifying conditions from the mean surprisal value of the object-modifying conditions. The surprisal value associated with the main effect of RC type was calculated in a similar manner. The predicted interaction was calculated by subtracting the mean of the surprisal value of subject-modifying ORs and object-modifying SRs from the mean surprisal of subject-modifying SRs and object-modifying ORs. The predicted effects of RC type in subject-modification and object-modification were calculated by subtracting the surprisal value associated with ORs from the surprisal value associated with SRs within each modification type separately. The same contrasts are applied in the statistical analyses of the reading times. A summary of the surprisal values associated with each of these effects is provided in Table 4.4.

	Adv	**V+N/N+V**	**RC head**
Main effect RC Type	*not applicable*	3.16	-0.01
Main effect Mod. Type	-0.63	-0.55	-0.01
RC Type× Mod. Type	*not applicable*	-0.36	0.01
RC Type [Subj-mod]	*not applicable*	3.52	-0.01
RC Type [Obj-mod]	*not applicable*	2.80	0.00

TABLE 4.4: Effects predicted by surprisal at the adverb, the RC-region (V+N/N+V) and the RC head noun. The predicted effects are calculated from the sentence completion-based surprisal values presented in Table 4.3. A positive sign associated with an effect of modification type or RC type means that subject-modifications or SRs are easier to process than object-modifications or ORs, respectively.

At the adverb, it becomes clear to the reader that an RC is being read. The sentence completion data show that conditional probability of an RC appearing in this position (i.e., after a Det+Cl) is higher in object-modifying contexts. This leads to a lower surprisal value in object-modifying conditions. Therefore, a main effect of modification type with faster reading times in object-modifications is predicted.

At the RC-region (V+N/N+V), the conditional probability of an SR is much higher than the conditional probability of an OR, i.e., surprisal is higher in ORs than in SRs. Thus, the expectation-based account predicts a relatively strong main effect of RC type with faster reading times in SRs. This SR advantage is predicted to be stronger in subject-modifications, i.e., a small interaction between RC type and modification type is predicted. Moreover, the higher surprisal values in subject-modifications compared to object-modifications predict a main effect of modification type (faster reading times in object-modifications).

At the RC head noun, no effect is predicted since across all conditions the head noun is highly expected leading to no difference in surprisal between RC types and between modification types.

Calculating surprisal based on the conditional probabilities from corpus counts rather than on the sentence-completion data would result in similar, but not completely identical predictions.[11] No effect of modification type at any region and no interaction between RC type and modification

[11] See Chen et al. (2012a) for an earlier attempt of using corpus frequency derived surprisal to account for the subject relative advantage in Chinese.

type would be predicted. We will focus on the surprisal predictions derived from sentence completion data because (i) they more closely match the experimental materials, (ii) they allow us to directly link data from the production experiment to comprehension data, and (iii) the absence of an effect of modification type and its interaction with RC type in the corpus is inconclusive given the relatively small size of the corpus. We believe that, due to its experimental setting, the sentence completion study is likely to have higher sensitivity than the corpus analyses.

4.3.1.2 The working-memory based account

We turn next to the predictions of the DLT memory cost metrics (Gibson, 2000), beginning with the DLT's storage cost metric. In the pre-head region, DLT's storage cost predicts that ORs and SRs are equally hard to process; this is because in the experimental materials it is clear that the sentences are relative clauses, and therefore an equal number of incomplete dependencies remain when the V+N/N+V region is processed (see page 6 of Hsiao and Gibson 2003). When reading the Det+Cl sequence one nominal head is predicted in both SRs and ORs. When reading the subsequent adverb, a relative clause, i.e., a VP and the relativizer *de* are predicted. When reading the next word (the RC verb in SRs and the RC subject in ORs) the RC object is predicted in SRs, while in ORs the RC verb (that was already predicted from the adverb on) is still predicted. In sum, at every word of the relative clause, an equal number of heads is predicted resulting in the same DLT storage cost predicted for SRs and ORs. The predictions of DLT storage cost are summarized in Table 4.5.

The storage cost metric also predicts that at the adverb, in the relative clause region and at the head noun, object-modifications should be easier to process than subject-modifications. This is because in object-modifications the main clause verb precedes the relative clause, whereas in subject-modifications the main clause verb follows the RC head, which leads to higher storage costs in the object-modifying conditions; see Table 4.5 for a detailed presentation of the storage cost for each region.

In terms of the DLT's integration cost metric (Gibson, 2000), at the head noun and possibly beyond (due to spillover), an OR advantage is predicted, because of the greater gap-head (or RC verb-head) distance in SRs compared to ORs. Integration cost is predicted to be higher in SRs due to a higher number of intervening discourse referents (the RC object and possibly the RC verb, depending on whether the verb or the gap preceding it is assumed to be retrieved). Similarly, the ACT-R based model of parsing (Lewis and Vasishth, 2005) predicts lower activation of the constituent that has to be retrieved (the gap or the relative clause verb) in SRs compared to ORs because the increased dependency length is predicted to lead to decay and interference.

Regarding the effect of modification type at the head noun, the integration cost metric predicts the opposite of the prediction of the storage cost account. Integration cost predicts that object-modification should be *harder* than subject-modification at the head noun, since in object-modifications there is one additional dependency to be built when reading the head noun: the dependency between the main clause verb and the main clause object (see Table 4.6). Similar to DLT integration cost, the ACT-R based model (Lewis and Vasishth, 2005) predicts a slowdown

	N	**V_{MC}**	**Det-Cl**	**Adv**	**V_{RC}/N_{RCsubj}**	**N_{RCobj}/V_{RC}**	**FreqP**	**de**	**N**
$SR_{Subj\text{-}mod}$									
Predicted heads	-	-	N,V_{MC}	N,V_{MC},V_{RC},de	N,V_{MC},de,N_{RCobj}	N,V_{MC},de	N,V_{MC},de	N,V_{MC}	V_{MC}
Storage cost			2	4	4	3	3	2	1
$OR_{Subj\text{-}mod}$									
Predicted heads	-	-	N,V_{MC}	N,V_{MC},V_{RC},de	N,V_{MC},V_{RC},de	N,V_{MC},de	N,V_{MC},de	N,V_{MC}	V_{MC}
Storage cost			2	4	4	3	3	2	1
$SR_{Obj\text{-}mod}$									
Predicted heads	V_{MC}	N	N	N,V_{RC},de	N,de,N_{RCobj}	N,de	N,de	N	-
Storage cost	1	1	1	3	3	2	2	1	0
$OR_{Obj\text{-}mod}$									
Predicted heads	V_{MC}	N	N	N,V_{RC}, *de*	N,V_{RC},de	N,de	N,de	N	-
Storage cost	1	1	1	3	3	2	2	1	0

TABLE 4.5: Predicted syntactic heads with DLT storage cost for each regions of interest by experimental condition.

	GAP (V_{RC})	**Det-Cl**	**V_{MC}**	***total***
$SR_{Subj\text{-}mod}$	2 (1)	3	-	5(4)
$OR_{Subj\text{-}mod}$	0	3	-	3
$SR_{Obj\text{-}mod}$	2 (1)	3	3	8(7)
$OR_{Obj\text{-}mod}$	0	3	3	6

TABLE 4.6: Predicted DLT structural integration cost at the RC head noun with respect to each constituent that needs to be retrieved. The numbers in parenthesis refer to the structural integration cost under the assumption that not the gap site inside the relative clause but rather the RC verb itself is retrieved.

in object-modifications. An overview of the constituents retrieved at the head noun and their predicted DLT structural integration cost is provided in Table 4.6.

4.3.2 Participants

49 college students at the Dalian University of Technology, China participated in the experiment as volunteers. All participants were native speakers of Mandarin Chinese.

4.3.3 Stimuli

The stimuli were as described above (see Example 4.2); there were 16 target items, which were adapted from the items in the Gibson and Wu (2013) study. In addition, 78 fillers consisting of various different syntactic structures (including eight relative clauses taken from Gibson and Wu (2013)'s filler items that were structurally different from the experimental sentences, which the authors generously provided to us). All stimuli and fillers are available from the first author.

4.3.4 Procedure

The experiment used the non-cumulative self-paced moving window method (Just et al., 1982). We presented stimulus items using Douglas Rohde's Linger software, version 2.88 (`http://tedlab.mit.edu/~dr/Linger/`). The target items were presented in Simplified Chinese characters (font size 18) using the standard Latin square design. The target items and fillers were pseudo-randomized for each subject such that each item was succeeded by at least one filler sentence. Each trial was followed by a *yes/no*-comprehension question designed to probe the correct understanding of both the relative clause and the main clause.

Each experimental session started with six practice trials which allowed participants to get used to the procedure. At the beginning of each trial, the participant saw a mask of hyphens that covered the upcoming sentence. Each hyphen represented the space-delimited words shown in Example 4.2. Participants were instructed to press the `f`-key in order to read the sentence;

successive presses of this key displayed the next word of a sentence and masked the preceding word. In order to respond to the comprehension questions, the `k`-key was used for a 'yes' response, and the `l`-key for a 'no' response. Reading times or RTs (in milliseconds) were taken as a measure of relative momentary processing difficulty.

4.3.5 Results

All analyses were carried out using linear mixed models using the `lme4` package version 1.0-6 in R (Bates et al., 2014). The analyses of reading times were carried out on log-transformed values in order to stabilize variance and to achieve approximately normal residuals (Box and Cox, 1964). Question-accuracies were analysed using generalized linear mixed models with a binomial link function. We defined two sets of contrasts. First, we tested for main effects of RC type (SRs coded as -0.5 and ORs as $+0.5$) and modification type (subject-modifications coded as -0.5 and object-modifications as $+0.5$) and an interaction between the two. Second, two sum contrasts nested within modification type (coded as in the first model) were defined so that SRs vs. ORs were each compared separately in the subject-modifying and object-modifying case. SRs were coded as -0.5 and ORs as $+0.5$. Residuals of linear mixed models were always checked to ensure that there were no serious deviations from the normality assumption. For linear mixed models, we took an absolute t-value equal to or above 2 to reach statistical significance at $\alpha = 0.05$. Throughout this paper, a full variance-covariance matrix was fit for participants and items (Gelman and Hill, 2007); whenever the model failed to converge or if the variance-covariance matrix of random effects was degenerate, we removed the relevant varying slopes for items or participants.

4.3.5.1 Comprehension questions

In the analyses of accuracy, no differences were found between SRs and ORs in either the subject-modifying or the object-modifying cases, and no difference was found between subject and object-modification. In subject-modifying cases, SRs had accuracy 85%, and ORs 85%; for object-modifying cases, the corresponding accuracies were 84% and 86%.

4.3.5.2 Reading times

The regions of interest were the adverb at the onset of the relative clause, the pre-head region following the Det+Cl+Adv sequence, the head noun, and the regions following it. We therefore analysed the adverb (Adv) (only testing for a main effect of modification type since conditions are identical across RC type), the V+N/N+V-region (SRs and ORs respectively), the region of the frequency phrase (FreqP) that followed the V+N/N+V region, the relativizer *de* before the head noun, the head noun, and the two regions following the head noun. The regions following the head noun differ lexically and syntactically across modification type (but not across RC type), therefore we only report models with pairwise comparisons of RC type nested within each level of modification type for the post-head regions.

Region	**Predictor**	***coef.***	***SE***	***t-value***
Adv	Modification Type	0.01	0.04	0.34
V+N/N+V	Modification Type	0.04	0.05	0.84
	RC Type	**0.11**	**0.04**	**2.51***
	Mod. Type × RC Type	0.00	0.03	-0.08
FreqP	Modification Type	0.03	0.04	0.74
	RC Type	-0.06	0.04	-1.57
	Mod. Type × RC Type	-0.01	0.04	-0.17
de	Modification Type	0.00	0.03	-0.05
	RC Type	0.02	0.02	0.97
	Mod. Type × RC Type	0.00	0.03	0.17
RC head	Modification Type	0.01	0.04	0.38
	RC Type	0.06	0.05	1.34
	Mod. Type × RC Type	-0.03	0.04	-0.87

TABLE 4.7: Experiment 1. Main effects of modification and RC type and their interaction by region of interest. The dependent variable is log-transformed reading time.

Region	**Predictor**	***coef.***	***SE***	***t-value***
V+N/N+V	RC Type [Subject mod]	**0.11**	**0.05**	**2.23***
	RC Type [Object mod]	0.11	0.06	1.83
FreqP	RC Type [Subject mod]	-0.06	0.06	-0.94
	RC Type [Object mod]	-0.07	0.05	-1.42
DE	RC Type [Subject mod]	0.02	0.03	0.52
	RC Type [Object mod]	0.03	0.03	0.76
RC head	RC Type [Subject mod]	0.10	0.06	1.56
	RC Type [Object mod]	0.03	0.06	0.45
head+1	RC Type [Subject mod]	0.09	0.05	1.56
	RC Type [Object mod]	0.04	0.05	0.86
head+2	RC Type [Subject mod]	**0.13**	**0.04**	**3.05***
	RC Type [Object mod]	0.03	0.04	0.89

TABLE 4.8: Experiment 1. The results of the RC type comparisons nested within each level of modification type by region of interest. The dependent variable is log-transformed reading time.

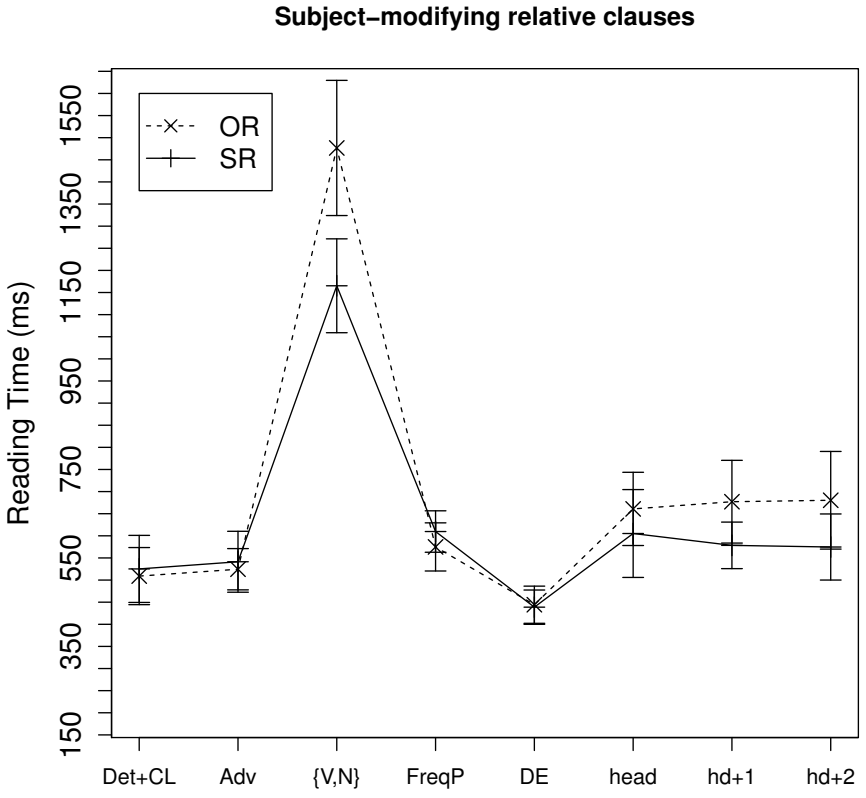

FIGURE 4.3: Experiment 1. Mean reading times of each region of interest in subject-modifying relatives, along with 95% confidence intervals.

The reading times for the different regions are summarized in Figures 4.3 and 4.4, and the results of the statistical analyses are shown in Tables 4.7 and 4.8.

At the adverb, no effect of modification type was observed. In the RC-region (V+N/N+V), a main effect of RC type with SRs being read faster than ORs was observed. In nested comparisons, this SR advantage reached significance in subject-modifications and was marginal in object-modifications. No effect of modification type and no interaction between modification and RC type were observed. No effects were observed at the frequency phrase following the (V+N/N+V) region, the relativizer, the RC head noun, and one word after the RC head. Two words after the head noun, we found faster reading times in SRs compared to ORs in subject-modifications, but not in object-modifications. No other effect reached significance in this region.

4.3.6 Discussion

The faster reading time in SRs in the pre-head region (V+N/N+V) is consistent with a sharpened expectation for the more frequently occurring SR. The data are consistent with the predicted

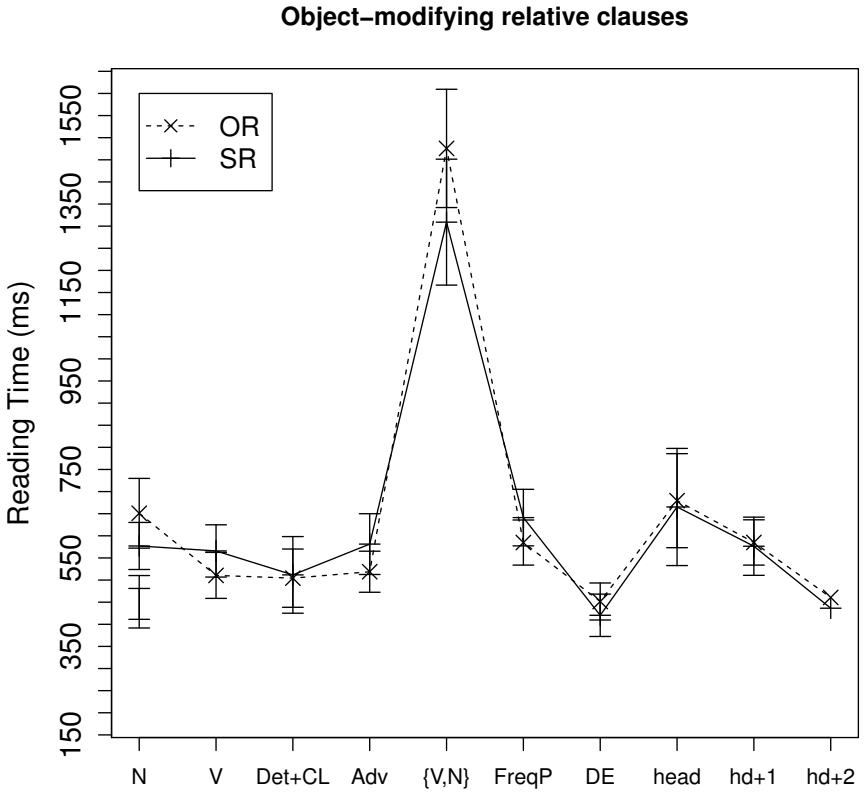

FIGURE 4.4: Experiment 1. Mean reading times of each region of interest in object-modifying relatives, along with 95% confidence intervals.

higher surprisal (Hale, 2001; Levy, 2008) cost in ORs compared to SRs in both subject and object-modifications. Surprisal can also explain the stronger SR advantage observed in subject-modifications compared to object-modifications since the difference in surprisal associated with SRs vs. ORs is smaller in object-modifications than in subject-modifications. Indeed, surprisal predicts a small interaction between RC type and modification type. In the reading data, however, this interaction is not significant. Surprisal might account for the absence of a significant interaction by the very small size of the predicted interaction compared to the size of the effect of RC type. The predictions of DLT storage cost (Gibson, 1998, 2000) (i.e., no effect of RC type), in contrast, are not consistent with the SR advantage observed at the V+N/N+V region.

At the head noun, we do not see evidence for a difference between processing difficulty associated with SRs vs. ORs. This absence of an effect is statistically inconclusive, but as predicted by surprisal and the DLT storage cost component. However, the observed null result cannot rule out retrieval-based memory accounts such as the DLT integration cost component (Gibson, 1998, 2000) or the ACT-R based model proposed by Lewis and Vasishth (2005), both of which predict an OR advantage at the head noun.

The SR advantage in subject-modifications in the spillover region after the head noun cannot be explained by either account under discussion. It is inconsistent with memory-based retrieval metrics like DLT integration cost (Gibson, 2000) and the ACT-R-based retrieval model by Lewis and Vasishth (2005), since they predict an effect in the opposite direction. It is also inconsistent with the storage cost component of DLT (Gibson, 2000) and surprisal (Hale, 2001; Levy, 2008) since both of them predict no effect of RC type at the head noun. Although one might interpret this effect as spillover from the RC region, in which case the effect would be consistent with surprisal, this explanation appears to be rather implausible given that the SR advantage had disappeared at the head noun and the following word. We suggest an explanation for this effect consistent with surprisal in the Discussion of Experiment 2.

The absence of an effect of modification type is inconclusive. We therefore do not discuss it any further here.

In sum, the evidence in Experiment 1 points in favor of the expectation-based account. However, given the conflicting results in the literature together with the SR advantage in subject-modifications two words after the head noun that cannot be explained by either account, and the null result with respect to the factor modification type, it is vital to attempt to replicate this result. Indeed, one of the biggest worries in psychology and linguistics today is the problem of non-replicable findings. Already Ronald Fisher, the founder of frequentist statistics, has advocated replication from the outset as the gold standard for science (Fisher, 1937, p. 16). Today, a growing number of methodologically and statistically concerned researchers emphasize the need for replication in experimental psychology (e.g., Asendorpf et al., 2013; Simmons et al., 2011).

4.4 Experiment 2 (eye-tracking)

This experiment extends Experiment 1 in two respects: first, the use of the eye-tracking method provides cross-methodological validation of the results in Experiment 1 and second, we doubled the number of items in order to increase statistical power. The same predictions hold as for Experiment 1.

4.4.1 Participants

This study was conducted at the eye-tracking lab of the Department of English at National Taiwan Normal University, Taipei. 49 students from that university participated in the experiment, each receiving payment of 250 NTD. All participants were native speakers of Mandarin and had normal or corrected-to-normal vision.

4.4.2 Design and materials

The experimental items, including comprehension questions, had the same design as in Experiment 1. We used all of the previous items and additionally created 16 new sets of items that had

the same structure. A minor difference between the old and the new items was that in the original items the sentence final materials sometimes varied across modification type, whereas in the new items they were identical across all four conditions. In addition, we made very minor lexical changes to the self-paced reading items to adapt them to the cultural environment of Taiwan, where this experiment was conducted. In contrast to the first experiment, all items were written in Traditional Chinese characters, the script officially used in Taiwan. In order to ensure that any difference in processing difficulty associated with the experimental manipulations is not due to a difference in acceptability between ORs and SRs that is particular to the present materials (i.e., due to the insertion of the Det+Cl+AdvP), we conducted a web-based acceptability rating study on the experimental materials to be used in the eye-tracking experiment with twenty native speakers of Mandarin. We did not find any evidence for a difference in acceptability between SRs and ORs induced by the present stimulus design.

4.4.3 Procedure

Eye movements were recorded with an SR Research Eyelink 1000 eyetracker at a sampling rate of 1000 Hz using a desktop-mount camera system (leveled, illuminator on the right) with a 35 mm lens to track participants' right eye. The participants' head was stabilized with a chin-rest. The camera and the presentation screen were installed on a table of 74 cm height, the chair on which participants were seated as well as the chin rest were adjustable in height. The camera-to-eye distance measured 52 cm, the eye-to-screen distance 62 cm. Stimuli were presented on an 19" monitor with a resolution of 1440×900 pixels; the stimuli were written in Traditional Chinese characters (font type SimSum, font size 20) in a black font on a light gray background.

The experiment was run using Experiment Builder software provided by SR Research. 32 experimental items, each with four conditions were presented in a Latin square. The same filler items as in Experiment 1 were used. Items were pseudo-randomized such that each relative clause item was preceded by at least one filler or one adverb item. Each trial was followed by a comprehension question to be answered with *yes* or *no* pressing a key on a response pad.

4.4.4 Results

Similar to Experiment 1, we used linear mixed effects models with a full variance-covariance matrix structure for participants and items when possible, applying two sets of contrasts: main effects and interaction and pairwise comparisons of RC type nested within modification type. Binomial dependent variables were analysed using generalized linear mixed effects models with a binomial link function.

4.4.4.1 Comprehension questions

The mean accuracy for subject-modifying SRs was 79%, for subject-modifying ORs 80%, object-modifying SRs 77%, and object-modifying ORs 79%. As in Experiment 1, none of the comparisons showed any statistically significant differences.

4.4.4.2 Eye movements

In eye-tracking data, the dependent measures can be partitioned into three broad classes: those that provide information about: (i) first-pass events; (ii) regression-related events (proportions of regressions and duration of regressive events); and (iii) second- and later pass events. Clifton et al. (2007) have shown in a large scale review on experiments relying on eye-tracking to measure sentence processing difficulty that it is still unclear in which eye-tracking measure to expect effects of syntactic processing. It is therefore common practice to report a wide range of dependent variables. However, since many of the eye-tracking measures are by definition correlated, this is statistically problematic since it increases the probability of a Type I error. We try to strike a balance between this statistical concern and the aim of providing a comprehensive picture of the data by selecting the most commonly reported dependent variable from each class of eye-tracking measures. As a representative first-pass measure, we report first-pass reading time (FPRT) (often also referred to as gaze duration), i.e., the sum of all fixations on a region before leaving it if and only if this region is entered progressively. As a measure for proportions of regressions, we analysed first-pass regression probability (FPRP), i.e., the proportion of trials in which a regression was initiated from a region when first entering this region, and regression-path duration (RPD), i.e., the sum of the time of all fixations starting from the first fixation on this region until leaving this region to the right including all fixations to the left of this region that fall into this time window, as a representative measure for regressive reading events. As a later-pass measure, we report total fixation times (TFT) which is defined as the sum of all fixations on a region. For all dependent variables, trials in which the region under consideration was skipped (i.e., in case the dependent fixation measure was 0) were excluded from analyses. Figures 4.5 and 4.6 provide a visual summary of relevant comparisons in TFT across all regions of the sentences up to two words after the head noun, along with 95% confidence intervals. The estimates, standard errors, and t-values of all linear mixed models for all regions of interest are shown in Tables 4.9 and 4.10.

At the adverb-region, a significant main effect of modification type was observed in both regression-based dependent variables (RPD and FPRP) with facilitated processing in object-modifications.

At the V+N/N+V region, we observed a significant main effect of RC type in FPRP, RPD, and TFT with faster reading times and less first-pass regressions in SRs compared to ORs. The pairwise comparisons nested within modification type showed that this main effect was driven by both, subject-modifying and object-modifying conditions, but the effect was stronger in subject-modifications: In subject-modifications, the SR advantage was significant in FPRP, RPD, and TFT. In object-modifications, the effect reached significance only in FPRP and was marginal in RPD. The main effect of modification type and the interaction between RC type and modification type did not reach significance in any measure.

At the frequency phrase (FreqP) that was inserted between the V+N/N+V region, none of the comparisons reached significance.

At the relativizer *de*, we found a marginal SR advantage in FPRP within subject-modifying conditions and a main effect of modification type in TFT (shorter fixations in object-modifications).

Region	DV	Modification Type			RC Type			Mod. Type×RC Type		
		coef.	*SE*	*t or z*	*coef.*	*SE*	*t or z*	*coef.*	*SE*	*t or z*
Adv	FPRT	0.02	0.03	0.67	*not applicable*			*not applicable*		
	FPRP	**-0.73**	**0.18**	**-3.95***	*not applicable*			*not applicable*		
	RPD	**-0.24**	**0.05**	**-4.99***	*not applicable*			*not applicable*		
	TFT	-0.02	0.04	-0.60	*not applicable*			*not applicable*		
V+N/N+V	FPRT	-0.07	0.04	-1.79	-0.05	0.04	-1.39	0.04	0.04	0.97
	FPRP	0.01	0.14	0.05	**0.76**	**0.22**	**3.51***	-0.15	0.15	-1.02
	RPD	**-0.10**	**0.04**	**-2.73***	**0.11**	**0.04**	**2.81***	-0.02	0.04	-0.51
	TFT	0.03	0.04	-0.91	**0.08**	**0.03**	**2.82***	-0.03	0.03	-0.92
FreqP	FPRT	-0.03	0.02	-1.12	-0.03	0.03	-1.13	0.00	0.02	0.03
	FPRP	0.27	0.16	1.67	0.01	0.18	0.06	0.12	0.16	0.75
	RPD	0.03	0.04	0.67	-0.02	0.04	-0.58	-0.01	0.04	-0.25
	TFT	-0.06	0.04	-1.32	-0.02	0.04	-0.53	-0.07	0.04	-1.74
de	FPRT	-0.03	0.03	-0.98	-0.01	0.03	-0.43	0.01	0.03	0.56
	FPRP	-0.04	0.20	-0.22	0.25	0.16	1.56	-0.29	0.22	-1.32
	RPD	0.04	0.05	0.81	0.01	0.05	0.22	-0.04	0.05	-0.87
	TFT	**-0.1**	**0.04**	**-2.34***	0.06	0.04	1.58	-0.04	0.04	-0.91
RC head	FPRT	-0.04	0.03	-1.63	0.01	0.02	0.23	**0.06**	**0.02**	**2.57***
	FPRP	-0.14	0.15	-0.95	0.20	0.16	1.28	-0.17	0.14	-1.23
	RPD	-0.07	0.04	-1.55	0.04	0.04	1.00	-0.01	0.04	-0.18
	TFT	**-0.18**	**0.03**	**-5.5***	0.07	0.04	1.81	-0.01	0.04	-0.22

TABLE 4.9: Experiment 2. Main effects of RC type and modification type and their interaction by region of interest for the dependent measures log-first-pass reading time, first-pass regression probability, log-regression path duration and log-total fixation time.

At the RC head noun, in FPRT, the interaction between modification type and RC type was significant. The pairwise comparisons revealed that this interaction was driven by a marginal SR advantage in object-modifications that was not present in subject-modifications. In TFT, we observed a main effect of modification type with shorter fixation times in object-modifications.

One word after the head noun, in subject-modifications only, a significant SR advantage was observed in TFT.

4.4.5 Discussion

The results of Experiment 2 confirmed the key findings of Experiment 1. We replicated the SR advantage at the RC region in both subject and object-modifications as well as the SR advantage in subject-modifications in the spillover region following the head noun. Moreover, we observed a main effect of modification type starting at the onset of the relative clause that was not present in Experiment 1.

The main effect of RC type (SR advantage) at the RC region (V+N/N+V) reached significance across all eye-tracking measures except for FPRT. This SR advantage is predicted by expectation-based accounts of parsing such as surprisal (Hale, 2001; Levy, 2008). As in the self-paced

Region	**DV**	**RC Type [subj-mod]**			**RC Type [obj-mod]**		
		coef.	*SE*	*t or z*	*coef.*	*SE*	*t or z*
V+N/N+V	FPRT	-0.08	0.05	-1.80	-0.01	0.04	-0.31
	FPRP	**0.91**	**0.28**	**3.27***	**0.61**	**0.24**	**2.51***
	RPD	**0.13**	**0.06**	**2.29***	0.09	0.05	1.81
	TFT	**0.10**	**0.04**	**2.62***	0.05	0.04	1.28
FreqP	FPRT	-0.03	0.04	-0.89	-0.03	0.03	-0.98
	FPRP	-0.11	0.26	-0.42	0.13	0.22	0.62
	RPD	-0.02	0.05	-0.29	-0.03	0.06	-0.57
	TFT	0.05	0.06	0.82	-0.09	0.06	-1.54
de	FPRT	-0.03	0.04	-0.72	0.00	0.04	0.09
	FPRP	0.55	0.28	1.96	-0.04	0.27	-0.15
	RPD	0.06	0.06	0.87	-0.03	0.08	-0.42
	TFT	0.10	0.05	1.81	0.02	0.06	0.42
RC head	FPRT	-0.05	0.03	-1.60	0.06	0.03	1.97
	FPRP	0.37	0.22	1.71	0.03	0.20	0.14
	RPD	0.05	0.06	0.91	0.04	0.06	0.64
	TFT	0.08	0.05	1.47	0.06	0.05	1.22
head+1	FPRT	0.04	0.03	1.22	0.02	0.03	0.47
	FPRP	0.15	0.22	0.69	0.17	0.21	0.83
	RPD	0.09	0.06	1.64	0.04	0.07	0.56
	TFT	**0.13**	**0.05**	**2.88***	0.07	0.05	1.39
head+2	FPRT	0.04	0.04	1.05	-0.01	0.04	-0.17
	FPRP	0.24	0.19	1.22	-0.06	0.17	-0.35
	RPD	0.10	0.07	1.41	-0.10	0.07	-1.31
	TFT	0.07	0.05	1.26	-0.06	0.06	-1.08

TABLE 4.10: Experiment 2. The results of the RC type comparisons nested within each level of modification type by region of interest for the dependent measures log-first-pass reading time, first-pass regression probability, log-regression path duration and log-total fixation time.

reading experiment, the effect was more pronounced in subject-modifications than in object-modifications. This can also be explained by surprisal, which predicts a bigger effect size in subject-modifications. Also similar to Experiment 1, the interaction between RC type and modification type did not reach significance in any dependent variable. Rather, the difference between the two modification types with respect to the effect of RC type manifested itself in the SR advantage reaching significance across more dependent variables in subject-modifications than in object-modifications. Note that although surprisal predicts an interaction driven by the stronger effect predicted for subject-modifications, the predicted size of this interaction is very small compared to the predicted size of the effect of RC type. Storage metrics (Gibson, 1998, 2000) which assume that processing difficulty depends on the number of predicted heads are

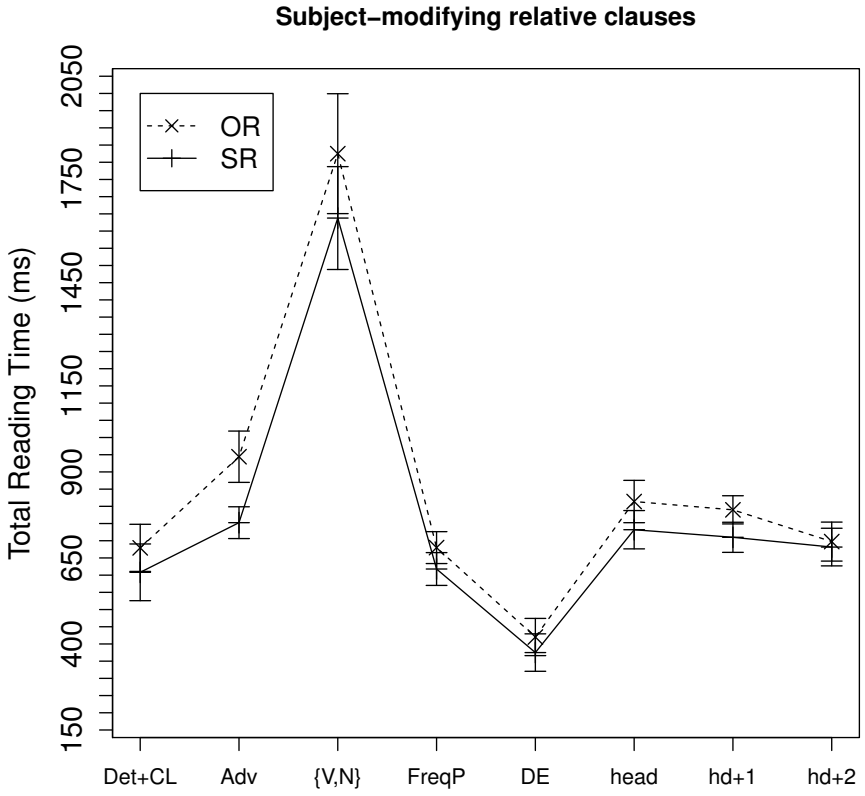

FIGURE 4.5: Experiment 2. Mean total fixation time of each region of interest in subject-modifying relatives, along with 95% confidence intervals.

inconsistent with the SR advantage in the relative clause region V+N/N+V since they predict the absence of an effect of RC type.

At the head noun, similar to Experiment 1, no main effect of RC type was observed. This is predicted by surprisal and DLT storage cost but statistically inconclusive. Although the integration cost component of DLT (Gibson, 2000) as well as the ACT-R based parsing model (Lewis and Vasishth, 2005) predict an OR advantage at the head noun, the observed null result cannot be interpreted as evidence against these theories. In contrast to Experiment 1, the interaction between RC type and modification type reached significance at the head noun (in FPRT only). This interaction was driven by a marginal SR advantage present only in object-modifications. None of the accounts under discussion can account for this interaction. It is inconsistent with retrieval-based memory accounts (DLT integration cost and the ACT-R based model of sentence processing) as they predict the absence of an interaction (they predict an OR advantage that is not modulated by modification type). The expectation-based account and DLT storage cost also predict the absence of an interaction since they predict an SR advantage that is not modulated by modification type either. However, it should be noted that this effect

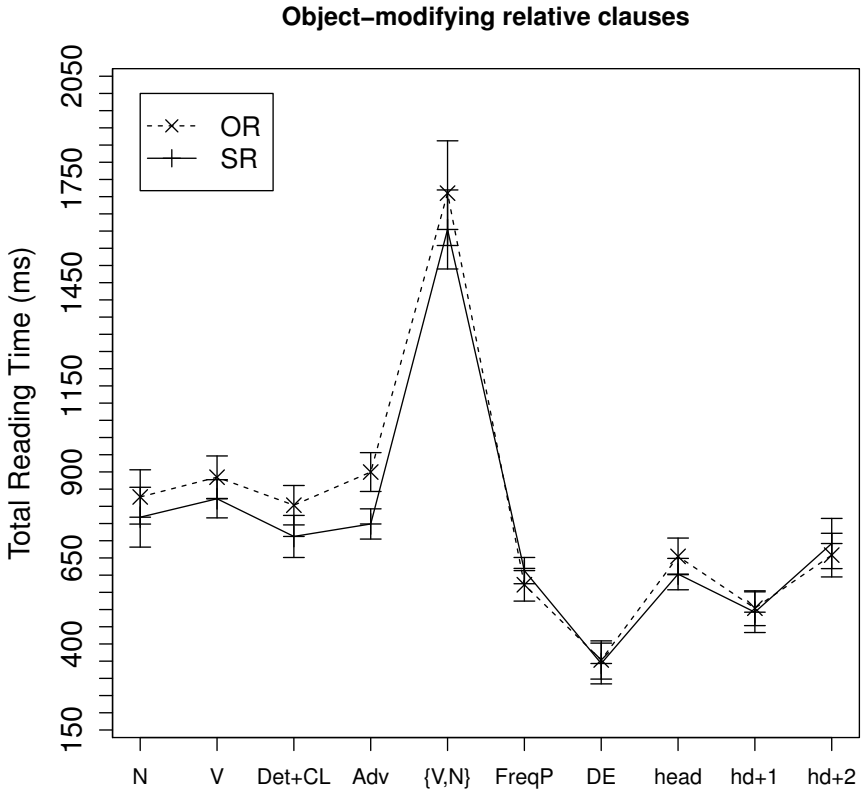

FIGURE 4.6: Experiment 2. Mean total fixation time of each region of interest in object-modifying relatives, along with 95% confidence intervals.

might not be reliable since it was only observed in a single eye-tracking measure (in which no other effects were observed in the experiment), it did not reach statistical significance in pairwise comparisons and it was not observed in Experiment 1. We will therefore not discuss this effect more in detail.

The SR advantage seen in subject-modifications one word after the head noun replicates the effect observed in the spillover region in Experiment 1. Compared to Experiment 1, the effect appeared one word earlier in the sentence. This earlier appearance of the effect is likely to be due to the higher temporal sensitivity of eye-tracking compared to self-paced reading. This effect cannot be accounted for by either of the accounts under discussion. It is inconsistent with the predictions made by retrieval-based accounts, e.g., the structural integration cost metrics of the DLT (Gibson, 1998, 2000) or the ACT-R based model (Lewis and Vasishth, 2005) as both predict an effect in the opposite direction. DLT storage cost (Gibson, 1998, 2000) is also incompatible with the SR advantage one word after the head noun since it predicts no effect at the head noun. Expectation-based accounts cannot explain this effect either since the sentence completion data and the corpus counts incorrectly predict no difference in surprisal at the head noun. However,

one plausible explanation in line with surprisal might be worth considering: The SR advantage in the spillover region might reflect a higher conditional probability of an overtly produced RC head noun (i.e., a headed RC) in SRs compared to ORs within subject-modifications. Our pre-tests might have failed to detect this difference in conditional probabilities due to some reason, e.g., low statistical power. The corpus analyses were restricted to RC tokens preceded by a Det+Cl+Adv sequence, which resulted in a total of 72 tokens considered in the estimation of the conditional probabilities of headless vs. headed RCs. A larger corpus search in which all tokens of subject- and object-modifying SRs/ORs were taken into account and not only RCs preceded by a Det+Cl+Adv sequence showed that in subject-modifications, 18% of the SR tokens (221 of 1244 tokens) and 38% of the OR tokens (220 of 582 tokens) are headless RCs. In object-modifications, in contrast, headless RCs are very rare independently of RC type, 2% in SRs (15 of 845 tokens) and 3% in ORs (8 of 281 tokens). Crucially, these numbers are based on a much larger amount of data (2952 RC tokens in total) compared to the original corpus analyses presented in the Pre-tests section. Thus, it might be related to the low statistical power that in the original corpus analyses and the sentence completion test, we did not find any headless RCs.

In contrast to Experiment 1, the effect of modification type reached significance at various regions. This can be attributed to the higher sensitivity of the eye-tracking method compared to SPR. The processing facilitation in object-modifications starting at the adverb and continuing up to the head noun is in line with the predictions of DLT storage cost. For the adverb and the V+N/N+V region, surprisal also explains this effect. At the head noun, surprisal predicts no effect, and therefore can account for the observed pattern only under the assumption that the effect is due to spillover from the previous regions. An alternative explanation in line with surprisal would be that in object-modifications, the conditional probability of the head noun being overtly realized is indeed higher compared to subject-modifications but the corpus search restricted to the RC tokens following a Det+Cl+Adv sequence and the sentence completion task did not have enough statistical power to detect this difference. This argument is supported by the more general corpus counts on headless vs. headed RCs that are not restricted to RC tokens preceded by Det+Cl+Adv (see above), where headless RCs occur more frequently in subject-modifications than in object-modifications.

Retrieval-based metrics for processing difficulty such as DLT integration cost (Gibson, 1998, 2000) or ACT-R (Lewis and Vasishth, 2005) are incompatible with the faster reading times in object-modifications at the head noun since they predict the opposite direction of the effect.

4.5 General discussion

We investigated whether subject relatives or object relatives (subject- and object-modifications) are easier to process in Chinese by creating experimental stimuli in which the left context strongly constrains the predicted structure to be a relative clause. Importantly, in the experimental materials, we eliminated several local ambiguities present in Chinese relative clauses that might have confounded previous studies. In Mandarin Chinese, SRs occur more frequently than ORs (Hsiao and Gibson, 2003; Vasishth et al., 2013). Therefore, if the left context leads the comprehender to posit a relative clause as the most likely continuation, expectation-based parsing accounts such

as surprisal (Hale, 2001; Levy, 2008) predict an SR advantage at the RC region (the verb and the object within SRs and the subject and the verb within ORs respectively) in both subject- and object-modifications.

In order to derive precise predictions of the expectation-based account for the materials to be used in the reading experiments, we assessed the conditional probability of SRs vs. ORs and subject- vs. object-modifications in a sentence completion experiment. The sentence completion data showed that in the experimental materials, the conditional probability of an SR continuation is higher than that of an OR continuation and this difference is larger in subject-modifications. Moreover, the conditional probability of a subject-modifying RC is lower than the conditional probability of an object-modifying RC. (The sentence completion-based higher conditional probability of SRs over ORs replicated corpus findings, whereas the differences with respect to modification type were not observed in corpus data.) Thus, the expectation-based account, and surprisal in particular, predicts an SR advantage at the relative clause region which is slightly stronger in subject-modifications. Moreover, it predicts faster reading times in object-modifications compared to subject-modifications starting at the onset of the relative clause.

In contrast to expectation-based accounts, storage-based working-memory accounts such as DLT storage cost (Gibson, 2000, 1998) assume that processing cost increases as a function of predictions to be kept track of. For the experimental materials tested here, storage cost predicts no difference between SRs and ORs in the relative clause region and at the head noun because an equal number of upcoming heads is predicted (Hsiao and Gibson, 2003, p. 6). Moreover, storage cost predicts processing facilitation in object-modifications since a smaller number of predicted syntactic heads have to be kept in memory.

Retrieval-based working memory accounts such as DLT integration cost (Gibson, 2000, 1998) or the ACT-R based model of sentence processing (Lewis and Vasishth, 2005) assume that processing difficulty is a function of the distance between the currently processed item and the item that needs to be retrieved for the successful integration of the current item into the parse constructed so far. For the experimental materials, retrieval-based working memory accounts predict an OR advantage at the head noun because of the greater gap-head distance in ORs compared to SRs. Moreover, they predict faster reading times in subject-modifications at the head noun since in object-modifications, an additional retrieval, namely the retrieval of the main clause verb, is triggered.

We conducted two experiments with similar materials but different methods (self-paced reading and eye-tracking). The evidence from both experiments was unequivocal: in both self-paced reading times and across eye-tracking measures, we found an SR advantage at the RC region V+N/N+V preceding the head noun. Nested comparisons showed that this effect was present in both subject- and object-modifications, but more pronounced in subject-modifications. Also in both experiments, in subject-modifications only, we found faster reading times in the materials following the head noun. As for the factor modification type, while no effect was found in the self-paced reading experiment, the eye-tracking data showed clear evidence for a processing facilitation in object-modifications. This facilitation was strongest at the very onset of the relative clause (the adverb) and continued to the head noun.

The predictions of expectation-based accounts of parsing are mostly in line with these results. The SR advantage at the RC region and the fact that this SR advantage was stronger in subject-modifications compared to object-modifications is predicted by surprisal. Moreover, surprisal can also account for the faster reading times in object-modifications starting at the very onset of the relative clause and continuing up to the head noun if we assume that the effect at the head noun is driven by spillover from the previous regions. In contrast, the SR advantage in the spillover region in subject-modifications is not in line with the surprisal predictions derived from sentence completion data or the corpus analyses restricted to relative clauses that have a similar structure as the experimental items. However, more general corpus counts that are not restricted to relative clauses appearing after a Det+Cl+Adv sequence indicate that the conditional probability of an overt head noun appearing after a relative clause is higher in subject-modifying SRs than in subject-modifying ORs while there is no such difference in object-modifications, where the RC head is almost always overtly produced. (This difference in conditional probabilities would also account for the effect of modification type at the RC head).

Memory-based accounts are not compatible with our results. Storage-based memory accounts such as the DLT storage cost metrics are compatible with the faster reading times in object-modifications but are inconsistent with the SR advantage observed at the RC region and, in subject-modifications, in the materials after the head noun. Our results are also inconsistent with the retrieval component of memory-based accounts. The faster reading times in object-modifications at the head noun are incompatible with both DLT integration cost (Gibson, 2000) and Lewis and Vasishth (2005)'s memory-based retrieval architecture as both predict an effect into the opposite direction. The SR advantage in the materials following the head noun is also inconsistent with DLT integration cost or the Lewis and Vasishth (2005) model since they predict an OR advantage at the head noun.

Note that although the effect of modification type and the modulation of the effect of RC type by modification type, i.e., the stronger SR advantage in subject-modifications, can be accounted for by surprisal, it might also reflect an underlying difference in the processing of the relative clause depending on which syntactic part of the sentence is being modified. The present experimental design does not allow us to draw any conclusions in this respect, but we believe that this issue deserves investigation in future research.

The SR advantage in Chinese adds to the growing body of cross-linguistic evidence showing that expectations for upcoming structure play a crucial role in determining processing difficulty. We have already mentioned the recent evidence from English (Staub, 2010) and Russian (Levy et al., 2013). In earlier work, Konieczny (2000) tested German verb-argument dependencies in verb-final structures and found evidence for processing facilitation as a function of increasing distance between the verb and its arguments. Similar results are reported for English (Jaeger et al., 2008), German (Vasishth and Drenhaus, 2011; Levy and Keller, 2013), and for Hindi (Vasishth and Lewis, 2006; Husain et al., 2014). Such anti-locality effects have been explained by Levy (2008) in terms of an increasing conditional probability (which effectively translates to higher predictability) of the verb given preceding context. Related work by Vasishth et al. (2010) has shown that English native speakers exhibit a counter-intuitive grammaticality illusion in reading

times but that Germans do not show this illusion; English speakers find double center embeddings with the middle verb missing easier to process than the grammatical counterpart, whereas German speakers find the ungrammatical version harder to process. This English grammaticality illusion effect has been argued to reflect the statistical infrequency of double embeddings in English; in German, double center embeddings occur relatively more often because all relative clauses are verb-final. In a follow-up study, Frank et al. (2015) showed that German and Dutch native speakers reading English double center embeddings track the structural probabilities of upcoming material, and that more proficient non-native English speakers show a greater grammaticality illusion effect in English, suggesting that they are more closely approximating the native-speaker knowledge of structural probabilities of upcoming structure. Finally, several large scale eye-tracking corpora developed for English and German have also been analysed using surprisal as a predictor (Boston et al., 2008, 2011; Demberg and Keller, 2008); in all cases surprisal has been shown to be a statistically significant predictor of various eye-tracking measures. The present study adds to this broad range of evidence, and provides new support for the idea that we generate and track predictions based on our knowledge of grammar, and that the degree of difficulty that we experience during sentence comprehension is determined, at least in part, by the extent to which our expectations for upcoming structure are met.

At the same time, a large number of studies have shown that increased distance between the two elements of a dependency leads to an increase in processing difficulty. For example, Grodner and Gibson (2005) and Bartek et al. (2011) present evidence that in English argument-verb dependencies, increasing the distance between the verb and the argument leads to slower reading times. A similar pattern is observed in argument-verb dependencies in Russian RCs (Levy et al., 2013), and this has been replicated in Hungarian (Kovács and Vasishth, 2013). These locality effects have been presented as evidence for decay or interference supporting memory-based accounts of sentence processing. In the light of this large body of evidence for memory-based accounts of sentence processing, it would be unreasonable to conclude that expectation rather than memory restrictions determine processing ease in general. However, it is possible that there is cross-linguistic variation in the extent to which one or the other factor dominates. Our data suggest that, at least in the case of Chinese relative clauses, it is expectation rather than memory cost that determines the relative processing ease of SRs vs. ORs.

It is worth noting that our findings are partially consistent with accounts based on the accessibility hierarchy (Keenan and Comrie, 1977) and accounts based on phrase-structure distance (O'Grady, 2007), which have been previously adopted to account for the SR advantage in East Asian languages (Kwon et al., 2006, 2010, 2013; Lin and Bever, 2006; Lin, 2008; Miyamoto and Nakamura, 2003; Ueno and Garnsey, 2008). According to the accessibility hierarchy, noun phrases of certain grammatical functions are easier to access (or extract) than noun phrases of other grammatical functions. The claim is that subject noun phrases, being highest on the hierarchy, are easier to process than object noun phrases. Regarding the phrase-structure distance of SRs and ORs, it has been proposed that processing SRs is less costly because a smaller number of phrasal nodes intervene between the head noun and a subject gap. Both accounts predict processing differences between SRs and ORs on and after the head noun rather than inside the RC regions. These accounts are therefore only consistent with one of the findings reported here, namely the SR advantage in the spillover region following the head noun in subject-modifications.

However, the SR advantage inside the RC region is best accounted for by the expectation-based account.

Another account that can partially explain our results is the so-called Perspective Shift account proposed by MacWhinney (2005). According to this account, processing difficulty increases as a shift in perspective is required. In the experimental materials, the perspective of the main clause and the relative clause are determined by the main clause agent (i.e., the main clause subject) and the RC agent (i.e., the RC subject), respectively.[12] We will assume that a subject takes the perspective as soon as the head noun of the subject NP is predicted. In our subject-modifying conditions, the SR shares the perspective of the main clause while in ORs a perspective-shift from the main clause subject to the RC subject is required when reading the RC subject. This is in line with the SR advantage at the N+V/V+N region in our data. For our object-modifying RCs, in contrast, no difference between SRs and ORs is predicted because in both cases a perspective shift from the main clause agent to the RC agent is required. This shift is predicted to happen at the RC region (at the RC subject in ORs and at the RC verb, i.e., when the RC head is predicted, in SRs). This prediction is not in line with the SR advantage at the N+V/V+N region observed in our object-modifying conditions. If, alternatively, we do not assume that a subject takes perspective as soon as its head is predicted but rather when its head is being encountered in the input, the pattern observed in our subject-modifying conditions cannot be explained by the Perspective Shift account, while the pattern observed in object-modifications might be partially explained. In subject-modifying SRs, the perspective of the main clause subject takes scope over the whole sentence, while in ORs, perspective needs to be shifted from the first encountered RC subject to the main clause subject. Therefore, an SR advantage is predicted at the RC head noun. Thus, the SR advantage observed in the spillover region of the head noun can be explained by the Perspective Shift account, but the even stronger SR advantage at the RC region cannot. In object-modifying conditions, perspective is shifted from the main clause agent to the RC agent when reaching the RC subject in ORs and when reaching the RC head in SRs. Therefore, an SR advantage is predicted at the RC region and an OR advantage at the head noun. Our data are consistent with the former prediction, but not with the latter. However, one could argue that in object-modifications, we do not observe any effect at the head noun because spillover of the SR advantage from the RC region and an OR advantage at the head noun are canceling each other out.

In sum, the Perspective Shift account may explain either the pattern we observed in subject- or in object-modifying conditions depending on the nature of the additional assumptions we make to derive predictions. Crucially, it cannot explain both subject- and object-modifications under the same set of assumptions.

One open issue that remains to be addressed is the role of animacy of the RC subject/object and the head noun. Wu et al. (2012) showed in three self-paced reading studies that SRs were

[12]For accusative-nominative languages it is assumed that language users take the perspective of the thematic agent while for ergative-absolutive languages it is assumed that the perspective of the thematic patient is taken. To derive the predictions of the Perspective Shift account for our materials, we will assume that Chinese, which does not have overt case marking, clusters with accusative-nominative languages in taking the agent's perspective.

read faster when they had animate heads and inanimate objects compared to the reverse animacy configuration. ORs, in contrast, were processed faster when they had inanimate heads and animate subjects. Moreover, they found an SR advantage in sentences with an inanimate subject and an animate object. However, this difference disappeared when the animacy configuration was reversed (animate subject and inanimate object). These findings reflect animacy preferences found in corpus counts. Therefore, the Wu et al. (2012) results are also consistent with expectation-based accounts of relative clause processing. An informative test case would be to replicate our study with animacy of the head noun and the relative clause subject/object as additional factors.

4.6 Conclusion

We present the first study comparing Chinese subject- and object-modifying subject and object relative clauses with materials that use syntactic cues to lead the comprehender to predict a relative clause as the upcoming structure. Two experiments show that the differential conditional probabilities of subject vs. object relative continuations, estimated using a sentence completion study, can predict the reading time difficulty experienced by readers. We found clear evidence for subject relatives being processed faster than object relatives at the region containing the relative clause verb and at the spillover region of the head noun noun in subject-modifying conditions. These data are consistent with a particular instantiation of an expectation-based account of processing, surprisal. We conclude from these results that expectation plays an important role in the processing of Chinese relative clauses. More generally, our data provide independent support for the idea that the human sentence comprehension system deploys its probabilistic knowledge of grammar to generate predictions about upcoming structure. When these predictions are met, processing is relatively easy, but when they are not, processing difficulty occurs. In sum, dashed expectations are costly.

Acknowledgements

We thank the Department of English at National Taiwan Normal University for allowing us to conduct our experiment in their eye-tracking lab. We are grateful to their staff, and to Shiaohui Chan and Sun-Chieh Liang in particular, for their assistance. We thank the audiences at CUNY 2012 and 2013 for comments. We are grateful to Ted Gibson for sharing his experimental items from a previous study on Chinese. Shravan Vasishth thanks the professors at the University of Sheffield's School of Mathematics and Statistics, in particular Dr. Jeremy Oakley and Dr. Kevin Walters, for valuable advice on statistical analysis using linear mixed models. Charles Lin thanks Indiana University's Office of the Vice Provost for Research for research support through its Faculty Research Support Program.

Chapter 5

Conclusion

In this thesis, I studied the working memory mechanisms underlying the processing of syntactic dependencies. In the first two articles, I investigated the properties of the working memory mechanisms subserving human sentence parsing and in the third paper, I compared the relative contribution to processing difficulty of working memory constraints and syntactic prediction. In the first paper, I presented two eye-tracking experiments on Mandarin Chinese reflexives aiming to tease apart syntactic-structure based accounts of memory access from cue-based, i.e., content-addressable, memory retrieval. The experimental results were interpreted as evidence against structure-based accounts of memory access in sentence processing (e.g., Nicol and Swinney, 1989; Sturt, 2003; Phillips et al., 2011). However, the experimental results were not consistent with the original cue-based ACT-R model of sentence processing proposed by Lewis and Vasishth (2005) either. In order to account for the Chinese data, an extension of the original ACT-R model was proposed. Two new principles, namely *prominence* and *cue confusion* were added to the model and it was shown how this extended ACT-R model can not only account for the Chinese data, but also for a wide range of previously unexplained patterns observed in earlier experiments testing other languages. Taken together, the experimental evidence and the computational modeling is interpreted in favor of a content-addressable memory architecture underlying human sentence parsing. However, it has been noted that encoding rather than retrieval interference might be an alternative explanation for interference effects in reflexives. If encoding interference causes the observed effects, it is impossible to distinguish between structure-based memory access and cue-based retrieval as proposed by Lewis and Vasishth (2005) since encoding interference and retrieval interference in a cue-based memory architecture make similar predictions. The second paper of this thesis therefore aimed at teasing apart encoding interference from retrieval interference in the processing of anaphors. Three experiments on German reflexives and Swedish pronominal and reflexive possessives show that there is no evidence for encoding interference affecting the processing of anaphor-antecedent dependencies. This supports the assumption that interference effects observed in reflexive processing arise at the moment of retrieval rather than at the encoding stage. Taken together with the results of the first paper of this thesis, these findings are interpreted as evidence in favor of a cue-based retrieval mechanism subserving the processing of reflexive-antecedent dependencies. As has been noted in the Introduction,

the case of reflexive-antecedent dependencies has been brought up in support of a structure-based memory access and against a content-addressable memory architecture. The results of the experiments presented in the first and the second paper of this thesis clearly challenge this view. On the contrary, the empirical evidence presented in this thesis provide support for a content-addressable memory architecture subserving human sentence parsing. These findings add to an increasing body of evidence from other types of linguistic dependencies favoring a cue-based retrieval mechanism. For example, McElree et al. (2003) presented evidence for cue-based retrieval in filler-gap dependencies, Martin and McElree (2008) for verb-phrase ellipsis, Vasishth et al. (2008) for the processing of negative-polarity items and Van Dyke and Lewis (2003), Van Dyke and McElree (2006), Van Dyke (2007), Wagers et al. (2009), Van Dyke and McElree (2011) and Dillon et al. (2013) for subject-verb dependencies. Thus, the empirical evidence presented in this thesis together with the proposed extension of the cue-based ACT-R model of sentence processing shows that the processing of reflexives can be explained with the same cue-based retrieval mechanism that has been invoked to explain syntactic dependency resolution in a range of other structures. This supports the view that the language processing system is located within a general cognitive architecture, with a general-purpose, i.e., not language-specific, content-addressable working memory system operating on linguistic expressions.

The third paper presented in this thesis investigated the relative contribution of working memory constraints and prediction to sentence processing difficulty using Chinese relative clauses. As has been noted in the Introduction, Chinese relative clauses are an important test case for teasing apart the predictions of expectation-based accounts of sentence processing from memory-based accounts. Although Chinese relative clauses have been tested in a considerable number of studies, the obtained results might not be valid. Due to the syntactic properties of Chinese, the materials of previous studies were confounded with the effects of various local ambiguities. In this thesis, two reading studies with syntactically unambiguous materials were presented. The predictions for the expectation-based account (e.g., Hale, 2001; Levy, 2008) were derived from a sentence completion experiment and cross-validated with corpus analyses. The results of both experiments are consistent with expectation-based accounts of sentence processing, but are not compatible with working-memory based accounts. Neither retrieval metrics (e.g., Stevenson, 1994; Gibson, 1998, 2000; Gordon et al., 2001, 2004; McElree, 2000; Lewis and Vasishth, 2005; Van Dyke and McElree, 2011) nor storage metrics (Gibson, 1998, 2000) can account for the observed results. From these findings, I conclude that syntactic prediction has an important influence on the ease of incremental human sentence parsing. Its relative contribution to processing difficulty can be even stronger than working memory constraints. This suggests that any theory of human sentence processing needs to take into account the power of predictive processes unfolding in the human mind.

References

Abney, S. P. (1989). A computational model of human parsing. *Journal of Psycholinguistic Research*, 18:129–144.

Alcocer, P. and Phillips, C. Using relational syntactic constraints in content-addressable memory architectures for sentence processing. Unpublished manuscript.

Altmann, G. T. M. and Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Journal of Memory and Language*, 31:685–712.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., and Qin, Y. (2004). An integrated theory of the mind. *Psychological Review*, 111(4):1036–60.

Anderson, J. R., Bothell, D., Lebiere, C., and Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, 38:341–380.

Anderson, J. R. and Lebiere, C. (1998). *Atomic Components of Thought*. Lawrence Erlbaum Associates, Inc., Hillsdale, NJ.

Anderson, M. C. and Neely, J. H. (1996). Interference and inhibition in memory retrieval. In Bjork, E. L. and Bjork, R., editors, *Memory*, chapter 8, pages 237–313. Academic Press, San Diego, CA.

Ariel, M. (1990). *Accessing Noun-Phrase Antecedents*. Routledge, London, UK.

Arnold, D. (2007). Non-restrictive relatives are not orphans. *Journal of Linguistics*, 43(2):271–309.

Arnold, J. E. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, 31:137–162.

Asendorpf, J. B., Conner, M., De Fruyt, F., De Houwer, J., Denissen, J. J., Fiedler, K., Fiedler, S., Funder, D. C., Kliegl, R., Nosek, B. A., Perugini, M., Roberts, B. W., Schmitt, M., Van Aken, M. A., Weber, H., and Wicherts, J. M. (2013). Recommendations for increasing replicability in psychology. *European Journal of Personality*, 27(2):108–119.

Baayen, R. H., Davidson, D. J., and Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language*, 59(4):390–412.

Badecker, W. and Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4):748–769.

Bartek, B., Lewis, R. L., Vasishth, S., and Smith, M. (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 37(5):1178–1198.

Bates, D., Maechler, M., Bolker, B., and Walker, S. (2014). *lme4: Linear mixed-effects models using Eigen and S4*. R package version 1.0-6.

Battistella, E. (1989). Chinese reflexivization: A movement to INFL approach. *Linguistics*, 27:987–1012.

Bonferroni, C. E. (1936). Teoria statistica delle classi e calcolo delle probabilità. *Pubblicazioni del R Istituto Superiore di Scienze Economiche e Commerciali di Firenze*, 8:3–62.

Boston, M. F., Hale, J. T., Patil, U., Kliegl, R., and Vasishth, S. (2008). Parsing costs as predictors of reading difficulty: An evaluation using the Potsdam Sentence Corpus. *Journal of Eye Movement Research*, 2(1):1–12.

Boston, M. F., Hale, J. T., Vasishth, S., and Kliegl, R. (2011). Parallel processing and sentence comprehension difficulty. *Language and Cognitive Processes*, 26(3):301–349.

Box, G. E. P. and Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society*, 26:211–234.

Brennan, S. E. (1995). Centering attention in discourse. *Language and Cognitive Processes*, 10(2):137–167.

Büring, D. (2005). *Binding Theory*. Cambridge Textbooks in Linguistics. Cambridge University Press, Cambridge, UK.

Cai, Q. and Brysbaert, M. (2010). SUBTLEX-CH: Chinese word and character frequencies based on film subtitles. *PloS ONE*, 5(6):e10729.

Caplan, D. and Waters, G. S. (1996). The capacity theory of sentence comprehension: Critique of Just and Carpenter (1992). *Psychological Review*, 103:761–772.

Caplan, D. and Waters, G. S. (1999). Verbal working memory and sentence comprehension. *Behavioral and Brain Sciences*, 22:77–126.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In Li, C. N., editor, *Subject and Topic*, pages 25–56. Academic Press, New York.

Chen, B. and Ning, A. (2008). The comparison of processing difficulty between Chinese subject-relative and object-relative clauses. *Yingyong Xinli Kexue [Chinese Journal of Applied Psychology]*, 14(1):29–34.

Chen, B., Ning, A., Bi, H., and Dunlap, S. (2008). Chinese subject-relative clauses are more difficult to process than the object-relative clauses. *Acta Psychologica*, 129(1):61–65.

Chen, Z., Grove, K., and Hale, J. T. (2012a). Structural expectations in Chinese relative clause comprehension. In Choi, J., Hogue, E. A., Punske, J., Tat, D., Schertz, J., and Trueman, A., editors, *Proceedings of the 29th West Coast Conference on Formal Linguistics*, pages 29–37, Somerville, MA. Cascadilla Press.

Chen, Z., Jäger, L., and Vasishth, S. (2012b). How structure-sensitive is the parser? Evidence from Mandarin Chinese. In Stolterfoht, B. and Featherston, S., editors, *Empirical Approaches to Linguistic Theory: Studies of Meaning and Structure*, Studies in Generative Grammar, pages 43–62. Mouton de Gruyter, Berlin, Germany.

Chomsky, N. (1981). *Lectures on Government and Binding*. Foris, Dordrecht, The Netherlands.

Chomsky, N. (1986). *Knowledge of Language: Its Nature, Origin, and Use*. Praeger, New York, NY.

Chow, W.-Y., Lewis, S., and Phillips, C. (2014). Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology*, 5(630).

Clackson, K., Felser, C., and Clahsen, H. (2011). Children's processing of reflexives and pronouns in English: Evidence from eye-movements during listening. *Journal of Memory and Language*, 65:128–144.

Clackson, K. and Heyer, V. (2014). Reflexive anaphor resolution in spoken language comprehension: Structural constraints and beyond. *Frontiers in Psychology*, 5(904).

Clifton, C. and Frazier, L. (1989). Comprehending sentences with long distance dependencies. In Carlson, G. and Tanenhaus, M., editors, *Linguistic Structure in Language Processing*, pages 273–317. Kluwer, Dordrecht, The Netherlands.

Clifton, C., Frazier, L., and Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica*, 11(1):11–39.

Clifton, C., Staub, A., and Rayner, K. (2007). Eye movements in reading words and sentences. In Van Gompel, R. P., Fischer, M. H., Murray, W. S., and Hill, R. L., editors, *Eye Movements: A Window on Mind and Brain*, chapter 15, pages 341–372. Elsevier, Oxford, UK.

Cole, P., Hermon, G., and Sung, L. (1990). Principles and parameters of long-distance reflexives. *Linguistic Inquiry*, 21(1):1–22.

Cole, P., Hermon, G., and Sung, L.-M. (1993). Feature percolation. *Journal of East Asian Linguistics*, 2(1):91–118.

Cole, P. and Sung, L.-M. (1994). Head movement and long-distance reflexives. *Linguistic Inquiry*, 25(3):355–406.

Cole, P. and Wang, C. (1996). Antecedents and blockers of long-distance reflexives: The case of Chinese *ziji*. *Linguistic Inquiry*, 27(3):357–390.

Cousineau, D. (2005). Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutorial in Quantitative Methods for Psychology*, 1(1):42–45.

Cowles, H. W., Walenski, M., and Kluender, R. (2007). Linguistic and cognitive prominence in anaphor resolution: topic, contrastive focus and pronouns. *Topoi*, 26:3–18.

Crocker, M. W. (1995). *Computational Psycholinguistics: An Interdisciplinary Approach to the Study of Language*. Kluwer, Dordrecht, The Netherlands.

Cunnings, I. and Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, 28(1-2):188–219.

Cunnings, I. and Sturt, P. (2014). Coargumenthood and the processing of reflexives. *Journal of Memory and Language*, 75:117–139.

DeLong, K., Urbach, T., and Kutas, M. (2005). Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. *Nature Neuroscience*, 8(8):1117–1121.

Demberg, V. and Keller, F. (2008). Data from eye-tracking corpora as evidence for theories of syntactic processing complexity. *Cognition*, 109(2):193–210.

Diana, R. A. and Reder, L. M. (2006). The low-frequency encoding disadvantage: Word frequency affects processing demands. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(4):1805–815.

Dillon, B. (2011). *Structured Access in Sentence Comprehension*. Phd thesis, University of Maryland, College Park, MD.

Dillon, B., Chow, W.-Y., Wagers, M., Guo, T., Liu, F., and Phillips, C. (2014). The structure-sensitivity of memory access: Evidence from Mandarin Chinese. *Frontiers in Psychology*, 5(1025).

Dillon, B., Clifton, C., Sloggett, S., and Frazier, L. (2015). Not all relative clauses interfere equally in filler-gap processing. In *28th Annual CUNY Conference on Human Sentence Processing*, page 281, Los Angeles, CA. University of Southern California.

Dillon, B., Mishler, A., Sloggett, S., and Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, 69:85–103.

Du Bois, J. W. (1987). The discourse basis of ergativity. *Language*, 63:805–855.

Du Bois, J. W. (2003). Argument structure: Grammar in use. In Du Bois, J. W., Kumpf, L. E., and Ashby, W. J., editors, *Preferred Argument Structure: Grammar as Architecture for Function*, volume 14 of *Studies in Discourse and Grammar*, pages 11–60. John Benjamins, Amsterdam, The Netherlands.

Dunn, J. O. (1959). Estimation of the medians for dependent variables. *The Annals of Mathematical Statistics*, 30(1):192–197.

Dunn, J. O. (1961). Multiple comparisons among means. *Journal of the American Statistical Association*, 56(293):52–64.

Ehrlich, S. E. and Rayner, K. (1981). Contextual effects on word perception and eye movements during reading. *Journal of Verbal Learning and Verbal Behavior*, 20:641–655.

Elbourne, P. D. (2005). *Situations and Individuals*, volume 41 of *Current Studies in Linguistics*. MIT Press, Cambridge, MA.

Engelmann, F., Jäger, L., and Vasishth, S. (2015). Cue confusion and distractor prominence can explain inconsistent interference effects. In *28th Annual CUNY Conference on Human Sentence Processing*, page 168, Los Angeles, CA. University of Southern California.

Everaert, M. (1986). *The Syntax of Reflexivation*. Foris, Dordrecht, The Netherlands.

Federmeier, K. and Kutas, M. (1999). A rose by any other name: Long-term memory structure and sentence processing. *Journal of Memory and Language*, 41:469–495.

Fedorenko, E., Gibson, E., and Rohde, D. (2006). The nature of working memory capacity in sentence comprehension: Evidence against domain specific resources. *Journal of Memory and Language*, 54:541–553.

Felser, C., Sato, M., and Bertenshaw, N. (2009). The on-line application of Binding Principle A in English as a second language. *Bilingualism: Language and Cognition*, 12:485–502.

Ferreira, F. and Clifton, C. (1986). The independence of syntactic processing. *Journal of Memory and Language*, 25:348–368.

Ferreira, F., Ferraro, V., and Bailey, K. G. D. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, 11:11–15.

Fisher, R. (1937). *The Design of Experiments*. Oliver and Boyd, Edinburgh, UK, 2nd edition.

Fox, D. (1998). Locality in variable binding. In Barbosa, P., Fox, D., Hagstrom, P., McGinnis, M., and Pesetsky, D., editors, *Is the Best Good Enough? Optimality and Competition in Syntax*, pages 129–155. MIT Press, Cambridge, MA.

Frank, S. L., Trompenaars, T., and Vasishth, S. (2015). Cross-linguistic differences in processing double-embedded relative clauses: Working-memory constraints or language statistics? *Cognitive Science*. In press.

Frazier, L. (1979). *On Comprehending Sentences: Syntactic Parsing Strategies*. Phd thesis, University of Massachusetts, Amherst, MA.

Frazier, L. (1987). Sentence processing: A tutorial review. In Coltheart, M., editor, *The Psychology of Reading*, volume 12 of *Attention and Performance*, pages 559–586. Erlbaum, Hillsdale, NJ.

Frazier, L. and Fodor, J. D. (1978). The sausage machine: A new two-stage parsing model. *Cognition*, 6(4):291–325.

Fukumura, K. and van Gompel, R. P. G. (2011). The effect of animacy on the choice of referring expression. *Language and Cognitive Processes*, 26(10):1472–1504.

Gao, L., Liu, Z., and Huang, Y. (2005). Ziji shi shei? Dui yueshu yuanze de shiyan yanjiu [Who is *ziji*? An experimental research on Binding Principles]. *Yuyan Kexue [Linguistic Sciences]*, 4(2):39–50.

Gast, V. and Haas, F. (2008). Reflexive and reciprocal readings of anaphors in German and other European languages. In König, E. and Gast, V., editors, *Reciprocals and Reflexives: Theoretical and Typological Explorations*, pages 307–346. Mouton de Gruyter, Berlin, Germany.

Gelman, A. and Hill, J. (2007). *Data Analysis using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, Cambridge, UK.

Gernsbacher, M. A. and Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, 27:699–717.

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

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In Marantz, A., Miyashita, Y., and O'Neil, W., editors, *Image, Language, Brain*, chapter 5, pages 95–126. MIT Press, Cambridge, MA.

Gibson, E. and Wu, H.-H. I. (2013). Processing Chinese relative clauses in context. *Language and Cognitive Processes*, 28(1-2):125–155.

Givón, T. (1983). *Topic Continuity in Discourse*. John Benjamins, Amsterdam, The Netherlands.

Gordon, P. C., Hendrick, R., and Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27(6):1411–1423.

Gordon, P. C., Hendrick, R., and Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, 51:97–114.

Gordon, P. C., Hendrick, R., Johnson, M., and Lee, Y. (2006). Similarity-based interference during language comprehension: Evidence from eye tracking during reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32(6):1304–1321.

Gordon, P. C., Hendrick, R., and Levine, W. H. (2002). Memory-load interference in syntactic processing. *Psychological Science*, 13(5):425–430.

Grodner, D. and Gibson, E. (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science*, 29:261–291.

Grodzinsky, Y. and Reinhart, T. (1993). The innateness of binding and coreference. *Linguistic Inquiry*, 24(1):69–101.

Grosz, B. J., Weinstein, S., and Joshi, A. K. (1995). Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics*, 21(2):203–225.

Gundel, J., Hedberg, N., and Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69:274–307.

Haegeman, L. (1994). *Introduction to Government and Binding Theory*. Blackwell, Oxford, UK, 2nd edition.

Hale, J. T. (2001). A probabilistic Earley parser as a psycholinguistic model. In *Proceedings of the 2nd Meeting of the North American Chapter of the Association for Computational Linguistics*, pages 159–166, Pittsburgh, PA. Association for Computational Linguistics.

Heim, I. and Kratzer, A. (1998). *Semantics in Generative Grammar*. Blackwell, Oxford, UK.

Heister, J., Würzner, K.-M., Bubenzer, J., Pohl, E., Hanneforth, T., Geyken, A., and Kliegl, R. (2011). dlexDB – eine lexikalische Datenbank für die psychologische und linguistische Forschung. *Psychologische Rundschau*, 62(1):10–20.

Holmes, P. and Hinchliffe, I. (1994). *Swedish. A Comprehensive Grammar*. Routledge, New York, NY.

Hsiao, F. and Gibson, E. (2003). Processing relative clauses in Chinese. *Cognition*, 90(1):3–27.

Hsiao, Y., Li, J., and MacDonald, M. (2014). Ambiguity affects Mandarin relative clause processing. In *27th Annual CUNY Conference on Human Sentence Processing*, Columbus, OH. The Ohio State University.

Hsiao, Y. and MacDonald, M. C. (2013). Experience and generalization in a connectionist model of Mandarin Chinese relative clause processing. *Frontiers in Psychology*, 4(767).

Hsu, C.-C. N., Hurewitz, F., and Phillips, C. (2006). Contextual and syntactic cues for head-final relative clauses in Chinese. In *19th Annual CUNY Conference on Human Sentence Processing*, New York, NY. City University of New York.

Hsu, C.-C. N., Phillips, C., and Yoshida, M. (2005). Cues for head-final relative clauses in Chinese. In *18th Annual CUNY Conference on Human Sentence Processing*, Tucson, AZ. University of Arizona.

Huang, C.-T. J. (1982). *Logical Relations in Chinese and the Theory of Grammar*. Phd thesis, Massachusetts Institute of Technology, Cambridge, MA.

Huang, C. T. J. (1989). Pro drop in Chinese: a generalized control approach. In Jaeggli, O. and Safir, K., editors, *The Null Subject Parameter*, pages 185–214. Kluwer, Dordrecht, The Netherlands.

Huang, C.-T. J. (2002). Distributivity and reflexivity. In Tang, S.-W. and Liu, C.-S. L., editors, *On the Formal Way to Chinese Languages*, pages 21–44. CSLI Publications, Stanford, CA.

Huang, C.-T. J., Huang, Y.-H., Teng, T.-H., and Tiedeman, R. (1984). Reflexives in Chinese and the teaching of Chinese. In *Proceedings of the First World Conference on Chinese Language*, pages 205–215, Taipeih, Taiwan. World Chinese Language Association.

Huang, C.-T. J. and Liu, C.-S. L. (2001). Logophoricity, attitudes, and *ziji* at the interface. In Cole, P., Hermon, G., and Huang, C., editors, *Long-Distance Reflexives*, volume 33 of *Syntax and Semantics*, pages 141–195. Academic Press, New York, NY.

Huang, C.-T. J. and Tang, C.-C. J. (1988). On the local nature of the long-distance reflexive in Chinese. In *Proceedings of NELS 18*, Amherst, MA. GLSA Publications.

Huang, C.-T. J. and Tang, C.-C. J. (1989). The local nature of the long-distance reflexive in Chinese. In *Proceedings of NELS 19*, pages 191–206, Amherst, MA. GLSA Publications.

Huang, C.-T. J. and Tang, C.-C. J. (1991). The local nature of the long-distance reflexive in Chinese. In Koster, J. and Reuland, E., editors, *Long-distance Anaphora*, pages 263–282. Cambridge University Press, Cambridge, UK.

Huang, Y.-H. (1984). Reflexives in Chinese. *Studies in English Literature and Linguistics*, 10:163–188.

Huettig, F. (2015). Four central questions about prediction in language processing. *Brain Research*. In press.

Husain, S., Vasishth, S., and Srinivasan, N. (2014). Strong expectations cancel locality effects: Evidence from Hindi. *PLoS ONE*, 9(7):e100986.

Ioannidis, J. P. A. (2005). Why most published research findings are false. *PLoS Medicine*, 2(8):e124.

Jaeger, F. T., Fedorenko, E., Hofmeister, P., and Gibson, E. (2008). Expectation-based syntactic processing: Antilocality outside of head-final languages. In *21th Annual CUNY Conference on Human Sentence Processing*, Chapel Hill, NC. University of North Carolina.

Jäger, L., Engelmann, F., and Vasishth, S. (2015). Retrieval interference in reflexive processing: Experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology*, 6(617).

Jurafsky, D. (1996). A probabilistic model of lexical and syntactic access and disambiguation. *Cognitive Science*, 20:137–194.

Just, M. A. and Carpenter, P. A. (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review*, 99(1):122–149.

Just, M. A., Carpenter, P. A., and Woolley, J. D. (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General*, 111(2):228–238.

Kaiser, E. (2003). Encoding (non)locality in anaphoric relations. In Nelson, D. and Manninen, S., editors, *Generative Approaches to Finnic and Saami Linguistics*, pages 269–293. CSLI Publications, Stanford, CA.

Kaiser, E. and Trueswell, J. C. (2004). The role of discourse context in the processing of a flexible word-order language. *Cognition*, 94(2):113–147.

Kamide, Y., Altmann, G. T. M., and Haywood, S. L. (2003). The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory and Language*, 49:133–156.

Kang, B.-M. (1988). Unbounded reflexives. *Linguistics and Philosophy*, 11:415–456.

Keenan, E. L. and Comrie, B. (1977). Noun phrase accessibility and universal grammar. *Linguistic Inquiry*, 8(1):63–99.

King, J., Andrews, C., and Wagers, M. (2012). Do reflexives always find a grammatical antecedent for themselves? In *25th Annual CUNY Conference on Human Sentence Processing*, page 67, New York, NY. The CUNY Graduate Center.

King, J. and Just, M. A. (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language*, 30:580–602.

Konieczny, L. (2000). Locality and parsing complexity. *Journal of Psycholinguistic Research*, 29(6):627–645.

Kovács, N. and Vasishth, S. (2013). The processing of relative clauses in Hungarian. In *Architectures and Mechanisms for Language Processing Conference*, Marseille, France. Aix-Marseille Université.

Kush, D. (2013). *Respecting Relations: Memory Access and Antecedent Retrieval in Incremental Sentence Processing*. Phd thesis, University of Maryland, College Park, MD.

Kush, D. and Phillips, C. (2014). Local anaphor licensing in an SOV language: Implications for retrieval strategies. *Frontiers in Psychology*, 5(1252).

Kwon, N., Kluender, R., Kutas, M., and Polinsky, M. (2013). Subject/object processing asymmetries in Korean relative clauses: Evidence from ERP data. *Language*, 89(3):537–585.

Kwon, N., Lee, Y., Gordon, P., Kluender, R., and Polinsky, M. (2010). Cognitive and linguistic factors affecting subject/object asymmetry: An eye-tracking study of pre-nominal relative clauses in Korean. *Language*, 86(3):546–582.

Kwon, N., Polinsky, M., and Kluender, R. (2006). Subject preference in Korean. In Baumer, D., Montero, D., and Scanlon, M., editors, *Proceedings of the 25th West Coast Conference on Formal Linguistics*, pages 1–14, Somerville, MA. Cascadilla Press.

Kwong, O. Y. and Tsou, B. K. (2003). Categorial fluidity in Chinese and its implications for part-of-speech tagging. In Copestake, A. and Hajic, J., editors, *Proceedings of the 10th Conference of the European Chapter of the Association for Computational Linguistics*, pages 115–118, Budapest, Hungary. Association for Computational Linguistics.

Laurinavichyute, A., Jäger, L., Akinina, Y., Benz, L., and Dragoy, O. (2015). Teasing apart retrieval and encoding interference in russian reflexives. In *28th Annual CUNY Conference on Human Sentence Processing*, page 140, Los Angeles, CA. University of Southern California.

Lebeaux, D. (1983). A distributional difference between reciprocals and reflexives. *Linguistic Inquiry*, 14(4):723–730.

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, 106:1126–1177.

Levy, R. and Andrew, G. (2006). Tregex and Tsurgeon: Tools for querying and manipulating tree data structures. In Calzolari, N., Choukri, K., Gangemi, A., Maegaard, B., Mariani, J., Odjik, J., and Tapias, D., editors, *Proceedings of the 5th International Conference on Language Resources and Evaluation*, pages 2231–2234, Genoa, Italy. European Language Resources Association.

Levy, R., Fedorenko, E., and Gibson, E. (2013). The syntactic complexity of Russian relative clauses. *Journal of Memory and Language*, 69:461–495.

Levy, R. and Keller, F. (2013). Expectation and locality effects in German verb-final structures. *Journal of Memory and Language*, 68:199–222.

Lewandowsky, S., Geiger, S. M., and Oberauer, K. (2008). Interference-based forgetting in verbal short-term memory. *Journal of Memory and Language*, 59:200–222.

Lewis, D. (1979). Attitudes de dicto and de se. *The Philosophical Review*, 88(4):513–543.

Lewis, R. L. and Vasishth, S. (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29(3):375–419.

Lewis, R. L., Vasishth, S., and Van Dyke, J. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10(10):447–454.

Li, N. (1991). *Perspective-Taking in Mandarin Discourse*. Phd thesis, State University of New York at Buffalo, Buffalo, NY.

Li, X. and Zhou, X. (2010). Who is *ziji*? ERP responses to the Chinese reflexive pronoun during sentence comprehension. *Brain Research*, 1331(1981):96–104.

Lin, C.-J. C. (2008). The processing foundation of head-final relative clauses. *Language and Linguistics*, 9:813–839.

Lin, C.-J. C. (2014). Effect of thematic order on the comprehension of Chinese relative clauses. *Lingua*, 140:180–206.

Lin, C.-J. C. and Bever, T. G. (2006). Subject preference in the processing of relative clauses in Chinese. In Baumer, D., Montero, D., and Scanlon, M., editors, *Proceedings of the 25th West Coast Conference on Formal Linguistics*, pages 254–260, Somerville, MA. Cascadilla Press.

Lin, C.-J. C. and Bever, T. G. (2011). Garden path and the comprehension of head-final relative clauses. In Yamashita, H., Hirose, Y., and Packard, J., editors, *Processing and Producing Head-final Structures*, pages 277–297. Springer, Dordrecht, The Netherlands.

Lin, Y. and Garnsey, S. (2011). Animacy and the resolution of temporary ambiguity in relative clause comprehension in Mandarin. In Yamashita, H., Hirose, Y., and Packard, J., editors, *Processing and Producing Head-final Structures*, pages 241–275. Springer, Dordrecht, The Netherlands.

Liu, L. (2010). A pragmatic account of anaphora: The cases of the bare reflexive in Chinese. *Journal of Language Teaching and Research*, 1(6):796–806.

Liu, T., Zhou, T., and Yang, Y. (2011). Universality of processing advantage for subject relative clause: An ERP study of Chinese relative clauses. *Yuyan Kexue [Linguistic Sciences]*, 10(1):1–20.

Liu, Z. (2009). The cognitive process of Chinese reflexive processing. *Journal of Chinese Linguistics*, 37(1):1–27.

Logačev, P. and Vasishth, S. (2015). A multiple-channel model of task-dependent ambiguity resolution in sentence comprehension. *Cognitive Science*. In press.

Longobardi, G. (1994). Reference and proper names: A theory of N-movement in syntax and logical form. *Linguistic Inquiry*, 25(4):609–665.

MacWhinney, B. (2005). The emergence of grammar from perspective. In Pecher, D. and Swann, R. A., editors, *Grounding Cognition: The Role of Perception and Action in Memory, Language, and Thinking*, pages 198–223. Cambridge University Press, New York, NY.

Manzini, M. R. and Wexler, K. (1987). Parameters, binding theory, and learnability. *Linguistic Inquiry*, 18:413–444.

Marslen-Wilson, W. (1973). Linguistic structure and speech shadowing at very short latencies. *Nature*, 244:522–523.

Marslen-Wilson, W. (1975). Sentence perception as an interactive parallel process. *Science*, 189:226–228.

Martin, A. E. and McElree, B. (2008). A content-addressable pointer mechanism underlies comprehension of verb-phrase ellipsis. *Journal of Memory and Language*, 58:879–906.

McDonald, S. A. and Shillcock, R. C. (2003). Low-level predictive inference in reading: The influence of transitional probabilities on eye movements. *Vision Research*, 43:1735–1751.

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29(2):111–123.

McElree, B. (2003). Accessing recent events. *Psychology of Learning and Motivation*, 46:155–200.

McElree, B. (2006). Accessing recent events. In Ross, B. H., editor, *The Psychology of Learning and Motivation: Advances in Research and Theory*, volume 46, pages 155–200. Elsevier, San Diego, CA.

McElree, B. and Dosher, B. (1989). Serial position and set size in short-term memory: The time course of recognition. *Journal of Experimental Psychology: General*, 118(4):346–373.

McElree, B., Foraker, S., and Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48:67–91.

McRae, K., Spivey-Knowlton, M. J., and Tanenhaus, M. K. (1998). Modeling the influence of thematic fit (and other constraints) in on-line sentence comprehension. *Journal of Memory and Language*, 38:283–312.

Miller, G. and Chomsky, N. (1963). Finitary models of language users. In Luce, R. D., Bush, R. R., and Galanter, E., editors, *Handbook of Mathematical Psychology*, volume 2, pages 419–491. John Wiley & Sons, New York, NY.

Mitchell, D., Cuetos, F., Corley, M. M. B., and Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (non-lexical) statistical records. *Journal of Psycholinguistic Research*, 24:469–488.

Mitchell, D. and Holmes, V. M. (1985). Specific information about the verb in parsing sentences with local structural ambiguity. *Journal of Memory and Language*, 24:542–559.

Miyamoto, E. T. and Nakamura, M. (2003). Subject/object asymmetries in the processing of relative clauses in Japanese. In *Proceedings of the 22nd West Coast Conference on Formal Linguistics*, pages 342–355, Somerville, MA. University of California, Cascadilla Press.

Morey, R. D. (2008). Confidence intervals from normalized data: A correction to Cousineau (2005). *Tutorial in Quantitative Methods for Psychology*, 4(2):61–64.

Nairne, J. S. (1988). A framework for interpreting recency effects in immediate serial recall. *Memory and Cognition*, 16(4):343–352.

Nairne, J. S. (1990). A feature model of immediate memory. *Memory and Cognition*, 18(3):251–269.

Nicol, J. and Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research*, 18(1):5–19.

Oberauer, K. and Kliegl, R. (2006). A formal model of capacity limits in working memory. *Journal of Memory and Language*, 55:601–626.

O'Grady, W. (2007). *Syntactic Development*. University of Chicago Press, Chicago, IL.

Pan, H. (1995). *Locality, Self-Ascription, Discourse Prominence, and Mandarin Reflexives*. Phd thesis, University of Texas at Austin, Austin, TX.

Pan, H. (1997). *Constraints on Reflexivization in Mandarin Chinese*. Garland Publishing, New York, NY.

Pan, H. (2000). Why the blocking effect? In Cole, P., Hermon, G., and Huang, C.-T. J., editors, *Long-Distance Reflexives*, volume 33 of *Syntax and Semantics*, pages 279–316. Academic Press, New York, NY.

Parker, D. and Phillips, C. (2014). Selective priority for structure in memory retrieval. In *27th Annual CUNY Conference on Human Sentence Processing*, page 100, Columbus, OH.

Patil, U., Vasishth, S., and Lewis, R. L. Retrieval interference in syntactic processing: The case of reflexive binding in English. Unpublished manuscript.

Phillips, C., Wagers, M. W., and Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In Runner, J., editor, *Experiments at the Interfaces*, volume 37 of *Syntax and Semantics*, pages 147–180. Emerald Group Publishing Limited, Bingley, UK.

Pica, P. (1987). On the nature of the reflexivation circle. In McDonough, J. and Plunkett, B., editors, *Proceedings of the 17th Annual Meeting of the North East Linguistic Society*, pages 483–499, Amherst, MA. GLSA Publications.

Pickering, M. J. and Barry, G. (1991). Sentence processing without empty categories. *Language and Cognitive Processes*, 6(3):229–259.

Pickering, M. J. and Garrod, S. (2007). Do people use language production to make predictions during comprehension? *Trends in Cognitive Sciences*, 11(3):105–110.

Pickering, M. J. and Van Gompel, R. P. G. (2006). Syntactic parsing. In Traxler, M. and Gernsbacher, M., editors, *Handbook of Psycholinguistics*, pages 455–503. Academic Press, Amsterdam, The Netherlands, 2nd edition.

Pollard, C. and Sag, I. A. (1992a). Anaphors in English and the scope of binding Theory. *Linguistic Inquiry*, 23(2):261–303.

Pollard, C. and Sag, I. A. (1992b). Binding theory. In Park, B.-S., editor, *Linguistic Studies on Natural Language*, volume 1 of *The Kyunghee Language Institute Monograph*, pages 49–78. Hanshin, Seoul.

Pollard, C. and Sag, I. A. (1994). *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago, IL.

Pollard, C. and Xue, P. (1998). Chinese reflexive *ziji*: Syntactic reflexives vs. nonsyntactic reflexives. *Journal of East Asian Linguistics*, 7:287–318.

Potts, C. (2005). *The Logic of Conventional Implicatures*. Oxford University Press, Oxford, UK.

Pritchett, B. L. (1992). *Grammatical Competence and Parsing Performance*. University of Chicago Press, Chicago, IL.

Qiao, X., Shen, L., and Forster, K. (2012). Relative clause processing in Mandarin: Evidence from the maze task. *Language and Cognitive Processes*, 27(4):611–630.

R Development Core Team (2009). *R: A Language and Environment for Statistical Computing*. Vienna.

Rayner, K., Carlson, M., and Frazier, L. (1983). The interaction of syntax and semantics during sentence processing: Eye moevements in the analysis of semantically biased sentences. *Journal of Verbal Learning and Verbal Behavior*, 22:258–274.

Reinhart, T. (2000). Strategies of anaphora resolution. In Bennis, H., Everaert, M., and Reuland, E. J., editors, *Interface Strategies*, pages 295–325. Royal Academy of Sciences, Amsterdam, The Netherlands.

Reinhart, T. and Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, 24(4):657–720.

Reuland, E. (2001). Primitives of binding. *Linguistic Inquiry*, 32(3):439–492.

Reuland, E. (2011). *Anaphora and Language Design*. MIT Press, Cambridge, MA.

Reuland, E. and Reinhart, T. (1995). Pronouns, anaphors and case. In Haider, H., Olsen, S., and Vikner, S., editors, *Studies in Comparative Germanic Syntax*, pages 241–269. Kluwer, Dordrecht, The Netherlands.

Simmons, J. P., Nelson, L. D., and Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, 22(11):1359–1366.

Staub, A. (2010). Eye movements and processing difficulty in object relative clauses. *Cognition*, 116(1):71–86.

Staub, A. and Clifton, C. (2006). Syntactic prediction in language comprehension: Evidence from *either... or*. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 32:425–436.

Staub, A. and Rayner, K. (2007). Eye movements and on-line comprehension processes. In Gaskell, M. G. and Altmann, G., editors, *The Oxford Handbook of Psycholinguistics*, pages 327–342. Oxford University Press, Oxford, UK.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, 153(3736):652–654.

Sternberg, S. (1969). Memory-scanning: Mental processes revealed by reaction-time experiments. *American Scientist*, 57:421–457.

Stevenson, S. (1994). Competition and recency in a hypbrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research*, 23:295–322.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48:542–562.

Swets, B., Desmet, T., Clifton, C., and Ferreira, F. (2008). Underspecification of syntactic ambiguities: Evidence from self-paced reading. *Memory and Cognition*, 36(1):201–216.

Tabor, W. and Tanenhaus, M. K. (1999). Dynamical models of sentence processing. *Cognitive Science*, 23:491–515.

Tang, C.-C. J. (1989). Chinese reflexives. *Natural Language and Linguistic Theory*, 7(1):93–121.

Traxler, M. J. (2007). Working memory contributions to relative clause attachment processing: A hierarchical linear modeling analysis. *Memory and Cognition*, 35(5):1107–1121.

Traxler, M. J. and Pickering, M. J. (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language*, 35:454–475.

Trueswell, J. C. (1996). The role of lexical frequency in syntactic ambiguity resolution. *Journal of Memory and Language*, 35:566–585.

Trueswell, J. C., Tanenhaus, M. K., and Kello, C. (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19:528–553.

Ueno, M. and Garnsey, S. M. (2008). An ERP study of the processing of subject and object relative clauses in Japanese. *Language and Cognitive Processes*, 23(5):646–688.

Van Berkum, J. J. A., Brown, C. M., Zwitserlood, P., Kooijman, V., and Hagoort, P. (2005). Anticipating upcoming words in discourse: Evidence from ERPs and reading times. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31:443–467.

Van Dyke, J. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology. Learning, Memory, and Cognition*, 33(2):407–430.

Van Dyke, J. and Lewis, R. L. (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language*, 49:285–316.

Van Dyke, J. and McElree, B. (2011). Cue-dependent interference in comprehension. *Journal of Memory and Language*, 65(3):247–263.

Van Dyke, J. A. and McElree, B. (2006). Retrieval interference in sentence comprehension. *Journal of Memory and Language*, 55(2):157–166.

Van Gompel, R. P. and Majid, A. (2004). Antecedent frequency effects during the processing of pronouns. *Cognition*, 90:255–264.

Vasishth, S., Bruessow, S., Lewis, R. L., and Drenhaus, H. (2008). Processing polarity: How the ungrammatical intrudes on the grammatical. *Cognitive Science*, 32:685–712.

Vasishth, S., Chen, Z., Li, Q., and Guo, G. (2013). Processing Chinese relative clauses: Evidence for the subject-relative advantage. *PLoS One*, 8(10):e77006.

Vasishth, S. and Drenhaus, H. (2011). Locality in German. *Dialogue and Discourse*, 1:59–82.

Vasishth, S. and Lewis, R. L. (2006). Argument-head distance and processing complexity: Explaining both locality and antilocality effects. *Language*, 82(4):767–794.

Vasishth, S., Suckow, K., Lewis, R. L., and Kern, S. (2010). Short-term forgetting in sentence comprehension: Crosslinguistic evidence from head-final structures. *Language and Cognitive Processes*, 25(4):533–567.

von der Malsburg, T. and Vasishth, S. (2012). Scanpaths reveal syntactic underspecification and reanalysis strategies. *Language and Cognitive Processes*, 28(10):1–34.

von der Malsburg, T. and Angele, B. (2015). The elephant in the room: False positive rates in standard analyses of eye movements in reading. Unpublished manuscript.

Wagers, M. W., Lau, E. F., and Phillips, C. (2009). Agreement attraction in comprehension: Representations and processes. *Journal of Memory and Language*, 61:206–237.

Watkins, O. C. and Watkins, M. J. (1975). Buildup of proactive inhibition as a cue-overload effect. *Journal of Experimental Psychology: Human Learning and Memory*, 104(4):442–452.

Wu, F. (2009). *Factors Affecting Relative Clause Processing in Mandarin: Corpus and Behavioral Evidence*. Phd thesis, University of Southern California, Los Angeles, CA.

Wu, F. (2011). Frequency issues of classifier configurations for processing Mandarin object-extracted relative clauses: A corpus study. *Corpus Linguistics and Linguistic Theory*, 7(2):203–207.

Wu, F. (2013). Will more heads increase processing difficulties? Evidence from Chinese. *Waiguoyu [Journal of Foreign Languages]*, 36(3):60–67.

Wu, F., Kaiser, E., and Andersen, E. (2010). Subject preference, head animacy, and lexical cues: A corpus study of relative clauses in Chinese. In Yamashita, H., Hirose, Y., and Packard, J., editors, *Processing and Producing Head-final Structures*, pages 173–194. Springer, Dordrecht, The Netherlands.

Wu, F., Kaiser, E., and Andersen, E. (2012). Animacy effects in Chinese relative clause processing. *Language and Cognitive Processes*, 27(10):1489–1524.

Wu, F., Luo, Y., and Zhou, X. (2014). Building Chinese relative clause structures with lexical and syntactic cues: Evidence from visual world eye-tracking and reading times. *Language, Cognition and Neuroscience*, 29(10):1205–1226.

Xiang, M., Dillon, B., and Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, 108(1):40–55.

Xu, L. (1993). The long-distance binding of *ziji*. *Journal of Chinese Linguistics*, 21:123–141.

Xue, N., Xia, F., Chiou, F.-D., and Palmer, M. (2005). The Penn Chinese TreeBank: Phrase structure annotation of a large corpus. *Natural Language Engineering*, 11(2):207–238.

Xue, P., Pollard, C., and Sag, I. (1994). A new perspective on Chinese *ziji*. In *Proceedings of the 13th Annual Meeting of the West Coast Conference on Formal Linguistics*, pages 432–447, Stanford, CA. CSLI Publications.

Yang, D.-W. (1983). The extended binding theory of anaphora. *Language Research*, 19:169–192.

Yoshida, M., Aoshima, S., and Phillips, C. (2004). Relative clause prediction in Japanese. In *17th Annual CUNY Conference on Human Sentence Processing*, College Park, MD. University of Maryland.

Yoshida, M. and Yoon, S. (2014). When are clause-final verbs facilitated in Korean? In *The 2nd East Asian Psycholinguistics Colloquium*, Chicago, IL. University of Chicago.

Yu, W. X. (1992). Challenging Chinese reflexive data. *The Linguistic Review*, 9(3):285–294.

Yu, X.-F. W. (1996). *A Study of Chinese Reflexives*. Phd thesis, University of London, London, UK.

Yun, J., Whitman, J., and Hale, J. T. (2010). Subject-object asymmetries in Korean sentence comprehension. In Ohlsson, S. and Catrambone, R., editors, *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society*, pages 2152–2157, Portland, OR. Cognitive Science Society.

Zaenen, A., Carletta, J., Garretson, G., Bresnan, J., Koontz-Garboden, A., Nikitina, T., O'Connor, M., and Wasow, T. (2004). Animacy encoding in English: why and how. In *Proceedings of the Association for Computational Linguistics Workshop on Discourse Annotation*, pages 118–125, Stroudsburg, PA. Association for Computational Linguistics.

Zhang, Q. and Jiang, H. (2010). Processing advantages in relative clauses and the local grammar complexity account. *Waiyu Yanjiu [Foreign Languages Research]*, 124(6):19–26.

Appendix

Appendix A

Retrieval interference in reflexives

A.1 Experiment 2. Pretest

Chosen antecedent of *ziji*	**Condition**		
	local	*non-local*	
local inanimate subj.	*n.a.*	15 (1.56%)	*sentence-internal antecedent*
non-local inanimate subj.	40 (4.17%)	*n.a.*	
cataphoric binding	40 (4.17%)	14 (1.46%)	
speaker	11 (1.15%)	4 (0.42%)	*sentence-external antecedent*
context	27 (2.81%)	7 (0.73%)	
author of media noun	19 (1.98%)	0	
recipient of media noun	2 (0.21%)	0	

Table A.1: Pretest of Experiment 2: Classification of the participants' answers in the 'incorrect' trials by experimental condition. Percentages refer to the total number of trials including the correct trials.

Appendix B

Teasing apart encoding and retrieval interference

B.1 Experiment 1

B.2 Experiment 2

B.3 Experiment 3

	Der/Die	*Dieb/-in*	*dem/der*	*der/die*	*Hehler/-in*	*befohlen*	*hat*	*zu stehlen*	*hat*	*überraschenderweise*	*sich*	*und die Kollegen*	*angezeigt*
gender-overlap - masc. ant.	448 (8)	547 (12)	460 (11)	412 (6)	505 (19)	523 (15)	499 (14)	669 (40)	525 (15)	482 (8)	444 (10)	1221 (16)	506 (13)
gender-overlap - fem. ant.	462 (12)	622 (17)	493 (10)	439 (8)	552 (16)	599 (28)	528 (16)	611 (24)	552 (19)	503 (13)	438 (6)	1233 (13)	530 (15)
no gender-overlap - masc. ant.	482 (15)	562 (22)	459 (8)	426 (10)	535 (14)	571 (17)	515 (17)	614 (23)	517 (15)	485 (9)	446 (8)	1224 (11)	535 (12)
no gender-overlap - fem. ant.	453 (10)	611 (15)	485 (9)	443 (9)	534 (13)	538 (12)	523 (19)	594 (22)	517 (16)	479 (9)	430 (6)	1218 (12)	539 (16)

TABLE B.1: Experiment 1. Means and standard errors of raw reading times in ms for each region by experimental condition. Between-participants variance has been removed using Cousineau (2005)'s normalization with Morey (2008)'s correction factor.

	Der/Die	*Dieb/-in*	*dem/der*	*der/die*	*Hehler/-in*	*befohlen*	*hat*	*zu stehlen*	*hat*	*überraschenderweise*	*sich*	*und die Kollegen*	*angezeigt*
FPRT													
gender-overlap - masc. ant.	257 (5)	351 (9)	288 (6)	275 (7)	333 (10)	344 (7)	272 (6)	431 (11)	301 (7)	350 (10)	286 (6)	246 (5)	348 (8)
gender-overlap- fem. ant.	244 (6)	481 (15)	338 (10)	293 (8)	372 (12)	347 (7)	273 (6)	424 (11)	295 (7)	357 (10)	300 (6)	251 (5)	341 (6)
no gender-overlap - masc. ant.	250 (5)	364 (13)	296 (6)	257 (8)	402 (11)	341 (7)	270 (6)	417 (10)	295 (8)	360 (9)	292 (5)	252 (6)	356 (7)
no gender-overlap - fem. ant.	243 (5)	468 (15)	342 (10)	308 (9)	314 (10)	344 (7)	283 (7)	421 (11)	305 (7)	351 (9)	299 (6)	252 (5)	347 (6)
RPD													
gender-overlap - masc. ant.	NA	408 (11)	336 (10)	379 (13)	434 (16)	389 (12)	384 (29)	748 (33)	373 (25)	487 (21)	375 (18)	348 (19)	449 (17)
gender-overlap - fem. ant.	NA	581 (17)	408 (12)	448 (16)	639 (22)	472 (20)	396 (27)	830 (43)	433 (32)	530 (27)	362 (14)	327 (12)	451 (19)
no gender-overlap - masc. ant.	NA	415 (13)	337 (9)	326 (12)	501 (15)	413 (13)	337 (15)	677 (28)	387 (28)	471 (16)	350 (15)	330 (11)	463 (18)
no gender-overlap - fem. ant.	NA	563 (16)	421 (16)	478 (16)	578 (22)	472 (19)	383 (24)	753 (38)	413 (37)	450 (17)	375 (15)	358 (19)	438 (15)
FPRP													
gender-overlap - masc. ant.	NA	15 (1)	7 (1)	22 (2)	17 (2)	6 (1)	12 (1)	28 (2)	8 (1)	20 (2)	11 (1)	15 (1)	14 (1)
gender-overlap. - fem. ant.	NA	19 (2)	13 (1)	26 (2)	29 (2)	12 (1)	11 (1)	28 (2)	10 (1)	18 (2)	10 (1)	15 (1)	14 (1)
no gender-overlap - masc. ant.	NA	17 (2)	8 (1)	16 (1)	15 (1)	10 (1)	10 (1)	26 (2)	7 (1)	17 (2)	9 (1)	14 (1)	13 (1)
no gender-overlap - fem. ant.	NA	21 (2)	13 (1)	30 (2)	32 (2)	13 (1)	10 (1)	29 (2)	9 (1)	16 (2)	11 (1)	16 (1)	13 (1)
TFT													
gender-overlap - masc. ant.	346 (10)	542 (16)	464 (14)	406 (16)	514 (19)	549 (16)	352 (11)	663 (18)	355 (12)	554 (16)	365 (11)	309 (8)	462 (11)
gender-overlap - fem. ant.	370 (11)	838 (25)	603 (21)	495 (19)	713 (23)	580 (15)	351 (12)	662 (18)	342 (12)	543 (17)	351 (9)	308 (9)	444 (11)
no gender-overlap - masc. ant.	336 (9)	535 (18)	473 (14)	324 (14)	618 (19)	530 (14)	346 (11)	617 (16)	353 (13)	528 (15)	358 (9)	306 (9)	465 (11)
no gender-overlap - fem. ant.	366 (10)	772 (23)	594 (20)	515 (19)	582 (22)	551 (16)	353 (11)	613 (17)	348 (10)	530 (17)	362 (9)	310 (8)	436 (10)

TABLE B.2: Experiment 2. Means and standard errors of raw first-pass reading time (FPRT), regression-path duration (RPD) and total-fixation time (TFT) in ms and first-pass regression probability (FPRP) in percentages for each region by experimental condition. From continuous dependent variables, between-participants variance has been removed using Cousineau (2005)'s normalization with Morey (2008)'s correction factor.

	Åke *Åke*	*säger* *som*	*att* *Alf/Ann*	*Alf/Ann* *tackade*	*jobbade med* *ringer*	*hans* *sina*	*sysslingar* *sysslingar*	*på helgerna.* *på kvällen.*
FPRT								
Pron - gender-overlap	381 (10)	403 (14)	257 (9)	292 (10)	421 (13)	287 (6)	327 (10)	650 (21)
Pron - no gender-overlap	393 (11)	388 (12)	277 (11)	288 (10)	393 (11)	285 (6)	334 (11)	673 (22)
Refl - gender-overlap	405 (11)	286 (10)	295 (10)	503 (17)	391 (12)	310 (11)	308 (11)	673 (21)
Refl - no igender-overlap	397 (12)	293 (10)	300 (9)	512 (16)	397 (13)	297 (8)	327 (12)	659 (24)
RPD								
Pron - gender-overlap	NA	553 (21)	421 (26)	432 (26)	505 (21)	400 (22)	549 (42)	3636 (133)
Pron - no gender-overlap	NA	513 (18)	405 (34)	413 (24)	521 (26)	421 (26)	525 (32)	3370 (115)
Refl - gender-overlap	NA	443 (20)	435 (22)	669 (31)	749 (42)	500 (37)	656 (56)	4120 (163)
Refl - no gender-overlap	NA	443 (19)	455 (25)	659 (28)	738 (40)	568 (49)	643 (55)	3882 (149)
FPRP								
Pron - gender-overlap	NA	19 (2)	20 (2)	16 (2)	11 (2)	16 (2)	17 (2)	80 (2)
Pron - no gender-overlap	NA	18 (2)	16 (2)	15 (2)	13 (2)	16 (2)	24 (2)	81 (2)
Refl - gender-overlap	NA	27 (2)	19 (2)	15 (2)	31 (2)	19 (2)	28 (2)	81 (2)
Refl - no gender-overlap	NA	25 (2)	21 (2)	14 (2)	31 (2)	22 (2)	26 (2)	79 (2)
TFT								
Pron - gender-overlap	856 (31)	1121 (40)	444 (23)	713 (29)	1010 (36)	601 (21)	685 (24)	1067 (30)
Pron - no gender-overlap	799 (33)	1047 (33)	468 (25)	634 (25)	917 (30)	579 (19)	675 (24)	1128 (34)
Refl - gender-overlap	1104 (46)	781 (32)	930 (38)	1412 (45)	1039 (34)	544 (20)	660 (24)	1133 (35)
Refl - no gender-overlap	991 (34)	777 (29)	890 (36)	1398 (49)	1051 (36)	532 (20)	663 (24)	1130 (38)

TABLE B.3: Experiment 3. Means and standard errors of raw first-pass reading time (FPRT), regression-path duration (RPD) and total-fixation time (TFT) in ms and first-pass regression probability (FPRP) in percentages for each region by experimental condition. From continuous dependent variables, between-participants variance has been removed using Cousineau (2005)'s normalization with Morey (2008)'s correction factor.

Appendix C

The subject-relative advantage in Chinese

C.1 Corpus study of Chinese Treebank 7.0

Structure following *zhe/na+Cl+Adv*		**Count**	**Frequency**
SR	*V (N) de N*	58	79.5%
OR (dropped subject)	*pro V de N*	13	17.8%
OR (overt subject)	*N V de N*	1	1.4%
Main clause VP	V N	1	1.4%

Table C.1: Corpus counts of structures following a zhe/na (*this/that*)+classifier+adverb sequence. All relative clause tokens are headed relative clauses. Headless relative clauses in this position have zero tokens in the corpus.

Noun modified by *zhe/na+Cl*	**Structure following *zhe/na+Cl+Adv***		**Count**	**Relative frequency**
Subject	SR	*V (N) de N*	31	77.5%
	OR (dropped subject)	*pro V de N*	8	20.0%
	OR (overt subject)	*N V de N*	0	0%
	*M*ain clause VP	V (N)	1	2.5%
Object	SR	*V (N) de N*	17	85.0%
	OR (dropped subject)	*pro V de N*	2	10.0%
	OR (overt subject)	*N V de N*	1	5.0%
	*M*ain clause VP	V (N)	0	0%
Other	SR	*V (N) de N*	10	76.9%
	OR (dropped subject)	*pro V de N*	3	23.1%
	OR (overt subject)	*N V de N*	0	0%
	*M*ain clause VP	V (N)	0	0%

Table C.2: Categorization of the corpus tokens presented in Table C.1 (structures following a zhe/na (*this/that*)+classifier+adverb sequence) by the syntactic role (subject, object or other) of the noun which the zhe/na+classifier phrase modifies. All relative clause tokens are headed relative clauses. Headless relative clauses in this position have zero tokens in the corpus.

Noun modified by *zhe/na+Cl*	**Structure following *zhe/na+Cl***	**Count**	**Relative frequency**
Subject	RC	163	13.9%
	other	1007	86.1%
Object	RC	95	12.3%
	other	677	87.7%

TABLE C.3: Corpus counts of structures following a zhe/na (*this/that*)+classifier sequence, categorized by whether the zhe/na+classifier phrase modifies the subject or the object of the main clause.

C.2 SENTENCE COMPLETION DATA

Sentence completion			**Count (subj-mod)**	**Count (obj-mod)**
RC	SR	*canonical*	27	48
		adjectival	0	0
		*passive (*bei*)*	2	2
		topicalized	1	2
	OR	*canonical*	3	9
		subj. drop	2	2
	Possessive RC		0	2
	Adjunct RC		0	0
not RC			117	87

TABLE C.4: Sentence completions produced after a Det+Cl sequence (subject-modification) and after an NP+V+Det+Cl (object-modification) sequence observed in the sentence completion study. Canonical RCs are defined as RCs with the canonical SVO word order and an overt RC subject in ORs.

Sentence completion			**Count (subj-mod)**	**Count (obj-mod)**
RC	SR	*canonical (headed)*	100	121
		canonical (headless)	1	0
		adjectival	1	3
		*passive (*bei*)*	5	3
		topicalized	0	1
	OR	*canonical (headed)*	3	8
		canonical (headless)	0	0
		subj. drop (headed)	16	8
		subj. drop (headless)	1	0
	Possessive RC		2	1
	Adjunct RC		1	0
not RC	diff. PoS for Adv		3	3
	elided NP betw. Det+Cl and Adv		15	0

TABLE C.5: Sentence completions produced after a Det+Cl+Adv sequence (subject-modification) and after an NP+V+Det+Cl+Adv sequence (object-modification) observed in the sentence completion study. Canonical RCs are defined as RCs with the canonical SVO word order and an overt RC subject in ORs.

Appendix D

The Mandarin reflexive *ziji*

The gender and number neutral Mandarin Chinese[1] reflexive *ziji* ('myself', 'yourself', 'himself', 'herself', 'itself', 'ourselves', 'yourselves', 'themselves')[2] has several cross-linguistically rather uncommon properties that have attracted considerable attention in the syntactic literature over the past three decades. There are a wide range of competing syntactic or pragmatic approaches of how to analyze *ziji*. I will first describe the syntactic properties of *ziji* independently from any syntactic framework and subsequently give a brief overview of the most prominent syntactic analyses that have been proposed so far. All Chinese example sentences will be provided in Simplified Chinese characters (i.e., the official script in mainland China) together with their Pinyin transcription (i.e., the official system of the Peoples Republic of China to phonetically transcribe Chinese characters to their Mandarin pronunciations using Latin characters). Note that virtually none of the publications from where I am citing Chinese example sentences provide Chinese characters and in some of the publications, the Latin transcriptions were not completely conform with the official Pinyin. Therefore, I have added Chinese characters to all sentences and in some examples also changed the Latin transcription.

D.1 The distributional properties of the bare reflexive *ziji*

Similar to many other languages including English, the Chinese reflexive *ziji* needs to be c-commanded by its antecedent.[3] However, the antecedent is not required to be contained in the same clause as the reflexive (see D.1) which poses a problem for Binding Principle A as originally formulated in (Chomsky, 1981).

[1] In the following, I will use the term 'Chinese' as synonym for 'Mandarin Chinese'.

[2] In addition to the mono-morphemic reflexive *ziji*, there is also a bi-morphemic reflexive consisting of pronoun+*ziji* in Chinese. The binding properties of these two forms differ in various ways. For the current purposes, I will only focus on the bare reflexive *ziji*.

[3] C-command refers to a tree-configurational relation between two nodes of a parse-tree which is defined as follows: A node A c-commands a node B if and only if A does not dominate B and the lowest branching node which dominates A also dominates B (Haegeman, 1994).

(D.1) *C-commanding antecedent*

男孩	觉得	母亲	不	在	家	的	时候	姐姐	应该
Nanhai$_i$	juede	muqin$_j$	bu	zai	jia	de	shihou	jiejie$_k$	yinggai
boy	think	mother	not	at	home	DE	time	elder sister	shall

照顾	好	自己。
zhaogu	hao	ziji$_{i/*j/k}$.
take care of	well	ziji

The boy thinks that his elder sister should take good care of him/herself when [their] mother is not at home.

In D.1, *ziji* can be either locally bound by *jiejie* ('elder sister') or long-distance bound by *nanhai* ('boy') since both c-command the reflexive. In contrast, *muqin* ('mother') does not qualify as an antecedent because it does not c-command the reflexive.

Another difference to languages like English is the subject-orientation of Chinese reflexives: No matter whether *ziji* is locally or long-distance bound, its antecedent is required to be a subject (Huang, 1984; see D.2 for an example taken from Huang and Liu, 2001).

(D.2) *Subject-orientation*

张三	送	给	李四	一	张	自己	的	相片。
Zhangsan$_i$	song	gei	Lisi$_j$	yi	zhang	ziji$_{i/*j}$	de	xiangpian.
Zhangsan	give	to	Lisi	one	CL	self	DE	picture

Zhangsan gave Lisi a picture of himself.

Not being a subject, *Lisi* is not a grammatical antecedent in spite of c-commanding the reflexive, leaving *Zhangsan* as the only possible antecedent.

Another cross-linguistically uncommon property of *ziji*, which is exploited in the designs of the experiments presented in Chapter 2 of this thesis, is that the antecedent of the reflexive is required to be animate (Tang, 1989).

(D.3) *Non-local animate antecedent*

男孩	说	这	篇	文章	批评	了	自己。
Nanhai$_i$	shuo	zhe	pian	wenzhang$_j$	piping	le	ziji$_{i/*j}$.
boy	say	this	CL	article	criticize	ASP	self

The boy says that this article criticized him.

(D.4) *Local animate antecedent*

这	篇	文章	说	男孩	批评	了	自己。
Zhe	pian	wenzhang$_i$	shuo	nanhai$_j$	piping	le	ziji$_{*i/j}$.
this	CL	article	say	boy	criticize	ASP	self

This article says that the boy criticized himself.

(D.5) *Local and non-local animate antecedents*

男孩	说	姐姐	批评	了	自己。
Nanhai$_i$	shuo	jiejie$_j$	piping	le	ziji$_{i/j}$.
boy	say	elder sister	criticize	ASP	self

The boy says [his] elder sister criticized him/herself.

D.3 and D.4 each have only one animate subject, which results in unambiguous sentences. D.5, however, has two animate subjects *nanhai* ('boy') and *jiejie* ('elder sister'), both c-commanding *ziji* and thus candidates for being its antecedent. This leads to a globally ambiguous sentence.

Although c-command, long-distance binding, subject-orientation and animacy of the antecedent appear to be the basic properties of *ziji* that are widely cited in the literature, the picture is even more complicated, since in certain syntactic or pragmatic contexts, there are exceptions to all of these phenomena. I will go through each of them in turn.

D.1.1 Exceptions to the c-command constraint

D.1.1.1 Sub-commanding antecedents

Tang (1989) first noted that under certain conditions, *ziji* can be bound by a non-c-commanding antecedent but by a noun phrase (NP) contained in a c-commanding phrase, a so-called sub-commanding noun-phrase. In case the c-commanding noun phrase itself is inanimate but contains an animate NP, this animate NP can bind *ziji* as shown in D.6 (example from Tang, 1989).

(D.6) *Sub-commanding antecedent 1*

张三	的	骄傲	害	了	自己。
[[Zhangsan$_i$	de]	jiaoao]$_j$	hai	le	ziji$_{i/*j}$.
Zhangsan	DE	pride	hurt	ASP	self

Zhangsan's pride hurt him.

In D.6, the lacking animacy feature of the c-commanding subject *Zhangsan de jiaoao* ('Zhangsan's pride') disqualifies the latter as antecedent. Therefore the sub-commanding animate NP *Zhangsan* is allowed to be the antecedent of *ziji*. Crucially, even in the sub-command configuration, the subject-orientation of *ziji* still holds, as can be seen from D.7 and D.8 (examples from Tang, 1989). In D.7, *Zhangsan* is a subject sub-commanding the reflexive: it is the subject of an embedded clause contained in the c-commanding NP *Zhangsan tou dongxi de shishi* ('the fact that Zhangsan stole things') that modifies the inanimate head *shishi* ('fact') of this NP. Therefore it is a grammatical antecedent of *ziji*. Similarly, in D.8, the third person pronoun *ta* ('he'), which is also the subject of an embedded clause contained in a c-commanding NP, can bind *ziji*. In contrast, although also being a sub-commander of the reflexive, the object of the embedded clause inside the c-commanding NP *ni* ('you') in D.8 is not a legal antecedent for the reflexive (Tang, 1989).

(D.7) *Sub-commanding antecedent 2*

张三	偷	东西	的	事实	被	自己	的	老板	发现
[[Zhangsan$_i$	tou	dongxi	de]	shishi]$_j$	bei	ziji$_{i/*j}$	de	laoban	faxian
Zhangsan	steal	things	DE	fact	PASS	self	DE	boss	discover
了。									
le.									
ASP									

The fact that Zhangsan stole things was discovered by his boss.

(D.8) *Sub-commanding antecedent 3*

他	打	了	你	的	证据	被	自己	的	爸爸	发现
[[Ta$_i$	da	le	ni$_j$	de]	zhengju]$_k$	bei	ziji$_{i/*j/*k}$	de	baba	faxian
he	beat	ASP	you	DE	evidence	PASS	self	DE	father	discover
了。										
le.										
ASP										

The evidence that he beat you was discovered by his father.

According to Tang (1989), in case the c-commanding subject lacks the animacy feature, only the structurally highest animate subject (or possessor) contained within this c-commanding subject qualifies as antecedent. In D.9 (example from Tang, 1989), two sub-commanding NPs, (*Zhangsan* and *Zhangsan de baba* 'Zhangsan's father') are present, but, being higher in the tree structure, only the latter is a legal antecedent.

(D.9) *Sub-commanding antecedent 4*[4]

张三	的	爸爸	的	钱	被	自己	的	朋友	偷走
[[[Zhangsan$_i$	de]	baba]$_j$	de	qian]$_k$	bei	ziji$_{*i/j/*k}$	de	pengyou	tou-zou
Zhangsan	DE	father	DE	money	PASS	self	DE	friend	stolen
了。									
le.									
ASP									

Zhangsan's father's money was stolen by his [the father's] friend.

Pollard and Xue (1998) noted that in a suitable pragmatic context, *ziji* can be also bound by a structurally lower sub-commanding antecedent even if an animate subject or possessor is available at a structurally higher position inside the c-commanding subject. Although when presented without a context, the sentence given in D.9 only allows the c-commanding subject (*Zhangsan de baba*) to bind the reflexive (Tang, 1989), according to Pollard and Xue (1998) the sub-commanding noun *Zhangsan* is the preferred antecedent when this sentence is presented in a context as the one provided in D.10 (example from Pollard and Xue, 1998).

[4]The Pinyin transcription of this example has been slightly changed.

(D.10) *Sub-commanding antecedent 5*[5]

张三	的	爸爸	的	钱	被	自己	的	朋友	偷走	了。
[[[Zhangsan$_i$	de]	baba]	de	qian]	bei	ziji$_i$	de	pengyou	touzou	le.
Zhangsan	DE	father	DE	money	PASS	self	DE	friend	stolen	ASP

Zhangsan's father's money was stolen by his [Zhangsan's] friend.

妈妈	的	书	也	被	自己	的	朋友	偷走	了。
[[Mama$_i$	de]	shu]$_j$	ye	bei	ziji$_{*i/j/*k}$	de	pengyou	touzou	le.
mother	DE	book	also	PASS	self	DE	friend	stolen	ASP

[His] mother's book was also stolen by his [Zhangsan's] friend.

他	急	得	哭	起来。
Ta	ji	de	ku	qilai.
He	worry	such that	cry	start

He worried so much that he started crying.

In contrast to Huang and Tang (1991) who had claimed that a sub-commanding antecedent is only possible in case the reflexive is locally bound but not if the sub-commanding NP is contained in a non-local c-commander, Xue et al. (1994) provided examples for a non-local sub-commanding antecedent (see D.11, example from Xue et al., 1994).

(D.11) *Non-local sub-commanding antecedent*[6]

张三	的	信	表明	那	本	书	害	了	自己。
[[Zhangsan$_i$	de]	xin]$_j$	biaoming	na	ben	shu$_k$	hai	le	ziji$_{i/*j/*k}$.
Zhangsan	DE	letter	indicate	this	CL	book	harm	ASP	self

Zhangsan's letter indicates that the book harmed him.

D.1.1.2 Antecedents outside the current sentence

Besides the sub-command configuration, there is another case in which *ziji* can co-refer with a non-c-commanding entity. Li (1991) provided an example for a sentence-free unbound instance of *ziji* (see D.12).

(D.12) *Unbound* ziji[7]

自己	能	去	那儿	吗?
Ziji	neng	qu	nar	ma?
self	can	go	there	Q

Can I go there?

[5]The English glosses of this Chinese sentence have been slightly changed.

[6]Xue et al. (1994) mark the co-indexation of *Zhangsan* and *ziji* with a question mark. Pan (2000), however, removes the question mark when citing this sentence implying that the co-indexation of the sub-commanding noun *Zhangsan* with the reflexive is fully grammatical. Note also that, in contrast to Xue et al. (1994), Pan (2000) also allows the co-indexation of the inanimate local subject *shu* ('book') with the reflexive, which is likely to be a typo.

[7]The English glosses of this Chinese sentence have been slightly changed.

According to Li (1991), an unbound instance of *ziji* automatically refers to the speaker of the utterance. However, Pan (2000) pointed out that in D.12, *ziji* may also refer to the addressee or even to a third entity salient in the discourse. Furthermore, he noted that *ziji* can generally refer to the speaker even if it is in object position and not only if it is in subject position as suggested by Li (1991).

D.1.1.3 Cataphoric binding

Huang and Liu (2001) indicated that there is another exception to the c-command constraint. In case the reflexive is part of an adjunct clause that precedes the matrix clause, it can be cataphorically bound by the subject of this matrix clause even though it is not its c-commandee (see D.13, example from Huang and Liu, 2001).

(D.13) *Cataphoric antecedent*

因为	李四	批评	自己,	所以	张三	很	生气。
Yinwei	Lisi	piping	ziji$_i$,	suoyi	Zhangsan$_i$	hen	shengqi.
because	Lisi	criticize	self	so	Zhangsan	very	angry

Because Lisi criticized him, Zhangsan was very angry.

D.1.2 Exceptions to subject-orientation

The subject-orientation of the reflexive *ziji* was first observed by Huang (1982). In later work, Huang and Tang (1991) noted that in the case of psychological verbs, not only the subject but also the experiencer argument can be an antecedent of *ziji* (see D.14, example from Huang and Tang, 1991). In addition to going against subject-orientation, the c-command constraint is violated in this configuration too. This kind of binding by a non-c-commanding experiencer argument is also possible when *ziji* is long-distance bound.

(D.14) *Experiencer argument of a psychological verb as antecedent*

自己	的	小孩	没	得	奖	的	消息	使	李四	很	难过。
[Ziji$_i$	de	xiaohai	mei	de	jiang	de	xiaoxi]	shi	Lisi$_i$	hen	nanguo.
self	DE	child	not	get	prize	DE	news	make	Lisi	very	sad

The news that his own child did not get a prize made Lisi very sad.

Another exception to both the subject-orientation of *ziji* and the c-command-constraint was first observed by Yu (1992) who noted that in passive constructions and so-called *ba*-constructions the antecedent can be the argument of a co-verb, more precisely, the argument of the passive marker *bei* (see D.15, example from Yu, 1992) or the pre-verbal object marker *ba* (see D.16, example from Yu, 1992), which constitutes not only a counter-example for the subject-orientation of *ziji* but also a violation of the c-command constraint.[8]

[8]Cole and Wang (1996) provided a formal analysis involving head-movement of the reflexive in which they claim that in *ba*- and *bei*-constructions, *ziji* is c-commanded by its antecedent at LF.

(D.15) *Argument of* bei *as antecedent*

书	被	李四	送给	自己	的	儿子	了。
Shu	bei	Lisi$_i$	song-gei	ziji$_i$	de	erzi	le.
book	BEI	Lisi	give-to	self	DE	son	ASP

The book was given by Lisi to his own son.

(D.16) *Argument of* ba *as antecedent*[9]

菲欧	把	女儿	送进	自己	的	房	里。
Feiou	ba	nüer$_i$	song-jin	ziji$_i$	de	fang	li.
Feiou	BA	daughter	send-into	self	DE	room	inside

Feiou sent [her] daughter into her own room.

D.1.3 Blocking of long-distance binding

Huang (1984) first noted that *ziji* can be long-distance bound (i.e., from outside the clause containing the reflexive). Tang (1989) showed that only remotest third person NPs can be an antecedent of *ziji*, non-local binding by a third person NP intervening between the local and the structurally highest clause appears to not be possible (see D.17, from Tang, 1989).

(D.17) *Long-distance binding only to remotest antecedent*

张三	知道	李四	觉得	王五	对	自己	没	信心。
Zhangsan$_i$	zhidao	[Lisi$_j$	juede	[Wangwu$_k$	dui	ziji$_{i/*j/k}$	mei	xinxin]].
Zhangsan	know	Lisi	think	Wangwu	to	self	no	confidence

Zhangsan knew that Lisi thought that Wangwu had no confidence in himself/him [Zhangsan].

While Huang and Tang (1991) had claimed that only c-commanding subjects and experiencer arguments of psychological verbs qualify as long-distance antecedents for *ziji*, Xue et al. (1994) provided examples for sub-commanding long-distance antecedents (see D.11).

Huang (1984) also first observed that the long-distance binding of *ziji* can be blocked by certain kinds of NPs intervening between the antecedent and the reflexive. In early work on Chinese reflexives, it was proposed that any intervening c-commanding subject NP with phi-features different from a potential long-distance antecedent blocked binding of *ziji* by the latter (Huang, 1984; Tang, 1989) (see D.18, example from Tang, 1989).

(D.18) *Blocking of long-distance binding by an intervening subject*

张三	觉得	我／你	对	自己	没	信心。
Zhangsan$_i$	juede	[wo/ni$_j$	dui	ziji$_{*i/j}$	mei	xinxin].
Zhangsan	think	I/you	to	self	no	confidence

Zhangsan thought that I/you had no confidence in myself/yourself.

[9] Although not mentioned by Yu (1992), in this example co-indexation of the reflexive with the subject *Feiou* is also grammatical.

Huang and Tang (1991) provided evidence that not only c-commanding subjects induce blocking effects on long-distance binding. A sub-commanding NP intervening between the reflexive and a remote c-commanding subject also blocks long-distance binding by the latter in case it displays person features different from that remote NP (see D.19, example from Huang and Tang, 1991).

(D.19) *Blocking by an intervening sub-commanding NP*

张三	说	我	的	骄傲	害	了	自己。
Zhangsan$_i$	shuo	[wo$_j$	de	jiaoao	hai	le	ziji$_{*i/j}$].
Zhangsan	say	I	DE	pride	hurt	ASP	self

Zhangsan said that my pride hurt myself.

In Example D.19, the long-distance binding by the third person NP *Zhangsan* is blocked by the intervening first person pronoun *wo* ('I') that sub-commands the reflexive.

Moreover, Huang and Tang (1991) also showed that an experiencer non-subject also blocks long-distance binding of *ziji* (see D.20, example from Huang and Tang, 1991).

(D.20) *Blocking by an experiencer non-subject*

张三	对	自己	没	信心	的	事	使	我	很
[[Zhangsan$_i$	dui	ziji$_{i/*j/*k}$	mei	xinxin	de	shi]	shi	wo$_i$	hen
Zhangsan	to	self	no	confidence	DE	fact	make	me	very

难过	的	消息	使	李四	很	意外。
nanguo	de	xiaoxi]	shi	Lisi$_k$	hen	yiwai.
sad	DE	news	make	Lisi	very	surprised

The news that I was saddened by the fact that Zhangsan had no confidence in himself surprised Lisi.

Example D.20 is a double embedding with an experiencer argument in the matrix clause (*Lisi*) and another experiencer argument in the first embedding (*wo* 'me'). The long-distance binding of *ziji* by the structurally highest experiencer argument *Lisi* is blocked by the intervening first person experiencer argument *wo*. Crucially, the blocker *wo* is not a subject. From these data Huang and Tang (1991) concluded that the set of potential blockers of a certain long-distance binder is the set of potential local or less remote antecedents. Long-distance binding is not blocked if and only if the non-local antecedent in question agrees in phi-features with all closer potential antecedents. However, Xue et al. (1994) noticed that direct objects and obliques (i.e., NPs that do not qualify as potential antecedents) can also induce blocking effects (see D.21, example from Xue et al., 1994).

(D.21) *Blocking by an object*

张三	告诉	我	李四	恨	自己。
Zhangsan$_i$	gaosu	wo$_j$	[Lisi$_k$	hen	ziji$_{*i/*j/k}$].
Zhangsan	tell	me	Lisi	hate	self

Zhangsan told me that Lisi hates himself.

In this example, *wo* ('me'), not being a subject, is not a potential antecedent but nevertheless blocks binding by the matrix subject *Zhangsan*.

Pan (2000) provided a thorough analysis of blocking effects induced on long-distance binding of *ziji*. He demonstrated that there is an asymmetry in the blocking effect induced by non-matching person features: while intervening 1st and 2nd person NPs can block long-distance binding by a 3rd person NP (see D.18), an intervening 3rd person NP does not block binding by non-local 1st or 2nd person NP (see D.22, example from Pan, 2000).

(D.22) *Person asymmetry in the blocking effect*

我	知道	李四	不	喜欢	自己。
Wo$_i$	zhidao	[Lisi$_j$	bu	xihuan	ziji$_{i/j/}$].
I	know	Lisi	not	like	self

I knew that Lisi did not like me/himself.

Tang (1989) had already noted an asymmetry in the blocking effect induced by non-matching number features (cf. Xu, 1993). Tang noted that while a plural local subject does not block a long-distance singular antecedent (see D.23, example from Tang, 1989), a local singular subject blocks binding by a remote plural antecedent (see D.24, example from Tang, 1989).

(D.23) *No blocking of a singular long-distance antecedent by a plural local NP*

张三	说	他们	批评	了	自己。
Zhangsan$_i$	shuo	tamen$_j$	piping	le	ziji$_{i/j}$.
Zhangsan	say	they	criticize	ASP	self

Zhangsan said that they criticized him/themselves.

(D.24) *Blocking of a plural long-distance antecedent by a singular local NP*

他们	说	张三	批评	了	自己。
Tamen$_i$	shuo	Zhangsan$_j$	piping	le	ziji$_{*i/j}$.
they	say	Zhangsan	criticize	ASP	self

They said that Zhangsan criticized himself.

Huang (2002) noted that the blocking of a plural long-distance antecedent induced by a local singular subject can be overcome if forcing a distributive reading of the matrix predicate by adding the distributive marker *ge* ('each') or *dou* ('all') (see D.25, example from Huang, 2002).

(D.25) *Distributive reading of non-locally bound* ziji

他们	都	说	张三	批评	了	自己。
Tamen$_i$	dou	shuo	Zhangsan$_j$	piping	le	ziji$_i$/j.
They	all	say	Zhangsan	criticize	ASP	self

Every one of them said that Zhangsan criticized him/himself.

Therefore, Huang (2002) explained the sensitivity of *ziji* to number mismatch not with blocking effects but rather with certain properties of Chinese reflexives regarding distributivity. In English,

sentences like *John and Mary criticized themselves* are ambiguous between a distributive reading in which each John criticizes himself and Mary criticizes herself, and a collective reading in which they both criticize their collective self saying, e.g., 'we are wrong'. In contrast, in the case of the Chinese *ziji* only the distributive reading is possible. In a sentence like D.26 only the reading according to which each *Lisi* is criticizing himself and *Zhangsan* is criticizing himself is available (Huang, 2002).

(D.26) *Distributive interpretation of a plural antecedent in the case of local* ziji

张三	和	李四	在	批评	自己。
Zhangsan	he	Lisi	zai	piping	ziji.
Zhangsan	and	Lisi	at	criticize	self

Zhangsan and Lisi are criticizing themselves.

While in the case of locally bound *ziji* it appears to be not necessary to overtly mark the distributive interpretation of the plural antecedent, in the case of long-distance bound *ziji*, the distributive reading of the antecedent has to be overtly marked by inserting a distributive marker such as *ge* ('each') or *dou* ('all'). Therefore, the long-distance binding is possible in D.27. Huang (2002) further provided evidence against the claim that mismatching number induces blocking effects by comparing sentences with a local and a long-distance antecedent, both being plural, with and without a plural marker for the remote antecedent. He shows that, as expected under his account, the long-distance binding of *ziji* is indeed only possible if the remote subject is overtly marked for distributivity. Hence, the long-distance binding is possible in D.27 (example from Huang, 2002), in which the matrix subject is followed by the distributive marker *dou* ('all'), but not in D.28 (example from Huang, 2002), in which the long-distance antecedent is not overtly marked for distributivity.

(D.27) *Plural long-distance antecedent with a distributive marker*

他们	都	说	他们	常	批评	自己。	
Tamen$_i$	dou	shuo	tamen$_j$	chang	piping	ziji$_{i/j}$.	(i≠j)
They	all	say	they	often	criticize	self	

Each of them said that they often criticize him/themselves.

(D.28) *Plural long-distance antecedent without a distributive marker*

他们	说	他们	常	批评	自己。	
Tamen$_i$	shuo	tamen$_j$	chang	piping	ziji$_{*i/j}$.	(i≠j)
They	say	they	often	criticize	self	

They$_i$ said that they$_j$ often criticized themselves$_{*i/j}$.

To conclude, in Example D.28 the local and and the remote antecedent share the same number feature, therefore the unavailability of a long-distance binding cannot be attributed to a blocking effect induced by number mismatch (Huang, 2002).

D.1.4 Exceptions to the animacy constraint

Already Tang (1989) had pointed out that the reflexive *ziji* requires its antecedent to be animate. Although it is widely agreed on in the literature that the animacy constraint imposed on *ziji*'s antecedent is a lexical property inherent to the reflexive (e.g., Tang, 1989; Xue et al., 1994; Huang and Liu, 2001), it has also been proposed that the animacy constraint results from the context rather than being a property inherent to the reflexive itself (Pan, 1995). However, there appear to be some exceptions to the animacy constraint: Tang (1989) indicated that if *ziji* is not used as an anaphor but as an intensifier (similar to the use of the English *himself* in a sentence like *He himself wrote the letter*), there is no animacy constraint imposed on the antecedent; the antecedent of emphatic *ziji* can be an animate or inanimate (concrete or abstract) NP. In the case of anaphoric *ziji*, according to Tang (1989), the only exception where an inanimate antecedent is allowed is the metaphoric use of an inanimate NP. In D.29 (example from Tang, 1989) the inanimate noun *yueliang* ('moon') can be the antecedent of *ziji* because in the given context it is understood as a personification.

(D.29) *Inanimate antecedent: Metaphoric extension*

月亮	拿	乌	云	来	遮盖	自己。
Yueliang$_i$	na	wu	yun	lai	zhegai	ziji$_i$.
moon	take	dark	cloud	come	cover	self

The moon covered herself with dark clouds.

In contrast to Tang (1989), Pan (1995) proposed that animacy is not a requirement *ziji* imposes on its antecedent. He provided several examples in which the antecedent of *ziji* is clearly inanimate (see D.30-D.32, examples taken from Pan, 1995).

(D.30) *Inanimate antecedent: organisation/group of humans?*[10,11]

中国	共产党	还	能	靠	谁?	要	创造
[Zhongguo	gongchandang]$_i$	hai	neng	kao	shei?	Yao	chuangzao
Chinese	communist party	still	can	depend	who	want	create

共产党	的	未来,	恐怕	也	只	能	靠	它自己。
gongchangdang	de	weilai,	kongpa	ye	zhi	neng	kao	[ta-ziji]$_i$.
communist party	DE	future	perhaps	also	only	can	depend	itself

Who else can the Chinese Communist Party depend on? Perhaps it can only depend on itself to create its future.

[10] In the Pinyin transcription of this sentence *chuangchao* has been corrected to *chuangzao*.

[11] Pan (1995) does not provide Chinese characters for this example sentence. Consultation with native speakers revealed that the bi-morphemic reflexive *ta-ziji* is preferable written as 他自己 ('himself') rather than 它自己 ('itself'). This indicates that speakers of Chinese perceive the antecedent *Zhongguo gongchandang* as animate rather than inanimate.

(D.31) *Inanimate antecedent:* ziji *followed by the possessive marker* DE[12]

但	见	一	道	精巧	的	白	石	拱桥	在	静止	的
Dan	jian	[yi	dao	jingqiao	de	bai	shi	gongqiao]$_i$	zai	jingzhi	de
but	see	one	CL	beautiful	DE	white	stone	arch-bridge	at	still	DE

水面	上	投下	自己	的	倒影。
shuimian	shang	tou-xia	ziji$_i$	de	daoying.
water surface	on	throw down	self	DE	reverse shape

Suddenly I saw a beautiful white stone arch-bridge throwing its own mirror image on the water.

(D.32) *Inanimate antecedent:* ziji *followed by possessive marker* DE

每	一	个	公园	都	有	自己	的	冬天。
[Mei	yi	ge	gongyuan]$_i$	dou	you	ziji$_i$	de	dongtian.
every	one	CL	park	all	have	self	DE	winter

Every park has its own winter.

Although the data provided by Pan (1995) challenge the claim that *ziji* requires an animate antecedent, a closer investigation of the examples puts into question the generalizability of Pan (1995)'s claim. In Example D.30, the compound reflexive *ta-ziji*, which besides is known to behave differently from the bare reflexive *ziji*, is anteceded by *Zhongguo gongchandang* ('Chinese Communist Party'). If one considers the Communist Party as referring to the whole of its human members, the animacy requirement could be considered as fulfilled. This analysis is in line with Zaenen et al. (2004)'s proposal to categorize organizations as animate referents.

In the other examples provided by Pan (1995), *ziji* is not in argument position but — as it is followed by the possessive marker DE — in specifier position. The sequence *ziji*+DE corresponds to a possessive pronoun like *his* in English. Pan (1995) does not provide any example in which the bare reflexive *ziji* is in argument position and is bound by an inanimate NP. This is crucial because Tang (1989) had already noted that only *ziji* in argument position can be regarded as an anaphoric instance of *ziji*; if *ziji* occurs in a non-argument position it should be considered as an instance of intensifying *ziji* which allows inanimate antecedents.

D.2 Proposed analyses of *ziji*

There are a wide range of different analyses of Chinese reflexives which can be classified into purely formal approaches on the one hand and functional accounts that try to explain long-distance bound *ziji* with pragmatic factors on the other hand.

[12]In the original example provided by Pan (1995), it says *shuimiao* instead of *shuimian* and *jinzhi* instead of *jingzhi*. I believe that both *shuimiao* and *jinzhi* are typographical errors in Pan (1995), but alternatives cannot be fully excluded since Pan does not provide the Chinese characters. Moreover, the English glosses of this Chinese example have been slightly changed.

D.2.1 Formal accounts

Two major lines of purely syntactic approaches to long-distance bound reflexives in general and to Chinese *ziji* in particular can be distinguished. The first class of account explains long-distance binding assuming a parametrization of the binding domain: While in English the binding domain is, roughly speaking, the clause, in Chinese it is the whole sentence (Yang, 1983; Manzini and Wexler, 1987). In the more recent literature, this approach has been largely abandoned. The second line of analyses does not assume a parametrization of the binding domain but rather regards long-distance reflexives as being covertly local in nature (Cole and Sung, 1994). Although all the approaches belonging to this line of research share their basic claim — the inherent local nature of long-distance bound *ziji* —, they differ substantially in the mechanisms they assume to connect *ziji* with its antecedent. In the following I will briefly list the most prominent of these accounts.

Tang (1989) proposed to account for the possibility of long-distance bound *ziji* by assuming language-specific Feature-Copying and Re-indexing Rules. Under this account, *ziji* is base generated as pro-*ziji*, i.e., as a compound reflexive with an empty pro. In the case of long-distance bound *ziji*, this pro transfers its phi-features to -*ziji* after Binding Theory has applied, which results in long-distance binding of the reflexive (optional Feature-copying Rule). The iterative and obligatory Re-indexing Rule requires that the long-distance reflexive is re-indexed with the c-commanding animate subject of the next higher governing category.

Following Lebeaux (1983) who had suggested that all anaphors move at the logical form (LF) level, it was proposed that *ziji* undergoes covert head movement to the local I and, in the case of long-distance binding, to a higher I via cyclic head movement (Battistella, 1989; Pica, 1987; Cole et al., 1990; Cole and Sung, 1994; Cole and Wang, 1996). Huang and Tang (1988, 1989, 1991) developed another approach that also assumes covert successive cyclic movement at LF, but in which *ziji* undergoes A'-movement rather than head movement. In their account long-distance *ziji* is derived by successive-cyclic adjunction of the reflexive to IP.

In the framework of Generalized Phrase Structure Grammar, the long-distance binding of *ziji* is explained with the percolation of the foot feature [REFL] from the reflexive to the domain of its antecedent (Kang, 1988). Cole et al. (1993) incorporate the idea of feature percolation of *ziji* in the Government and Binding framework.

D.2.2 Pragmatic and non-uniform accounts

Huang et al. (1984) have proposed a functional approach to long-distance bound *ziji*. They suggested that the long-distance *ziji* should be regarded as an anaphoric pronoun rather than a real anaphor in the sense of Binding Theory (Chomsky, 1981). They suggested that the non-locally bound *ziji* referred to the subject of the matrix clause which is argued to be, in an underlying representation, the "speaker" of the embedded clause containing *ziji*. The embedded clause itself is assumed to be a direct quote originating from the antecedent in that underlying representation.

Yu (1992) proposed that sentence-free *ziji* might be a logophor, an idea he further pursued in his dissertation (Yu, 1996). Huang and Liu (2001) developed this account clearly distinguishing between the locally bound *ziji* and the long-distance bound *ziji*. The local *ziji* is analyzed as a syntactic anaphor in the sense of Binding Theory (Chomsky, 1981) whose binding domain is the traditional Governing Category (Chomsky, 1986). For this locally bound *ziji*, a head movement analysis is proposed in which the reflexive is raised and left-adjoined to V (Huang and Liu, 2001; Huang, 2002). Long-distance bound *ziji*, in contrast, is not subject to Binding Theory, but treated as a logophor. Huang and Liu (2001) define the term logophor as follows:

> [...]a logophor refers to a person whose (a) speech or thought, (b) attitude or state of consciousness, and/or (c) point of view, or perspective, is being reported. This person may be the speaker (the external Source, Self, or Pivot) or an internal protagonist denoted by an argument of the sentence (e.g., the matrix subject).

In other words, when *ziji* is not locally bound, it is subject to conditions of logophoricity, i.e., it is required to be contained in a description of a certain property which the referent of its antecedent explicitly or implicitly self-ascribes or which he self-ascribes via the help of the speaker (Huang and Liu, 2001; for a similar approach c.f. Liu, 2010). In a related account, Pan (1997) suggests to analyze long-distance bound *ziji* as an element that expresses *de se* beliefs (Lewis, 1979). For the difference between Pan's approach and the analysis of *ziji* as a logophor q.v. Huang and Liu (2001).

Another pragmatic approach to long-distance *ziji* was proposed by Xue et al. (1994) and Pollard and Xue (1998). In contrast to the above described accounts, they explain the difference between what they call a 'syntactic reflexive' (i.e., locally bound *ziji*) and a 'non-syntactic' reflexive (i.e., long-distance *ziji*)[13] not with lexical ambiguity between an anaphor and some kind of discourse pronoun, but propose that there is only one type of *ziji* that can be related to its antecedent either via syntactic binding or via discourse co-reference. In case *ziji* is not syntactically bound, its interpretation is subject to pragmatic or discourse constraints. In contrast, when *ziji* is syntactically bound, it is exempt from these constraints (Pollard and Xue, 1998). Another difference to Huang and Liu (2001)'s account is their definition of binding domain. Following Pollard and Sag (1992a,b, 1994) they propose the use of relative obliqueness of grammatical relations to define the binding domain of reflexives rather than purely tree-configurational relations such as c-command.

[13]Cases in which *ziji* is bound by a sub-commanding antecedent (local or remote) are treated as instances of 'non-syntactic' reflexives.