ABSTRACT

Title of Document: THE COGNITIVE BASIS FOR ENCODING AND NAVIGATING LINGUISTIC STRUCTURE

Daniel J. Parker, Doctor of Philosophy, 2014

Directed By: Professor Colin Phillips, Department of Linguistics

This dissertation is concerned with the cognitive mechanisms that are used to encode and navigate linguistic structure. Successful language understanding requires mechanisms for efficiently encoding and navigating linguistic structure in memory. The timing and accuracy of linguistic dependency formation provides valuable insights into the cognitive basis of these mechanisms. Recent research on linguistic dependency formation has revealed a profile of selective fallibility: some linguistic dependencies are rapidly and accurately implemented, but others are not, giving rise to “linguistic illusions”. This profile is not expected under current models of grammar or language processing. The broad consensus, however, is that the profile of selective fallibility reflects dependency-based differences in memory access strategies, including the use of different retrieval mechanisms and the selective use of cues for different dependencies. In this dissertation, I argue that (i) the grain-size of variability is not dependency-type, and (ii) there is not a homogenous cause for linguistic illusions. Rather, I argue that the variability is a consequence of how the grammar interacts with general-

purpose encoding and access mechanisms. To support this argument, I provide three types of evidence. First, I show how to “turn on” illusions for anaphor resolution, a phenomena that has resisted illusions in the past, reflecting a cue-combinatorics scheme that prioritizes structural information in memory retrieval. Second, I show how to “turn off” a robust illusion for negative polarity item (NPI) licensing, reflecting access to the internal computations during the encoding and interpretation of emerging semantic/pragmatic representations. Third, I provide computational simulations that derive both the presence and absence of the illusions from within the same memory architecture. These findings lead to a new conception of how we mentally encode and navigate structured linguistic representations.

THE COGNITIVE BASIS FOR ENCODING AND NAVIGATING LINGUISTIC STRUCTURE

By

Daniel J. Parker

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park, in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2014

Advisory Committee:
Professor Colin Phillips, Chair
Professor Norbert Hornstein
Assistant Professor Ellen Lau
Professor Jeffrey Lidz
Associate Professor Michael Israel, Dean's Representative

Preface

All of the work reported in this dissertation was conducted under the direction of Colin Phillips. Chapter 3 reports research that was jointly conducted with Sol Lago. Appendix B reports research that was jointly conducted with Alan Du. The title of this dissertation is a nod to Bever's (1970) article "The cognitive basis for linguistic structures" in R. Hayes (Ed.) *Cognition and language development*. New York: Wiley.

Dedication

To my parents and Luiza

Acknowledgements

I am extremely grateful to all of the colleagues, friends, and family members that have given me their time, energy, and support in helping me get to this point. I value my experiences with all of you and I want to express my sincerest gratitude for shaping me both as an academic and as a person.

First and foremost, I would like to thank Colin Phillips, my advisor and committee chair. I have gained so much insight working with Colin over the past five years. I probably lost a lot too, but that would be in the amount of sweat that I lost training for Ragnar during the hot and humid MD summers. Colin is extremely generous with his time, week after week. He always listened to whatever half-baked idea that I had at the time and he would always steer me aright by asking the difficult questions that made me quickly realize what worked, what didn't, and why. My research has greatly improved from his feedback and I am very thankful that he gave me the opportunity to work with him.

I also want to thank my committee members, Michael Israel, Ellen Lau, Jeff Lidz, and Norbert Hornstein. They have helped me through all stages of my research at lab meetings, presentations, and other events. I have benefited from their enthusiasm and insights about language research and beyond. The cookies from Norbert also helped, particularly the maple leaf ones. He is right that cognitive science runs on its stomach.

Thanks to Valentine Hacquard and Alexander Williams for allowing psycholinguists to participate in semantics valueball. It provided some much needed stress relief.

When I first came to Maryland, my professors told me that I would learn the most from my fellow students. I value all of the interactions I've had with my peers. To my first year cohort, Kenshi Funakoshi, Angela He, Yakov Kronrod, Sol Lago, Darryl McAdams, Megan Sutton, and Alexis Wellwood, thanks for building a support foundation. I enjoyed our dinners, working together over Howard's first year syntax assignments, designing experiments for Colin's psycholinguistics course, and stressing out over those long derivations for Paul's semantics course.

To Sol Lago, who is always willing to call me out on my superfluous BS. Sol has an great ability to see what I'm trying to say and to get me to say it in the most straightforward way. I am very happy that we have become good friends.

To Yakov Kronrod, who always reminded me not to sweat the little things. And thanks for bringing me back to disc golf.

I have also had the pleasure of sharing thoughts, exchanging ideas, and conversing with many other students who are much sharper and quicker than I am. In particular, I would like to thank Aaron Steven White. Aaron always kept me on my toes, leading me to new, interesting, and diverse ideas. He also led me to new and interesting musical experiences, whether that be in his car, in the office, or in the Ragnar van. He, himself, is also a top-notch musician to jam with. He is my main bro to geek out with. I have also enjoyed my conversations and experiences with Shannon Barrios, Dustin Chacón, Wing Yee Chow, Mike Fetters (we still need to discuss ellipsis some time), Annie Gagliardi, Brad Larson, and Shevaun Lewis.

The work in this dissertation wouldn't be possible if it wasn't for the students ahead of me. This work has benefited greatly from comments and feedback from Pedro Alcocer, Brian Dillon, Dave Kush (a great conference buddy), Akira Omaki, Matt Wagers, Ming Xiang, and Masaya Yoshida.

Beyond Maryland, I would like to thank my MA advisor from Eastern Michigan University, Daniel Seely. Daniel is the most encouraging person I know. I value the help and guidance that he provided to get me on my current path. I would also like to thank Helen Aristar-Dry and Anthony Aristar. They gave me an excellent opportunity to work at The LINGUIST List. In addition, I would like to acknowledge Shravan Vasishth for his comments and feedback on many aspects of this work.

I also want to thank my two best friends from Toledo, Ross Metusalem and Steve Sommer. These two always encouraged me to study and explore the strangest things that I could find. For this dissertation, that turned out to be linguistic illusions.

To my parents, Mark and Sandy Parker. They gave me the freedom to pursue my interests, even if that meant changing courses at the last minute.

Last but not least, to my better half, my wife, and best friend, Luiza Newlin-Łukowicz. Luiza is an endless source of encouragement and support. She has helped me work through various ideas, rants, and frustrations, even when they were uncalled for. Luiza is the most thoughtful, selfless, and loving person I know. Thanks for being by my side.

Table of Contents

List of Tables

List of Figures

Chapter 1 Encoding and navigating linguistic structure

1.1 Introduction

Successful language comprehension requires the ability to efficiently encode and navigate complex linguistic representations on a rapid time-scale. This ability is demonstrated perhaps most clearly in cases where the parser must establish linguistic relations between temporally distant words, phrases, and sentences. These non-adjacent relations can span a potentially large amount of material, creating 'long-distance dependencies' that rely on working memory to rapidly recover the necessary information from the preceding context. For example, upon encountering the verb *purchased* in (1), a representation of the noun phrase (NP) *the painting by Van Gogh* must be recovered from memory to integrate it as the underlying object of the verb.

(1) It was the painting by Van Gogh that the curator from the museum purchased.

Other common linguistics dependencies include the relationship between a pronoun and its referent, subject-verb agreement, and filler-gap relations. For each of these dependencies, there are both syntactic and semantic constraints on the relationship between the dependent element and the licensor. These constraints can be thought of as a set of instructions for what information must be recovered from memory and where that information is located in the linguistic or discourse representation. For example, in the case of the local reflexive anaphors in English, like *herself* in (2), there are syntactic constraints on which positions the licensor can occupy in the syntactic hierarchy: according to Principle A of the binding theory (or its analogs), the licensor of the reflexive must be the subject of the same clause (Chomsky, 1981; Pollard & Sag, 1992; Reinhart & Reuland, 1993). In addition to this syntactic constraint, there is a formal constraint of feature agreement between the reflexive and its licensor, which requires the morphological features of the reflexive to be matched against the corresponding features on the licensor. In (2), the NP *Mary* is the target of the memory retrieval for the reflexive because it is in the correct syntactic position and it bears matching morphological features, i.e., it matches both gender and number. By contrast, the distractor NP *the woman* is not a suitable licensor for the reflexive. Despite having the appropriate morphological features, it is in a structurally irrelevant position.

(2) The woman$_1$ said that Mary$_2$ hurt herself$_{*1/2}$ on the playground.

In order to ensure successful dependency formation, the processing architecture must implement data structures that are capable of reliably preserving the representational properties licensed by the grammar, in particular information about feature composition and structural relations. This may be achieved in a couple of ways. One way would be to actively maintain this information concurrently with incoming material. Another way would be to encode and store this information in memory to make it available later via retrieval. Additionally, the mechanisms that are used to access and recover linguistic information in memory must be able to reference featural and structural information to accurately distinguish relevant and irrelevant material in accordance with the grammar. This challenge, reconciling the primitives of the grammar with what we know about the constraints of the memory architecture on the representation and retrieval of linguistic information in memory, is the focus of this dissertation. In particular, the aim of this dissertation is to explicitly characterize the relationship between the representations and computations of the grammar and the two memory-related components described below:

i) **Encoding operations**
How are hierarchical linguistic representations encoded and stored in memory to make previously processed information readily accessible for ongoing parsing operations?

ii) **Access operations**
How does the parser target specific information in memory, and how is that information accessed and extracted from the encodings to elaborate the representation?

1.2 Diagnosing encoding and access mechanisms

A number of recent studies on memory in language processing have relied on the timing and accuracy of real-time constraint application during dependency resolution to make inferences about the nature of the representations that are encoded in memory and the ways in which information in those representations is accessed for ongoing parsing operations. Constraint application could proceed in one of two ways: constraint application could be immediate and accurate, such that memory access operations recover only those antecedents that satisfy the formal requirements of the dependency, or it could be late and/or inaccurate, such that memory retrieval is sensitive to grammatically illicit material.

Findings of accurate and inaccurate constraint application are equally informative about the underlying nature of the encoding and access mechanisms. Many recent studies on the role of memory in sentence comprehension have adopted the following logic: evidence that a linguistic constraint, e.g., the locality requirement on English reflexives, is immediately and accurately applied during real-time processing, then this suggests that the encoding and access mechanisms can refer to that constraint. On the other hand, evidence that a linguistic constraint is not immediately or accurately respected during real-time processing, then that could reflect a failure either on the part of the encoding mechanisms or the access mechanisms, or both. In cases of failure, protocol has been to test a broad range of dependencies that are governed by that same constraint or some analogue of it. If the processing of some dependencies respects the constraint in question, then this suggests that the encoding of the representation in memory and grammar are intact. In this case, evidence for the selective implementation of a constraint would suggest instead an error on the part of the access mechanisms (Phillips, 2013).

Recent research on the real-time status of constraint application has revealed a selective profile. Some constraints appear to be implemented rapidly and accurately during real-time processing. For example, a number of studies have shown that real-time processing mechanisms respect island constraints on filler-gap dependencies and the structural constraints on forwards and backwards anaphora and reflexive anaphors (Aoshima, Yoshida, & Phillips, 2009; Clifton, Frazier, & Deevy, 1999; Dillon, Mishler, Slogget, & Phillips, 2013; Kazanina, Lau, Lieberman, Yoshida, & Phillips, 2006; Kazanina & Phillips, 2010; Kush, 2013; Lukyanenko, Conroy, & Lidz, 2014; Nicol & Swinney, 1989; Phillips, 2006; Runner, Sussman, & Tanenhaus, 2006; Stowe, 1986; Sturt, 2003; Traxler & Pickering, 1996; Wagers & Phillips, 2009; Xiang, Dillon, & Phillips, 2009; Yoshida, 2006). In each of these cases, the parser is able to reliably avoid interference ('attraction') from similar morphologically or semantically compatible material in structurally irrelevant positions.[1] These findings of grammatical precision in online parsing are important for our understanding of how we encode and navigate linguistic structure in real time, because evidence of rapid and reliable sensitivity to the constraints of the grammar is a clear indicator that the processor is capable of encoding and access complex representations in memory in accordance with the grammar.

However, the argument for grammatically accurate constraint application has recently been challenged by a growing number of cases involving grammatically inaccurate constraint application in online processing. For example, there is growing evidence that the parser can be mislead by morphologically or semantically compatible material that is in structurally irrelevant positions when processing dependencies such as those involving subject-verb agreement and negative polarity items (NPIs) (Bock & Miller, 1991; Clifton et al., 1999; Dillon et al., 2013; Drenhaus, Saddy, & Frisch, 2005;

[1] The term 'attraction' refers to a specific type of memory retrieval interference that is indicative

Eberhard, Cutting, & Bock, 2005; Franck, Vigliocco, & Nicol, 2002; Pearlmutter, Garnsey, & Bock, 1999; Staub, 2010; Vasishth, Brüssow, Lewis, & Drenhaus, 2008; Wagers, Lau, & Phillips, 2009; Xiang et al., 2009) .

In very specific configurations, interference from structurally irrelevant material can give rise to 'linguistic illusions' or illusions of grammaticality (Phillips, Wagers, & Lau, 2011). Linguistic illusions are cases where comprehenders temporarily accept or mis-judge ungrammatical sentences in online or timed measures (e.g., speeded-acceptability judgments, self-paced reading, eye-tracking, ERPs), but later judge those same sentences as unacceptable after more reflection in untimed tasks (e.g., offline acceptability rating tasks). In this dissertation, I use linguistic illusions to uncover the micro-structure of the representations that are constructed in memory during real-time language comprehension and the computations that are performed over those representations.

This profile of 'selective fallibility' is not predicted by existing models of grammar or language processing, and it has motivated various proposals about the memory architecture for the parser. In particular, there is a growing consensus that the selective profile of linguistic illusions reflects dependency-based differences in how we access linguistic information in memory. Recent proposals suggest that the profile of selective fallibility reflects the use of qualitatively different retrieval mechanisms or the selective use of non-structural cues for different dependencies (e.g., Alcocer & Phillips, 2012 ms.; Dillon et al., 2013; Phillips et al., 2011). In short, many accounts claim, either explicitly or implicitly, that the candidate access mechanisms are directly linked to specific types of grammatical dependencies. Furthermore, similarity across the profile of failures (e.g., subject-verb agreement and NPIs) has led to the attractive idea that there is a homogenous cause for the failures, namely the mis-retrieval of structurally irrelevant material in memory. The experimental and computational findings that I report in this dissertation challenge these claims, leading us to reconsider the source and scope of linguistic illusions in comprehension.

The profile of selective fallibility also plays an important role in characterizing the relationship between the parser and the grammar. The central question here is whether the parser and grammar describe separate cognitive systems with distinct representational and computational capacities or different aspects of the same system. For example, cases of failure to immediately and reliably implement grammatical constraints, as in the case of linguistic illusions, could reflect a misalignment between the representations that are constructed in the initial stages of real-time processing and those that are licensed by the grammar. Such misalignments could then be interpreted as evidence for a dual system model of language processing that analyzes the input using different structure-building systems at different points in time. In particular, misalignments have lead to various proposals about the cognitive architecture of language, including the use of task-specific 'rough-and-ready' or 'good enough' representations that are not licensed by the grammar (Ferreira, 2003; Ferreira & Patson, 2007; Townsend & Bever, 2001), or the use of violable structural constraints in the early stages of processing (Tabor, Galantucci, & Richardson, 2004). Dual-system models have direct consequences for the memory architecture for the parser, as they imply that the memory mechanisms that are recruited for real-time processing either fail to encode the fine-grained structural and featural information that is required to successfully compute a linguistic dependency or fail to use this information to accurately guide retrieval in accordance with the grammar.

However, it has recently been argued that the profile of selective fallibility does not reflect the properties of a parser that is separate from the grammar, but rather the properties of a single, task-independent grammar and general-purpose memory and cognitive control mechanisms (e.g., S. Lewis & Phillips, in press; Phillips & Lewis, 2013). According to the single-system view, there is a single representational/structure-building system that is embedded in a general cognitive architecture. The argument for a single-systems view is based on two key factors. First, the numerous studies that show rapid online detection of grammatical anomalies or the avoidance of ungrammatical parses, as observed for reflexive anaphors, for example, suggest a close alignment between the parser and grammar, in which online processes construct richly structured representations that are the same as those licensed by the grammar. These cases of success are not expected under dual-system models, which rely on the use of heuristics or impoverished representations to guide real-time processing. Second, because the single, procedural grammar is embedded in a general cognitive architecture, cases of apparent misalignment could arise from the limitations of domain-general memory and cognitive control mechanisms that are used to implement linguistic dependencies (e.g., Vasishth et al., 2008; Wagers et al., 2009; Xiang et al., 2009). Under this view, linguistic illusions are understood as a misalignment between the constraints of a task-independent grammar and the constraints that the general cognitive architecture places on how we encode and navigate structured representations, rather than as a misalignment between multiple structure-building systems.

In short, cases of successful real-time constraint application are unexpected under models that assume multiple structure-building mechanisms, but cases of failure resulting in linguistic illusions do not fit well with models that assume a transparent mapping between the grammar and the parser. One issue that I take up in this dissertation is how to capture the profile of successes and failures in a single, procedural system that relies on a noisy memory architecture. More generally, given that the role of domain-general mechanisms like memory and cognitive control factor so heavily into the debate over the relationship between the parser and the grammar, developing a detailed model of how we encode and navigate linguistic information in memory—the focus of this dissertation—is thus likely to make an important contribution to our understanding of the broader cognitive architecture of language.

1.3 The main claims of the dissertation

In this dissertation, I argue that (i) the grain-size of selective fallibility is not dependency-type, and (ii) there is not a homogenous cause for linguistic illusions. Rather, I argue that the variability is a consequence of independently motivated principles of general cognition, including combinatory cue-based memory retrieval and temporal feature integration, without invoking multiple structure building systems. To support this argument, I provide converging evidence from two case studies. In the first case study, I show that for anaphora resolution, a phenomena that resisted illusory licensing effects in the past, we can systematically turn the illusion on, reflecting the use of a weighted cue-combinatorics scheme in antecedent retrieval. In the second case study, I show that for negative polarity items (NPIs), a phenomena that is highly susceptible to illusory licensing effects, we can systematically turn the illusion off, reflecting access to the intermediate stages of linguistic feature integration in the encoding of hierarchical structure.

Below, I provide a brief synopsis of the findings reported in this dissertation and their consequences for the cognitive architecture of language.

Selective priority for structure in memory retrieval

The role of structural information in guiding memory retrieval operations in sentence comprehension is controversial. Evidence from multiple tasks and phenomena has been used to argue that structural information has a privileged status in retrieving the heads of dependencies (Dillon et al., 2013; Kush, 2013; Sturt, 2003; Van Dyke & McElree, 2011). However, widespread evidence of interference and linguistic illusions from morphologically compatible, but structurally irrelevant antecedents suggests that structural priority is much more limited (Vasishth et al., 2008; e.g., Wagers et al., 2009). This contrasted has motivated various proposals regarding the memory architecture for the parser, including the use of qualitatively different access mechanisms or the selective use of non-structural information in retrieval for different dependencies.

I show that for anaphor resolution, a phenomenon that has resisted effects of non-structural interference and illusory licensing in the past, we can systematically control where illusions do and do not occur. These effects, which are not predicted by existing accounts, are observed for two types of anaphoric dependencies involving null subjects and reflexive anaphors in English. Furthermore, I show based on computational simulations that it is possible to derive both the presence and absence of the illusion from the same retrieval mechanism using a weighted cue-combinaorics retrieval scheme (e.g., Kush, 2013; Van Dyke & McElree, 2011).

I then propose an account of the selective profile, focusing on the question of why it is harder to obtain illusory licensing effects for anaphora than for other dependencies, e.g., subject-verb agreement. This account is based on whether retrieval is triggered by error correction or normal resolution. Specifically, I argue that whereas structural cues are prioritized in retrieval in the normal course of processing unpredictable dependencies, such as anaphora or thematic binding (Van Dyke & McElree, 2011), they are not prioritized when retrieval is triggered by an unfulfilled prediction, e.g. subject-verb agreement violations, leading the parser to be uncertain about the accuracy of its existing structural encoding. In support of this account, I provide eye-tracking data which shows different timing profiles for illusory licensing, which is observed in the early stages of processing for anaphora, but only in later stages for subject-verb agreement.

I argue that these results supports recent proposals that sentence comprehension relies on a single direct-access memory mechanism that deploys all available information using a linear, weighted cue-combinatorics scheme (e.g., Van Dyke & McElree, 2011; Kush, 2013). More generally, these results suggest that variability in illusory licensing effects is, in part, a consequence of how different sources of linguistic information are combined by the memory architecture to guide memory access.

Changing encodings and interpretations

In previous research on linguistic illusions, there has been a consensus that the time profile of the illusory licensing effects varies as a function of the amount of time after the introduction of the tail of the dependency (the licensee) (Vasishth et al., 2008; Wagers et al., 2009; Xiang et al., 2009). These accounts predict that the representation of existing structure is stable over time, and that variation in the time or position where the licensee is introduced in the sentence should not drastically impact the illusion. I present data from self-paced reading and speeded-acceptability judgments which show that one type of illusion involving negative polarity items (NPIs), which depend on semantic and pragmatic licensing mechanisms, shows a fleeting time profile, such that it is present or absent depending on the timing of when the licensee is introduced relative to the licensor. But I also find that the same time profile does not extend to another type of illusion involving subject-verb agreement (agreement attraction), which depends on a morpho-syntactic licensing mechanism.

This contrast supports the claim in the linguistics literature that these dependencies are qualitatively different, as distinguished by their formal licensing mechanisms (e.g., Chierchia, 2006; Giannakidou, 1998; Kadmon & Landman, 1993; Krifka, 1995; Ladusaw, 1979; Linebarger, 1987, among others). However, the fleeting profile for illusory NPI licensing effects is not predicted by existing accounts. Based on these findings, I argue that the selective nature of the NPI illusion is a consequence of changes over time in the encoding of emerging semantic representations, such that semantic licensing features are no longer independently accessible from their position in the hierarchical structure, preventing further illusions from semantically compatible licensors that are in a structurally irrelevant position. Under this account, the fleeting NPI illusion reflects access to the intermediate stages of linguistic feature integration processes during the encoding of hierarchical structure. In support of this contrast, I provide computational simulations in which the feature bindings for a representation change over time from a transparent to an opaque encoding.

The finding that NPI and subject-verb agreement show different time profiles suggests that there is not a homogenous cause for linguistic illusions. Based on their alignment with distinct licensing mechanisms as described by the grammar, I argue that variability in the ability of the parser to accurately implement linguistic constraints in real-time processing is, in part, a consequence of processing mechanisms that distinguish the encoding and interpretation of emerging syntactic and semantic representations.

Fleeting illusions reflect a grammar embedded in a noisy cognitive architecture

One challenge that my proposal faces is that the mechanisms that I invoke to capture the fleeting illusions could be argued to reflect a task-specific "grammar of access and encoding" or the properties of a parser that is separate from the grammar and separate from general-purpose memory and cognitive control mechanisms. I argue that the current findings do not necessitate a separate grammar of encoding and access. Specifically, I suggest that the current findings actually fit naturally with the recent proposal that real-time linguistic computation is the product of a single, task-independent grammar that is embedded in a noisy cognitive architecture (e.g., Lewis and Phillips, in press; Phillips & Lewis, 2013). Under this view, the fleeting illusions reflect a misalignment between the constraints of a task-independent grammar and the representational and computational constraints that working memory and cognitive control mechanisms place on how we mentally encode and navigate linguistic structure. Specifically, I argue that fleeting illusions reflect two types of misalignment under a one-system view, as outlined in Lewis and Phillips (in press): (i) linguistic computations that are completed, but inaccurately, due to noisy, general-purpose memory architecture, and (ii) access to the internal stages of encoding and interpretative computations that have not yet completed.

1.1. Parallels to research on optical illusions

Before we continue, it is important to note that the approach I adopt to systematically investigate linguistic illusions borrows heavily from parallel research on optical illusions in research on vision cognition. Linguistic illusions and optical illusions describe very similar phenomena. For example, just as optical illusions reflect conflicting visual percepts within the visual system, linguistic illusions reflect conflicting linguistic percepts (specifically either explicit or implicit judgments about the well-formedness of a sentence) within the linguistic system at different points in time.

In vision research, there is a rich literature on how to systematically turn optical illusions on and off (e.g., Trommershäuser, Körding, & Landy, 2011). A parade case in vision involves the scintillating grid illusion, shown in Figure 1.1 (Schrauf, Lingelbach, Lingelbach, & Wist, 1995):

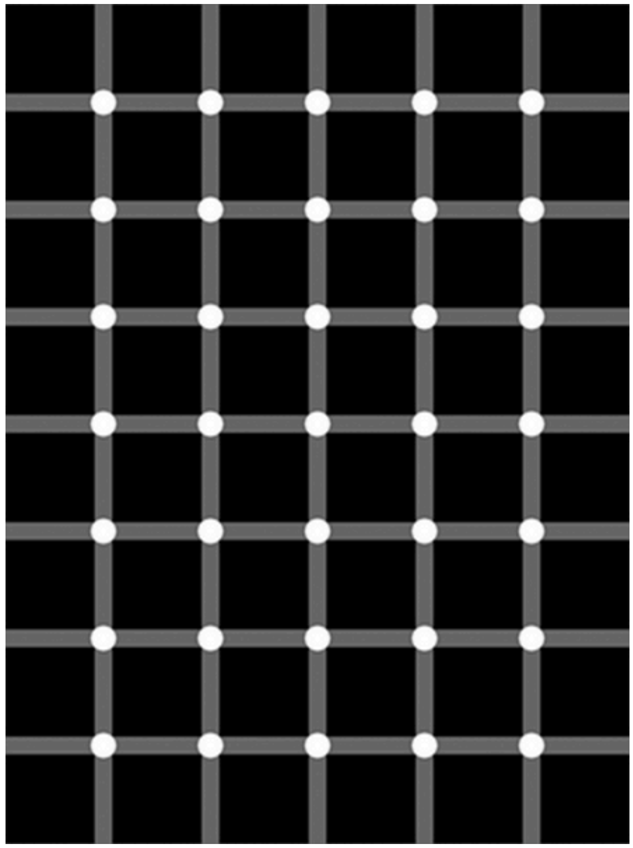


Figure 1.1: Scintillating grid illusion (Schrauf et al., 1995). Phantom dots rapidly flicker at the intersections. The number of phantom dots appear to increase as the eye is scanned across the image.

The scintillating grid illusion is a variant of the Hermann grid illusion in which black dots appear to flicker (rapidly form and vanish) at the intersections of the gray lines. The standard explanation of the illusion is that the phantom dots reflect lateral inhibition in the low-level visual system (specifically: a center/surround antagonism response of the retinal ganglion cells within the ON-OFF or OFF-ON receptive fields (Baumgartner, 1960). Strangely, however, it was recently discovered by Geier and colleagues (Geier, Bernáth, Hudák, & Séra, 2008) that both the scintillating grid illusion and Hermann grid illusions totally disappear when the lines of the grid are distorted by a sine curve that is specifically less than 10% of its wavelength, as shown in Figure 1.2.

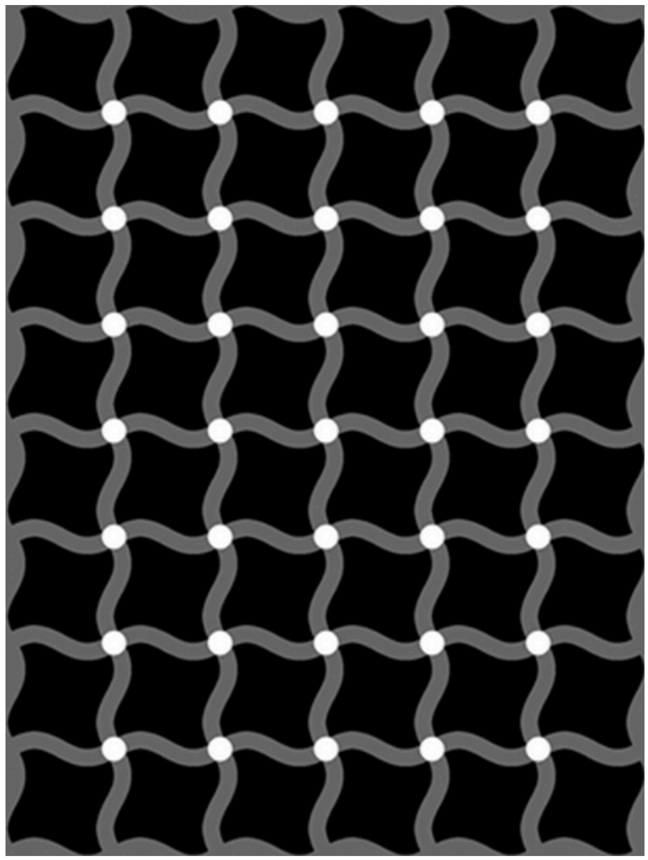


Figure 1.2: Stopping the scintillating grid illusion by sinusoid variation (Geier et al., 2008). Phantom dots completely disappear when the lines of the grid are distorted with a sine curve. The disappearance of the illusion is not predicted by standard accounts.

The minimal manipulation that systematically controls the presence and absence of the illusion provides a valuable tool for understanding how the visual system processes information because the disappearance is not predicted under existing accounts. Geier and colleagues argue, for example, that if Baumgartner's hypothesis were correct, then phantom spots should also appear in the sinusoid grid. However, this is not the case. This simple manipulation demonstrates that while the original illusion itself can be very informative (e.g., it generated years of productive research following Baumgartner's work in the 1960's), it is the nature of the minimal change that can make the illusion disappear that is even more informative for vision scientists because it constrains the resulting theory. In the case of the scintillating grid illusion, the unexpected disappearance of the illusion rules out a prominent theory of how the brain processes visual system.

In this dissertation, I will show that there are ways of manipulating linguistic illusions in a similar fashion, making them systematically appear and disappear, in a way that is informative about how we mentally encode and navigate structured representations. In particular, I will argue that while we can learn a great deal from where people do and do not fall for linguistic illusions, the ability to selectively make the illusion appear and disappear makes the tests even more valuable.

1.5 Outline of the dissertation

1.5.1 Chapter 2

Chapter 2 summarizes the profile of selective fallibility, focusing on three dependencies involving anaphora, subject-verb agreement, and NPIs. I begin by discussing the candidate retrieval mechanisms and review their respective advantages and disadvantages for accessing linguistic information in memory. I then describe the profile of selective fallibility, which has motivated the use of these different mechanisms in real-time sentence processing. I argue in the conclusion of Chapter 2 that the reasons for the contrasts remain unresolved, motivating a deeper investigation into the source and scope of linguistic illusions.

1.5.2 Chapter 3

Chapter 3 investigates the licensing of null subjects (PRO) in adjunct clauses in English to better understand what is responsible for the selective profile. Null subject licensing has received little attention in the psycholinguistics literature, but it provides a good test of the candidate memory access strategies because it shares properties with both subject-verb agreement and reflexive licensing. Results from three studies using off-line judgments and self-paced reading confirm the structural constraints on null subject licensing, but show an online illusory licensing profile similar to subject-verb agreement, suggesting the use of non-structural cues for retrieving the licensor of a null subject. These results provide evidence of illusory licensing for anaphoric dependencies, motivating a revision of the generalization that the profile of selective fallibility reflects dependency-based differences in memory access. The findings from this study suggest instead that that the profile of selective fallibility may be a consequence of the role of specific content features in retrieval (such as animacy) or more generally, the degree of similarity between the retrieval probe and the target.

1.4.3 Chapter 4

Chapter 4 tests whether that the profile of selective fallibility is a consequence of the role of specific features or the degree of similarity between the retrieval probe and the target using reflexive anaphors. In English, there is a formal requirement of feature concord between the reflexive anaphor and its antecedent, which includes animacy match along with gender and number. Previous studies on reflexive anaphors have failed to consistently find illusory licensing effects, but they have only tested configurations in which the reflexive-antecedent dependency involved a single feature manipulation, e.g., mismatches on gender (Sturt, 2003) or number (Dillon et al., 2013). Across three eye-tracking experiments, I compare configurations involving multiple feature mismatches, including gender+animacy, number+animacy, and gender+number. Additional conditions involving agreement attraction provided a baseline measure of illusory licensing. Results show that reflexives are indeed susceptible to illusory licensing, but only selectively: when the reflexive mismatches the true subject in just one feature, there is no illusory licensing effects, but when it mismatches in two features, strong illusory licensing effects are found, comparable in size to agreement. These findings suggest that the profile of selective fallibility is due to the role of specific features, since the illusions for reflexive are also observed for other feature combinations. More generally, these findings provide mounting evidence that the profile of selective fallibility cannot be a consequence of dependency-based differences in memory access as assumed in previous work. Instead, these findings imply that the variability is due to a more general effect of the degree of similarity between the target and the retrieval probe.

Computational simulations presented in this chapter suggest that the contrasting illusory licensing effects observed in previous studies on reflexives and subject-verb agreement may reflect a weighted cue-combinatorics scheme in which structural cues are weighted more strongly in retrieval than morphological cues. These findings are consistent with previous proposals about the use of cue-weighting in memory retrieval (Van Dyke & McElree, 2011; Kush, 2013). Crucially, computational simulations revealed weights that can capture both the presence and absence of illusions from the same model. However, this raises the question of why subject-verb agreement shows strong illusions even in one-feature mismatch configurations, where the target mismatches only on number. These findings imply that structure is not weighted more strongly for agreement. To account for this asymmetry, I suggest that whereas structural cues are prioritized in retrieval in the normal course of processing unpredictable dependencies, such as those involving anaphora (e.g., null subjects and reflexives) and thematic binding (Van Dyke & McElree 2011), they are not prioritized when retrieval is triggered by an unfulfilled prediction (e.g., agreement violations), leading the parser to be uncertain or less confident about the accuracy of its existing structural encoding. In support of this account, I present eye-tracking data that reveals different timing profiles for illusory licensing effects, which is observed in the early stages of processing for reflexives, but only in later stages for subject-verb agreement.

1.4.4 Chapter 5

Chapter 5 investigates a family-related phenomenon involving illusory NPI licensing, with a comparison to agreement attraction. In previous research, there has been a consensus that the time profile of illusory licensing effects observed for NPIs and subject-verb agreement varies as a function of the amount of time after the dependent element (i.e., the NPI or the agreeing verb) is introduced in the sentence. These accounts predict that the representation of existing structure is stable over time, and that variation in the time or position where the dependent element is introduced in the sentence should not drastically impact the illusion. In Chapter 5, I present results from a series of studies using untimed acceptability judgments, self-paced reading, and speeded-acceptability judgments, that show that illusory licensing effects for NPIs, which depend on semantic/pragmatic licensing mechanisms, shows a fleeting time profile, such that it is present or absent depending on the timing of when the NPI is introduced in the sentence relative to a licensor. But, I find that the same time profile does not extend to illusory licensing effects for subject-verb agreement, which depend on a morpho-syntactic licensing mechanism. These findings provide even more evidence that the selective profile is not a consequence of dependency-based differences in memory access.

The contrast between NPIs and subject-verb agreement supports the claim that these dependencies are qualitatively different, but the fleeting profile for NPI illusions is not predicted by existing accounts. I argue that the contrasting profiles reflect access to internal computations during the encoding and interpretation of emerging semantic/pragmatic representations. In support of this contrast, I provide computational simulations in which the feature-bindings in a representation change over time from a transparent to an opaque format.

1.4.5 Chapter 6

In Chapter 6, I discuss the status of the proposed mechanisms in the overall architecture of language and address the broader implications for the parser-grammar relation. In particular, I ask whether the mechanisms that I proposed to capture the fleeting illusions require additional, task-specific encoding and access strategies that cannot be attributed to a domain-general cognitive architecture or a task-independent grammar (e.g., a separate 'grammar of encoding and access'). I first review several other properties that have been attributed to parser-specific heuristics and then discuss recent accounts that have attributed those properties to a task-independent grammar that is embedded in a general cognitive architecture. I then step through the argument of why my proposed mechanisms do not reflect a separate, task-specific "grammar of encoding and access", and spell out the details of how to capture the effects using the properties of a task-independent grammar and general-purpose memory and cognitive control structures. Lastly, I discuss several key predictions and suggest several areas for future research.

Chapter 2 The memory architecture for the parser

2.1 Looking back: Memory in the psycholinguistics enterprise

Language comprehension involves a number of complex cognitive operations, some of which engage the memory system. For example, interpreting words requires accessing long-term lexical memory. Composing linguistic structure involves incrementally constructing, storing, and maintaining complex representations and relations between different words in phrases. Resolving ambiguity requires recovering alternative hypotheses about the input. Interpretation requires accessing and integrating information across a variety of representational levels into the developing discourse. Each of these processes takes time (even if only a few hundred milliseconds) and they may be costly in terms of the demands they place on the available memory resources.

There is a long tradition in psycholinguistic research to understand how the architecture of the memory systems impact language comprehension. For example, research on memory and language comprehension dating back to Miller and Chomsky in the late 1950s and early 1960s (e.g., Chomsky & Miller, 1963; Chomsky, 1965; Miller & Chomsky, 1963) focused primarily on questions of how memory capacity limitations impact language comprehension. More recently, the emphasis has shifted away from issues of capacity and limitations to issues of how (i) structural relations are encoded in memory and (ii) relevant and irrelevant information is distinguished in memory. One goal of current research on memory and language comprehension is to develop a detailed, computationally complete (in the sense of R. L. Lewis & Vasishth, 2005) model of mechanisms that are used to encode linguistic representations in memory and the dynamics of accessing specific pieces of information in those representations. The research that I report in this dissertation is part of a larger effort within the psycholinguistics enterprise to better understand how we encode and navigate linguistic structure in memory.

Recent research on the memory architecture for the parser has identified two ways in which linguistic information may be accessed in memory for language comprehension. One way of accessing linguistic information in memory involves a search process, which exploits the graph structure of a sentence representation by searching the representation node-by-node in a serial fashion. Another way of accessing linguistic information in memory involves a direct-access, content-addressable retrieval, which exploits the inherent properties of individual items in memory, rather than global structure, to access all memory items in parallel. The main advantage of a search mechanism is that memory access is structurally constrained: by relying on structural relations to guide representation navigation, the parser can avoid interference from morphologically or semantically similar items that are in structurally irrelevant positions. The downside, however, is that search is not executed in constant time: retrieval latencies grow as the structural distance between the target item and the point of retrieval increases. Content-addressable mechanisms, on the other hand, execute in constant time, but they are prone to interference from similar items in structurally irrelevant positions.. The tradeoff is accuracy for speed.

A primary goal in recent research on memory and language processing is to delineate the circumstances in which search or content-addressable access mechanisms are used for dependency resolution. A number of studies have shown effects of similarity-based interference and length-invariant processing dynamics, consistent with a content-addressable access mechanism (Martin & McElree, 2008; Martin & McElree, 2009; McElree, Foraker, & Dyer, 2003; Van Dyke & McElree, 2006; Van Dyke & McElree, 2011; Vasishth et al., 2008; Wagers et al., 2009). But, a growing number of studies have reported effects of non-interference from structurally irrelevant materials (Clifton et al., 1999; Dillon et al., 2013; Kush, 2013; Nicol & Swinney, 1989; Sturt, 2003; Xiang et al., 2009) and a limited number of studies have found effects of length-dependent processing dynamics (Dillon et al., submitted), consistent with a structure-guided search mechanism.

In this chapter, I review the nature of these access mechanisms and the evidence taken to support their use in sentence comprehension.

2.2 Encoding and accessing linguistic information in memory

Research on the cognitive science of memory retrieval has identified two ways in which information in memory can be accessed: (i) serial search, and (ii) direct-access, content-addressable retrieval. I discuss each of these access mechanisms in detail below.

Serial search / Tree-traversal

'Search' refers to a class of memory access procedures that execute pairwise evaluations of the match between the desired information and the items in memory. With a search procedure, the contents of memory are accessed by location rather than content-match (Gillund & Shiffrin, 1984). There are at least two ways in which search could implemented. One way involves a low-level serial search that operates over an unstructured, but ordered representation of the input. The ordered representation could be scanned either in a forwards or backwards method while sequentially evaluating each item, one at time, for its match to the search criteria. This type of search has been argued to mediate the recovery of certain types of relational information, such as temporal and spatial order and positional information (Grodner & Gibson, 2005; McElree & Dosher, 1989; McElree & Dosher, 1993; McElree, 2001; McElree, 2006). For example, McElree and Dosher (1993) examined the retrieval of temporal order information in a judgment of recency task. In this task, participants were presented with two test probes from a short word list that they memorized and they were asked to select the word that occurred more frequently in the list. McElree and Dosher found that both accuracy and speed decreased as additional items were interpolated between the study phase and test phase. McElree (2001) observed similar effects using an *n*-back task, in which participants responded to information about ordinal position in a sequence. McElree found that retrieval speed decreased as the amount of interpolated words increased. These findings suggest that ordered representations in memory are scanned in a serial fashion, evaluating the most recent item first, then iteratively moving backwards in time.

A more sophisticated implementation of search would operate over a structured representation of the input, which could restrict access to constituents in certain positions, along a particular path, or within some domain defined in terms of relational properties (Berwick & Weinberg, 1984; McElree et al., 2003). For example, a hierarchically structured representation of the input could be searched using a tree-traversal algorithm that iteratively evaluates the contents of the representation node-by-node using the dominance relations between nodes to determine the search path (Knuth, 1965).

The main advantage of search is that the recovery of linguistic information from memory is grammatically constrained. For example, if a hierarchical representation of the sentence is accurately encoded in memory, the structural constraints that govern subject-verb agreement or reflexive licensing can be straightforwardly implemented by proceeding up the dominance path to the top of the current clause and evaluating the features of the left-hand daughter of the top node, i.e. the local the subject, in accordance with the grammar. The ability to execute a structure-guided search thereby allows the processor to selectively target information in specific structural positions, avoiding interference from feature compatible elements that are in structurally irrelevant positions. However, the disadvantage of this type of memory access procedure for real-time language comprehension is that it can be slow, depending on the size of the representation.

Content-addressable retrieval

Many models of memory assume that the default access procedure involves direct-access, content-addressable retrieval (Anderson & Lebiere, 1998; Anderson et al., 2004; Clark & Gronlund, 1996; Gillund & Shiffrin, 1984; Hintzman, 1984; Hintzman, 1988; Kawamoto, 1988; Kohonen, 1980; R. L. Lewis & Vasishth, 2005; R. L. Lewis et al., 2006; McElree, 2000; Murdock, 1982; Murdock, 1993; Nairne, 1990; Plaut, 1997; Vasishth et al., 2008). The key property of content-addressable retrieval is that all items in memory are probed in parallel using a set of cues to the relevant memory representations. This type of retrieval allows direct contact with memory items based on content, rather than linear or structural position, obviating the need to search through irrelevant items. The term 'content' refers to the inherent properties of an individual item in memory, including one-place predicates such as 'subject', 'nominative' or 'plural', but not two-place predicates that encode configurational relations between elements in a representation, like c-command. In a content-addressable memory architecture, content cues are combined (either linearly/additively or multiplicatively) to form a retrieval probe, which is used to compute the degree of similarity between each cue and the items in memory.[2] The probability of retrieving a particular memory item is an increasing function of the degree to which the content features of the memory item match the content cues of the retrieval probe, providing a 'probe-to-item' strength that is often described in terms of activation (Lewis and Vasishth, 2005). Crucially, however, because content-addressable retrieval is driven by

[2] I discuss issues of cue-combinatorics in more detail in Chapter 4.

probe-to-item similarity, the probability of retrieving a particular memory item will be reduced as a function of the degree to which the retrieval cues match other items in memory. One consequence of content-addressable access is that items that partially match the retrieval cues may be incorrectly retrieved, especially when the probe-to-distractor strength is similar to the probe-to-target strength.

2.3 Empirical evidence for the candidate memory mechanisms

Serial search and content-addressable access mechanisms make contrasting predictions about (i) the temporal dynamics of memory access, and (ii) the impact of similar but structurally irrelevant material on dependency resolution. Below, I discuss the empirical findings in support of these predictions.

2.3.1 Evidence from temporal dynamics

Serial search mechanisms

Serial search predicts that retrieval latencies should grow as the structural distance between the target and the point of retrieval increases. While there is a substantial amount of evidence for the use of serial search in other cognitive domains, such as vision (Herd & O'Reilly, 2005; Steinman, 1987; Treisman & Gelade, 1980; Treisman & Schmidt, 1982) there is relatively little evidence for the use of a serial search mechanism in language processing. There is only one study that I am aware of that provides time-course evidence consistent with the predictions of a serial access mechanism, namely Dillon et al. (submitted). Dillon and colleagues examined the processing of the Mandarin Chinese anaphor *ziji*, which permits both local and non-local antecedents, using a speed-accuracy tradeoff (SAT) procedure. The SAT procedure probes participants for a binary sensicality judgment at cued intervals after the onset of the dependent element, in this case the anaphor ziji. The result is a curve that describes the growth in sensitivity to semantic congruity as a function of time. Unlike standard reaction time paradigms, such as self-paced reading, SAT does not conflate differences between speed and the success of interpretation into a single mean reaction time. SAT measures the rate that sensitivity to semantic congruity rises to asymptotic accuracy, separating the speed of processing and accuracy of completing the task into two independent parameters. Dillon and colleagues compared the processing dynamics for sentences like those in (1), where the antecedent was in the same clause as the anaphor (1a) or separated from the anaphor by one clause (1b).

(1) a. Local antecedent
Auto-biography say [coach Zhang [when team not perform well time] underestimate ziji].
'The auto-biography says that Coach Zhang underestimated self when the team was doing poorly.

b. Distant antecedent.
Coach Zhang say [that report [when team not perform well time] underestimate ziji].
'Coach Zhang says that the report underestimated self when the team was doing poorly.

SAT temporal dynamics from their study revealed a faster rate of access for sentences with a local antecedent relative to sentences with a distant antecedent. These data could be taken as evidence for the use of a serial, structure-guided search mechanism that is used to recover a referential antecedent. However, these data are also consistent with the use of a direct-access, content-addressable access mechanism that executes distinct retrievals via serial chaining, a content-based recovery process that uses the last node as a cue to recover the next node until the required information is found (cf. Lewandowsky & Murdock, 1989). I discuss this alternative in more detail below.

Content-addressable access mechanisms

Content-addressable access, in contrast to serial search, executes in constant time, predicting size- and length-invariant retrieval dynamics. There is considerable evidence for the use of content-addressable access mechanisms in language comprehension. For example, McElree and Dosher (1989) used an SAT procedure to examine the temporal dynamics and accuracy of memory retrieval for short-term item recognition. They found that the number of items in a list and the number of items interpolated between the study phase and test phase impacted the accuracy of recognition judgments, as measured by the SAT asymptote parameter. In particular, the accuracy of recognition judgments decreased as the set size increased. Crucially, however, they found that neither the set size nor recency of the test probe affected temporal dynamics, as measure by the SAT rate and intercept parameters. Based on these findings, McElree and Dosher argued that access to an item's representation in memory is direct.

In the domain of sentence processing, McElree, Foraker, and Dyer (2003) examined the processing of sentences involving filler-gap dependencies, such as (2a) and subject-verb thematic binding, such as in (2b), using SAT. In (2a), the filler NP *the book* must be associated with the gap in the direct object position of the verb *admired*. In (2b), the NP *the editor* must be thematically bound to the verb *laughed*, across an intervening relative clause.

(2) a. This was the book that the editor admired.
b. The editor that the book amused laughed.

McElree and colleagues also manipulated the amount of material that was interpolated between the non-adjacent dependencies in (2). Across three experiments, they found processing accuracy, as measured by the SAT asymptote parameter, decreased as more material was interpolated. Crucially, however, processing speed, as measured by SAT rate and intercept parameters, did not change with the amount of interpolated material. McElree and colleagues took these results to suggest that a content-addressable memory access system mediates dependency formation in language comprehension, in which syntactic and semantic constraints provide the necessary retrieval cues to directly access to the head of the dependency.

Van Dyke and McElree (2007) conducted a similar study investigating the processing of sentences like those in (3) using SAT. In (3a), the subject of the verb *objected*, i.e., *the client*, is in sentence initial position, but in (3b), it is embedded.

(3) a. The client who the assistant forgot had said that the visitor was important objected.
b. The assistant who had said that the visitor was important forgot that the client at the office objected.

Van Dyke and McElree reasoned that if a backwards serial search was engaged to recover the subject for thematic binding, then processing speed should be slower for (3a) relative to (3b) due to the increased amount of intervening material between the dependent elements. They found a difference in accuracy, as measured by the SAT asymptote parameter, with lower levels of accuracy observed for sentences like (3a) as compared to (3b). Crucially, however, they found no evidence for a difference in processing speed between the sentences, as measured by the SAT rate and intercept parameters. These results, as Van Dyke and McElree argued, are inconsistent with a serial search, but compatible with a direct-access, content-addressable retrieval mechanism.

Martin and McElree (2008; 2009; 2011) found similar time-course profiles for the processing of (VP and IP) ellipsis. In particular, they examined the effects of distance between the ellipsis and its antecedent on the speed and accuracy of comprehension using SAT, e.g., (4).

(4) a. Near antecedent
The editor admired the author's writing, but the critics did not.

b. Far antecedent
The editor admired the author's writing, but everyone at the publishing house was shocked to hear that the critics did not.

Consistent with the other studies conducted by McElree and colleagues, Martin and McElree found that asymptotic accuracy was lower for longer distances. However, they found that the distance did not impact the speed of interpreting ellipsis, as measured by SAT temporal dynamics. They presented these data as evidence that antecedent representations are content-addressable and retrieved from memory using a cue-based, direct-access retrieval process, rather than a serial search process (but cf. Phillips & Parker, 2014, for critical discussion).

2.3.2 Evidence from (non-)interference

Recent research on interference effects in sentence comprehension has revealed a mixed profile, or a profile of 'selective fallibility' (Phillips et al., 2011) For example, some studies have reported profiles of interference-free dependency resolution that are expected if the parser engages a structure-guided access strategy (Clifton et al., 1999; Dillon, 2011; Dillon et al., 2013; Dillon et al., submitted; Kush, 2013; Nicol & Swinney, 1989; Sturt, 2003; Xiang et al., 2009). However, other studies have reported evidence of interference effects that are expected if the parser engages a direct-access, content-addressable retrieval mechanism (Dillon et al., 2013; Drenhaus et al., 2005; Vasishth et al., 2008; Wagers et al., 2009; Xiang et al., 2009).

This subsection assesses the evidence for interference effects in sentence processing, focusing on three dependencies involving anaphors, subject-verb agreement, and NPIs. But before we begin, it is important to distinguish two types of interference that are often conflated or confused in the literature: *inhibitory* and *facilitatory* interference. These two types of interference are often collapsed together, since they both occur when an overlap in feature match impacts memory access. Thus the general term for these types of interference is 'similarity-based interference’. But, they have different time profiles and they license different conclusions about the nature of memory encoding and access. The contexts that inhibitory vs. facilitatory interference arise can be distinguished based on two types of feature match. Inhibitory interference occurs in ‘multiple match’ contexts, where the target memory perfectly matches the retrieval probe, but a structurally irrelevant item that overlaps in some features with the target disrupts retrieval, leading to slower retrieval latencies, hence the inhibitory description. In a multiple match configuration, processing difficulty may arise during retrieval because multiple feature matched items either disrupt encoding, via feature overwriting for example (e.g., Nairne, 1988; Nairne, 1990) or decrease cue diagnosticity during retrieval (Anderson, 1974; Anderson & Reder, 1999; R. L. Lewis & Vasishth, 2005). Facilitatory interference, by contrast, arises in ‘partial match’ contexts, where neither the target nor the distractor is a perfect match to the retrieval cues, and the best-matching item has only a subset of the desired features, i.e. facilitatory interference arises in ungrammatical contexts. In a partial match context, the presence of a structurally irrelevant but partially matched item can reduce the disruption associated with the lack of a perfectly matched target. In this configuration, partially activated, but structurally irrelevant items can temporarily mislead the parser into considering an grammatically illicit dependency, leading to an ‘illusion of grammaticality’ (Phillips et al., 2011). For this reason, facilitatory interference is often called ‘attraction’ or ‘intrusion’, referring to the attraction from or the intrusion of grammaticality illicit material on the processing of linguistic dependencies (see Dillon 2011 for further discussion of the distinctions between inhibitory and facilitatory interference). For consistency, I will refer to instances of facilitatory interference as 'attraction'.

Inhibitory and facilitatory interference license different conclusions about the nature of the encoding and retrieval mechanisms. Inhibitory effects arise in contexts where the inhibiting feature is not actually used as a retrieval cue (see Gordon, Hendrick, & Johnson, 2001; Gordon, Hendrick, & Levine, 2002; Gordon, Hendrick, & Johnson, 2004; Gordon, Hendrick, Johnson, & Lee, 2006) and they do not entail that the distractor was actually retrieved or considered. Facilitatory interference effects, by contrast, do license the conclusion that illicit material is retrieved or considered, as it is a key prediction of a direct-access, content-addressable retrieval mechanism that recovers the memory item with the highest degree of match to the retrieval probe, regardless of grammatical fit. For the rest of this dissertation, I will focus only on facilitatory interference (attraction) effects because unlike inhibitory effects, facilitatory effects license clear conclusions about the nature of the retrieval mechanisms that are used to recover linguistic information from memory (see Dillon 2011 for further discussion).

Non-illusions of grammaticality

A number of studies have tested whether retrieval processes show sensitivity to only grammatically licit antecedents upon encountering a reflexive anaphor. These studies have consistently failed to find clear evidence of attraction from grammatically illicit antecedents, suggesting that the syntactic requirements that govern the grammatical relation between a reflexive and its antecedent act as a hard constraint on the nouns that are considered for participation in the dependency, consistent with the predictions of a structure-guided search mechanism.

The argument for grammatical sensitivity in the processing of reflexive anaphors is based on tests of (non-)interference from grammatically illicit antecedents. To understand the nature of these tests, we must first distinguish the notions of target vs. distractor. In the case of reflexive licensing, the target is the subject of the same clause that contains the reflexives, i.e. the local subject, in accordance with the grammar, and the distractor is some other item in the sentence that has the desired morphological feature composition, but is not in the correct position in the hierarchical representation of the sentence. There are two configurations for probing for attraction effects from structurally irrelevant material in reflexive licensing. The first configuration involves a proactive interference paradigm, where the distractor appears in a position before the target, usually as the subject of the matrix clause, e.g., (5). The second configuration involves a retroactive interference paradigm, where the distractor appears in a position after the target, usually as the subject or object of subject-modifying relative clause, e.g., (6). In both configurations, the target is often said to be 'structurally accessible' because it occupies the correct local subject position in the syntactic hierarchy, and the distractor is said to be structurally inaccessible because it does not occupy this position.[3]

[3] The notion of accessibility here is not a claim about accessibility in terms of memory retrieval. It is only a claim about accessibility in terms of structural hierarchy.

(5) Proactive interference paradigm
The man said that [**the boy** hurt himself]

DISTRACTOR TARGET

(6) Retroactive interference paradigm
[The man [that **the boy** saw]] hurt himself

TARGET DISTRACTOR

Using configurations such as those in (5) and (6), many tests of memory interference vary the feature composition of the target, the distractor, or the reflexive to determine whether there are any conditions under which the processor is sensitive to the feature match of the structurally irrelevant distractor, as measured by the processing dynamics at the reflexive. The general argument behind the studies that I review below has been that the processing of the reflexive is impacted only by the feature match of the target, then this suggests that memory access was restricted to the position of the target, in accordance with the grammar

For example, Sturt (2003) tested sentences like those in (7) and (8). He manipulated several features of the sentence. He manipulated grammaticality by varying the gender of the reflexive, such that the target matched the gender of the reflexive in the grammatical conditions and mismatched the gender of the reflexive in the ungrammatical conditions. He also independently manipulated the (stereotyped) gender match of the distractor such that is either matched or mismatched the gender of the reflexive. He was primarily interested in the impact of a structurally irrelevant but feature matched distractor on reflexive processing.

(7) Sturt (2003) Experiment 1 (proactive interference paradigm)

Jonathan was pretty worried at the City Hospital.
a. He remembered that the surgeon pricked himself with a needle.
b. *He remembered that the surgeon pricked herself with a needle.

Jennifer was pretty worried at the City Hopsital.
c. She remembered that the surgeon pricked himself with a needle.
d. *She remembered that the surgeon pricked himself with a needle.

(8) Sturt (2003) Experiment 2 (retroactive interference paradigm)

Jonathan was pretty worried at the City Hospital.
a. The surgeon who treated Jonathan had pricked himself with a needle.
b. *The surgeon who treated Jonathan had pricked herself with a needle.

Jennifer was pretty worried at the City Hospital.
c. The surgeon who treated Jennifer had pricked himself with a needle.
d. *The surgeon who treated Jennifer had pricked herself with a needle.

Sturt used eye-tracking to determine whether or not a morphological feature match between the reflexive and the distractor impacted the processing of the reflexive. Sturt found that initial reading times for the reflexives were affected by the (mis-)match of the target, but not by the (mis-)match of the distractor. Based on these findings, Sturt concluded that structure act as a hard constraint on the initial stages of reflexive resolution, consistent with the predictions of a structure-guided memory access mechanism.

This finding of grammatical accurate constraint application for reflexives was replicated Xiang et al. (2009) using event-related-potentials (ERPs) and by

Dillon et al. (2013) using eye-tracking. Xiang and colleagues tested sentences like those in (9), which were based on those used by Sturt (2003).[4]

(9) a. The tough soldier that Fred treated in the military hospital introduced himself to all the nurses.
b. *The tough soldier that Katie treated in the military hospital introduced herself to all the nurses.
c. *The tough soldier that Fred treated in the military hospital introduced herself to all the nurses.

The ERP component of interest to Xiang and colleagues was the P600, which is a positive-going component that occurs ~600ms post-stimulus. Xiang and colleagues found that the P600 effect elicited in response to the mismatch between the (stereotyped) gender of the local subject and the reflexive was not attenuated by the presence of the structurally irrelevant but feature matched distractor.

Dillon et al. tested sentences like those in (10), which had a similar structure to those tested by Sturt (2003) and Xiang et al. (2009), but instead manipulated number match rather than gender match to the reflexive.[5]

(10) a. The new executive who oversaw the middle manager apparently doubted himself on most major decisions.
b. The new executive who oversaw the middle managers apparently doubted himself on most major decisions.
c. *The new executive who oversaw the middle manager apparently doubted themselves on most major decisions.
d. *The new executive who oversaw the middle managers apparently doubted themselves on most major decisions.

[4] Xiang and colleagues contrasted these sentences with a parallel set of sentences that tested the impact of structurally illicit, feature matched licensors on the processing of NPIs. I will discuss these findings in the next subsection on grammatical illusions.

[5] Dillon and colleagues contrasted these sentences with parallel sentences involving subject-verb agreement, which I discuss in detail in the next subsection, as they induced attraction effects.

Dillon found that the processing of the reflexive, as measured by re-reading times in eye-tracking, was impacted by the number match to the target, but not by the corresponding match to the distractor. These findings support the emerging consensus that memory access for reflexives, across a range of feature manipulations, is grammatically constrained as expected by the use of a structure-guided search mechanism.

Similar effects of grammatically accurate constraint application for reflexive anaphors have been observed across a wide range of measures and methodologies, syntactic configurations, and languages, and they have been corroborated for other anaphoric dependencies, such as those involving reciprocals and null subjects. For example, Nicol (1988) tested sentences like those in (11) using a cross-modal lexical priming (CMLP) paradigm to probe for the priming of semantic associates to *boxer*, *skier*, and *doctor*.[6]

(11) The boxer told the skier that the doctor for the team would blame himself * for the injury.

In her experiment, participants listened to sentences like (11) and performed a lexical decision task at the point indicated by the asterisk. Nicol observed priming to the semantic associates of the local subject/target *doctor* (e.g., *nurse*), but not the semantic associates of other nouns in the sentence (*boxer*, *skier*), as measured by faster lexical decisions to visually presented associates. In addition, no such activation was found when the lexical decision task immediately

[6] Discussed also in Nicol and Swinney 1989.

preceded the reflexive. These findings suggest that (i) the processor does not maintain potential antecedents in memory, implicating antecedent retrieval upon encountering a reflexive, and (ii) the syntactic constraints that govern the reflexive-antecedent dependency are applied immediately as a hard constraint on antecedent retrieval. These findings were replicated with across different grammatical structures and languages, e.g. Mandarin Chinese (Liu, 2009).

Clifton et al. 1999 reported similar results of grammatically accurate constraint application in two self-paced reading experiments with sentences like those in (12).

(12) a. The waitress (with the handsome friend) burned herself after spilling the soup
b. The son (of the fireman) hurt himself in a bad accident.

In both experiments, reading times at the reflexive were not modulated by the presence of a structurally irrelevant feature matched distractor. In a similar study that also used self-paced reading measures, Badecker and Straub (2002) failed to find evidence that a genitive possessor or an experiencer argument of a raising predicate influenced the processing of a reflexive. Just as in the previous studies, these findings have been taken to suggest that comprehenders access only the grammatically licit antecedent when constructing the reflexive-antecedent dependency.

Similar conclusions about the grammatical accuracy of real-time constraint application for reflexive anaphors have been reached based on visual world data. Runner and colleagues (2006) tested reflexives contained inside a possessed picture noun phrase (PPNP) while monitoring the eye-movements of participants using a visual display. Participants were given the instructions in (13).

(13) Pick up Joe. Look at Ken. Have Joe touch Harry's picture of himself.

In their experiments, Runner and colleagues found evidence of early looks to both the local subject, e.g., *Joe*, and the possessor, e.g., *Harry*, but the looks to the grammatically illicit antecedent, e.g., *Ken*, did not increase.[7]

Cunnings and Sturt (2012) tested similar constructions in which the reflexive was contained inside a picture NP (PNP), like those in (14), using eye-tracking while reading.

(14) Jonathan was walking through the military barracks.
  a. He heard that the solider had a picture of himself in the middle of the mess hall.
  b. *He heard that the solider had a picture of herself in the middle of the mess hall.
  Jennifer was walking though the military barracks.
  a. She heard that the solider had a picture of herself in the middle of the mess hall.
  b. *She heard that the solider had a picture of himself in the middle of the mess hall.

Cunnings and Sturt replicated the findings from Sturt (2003). They found no evidence of sensitivity to the gender match of the structurally irrelevant distractor, which they presented as further evidence of grammatically accurate

[7] Runner and colleagues argued that reflexives contained inside PPNPs are binding theory exempt anaphors that can be bound by either the local subject or the possessor (but cf. Chomsky, 1981, Reinhart & Reuland, 1993).

real-time constraint application. Clackson and colleagues (2011) found similar results using sentences that contained a benefactive reflexive. Using visual world eye-tracking, they tested for sensitivity to the feature match of the structurally irrelevant main clause subject in sentences like (15).

(15) a. Peter was waiting outside the corner shop. He watched as Mr. Jones bought a huge box of popcorn for himself over the counter.
b. Susan was waiting outside the corner shop. She watched as Mr. Jones bought a huge box of popcorn for himself over the counter.

Clackson and colleagues failed to find reliable evidence of increased looks to the structurally irrelevant, but gender-matched antecedent, e.g., *Mr. Jones*, in the multiple match conditions (e.g., 15a) relative to the single match conditions (e.g., 15b) at the reflexive, indicating grammatically accurate constraint application.

Similar effects of grammatically accurate constraint application have been reported for other types of anaphoric dependencies. For example, Kush and colleagues (2012, as reported in Kush, 2013) reports evidence of grammatically accurate processing of Hindi reciprocals. Reciprocals in Hindi cannot be bound by non-local, non-c-commanding items, just as reflexive anaphors and reciprocals in English. Kush and colleagues tested sentences like those in (16) using self-paced reading.

(16) a. Grammatical, no intervener

DoctoroN-ne mariz-ki dekhbaal-karne wali nure-ke steSan
Doctors-ERG patient-GEN care-doing wala nurse-GEN station
me ek-durse ke-bare-me gupt-rup-se baat kii.
in one-another about secretly chat did
'The doctors talked about each other in the station of the nurse caring for (the) patient.'

b. Grammatical, intervener

DoctoroN-ne marizo-ki dekhbaal-karne wali nure-ke steSan
Doctors-ERG patients-GEN care-doing wala nurse-GEN station
me ek-durse ke-bare-me gupt-rup-se baat kii.
in one-another about secretly chat did
'The doctors talked about each other in the station of the nurse caring for (the) patients.'

c. Ungrammatical, no intervener

Doctor-ne mariz-ki dekhbaal-karne wali nure-ke steSan
Doctors-ERG patient-GEN care-doing wala nurse-GEN station
me ek-durse ke-bare-me gupt-rup-se baat kii.
in one-another about secretly chat did
'The doctor talked about each other in the station of the nurse caring for (the) patient.'

d. Ungrammatical, intervener

DoctorN-ne marizoN-ki dekhbaal-karne wali nure-ke steSan
Doctor-ERG patients-GEN care-doing wala nurse-GEN station
me ek-durse ke-bare-me gupt-rup-se baat kii.
in one-another about secretly chat did
'The doctor talked about each other in the station of the nurse caring for (the) patients.'

In their experiment, reading times at the reciprocal region were affected by the (mis-)match with the number match to the local subject, *Doctor(oN)-ne* but not by corresponding (mis-)matches to the structurally irrelevant distractor, *mariz(oN)-ki*.

Alcocer and colleagues (2010; see also Alcocer, 2011) tested the processing of null subjects in Brazilian Portuguese, which require access to the subject of the next higher clause. Using number agreement as a diagnostic for the accuracy of constraint application, as in (17), self-paced reading times from their study show that comprehenders are sensitive to the number (mis-)match to the subject of the next higher clause, *O(s) turista(s)*, but not the number (mis-)match to the structurally irrelevant distractor, *o(s) ladr{ã|oes}*.

(17) a. Os turistas que o ladrão enganou na rua deserta
DET.PL tourist.PL that DET.SG thief.SG fooled.SG in-the street deserted
perceberam que *pro* extavam numa area ruim da cidade.
noticed.PL that *pro* were.PL in-a area bad of-the city.
'The tourists that the thief fooled on the deserted street noticed that they were in a bad part of the city.'

b. O turista que o ladrão enganou na rua deserta
DET.SG tourist.SG that DET.SG thief.SG fooled.SG in-the street deserted
perceberam que *pro* extavam numa area ruim da cidade.
noticed.PL that *pro* were.PL in-a area bad of-the city.
'The tourist that the thief fooled on the deserted street noticed that they were in a bad part of the city.'

c. O turista que o ladrões enganou na rua deserta
DET.SG tourist.SG that DET.SG thief.PL fooled.SG in-the street deserted
perceberam que *pro* extavam numa area ruim da cidade.
noticed.PL that *pro* were.PL in-a area bad of-the city.
'The tourist that the thieves fooled on the deserted street noticed that they were in a bad part of the city.'

There are at least two exceptions to the finding of grammatically accurate constraint application for reflexive anaphors. King and colleagues (2012) observed sensitivity to structurally irrelevant, but feature matched antecedents in sentences involving a reflexive embedded inside a post-verbal prepositional phrase (PP). Following Sturt (2003), King et al. manipulated the gender match of the target and distractor to the reflexive. But they also manipulated the distance between the verb and the reflexive using a dative/benefactive alternation. In the predicate-adjacent conditions, the reflexive appeared immediately following the verb as part of a double-object construction (18). In the predicate-separated conditions, the reflexive was embedded inside a prepositional object that was separated from the verb by the direct object and the preposition (19).

(18) Predicate-adjacent reflexive

c. The bricklayer who employed Gregory shipped himself sacks of mortar.

d. The bricklayer who employed Helen shipped himself sacks of mortar.

e. *The bricklayer who employed Helen shipped herself sacks of mortar.

f. *The bricklayer who employed Gregory shipped himself sacks of mortar.

(19) Predicate-separated reflexive

a. The bricklayer who employed Gregory shipped sacks of mortar to himself…

b.The bricklayer who employed Helen shipped sacks of mortar to himself…

c.*The bricklayer who employed Helen shipped sacks of mortar to himself…

d.*The bricklayer who employed Gregory shipped sacks of mortar to himself…

King and colleagues tested the sentences in (18) and (19) using eye-tracking. In the predicate-adjacent conditions, reading times were affected by the gender (mis-)match to the target, but not by the corresponding (mis-)match to the distractor. However, a different profile was observed in the predicate-separated conditions, where reading were affected by the gender match of the structurally illicit distractor. In predicate-separated conditions with a gender-matched distractor, King and colleagues observed reading times at the reflexive that were comparable to the grammatical counterparts, indicating sensitivity to the structurally illicit but gender-matched distractor. However, it is important to note that a reflexive in a PP may be associated with different licensing conditions than a reflexive in a direct object position, and further research is needed to better understand the licensing conditions for predicate-separated reflexives. I discuss these findings in more detail in section 2.3.1.

Cunnings and Felser (2012) also report evidence of sensitivity to grammatically illicit antecedents in the processing of reflexives. They investigated whether working memory capacity influences the degree to which comprehenders consider grammatically irrelevant antecedents when processing a reflexive. They tested items similar to those in Sturt (2003) using eye-tracking. In Experiment 1 of their study, Cunnings and Felser found that comprehenders with high working memory span accurately apply the structural constraints on reflexives in the early stages of processing. But, a different profile was observed in the data for comprehenders with low working memory span, as low span comprehenders showed immediate sensitivity to the feature match of a linearly closer but structurally inaccessible distractor. However, this profile did not replicate in Experiment 2, but rather reversed, where high span comprehenders showed immediate sensitivity to the feature match of the structurally illicit distractor.

Overall, the results from these studies converge on the same conclusion: the structural constraints that govern anaphoric dependencies are applied immediately and accurately. The effects of grammatically accurate constraint application are consistent and robust across different paradigms and methodologies (e.g, proactive and retroactive interference paradigms, eye-tracking while reading, visual world eye-tracking, self-paced reading, cross-modal lexical priming), feature manipulations (e.g., number and gender match), languages (e.g., English, Mandarin Chinese, Hindi, Brazilian Portuguese), and anaphor types (e.g., argument reflexives, benefactive reflexives, possessed picture noun phrase reflexives, picture noun phrase reflexives, reciprocals, and null subjects). Taken together, these results provide strong evidence for the use of a structure-guided memory access strategy for anaphoric dependencies.

Illusions of grammaticality

In contrast to the grammatical constraints for anaphora, which appear to be implemented rapidly and accurately, a number of other constraints appear to be implemented with much less precision during real-time comprehension, leading to cases of linguistic illusions. This subsection reviews two dependencies that have been shown to be highly susceptible to illusions: subject-verb agreement and NPIs.

Illusory subject-verb agreement licensing (agreement attraction)

Subject-verb agreement in English and many other languages is subject to a relatively straightforward constraint: the morphological features of the verb must agree with the corresponding features of the local subject, as shown in (20).

(20) a. The books$_{[+pl]}$ were$_{[+pl]}$ on the table
b. *The books$_{[+pl]}$ was$_{[+sg]}$ on the table.

As such, the featural and structural requirements on subject-verb agreement are identical to those that govern reflexive licensing in English. The similarity between these two sets of constraints makes the finding that the parser is highly to susceptible to errors in the implementation of subject-verb agreement all the more surprising.

Many studies have shown that the implementation subject-verb agreement is sensitive to structurally irrelevant, but feature matched distractors, and effect termed 'agreement attraction'. Agreement attraction involves facilitatory interference or 'intrusion' from structurally irrelevant material, and the effect qualifies as a linguistic illusion because speakers or comprehenders typically fail to initially notice the error. More specifically, agreement attraction occurs when the morphological features of the verb agree with the corresponding features of another constituent in the sentence that is not the local subject. Such errors are found frequently in natural speech and writing. For example, Bock and Miller (1991) presented sentence fragments like those in (21) to participants in a sentence production study.

(21) a. The key to the cabinet . . .
b. The key to the cabinets . . .

Bock and Miller found that when the object of the subject-modifying PP was plural, e.g., *cabinets*, there was a significant increase in the number of trials where participants produced a plural verb form, such as *were*, as compared to when the object of the PP was singular, e.g., *cabinet.* Similar effects have been observed for English (e.g., Eberhard et al., 2005; Franck et al., 2002; Gillespie & Pearlmutter, 2011; Staub, 2010; Vigliocco & Nicol, 1998) and other languages such as Dutch (Hartsuiker, Schriefers, Bock, & Kikstra, 2003), French (Franck et al., 2002), German (Hartsuiker et al., 2003), Italian (Vigliocco & Franck, 2001), Spanish (Antón-Mendex, Nicol, & Garrett, 2002)**,** and Slovenian (Badecker & Kuminiak, 2007).

Importantly, analogue effects of agreement attraction are observed in comprehension. For example, Pearlmutter et al. (1999) tested the processing of subject-verb agreement in sentences like those in (22) using both self-paced reading and eye-tracking. Note that the design in (22) has the same form as the paradigms that are typically used to probe for sensitivity to structurally illicit antecedents in studies on reflexive licensing. Across these studies, grammaticality is manipulated by varying a single feature of the agreeing element, e.g., the number of the agreeing verb or reflexive, such that it either matches or mismatches the corresponding feature of the target, e.g., the local subject, while independently varying a single feature of the distractor, e.g., number.

(22) a. The key to the cabinet was rusty from years of disuse.
b. The key to the cabinets was rusty from years of disuse.
c. *The key to the cabinets were rusty from years of disuse.
d. *The key to the cabinet were rusty from years of disuse.

Pearlmutter and colleagues observed consistent effects across both studies. Reading times after the verb were disrupted by agreement errors, as reflected by slower reading times in the ungrammatical conditions (22c-d) relative to their grammatical counterparts (22a-b). However, this disruption was reduced in the ungrammatical sentences due to the presence of a structurally irrelevant feature matched distractor (22c vs. d). The presence of a feature matched distractor also influenced reading times within the grammatical conditions, as reading times from the distractor region onward were read more slowly in the conditions that contained a feature matched distractor.

Wagers and colleagues (2009) showed that similar results obtain across a range of configurations and methodologies. For example, they tested both retroactive and proactive configurations, with sentences like those in (23) and (24), using self-paced reading and speeded-acceptability judgments.

(23) Retroactive interference paradigm
a. The key to the cell unsurprisingly was rusty from many years of disuse.
b. They key to the cells unsurprisingly was rusty from many years of disuse.
c. *The key to the cells unsurprisingly were rusty from many years of disuse.
d. *The key to the cell unsurprisingly were rusty from many years of disuse.

(24) Proactive interference paradigm
e. The musician who the reviewer praises so highly will probably win a Grammy
b. The musicians who the reviewer praises so highly will probably win a Grammy
c. *The musicians who the reviewer praise so highly will probably win a Grammy.
d. *The musician who the reviewer praise so highly will probably win a Grammy.

Consistent with the findings reported in Pearlmutter et al. (1999), Wagers and colleagues found that the presence of a morphological feature match with the structurally irrelevant distractor impacted the processing of the subject-verb agreement dependency. However, unlike Pearlmutter et al. (1999), Wagers and colleagues found that comprehenders were sensitive to the feature match of the distractor only in the ungrammatical conditions. This asymmetry was replicated across seven experiments involving self-paced reading and speeded-acceptability judgments. In self-paced reading measures, the presence of a plural distractor in the ungrammatical conditions eased the processing of the ungrammatical verb, as reflected by facilitated reading times for ungrammatical sentences with a plural distractor relative to ungrammatical sentences that lacked a plural distractor. In speeded-acceptability judgments, participants were more likely to accept ungrammatical sentences when a plural distractor was present. This increased rate of acceptance corresponds to an 'illusion of acceptability'.

Wagers et al. (2009) argued that the eased processing and increased acceptability of an illicit agreeing verb in the presence of a structurally irrelevant but feature matched distractor is expected in a direct-access, content-addressable memory architecture. Upon encountering an agreeing verb in the input, direct-access, content-addressable retrieval mechanisms probe all memory items in parallel for matches to the features of the verb, e.g., [+subject] and [+plural], as required by the grammar. In the configurations that give rise to linguistic illusions, the candidate NPs in memory match only one of the two retrieval cues. That is, illusions occur in partial-match contexts: the structurally defined licensor, i.e., the matrix subject, matches the structural cue [+subject] but mismatches the morphological number cue, e.g., [+plural], whereas the structurally irrelevant distractor mismatches the structural cue, but matches the morphological number cue. On some portion of trials, direct-access, content-addressable retrieval may mis-retrieve the partially activated distractor, misleading the comprehender into a false impression of grammaticality/acceptability. Under this account, the lack of interference in the grammatical conditions (which would result in an illusion of ungrammaticality) can be captured by the fact that the probability of mis-retrieval is greatly reduced in these conditions due to the presence of a target that perfectly matches the structural and morphological retrieval cues.[8]

Similar agreement attraction effects indicative of mis-retrieval within a content-addressable architecture have also been found in eye-tracking measures. Dillon et al. (2013) for example tested sentences like those in (25), in which the distractor was contained inside a subject-modifying relative clause.[9]

(25) a. The new executive who oversaw the middle manager apparently was dishonest about the company'
b. The new executive who oversaw the middle managers apparently was dishonest about the company'
c. *The new executive who oversaw the middle managers apparently were dishonest about the company'
d. *The new executive who oversaw the middle manager apparently were dishonest about the company'

[8] An alternative account for the grammatical asymmetry proposed by Wagers and colleagues is that a direct-access, content-addressable retrieval for subject-verb agreement is triggered only when the verb does not have the anticipated agreement features. In configurations where the verb has the anticipated, correct agreement, no retrieval of the subject features is required, preventing the possibility for retrieval error. In the ungrammatical configurations, by contrast, the verb does not have the anticipated agreement features, triggering a error-prone content-addressable retrieval. I discuss the role of prediction in subject-verb agreement processing further in Chapter 4.

[9] Experiment 1 of Dillon et al. (2013) directly subject-verb agreement and reflexives using parallel conditions like those in (10).

Like Wagers et al. (2009), Dillon and colleagues found that reading times were disrupted by agreement errors, but the presence of a plural distractor led to reduced disruption effects in total reading times.

Subject-verb agreement comprehension errors are also attested in languages other than English. For example, Lago, Alcocer, and Phillips (2011) showed that Spanish speakers are susceptible to the same attraction errors as English speakers, as based on evidence from self-paced reading measures and speeded-acceptability judgments. In an earlier study, Alcocer and Phillips (2009) showed that the attested grammatical asymmetry is present for agreement attraction in Spanish, but the effect is the exact opposite of the English effect, with increased errors of acceptance of agreement attraction configurations in grammatical conditions only. Lago, Shalom, Sigman, Lau, & Phillips (2014) also observe agreement attraction in both grammatical and ungrammatical configurations. Alcocer et al. (2010) also reported effects of agreement attraction in Brazilian Portuguese.

Before I discuss the implications of agreement attraction effects for our understanding of encoding and access mechanisms, it should be emphasized that agreement attraction effects do not merely reflect proximity concord or local coherence. For example, robust agreement attraction effects are observed in production and comprehension for non-local configurations where a relative clause verb agrees with the head of the main clause, rather than the subject of the relative clause, as for the proactive interference paradigm in (24). With this non-local configurations, there is not a locally coherent interpretation that could give rise to the observed pattern of results. Furthermore, agreement attraction is not a case of variable of incorrectly described grammar. When comprehenders are given sufficient time, they are highly sensitive to agreement errors and agree on their unacceptability (e.g., see Dillon et al., 2013 for supporting untimed acceptability judgments). Importantly, it is the mismatch between these measures of considered judgments and the effects that are observed during real-time processing, or when under a time pressure, that makes this effect qualify as a linguistic illusion. Lastly, agreement attraction is not a case of dialectal variation, as the same participants who show high sensitivity to agreement errors in untimed or off-line measures also fail to notice such errors in timed or online measures.

In sum, across a number of different configurations, methodologies, and languages, we find consistent errors of real-time constraint application for subject-verb agreement leading to linguistic illusions in exactly the same configurations that fail to show corresponding effects for anaphor processing. This is most clearly demonstrated by the within-subjects comparison made by Dillon et al. (2013), which showed that the same materials that yield errors of attraction for subject-verb agreement fail to yield corresponding errors for reflexive anaphors.

Illusory NPI licensing

NPIs are words like *ever*, *any*, *yet*, or phrases like *lift a finger* or *a damn thing* that are licensed in the scope of a downward entailing (DE) expression (Ladusaw, 1979). DE expressions license logical inferences from superset to subset relations within their scope, as they entail a more specific proposition. For example, the statement *Jack didn't eat an apple* entails that *Jack didn't eat a red apple*, whereas *Jack ate an apple* does not entail that *Jack ate a red apple*. Negative-like words or phrases such as *no*, *not*, *few*, *rarely*, and *doubt*, as well as conditionals and expressions of surprise are DE expressions that can license NPIs (Ladusaw, 1979; Von Fintel, 1999; Zwarts, 1995, inter alia).[10] Current accounts of negative polarity phenomena claim that the licensing conditions on NPIs reflects an interaction between the semantic features of NPIs and the semantic and pragmatic properties of the environments that host NPIs (Chierchia, 2006; Giannakidou, 2011; Horn, 2010; Israel, 2004; Kadmon & Landman, 1993; Krifka, 1995; Ladusaw, 1979).

Under these accounts, the scope requirement on NPI licensing can be viewed as an emergent property of the interaction between the semantics and pragmatics of NPIs and their host environments, rather than an explicit grammatical constraint. Nonetheless, for current purposes, the descriptive generalization holds: an NPI must appear in the scope of a DE expression in order to obtain the necessary semantic and pragmatic effect. For example, the NPI *ever* in (26a) is appropriately licensed because it appears in the scope of the negatively quantified NP *no diplomats*. The scope of negation for purposes of NPI licensing corresponds roughly to the c-command domain of negation, i.e., the structural sister of the negation and any element contained within the structural sister. [11]

---

[10] Specifically, DE expressions, including negation, introduce an operator that takes scope over an entire context.

[11] There are cases that call for an elaboration of the c-command generalization. For example, in the sentence *Nobody's mother has ever served ice cream for dinner*, the NPI *ever* is licensed even though it is not syntactically c-commanded by the negation. In this case, it appears that the entire NP *nobody's mother* counts as the relevant c-commander.

When negation is absent, as in (26b), or fails to scope over the NPI, as in (26c), the NPI is not licensed.

(26) a. No diplomats have ever supported a drone strike.
b. *The diplomats have ever supported a drone strike.
c. *The diplomats that no congressmen could trust have ever supported a drone strike.

Comprehenders are highly sensitive to the presence and location of a potential licensor in the context prior to the NPI, as sentences like those in (26b) and (26c) are reliably judged to be unacceptable in off-line tasks, where participants are given sufficient time to make their judgment. However, in time-sensitive measures, sentences like (26c) are occasionally processed as if they were actually acceptable, leading to illusory licensing effects. For example**,** Drenhaus and colleagues (2005) used speeded-acceptability judgments and ERPs to test native German speakers' sensitivity to the licensing conditions of the German NPI *jemals* in sentences like (27). The German NPI *jemals* shares identical licensing constraints with the English NPI *ever*.

(27) a. Grammatical licensor
Kein Mann, der einen Bart hatte, war jemals glüklieh.
'No man who had a beard was ever happy.'
b. Irrelevant licensor
*Ein Mann, der keinen Bart hatte, war jemals glüklieh.
'*A man who had no beard was ever happy.'
c. No licensor
*Ein Mann, der einen Bart hatte, war jemals glüklieh.
'*A man who had a beard was ever happy.'

The NPI *jemals* is appropriately licensed in (27a) because it appears in the scope of the negatively quantified NPI *kein mann* ('no man'). Both (27b) and (27c) are ungrammatical. The licensor in (26b) does not scope over of the NPI, and (27c) lacks a licensor. Results of a speeded-acceptability judgment task showed that the presence of an intrusive licensor in ungrammatical sentences like (26b) increased rates of acceptance relative to ungrammatical sentences like (26c) that did not have a licensor at all. ERPs showed that while both ungrammatical sentences elicited an N400 effect upon processing an illicit NPI, the amplitude of the N400 for sentences with an structurally irrelevant licensor was significantly reduced relative to the no licensor condition.

Vasishth and colleagues (2008) obtained similar results using eye-tracking while reading in German. Vasishth and colleagues tested sentences like those in (27) and observed facilitated reading times at the NPI for sentences involving a structurally irrelevant licensor relative to sentences with no licensor. These results have also been corroborated in English using speeded-acceptability and ERPs. For example, Xiang and colleagues (2006; 2009) tested sentences involving a wide range of NPI licensors, including *no*, and *few*, like those in (28).

(28) a. Grammatical licensor
{No | very few} bills that the Democratic senators have supported will ever become law.
b. Irrelevant licensor.
The bills that {no | very few} Democratic senators have supported will ever become law.
c. No licensor
The bills that the Democratic senators have supported will ever become law.

In speeded-acceptability measures, they observed increased rates of acceptance for sentences that contained a structurally irrelevant licensor, both for sentences involving the licensors *no* and *few*, relative to sentences that had no licensor at all. The impact of a structurally irrelevant licensor was also observed in ERPs, as reflected by a reduced P600 effect for sentences involving a structurally irrelevant licensor relative to sentences that lacked a licensor.

The key finding from these studies is that the disruption associated with the processing of an illicit NPI was significantly reduced in sentences that contain a semantically compatible licensor that is in a structurally illicit or non-c-commanding position. This effect suggests that in some portion of trials, comprehenders treated the illicit NPI as if it were acceptable, or on a par with the grammatical licensed cases. That is, the presence of a structurally irrelevant licensor can ease the processing of an illicit NPI during real-time comprehension, making an ungrammatical sentence with an otherwise unlicensed NPI temporarily appear well-formed.

Like agreement attraction, NPI illusions have been analyzed as mis-retrieval in a content-addressable memory architecture. For example, Vasishth and colleagues have presented illusory NPI licensing as evidence for a general dependency formation mechanism that is sensitive to partial matches between retrieval cues and memory encodings (Vasishth et al., 2008). Their account is based on recent findings of similarity-based interference effects in on-line dependency resolution, which suggest that at least some encodings of linguistic structure in memory allow matches to their sub-parts (e.g., Gordon, Hendrick, &

Johnson, 2001, 2004; Gordon, Hendrick, Johnson, & Lee, 2006; Gordon, Hendrick, & Levine, 2002; R. L. Lewis, 1996; Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006). According to Vasishth and colleagues, encountering an NPI initiates a retrieval for an item in memory that has the properties [+negative] and [+c-command]. Illusory licensing can arise when a non-commanding licensor is erroneously retrieved due to a partial match to [+negation], leading to facilitated processing of an otherwise unlicensed NPI.

The account proposed by Vasishth and colleagues is plausible under certain recent accounts of NPI licensing that assume that NPIs are licensed through a direct relation with particular interpretative features, such as the scalar feature proposed by Chierchia (2006). However, as described above, the diversity of NPIs and NPI licensors has led to the broad consensus that NPI licensing is driven by compositional interpretive mechanisms that rely on the semantic and pragmatic features of NPIs and entire propositions, rather than the type of item-to-item structural licensing mechanisms that mediate subject-verb agreement (e.g., Chierchia, 2006; Giannakidou, 1998; Israel, 2004; Kadmon & Landman, 1993; Krifka 1995; Ladusaw, 1992, but cf. Giannakidou 2006, 2011).

Based on this consensus that NPIs and subject-verb agreement rely on different licensing mechanisms, Xiang and colleagues have presented an alternative account of the NPI illusion which suggests that the illusion reflects over-active pragmatic accommodation, rather than mis-retrieval during item-to-item dependency formation (Xiang et al., 2009). Their account is based on the observation that in addition to direct semantic licensing through DE environments, NPIs may also be licensed through negative inferences. For instance, the NPI *ever* is licensed in a sentence like (29a) through a negative implicature (29b).

(29) a. I am surprised that John ever finished that assignment.
b. I expected that John would not finish that assignment.

According to Xiang and colleagues, in cases involving an illicit NPI, such as in (30), the parser may over-apply normal pragmatic licensing mechanisms in an attempt to make sense of the sentence. Specifically, they suggested that when parsing a sentence like (30), comprehenders may reasonably infer that the set denoted by the complex NP has some property *P*, which denotes the upcoming predicate, e.g. *the bills that no democratic senators voted for will P*. However, the embedded negation and the restrictive relative clause may also invite an additional negative inference about the contrasting set of referents, namely that *the bills that some democratic senators voted for will not have P*. Although this inference is pragmatically sensible, it is not logically necessary, and it cannot appropriately license the NPI because it does not have similar truth-conditions as the original proposition (which is a requirement for licensing NPIs by a pragmatic inference). Nevertheless, when the parser attempts to make sense of the illicit NPI in sentences like (30), the availability of such inferences may spuriously license the NPI or boost its acceptability.

(30) The bills that no democratic senators have voted for will ever become law.

The accounts of illusory NPI licensing proposed by Vasishth et al. (2008) and Xiang et al. (2009) differ in their views about the licensing conditions on NPIs, but agree that NPI licensing is a function of (i) the licensing conditions on NPIs, (ii) the encoding of the context, and (iii) noisy access mechanisms. Under both accounts, the illusion may be understood as a kind of partial-match effect, and suggests that the licensing mechanism can access semantic licensing features on-line, independently from the position of those features in the structured representation of the sentence. Importantly, these studies have tested for illusory licensing effects by probing comprehenders processing at different time points after the appearance of the NPI, and they have shown that sensitivity to the structural properties of a potential licensor grows as the time from when the NPI was encountered in the sentence increases.

Illusory NPI licensing and illusory subject-verb agreement licensing are qualitatively similar in several ways. The key similarity is that in both cases, the presence of a structurally irrelevant licensor temporarily misleads comprehenders into a false impression of acceptability. The illusions that are observed for both of these dependencies are robust across methodologies and languages. Furthermore, just like illusory agreement licensing, illusory NPI licensing, is not a case of variable or incorrectly described grammar, or a case where the licensor somehow takes scope outside of its clause. Comprehenders agree on the unacceptability of these sentences when they are given sufficient time. Crucially, cases of illusory NPI licensing and illusory agreement licensing reflect a mismatch between the behavior observed immediately after the appearance of the NPI in time-sensitive measures and the behavior observed at a much later point when comprehenders have had plenty of time to consider the sentence, and it is this mismatch that makes these effects qualify as a linguistic illusion. Furthermore, both types of illusions can be captured as mis-retrieval of the structurally irrelevant licensor in a content-addressable memory, an attractive, uniform account for both phenomena. In sum, the findings for subject-verb and NPIs have been presented as importance evidence for inaccurate on-line implementation of grammatical constraints. Taken together, these results have been argued to provide evidence for the use of a content-addressable access mechanism in real-time sentence comprehension (but cf. Xiang et al., 2009).

2.4 Leading generalizations

The on-line effects of anaphora, subject-verb agreement, and NPIs reviewed above yield a profile of selective fallibility. There is robust evidence across languages and methodologies that the structural and featural constraints on reflexive anaphors are implemented immediately and accurately during real-time language processing, with very limited evidence that the parser considers grammaticality irrelevant antecedents. The robust immunity from illusions for reflexive licensing is expected if a structure-guided access mechanism is used to recover the antecedent from memory. By contrast, there is robust evidence that the constraints on subject-verb agreement and NPIs have a much weaker effect on real-time dependency formation. Under certain circumstances, typically restricted to ungrammatical contexts, structurally irrelevant but semantically or featurally compatible licensors may be fleetingly considered, in violation of the relevant structural constraints on the dependency, giving rise to a linguistic illusion. In these cases, the illusions are predicted as mis-retrieval in a content-addressable memory architecture.

These findings have led to two generalizations. First, the candidate access mechanisms are directly linked to specific types of grammatical dependencies. For example, the evidence reviewed above suggests that anaphoric dependencies are tied to a structure-guided access mechanism, whereas dependencies involving subject-verb agreement and possibly NPIs are tied to a direct-access strategy. Second, similar illusory licensing profiles for subject-verb agreement and NPIs has led to the attractive idea that there is a homogenous cause for the failure, namely the mis-retrieval of structurally irrelevant material in memory.

However, it remains unclear under current accounts why different dependencies would engage different access strategies. Previous research has tried to explain why different dependencies would make use of different memory access strategies by appealing to underlying grammatical organization of the different dependencies. For example, Dillon (2011) discusses the possibility that the use of different memory access strategies could reflect the fact that anaphoric dependencies are interpreted whereas agreement is not. Under this view, the parser might engage a more conservative, structure-guided memory access strategy for anaphora to ensure proper interpretation in the case of anaphora. I explore this possibility further in Chapter 3.

In sum, existing findings show that subject-verb agreement and NPI licensing are susceptible to illusions, as expected with the use of a content-addressable access mechanism. Anaphor licensing is immune to illusions, as expected with the use of a structured-access strategy. These findings have given rise to the claim that the candidate access mechanisms are directly linked to specific dependencies, and that when illusions arise, they reflect mis-retrieval in a content-addressable memory architecture. Over the next three chapters, I will investigate further the source and scope of the effects of (non-)illusions, and I will present empirical and computational evidence that suggests that these generalizations are incorrect, motivating a new conception of how we encode and navigate linguistic structure.

Chapter 3 Null subject licensing in adjunct control*

3.1 Introduction

In Chapter 3, we test the hypothesis that the contrasting illusory licensing profiles seen for anaphora and agreement in previous studies reflect the fact that anaphoric dependencies are interpreted whereas agreement is not. To achieve this, we examine the processing of null subject licensing in English. Null subject licensing has not received much attention in the sentence comprehension literature, but it provides a good test to better understand what is responsible for the contrasting illusions because it shares properties with both agreement and anaphor licensing.

As discussed in Chapter 2, the contrasting illusion profiles observed for subject-verb agreement and reflexive anaphors is not predicted by existing accounts One possibility discussed by Dillon (2011; see also Dillon et al., 2013) is that differences in interpretative content between subject-verb agreement and anaphora drive the contrasting illusory licensing profiles. This explanation appeals to the intuitive notion that antecedent-reflexive dependencies are semantically interpreted, but subject-verb agreement is not. For example, subject-verb agreement is described as a low-level, morpho-syntactic process that has no immediate consequences for the meaning or interpretation of a sentence. By contrast, since proper reference resolution is necessary for correctly recovering the intended meaning or interpretation of the sentence, the parser might selectively engage a more structured, accurate memory access mechanism for

* Chapter 3 reports research that was jointly conducted with Sol Lago.

anaphor resolution. According to this hypothesis, structured memory access is a general property of interpreted anaphoric dependencies. If this hypothesis is correct, then structured access should be observed for all anaphoric dependencies, such as those involving reflexives and null subjects.

3.2 Null subject licensing

Null subject licensing in adjunct control is an anaphoric dependency that relies on a structural constraint: the licensor ('controller') must be the subject of the next higher clause. For example, in (1), the null subject represented as PRO is licensed by the NP *the little girl* because it is the subject of the next higher clause.[12] Other controllers are not possible, e.g., the NP *the mother*, for example, is too remote and it is not the subject of the next higher clause. The absence of a reading for (2a) analogous to (2b) indicating the antecedent must be sentence-internal.

(1) The $mother_1$ said [the little $girl_2$ fell asleep after $PRO_{*1/2}$ playing in the back yard].

(2) a. *The picture startled the little girl [after PRO_{arb} running in the room].
b. The picture startled the little girl [after *someone* ran into the room].

Null subject licensing provides a good test to better understand what is responsible for the contrasting attraction profiles between agreement and reflexives because it shares properties with both dependency types. First, both null subject licensing and reflexive licensing are interpreted anaphoric

[12] I represent the missing subject as PRO throughout the paper for ease of exposition. None of our discussion turns upon whether the missing subject is an empty category.

dependencies that require access to a full subject NP for interpretation. By 'interpretative', we mean that the processor must assign an interpretation to the dependent element (i.e. PRO or a reflexive), as this element lacks independent semantic content. Conversely, null subject licensing shares at least one important property with subject-verb agreement: retrieval for a licensor is triggered upon encountering a verb rather than by an independent anaphoric element. In sentences like (1), detection of the null subject in an adjunct control construction is based on encountering a gerundive verb with no preceding subject in the input, which should trigger a retrieval process to recover the subject of the verb from the next higher clause.

If the parser engages a structure-guided memory access strategy for reflexive dependencies because of their interpretative status, and if structured access is a general property of interpreted anaphoric dependencies, then memory retrieval for null subject licensing should proceed in a structured, grammatically accurate fashion. Therefore, we should see sensitivity only to the properties of the structurally licit licensor, the subject of the next higher clause, and insensitivity to the properties of other structurally irrelevant material. This should result in an attraction-free processing profile, as observed in previous studies of reflexive licensing. If, however, verbal dependencies engage a noisy, error-prone memory access strategy, then we should observe attraction from potential licensors in structurally irrelevant locations for the null subject licensing because subject retrieval is triggered by the gerundive verb. This would predict an attraction profile for null subject licensing similar to the one observed in previous studies of subject-verb agreement.

However, one challenge in devising a test to probe for attraction effects for null subject licensing is that gerundive verbs in adjunct control structures do not express any of the cues typically use to probe for attraction effects, such as number or gender, are not expressed by the gerundive verb in adjunct control structures nor are they required by the controller-controlee relationship. However, null subject licensing has been claimed to be sensitive to an animacy constraint. For example, Kawaski (1993) reports that sentences involving adjunct control are judged to be more acceptable when the controller is animate, as in (3a), as compared to when the controller is inanimate, as in (3b). This animacy preference might be a property of control constructions involving an adjunct subordinator, such *after*, and a gerundive verb, as the contrast between (3a) and (3b) is neutralized when the same nouns function as overt subjects for the verb, as shown in (4) (see Landau, 2001, for supporting judgments).

(3) Adjunct control structures

a. The doctor was certified after PRO debunking the urban myth.
b. The discovery was certified after debunking the urban myth.

(4) Overt subject structures

a. The journalist was surprised that the doctor debunked the urban myth.
b. The journalist was surprised that the discovery debunked the urban myth.

Previous studies have shown that various dependencies make use of animacy information in memory retrieval for range of dependencies (Van Dyke & Lewis, 2003; Van Dyke & McElree, 2006; Van Dyke & McElree, 2011). Based on the results from these studies, taken together the judgments reported for sentences like those in (5), we employed a similar test of animacy in the present study. Specifically, we hypothesized that comprehenders might exploit the animacy preference described above during memory access for the licensor of a null subject in adjunct control constructions.

3.3 Experiment 1: Off-line acceptability ratings

The goal of Experiment 1 was to confirm that the licensor of a null subject in adjunct control structures is preferentially interpreted as animate using an untimed, off-line acceptability rating study. If comprehenders prefer an animate licensor for adjunct control structures, animacy could be used as a cue in memory retrieval for null subject licensing.

3.3.1 Participants

Twenty-four participants were recruited using Amazon's Mechanical Turk crowdsourcing web-service (http://aws.amazon.com/mturk/). All participants provided informed consent, and were paid $2.00 for their participation. The experiment lasted approximately 10 minutes.

3.3.2 Materials

Twenty-four sets of items like those in (5-6) were constructed. Two experimental factors were manipulated: subject animacy (animate vs. inanimate) and construction (overt vs. covert subject). The 24 item sets were equally distributed across four lists in a Latin Square design. Within each list, the 24 target sentences were mixed with 48 filler sentences of similar length and complexity, for a total of 98 sentences. The order of the sentences was randomized for each participant. The ratio of grammatical-to-ungrammatical sentences was 1:1. The anomalies in the ungrammatical fillers involved several different types of grammatical errors, including unlicensed NPIs, subject-verb agreement mismatches, and unlicensed verbal morphology, e.g., *-ed* → *-ing*.

3.3.3 Procedure

Participants were asked to rate the acceptability of the sentences along a 7-point Likert scale ('7'=most acceptable, '1'=least acceptable), according to their perceived acceptability in informal, colloquial speech. Ratings were untimed, but participants were required to complete the task in 30 minutes.

3.3.4. Analysis

Statistical analysis was performed using mixed-effect linear regressions to assess the reliability of the effects associated with the experimental factors, with fixed factors for experimental manipulations and their interaction. Models were estimated using the *lme4* package (Bates, Maechler, & Bolker, 2011) in the R software environment (R Development Core Team, 2014). We used orthogonal contrast coding for experimental fixed effects (subject animacy and construction ±0.5 for each factor) and their interaction. The fixed effect for animacy was coded as +0.5 for animate subject conditions and -.05 for inanimate subject conditions. The fixed effect for construction was coded as +0.5 for control constructions and +0.5 for overt subject constructions. In addition to these fixed effects, items and participants were crossed as random effects (following Baayen, Davidson, & Bates, 2008; Baayen, 2008). To determine whether inclusion of random slopes was necessary, we compared a model with a fully-specified random effects structure, which included random intercepts and slopes for all fixed effects and their interaction by items and by participants (Baayen et al., 2008; Barr, Levy, Scheepers, & Tily, 2013) to an intercept-only model. A log-likelihood ratio test between the two models revealed that the maximal model provided a significantly better fit to the data ($X^2_{(2)}$=67.36, p<0.001;). Therefore, we adopted the model with a maximally specified random effects structure. For all statistical analyses, an effect was considered significant if its absolute *t* or *z*-value was greater than 2 (Gelman & Hill, 2007).

3.3.5 Results

The results of the acceptability rating study are presented in Figure 3.1. The results revealed that in the adjunct control conditions, adjunct control sentences with an inanimate subject received a mean rating of 4.81, whereas the corresponding sentences with an animate subject received a mean rating of 6.09.

In contrast, sentences with an overt subject received similar ratings, with a mean rating of 6.43 for sentences with an inanimate subject and a mean rating of 6.40 for sentences with an animate subject. The statistical analysis revealed a main effect of subject animacy ($\hat{\beta}$=-0.64, SE=0.18, t=-3.57) a main effect of construction ($\hat{\beta}$=-0.95, SE=0.19, t=-4.95), and a significant interaction between subject animacy and construction ($\hat{\beta}$=-1.24, SE=0.32, t=-3.87). Planned pairwise comparisons revealed that this interaction was driven by a significant difference due to subject animacy in the adjunct control sentences, in which sentences with an inanimate subject were rated significantly lower than control sentences with an animate subject ($\hat{\beta}$=-1.26, SE=0.29, t=-4.32), whereas no such effect was found for the overt subject sentences.

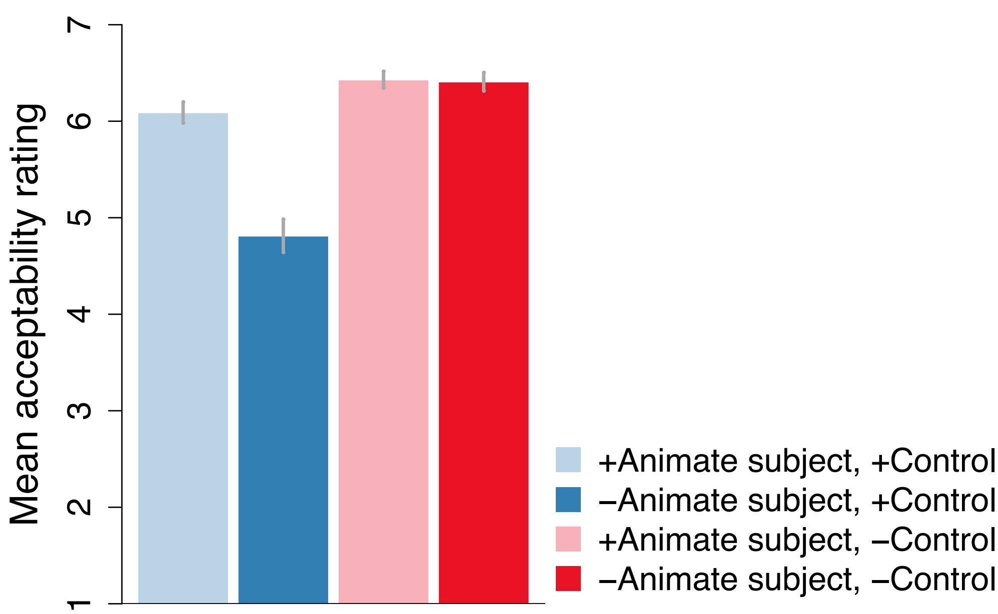


Figure 3.1: Mean ratings and standard error by participants for Experiment 1 acceptability rating study. Values are on a 7-point Likert scale, with '7' being most acceptable, and '1' the least acceptable.

3.3.6 Discussion

The results of Experiment 1 confirm that comprehenders prefer an animate controller to an inanimate one, but they suggest that licensing by an animate subject is not a hard grammatical constraint on the dependency between the controller and controlee in adjunct control structures, as participants still rated sentences involving an inanimate controller as relatively acceptable. The preference for an animate subject stands in contrast with other hard constraints that have been tested in previous studies, such as the number or gender requirement on the relationship between a reflexive and its antecedent.

The finding that the overt subject sentence received similar ratings for animate and inanimate subjects suggests that the animacy preference is not a lexical restriction of the adjunct verb. Instead, the present results suggest that the animacy preference may be a consequence of the compositional interpretation of adjunct control constructions involving a subordinator followed by a gerundive verb. While further research is needed to understand the source of this preference, it is possible that comprehenders could use animacy as a cue in memory retrieval for null subject licensing.

3.4 Experiment 2

Experiment 2 directly compared the processing of null subject dependencies and subject-verb agreement in English using self-paced reading. The main question for Experiment 2 was whether the attraction effects commonly observed for subject-verb agreement would extend to adjunct control. To address this question, we used the animacy preference confirmed in Experiment 1 to investigate the impact of structurally illicit but feature matched distractors on the processing of adjunct control. If comprehenders deploy a structure-guided memory access mechanism for all anaphoric dependencies, as predicted by the interpretative content hypothesis described in the introduction of this chapter, then we should observed contrasting illusory licensing profiles for agreement and null subject licensing, with an absence of illusory licensing effects for null subject licensing.

3.4.1 Participants

Thirty-two members of the University of Maryland community participated in Experiment 1. All participants gave informed consent, and received either course credit or payment of $10 for their participation. The experimental session lasted approximately 45 minutes.

3.3.2 Materials

Two experimental factors were manipulated for adjunct control and subject-verb agreement: grammaticality and attraction. Within the subject-verb agreement conditions, grammaticality was manipulated by varying the number of the agreeing verb, which was singular in the grammatical conditions and plural in the ungrammatical conditions. Attraction was manipulated by varying the number of the distractor NP such that it either matched the agreeing verb (+distractor conditions) or mismatched (-distractor conditions). The main clause subject was always singular. Within the adjunct control conditions, grammaticality was manipulated by varying the animacy of the main clause subject, which was animate in the grammatical conditions and inanimate in the ungrammatical conditions. Attraction was manipulated by varying the animacy of the distractor NP. The main clause subject and distractor NPs were always semantically appropriate as potential subjects of the adjunct verb to prevent further biases in subject retrieval.

Forty-eight item sets of the items like those in Table 3.1 were constructed. In all conditions, the main clause subject was modified by an object relative clause that contained the distractor NP in subject position. The relative clause verb never expressed agreement to prevent biases in retrieval. A post-verbal adverbial was used in all conditions to signal the end of the relative clause. The main clause was passivized for all conditions. Passive structures were used for two reasons. First, they avoid introduction of an additional distractor NPs. The use of intransitive structures was not possible due to the limited number of main clause verbs that could naturally allow both an animate and inanimate subject. Second, passive structures provided a clear attachment site for the adjunct clause to the main clause VP, avoiding potential attachment ambiguities.

In the adjunct control conditions, an emphatic reflexive that was embedded inside the adjunct clause served as a critical probe to determine whether or not an animate NP was retrieved as the subject of the adjunct clause. Although it is not the case for all emphatic reflexives, the emphatic reflexive that we used in our materials required an animate, gender-matching antecedent as the local subject, i.e. the null subject PRO. As such, the processing profile of the emphatic reflexive should provide a passive reflection of what was retrieved as the subject of the adjunct clause.

Table 3.1: Summary of experimental items for Experiment 2.

Subject-verb agreement conditions

Grammatical, no attraction

1. The doctor that the researcher described meticulously was certified after debunking the urban myth in the new scientific journal.

Grammatical, attraction

2. The doctor that the reports described meticulously was certified after debunking the urban myth in the new scientific journal.

Ungrammatical, attraction

3. The doctor that the researchers described meticulously were certified after debunking the urban myth in the new scientific journal.

Ungrammatical, no attraction

4. The doctor that the report described meticulously were certified after debunking the urban myth in the new scientific journal.

Adjunct control conditions

Grammatical, attraction

5. The doctor that the researcher described meticulously was certified after debunking the urban myth himself in the new scientific journal.

Grammatical, no attraction

6. The doctor that the report described meticulously was certified after debunking the urban myth himself in the new scientific journal.

Ungrammatical, attraction

7. The discovery that the researcher described meticulously was certified after debunking the urban myth himself in the new scientific journal.

Ungrammatical, no attraction

8. The discovery that the report described meticulously was certified after debunking the urban myth himself in the new scientific journal.

Each sentence was followed by a comprehension question. Comprehension questions addressed various parts of the sentence to prevent the possibility that participants might develop superficial reading strategies whereby they extracted only the information necessary to answer the comprehension question without reading the entire sentence.

3.3.3 Predictions

In the subject-verb agreement conditions, based on previous studies we expected to observe attraction effects, as reflected by facilitated reading times for ungrammatical sentences with a plural distractor relative to ungrammatical sentences with a singular distractor. No difference between the grammatical conditions was expected.

In the case of adjunct control, two scenarios are possible. If subject retrieval proceeds in a structure-guided fashion, as predicted by the interpretative content hypothesis, then we should observe only a main effect of grammaticality at both the gerundive and reflexive regions, as reflected by a slowdown in the ungrammatical conditions. If, on the other hand, an animacy match to the structurally irrelevant distractor can induce attraction due to mis-retrieval, then we should observe attraction effects both at the gerundive and the reflexive regions, as reflected by facilitated reading times for ungrammatical sentences with an animate distractor relative to ungrammatical sentences with an inanimate distractor.

3.3.4 Procedure

Sentences were presented on a desktop PC with a moving-window, non-cumulative, self-paced reading display using the Linger software package (Rohde, MIT). Sentences were initially masked by dashes, with spaces and punctuation intact. Participants used the space bar to reveal each word. Each sentence was followed by a ‘yes/no’ comprehension question, with onscreen feedback for incorrect responses. The order of presentation was randomized for each participant.

3.3.5 Analysis

Self-paced reading times were examined region-by-region, with the regions consisting of single words. Reading times that exceeded 2500 ms. were excluded from analysis, and raw reading times were log-transformed to reduce non-normality (Hofmeister; Vasishth & Drenhaus, 2011).[13] This trimming method affected less than 1% of the data. Participants with less than 70% accuracy on the comprehension questions were also excluded from analysis. Reading time data was analyzed using linear mixed-effects models. We used orthogonal contrast coding for experimental fixed effects (grammaticality and attraction; ±0.5 for each factor) and their interaction. The fixed effect for grammaticality was coded as +0.5 for ungrammatical conditions and -.05 for grammatical conditions. The fixed effect for attraction was coded as +0.5 for non-attraction and +0.5 for attraction. In addition to these fixed effects, items and participants were crossed as random effects (following Baayen et al., 2008). To determine whether inclusion of random slopes was necessary, we compared a model with a fully-specified random effects structure, which included random intercepts and slopes for all

[13] Retention of all data in the data analysis resulted in a marginal interaction between grammaticality and attraction at the reflexive spill-over region for the adjunct control conditions (t=-1.93), rather than the significant interaction (t<2.00), as reported below.

fixed effects and their interaction by items and by participants to an intercept-only model. Model comparison was carried out over the critical verb regions for both the subject-verb agreement conditions and the adjunct control conditions. A log-likelihood ratio test between the two models revealed that the maximal model did not provide a significantly better fit to the data (subject-verb agreement: $X^2_{(10)}$=4.39, p=0.92; adjunct control: $X^2_{(18)}$=9.98, p=0.93;). Therefore, we adopted the intercept only model, and for consistency, we applied the same model to all regions of interest.

3.3.6 Results

Comprehension question accuracy

Mean comprehension question accuracy was 91.6%. Accuracy across conditions ranged from 85% to 95%, indicating that participants successfully comprehended the sentences. For both dependencies, the lowest accuracy scores were observed in the ungrammatical conditions with a feature-matched distractor.

Subject-verb agreement

The critical verb region showed no significant effects or an interaction. The immediate post-verbal region showed a main effect of attraction ($\hat{\beta}$=0.06, SE=0.02, t=2.22) and a significant interaction between grammaticality and attraction ($\hat{\beta}$=-0.17, SE=0.05, t=-2.96). Mean reading times for the post-verbal region are shown in Figure 3.2. Planned pairwise comparisons revealed that the interaction was driven by a significant difference due attraction in the ungrammatical conditions ($\hat{\beta}$=-0.15, SE=0.04, t=-3.56), as reflected by facilitated reading times for ungrammatical sentences with a plural distractor compared to ungrammatical sentences without a plural distractor. No such difference was observed for the grammatical sentences. The lack of a main effect of grammaticality at the gerundive verb region likely reflects the fact that the ungrammatical condition with a plural distractor was read faster than the grammatical conditions, resulting in a negatively skewed mean for the ungrammatical conditions in the comparison.

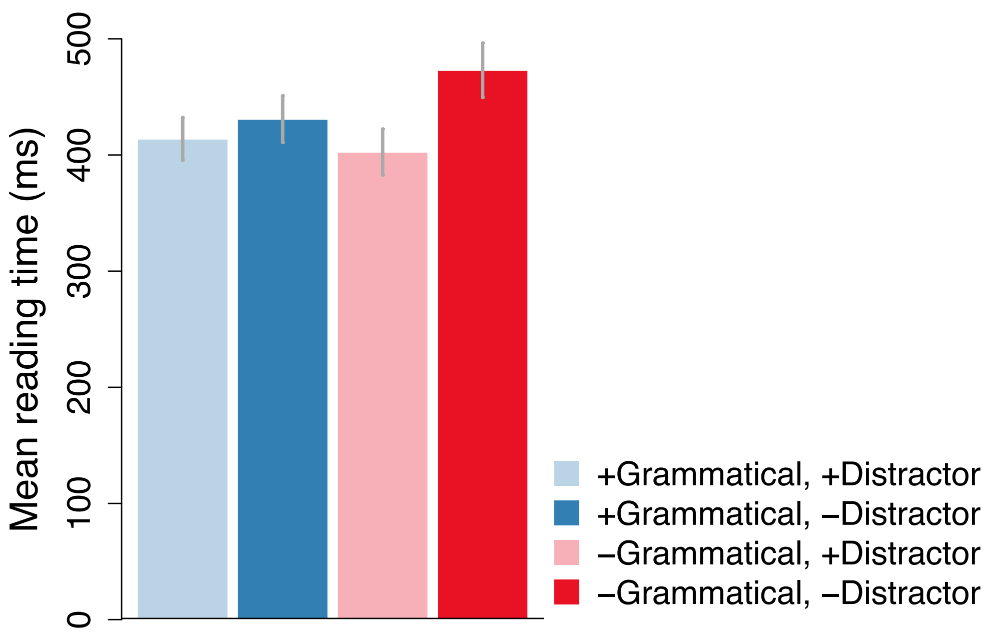


Figure 3.2: Self-paced reading times for Region 9 (first spill-over region following the agreeing verb). Error bars indicate standard error of the mean.

Adjunct control

The control dependency

The pre-critical adjunct subordinator region (e.g., *after*) showed no significant effects or an interaction. The critical adjunct verb region showed a significant interaction between grammaticality and attraction ($\hat{\beta}$=-0.11, SE=0.04, t=-2.48).

Mean reading times for the critical verb region are shown in Figure 3.3. Planned pairwise comparisons revealed that the interaction was driven by a significant difference due attraction in the ungrammatical conditions ($\hat{\beta}$=-0.06, SE=0.03, *t*=-2.03), as reflected by facilitated reading times for ungrammatical sentences with an animate distractor compared to ungrammatical sentences without an animate distractor. No such difference was observed for the grammatical sentences. As in the subject-verb agreement conditions, the lack of a main effect of grammaticality likely reflects the fact that the ungrammatical condition with an animate distractor was read faster than the grammatical conditions, resulting in a negatively skewed mean for the ungrammatical conditions in the comparison.

The reflexive dependency

The critical reflexive region showed no significant effects or an interaction. The immediate post-reflexive region showed a main effect of grammaticality ($\hat{\beta}$=-0.06, SE=0.02, *t*=-3.02) and a significant interaction between grammaticality and attraction ($\hat{\beta}$=-0.10, SE=0.04, *t*=-2.23). Mean reading times for the post-reflexive region are shown in Figure 3.4. Planned pairwise comparisons revealed that the interaction was driven by a significant difference due attraction in the ungrammatical conditions ($\hat{\beta}$=-0.07, SE=0.03, *t*=-2.02), as reflected by facilitated reading times for ungrammatical sentences with an animate distractor compared to ungrammatical sentences without an animate distractor. No such difference was observed for the grammatical conditions.

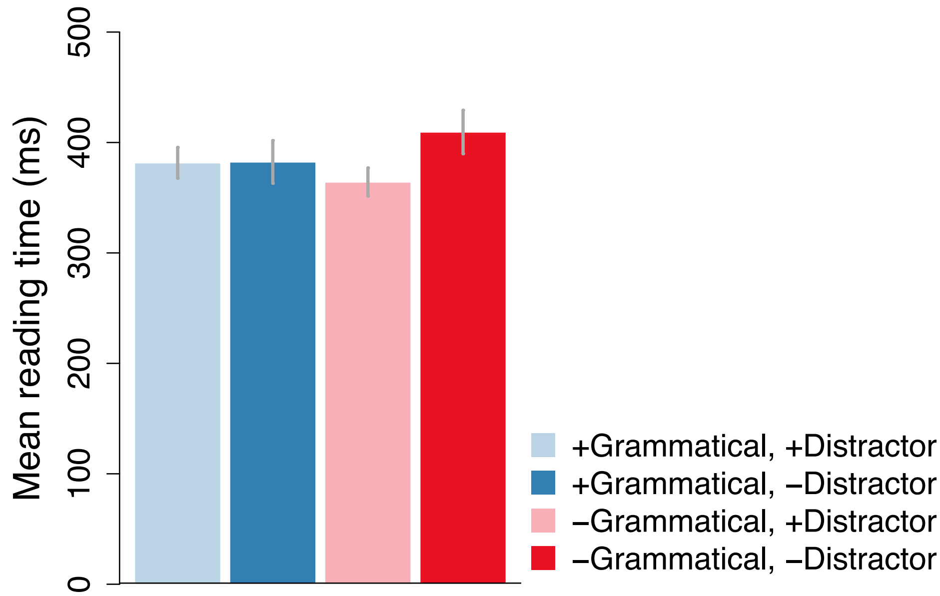


Figure 3.3: Self-paced reading times for Region 11 (gerundive verb). Error bars indicate standard error of the mean.

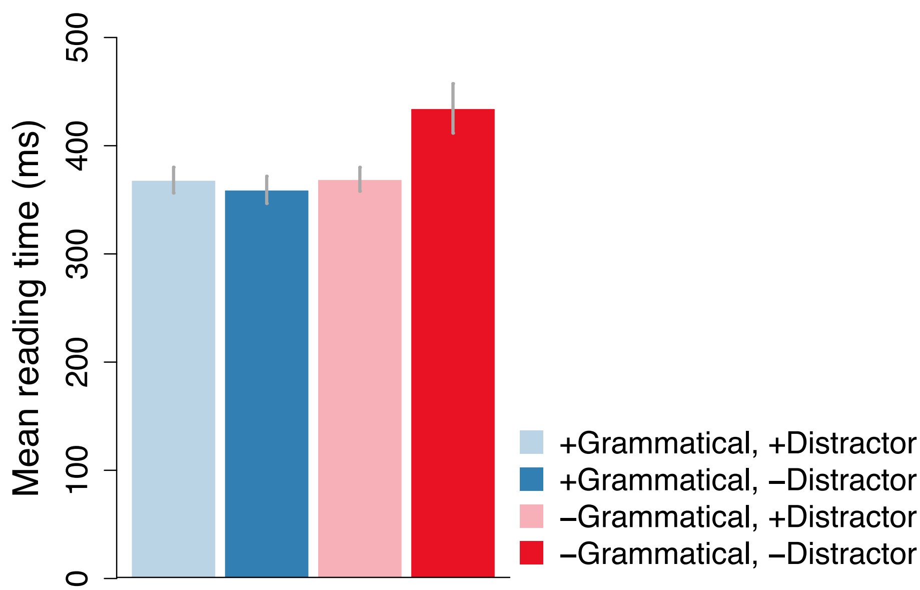


Figure 3.4: Self-paced reading times for Region 16 (first region following the reflexive). Error bars indicate standard error of the mean.

3.3.7 Discussion

Reading times from Experiment 2 showed qualitatively similar attraction profiles for subject-verb agreement and adjunct control sentences. Reading times for subject-verb agreement showed a clear attraction effect reflected by reduced disruptions for ungrammatical verbs in the presence of a plural distractor NP. In the adjunct control conditions, reading times at the adjunct verb showed an attraction effect that was qualitatively similar to that shown for subject-verb agreement, reflected by reduced disruptions for the ungrammatical conditions in the presence of an animate distractor NP. Importantly, this attraction effect was mirrored at the post-reflexive region.

These results suggest that the type of attraction effects commonly observed for subject-verb agreement extend to interpreted anaphoric dependencies. In particular, we take the attraction effects observed in the adjunct control sentences to indicate that the memory access mechanisms were able to retrieve the structurally irrelevant animate distractor as the subject of the adjunct clause, allowing the reflexive to be licensed without detection of the ungrammaticality. That is, mis-retrieval of the structurally irrelevant distractor induce illusions of grammaticality both for the control dependency as well as the reflexive dependency.

Importantly, the attraction effects that we observed for null subject licensing suggests that the contrasting attraction effects seen for subject-verb agreement and anaphora in previous studies (e.g., Dillon et al., 2013) is unlikely to be a consequence of dependency type or the interpretative status of the dependency. That is, immunity to attraction as the result of a structure-guided access strategy is not a property specific to interpreted anaphoric dependencies. Rather, these results suggest that that contrasts in the susceptibility to attraction effects in comprehension is a consequence of either (i) the role of specific features like animacy in memory retrieval, such that certain features give rise to attraction effects for some dependencies, but not others, or (ii) the degree of match between the target of memory access and the retrieval probe, such that when the target is an extremely poor fit, as in the case of an animacy mismatch, comprehenders become more sensitive to the feature match of structurally irrelevant distractors. We return to a fuller discussion of these possibilities and how they capture the observed contrasts in the general discussion.

3.5 Experiment 3

The results from Experiment 2 suggest that the parser engages a noisy, error-prone memory access mechanism for null subject licensing, leading to an illusion of grammaticality from structurally irrelevant but feature compatible material. These results contrast with previous findings of non-attraction effects anaphoric dependencies, which have been taken to suggest that the parser selectively engages a structure-guided memory access procedure for interpreted anaphoric dependencies.

My interpretation of the processing profile observed at the reflexive regions relied on the assumption that the reflexive was a passive reflection of what was retrieved as the subject of the adjunct clause. However, an alternative explanation is that the attraction effect observed at the reflexive regions resulted in mis-retrieval initiated by the reflexive itself. That is, it is possible that the reflexive was not a reflection of what was retrieved as the subject of the adjunct clause, but instead, that upon encountering the reflexive, participants initiated a new retrieval where previous material in both the adjunct and main clauses were probed in parallel. As a result, one might expect sensitivity to structurally irrelevant but feature compatible material in the next higher clause.

Recent findings on the processing oblique reflexives are consistent with this alternative explanation. For example, King, Andrews, and Wagers (2012) found that reflexives in prepositional object position give rise to attraction effects, in contrast with reflexives that are adjacent to verbs. They observed facilitation at the reflexive region from gender matching but structurally-illicit nouns in sentences like (5a) in contrast with verb-adjacent reflexives such as in (5b). They suggested that in the verb-separated configuration, the intervening direct objet NP displaces critical information about the verb's argument structure from the focus of attention (i.e., the current processing state), triggering the use of an error-prone memory retrieval procedure to access the antecedent.

(5) a. The mechanic who spoke to John/Mary sent herself a package…
b. The mechanic who spoke to John/Mary sent a package to herself…

The adjunct control sentences tested in Experiment 2 were superficially similar to those tested by King and colleagues. In our experimental materials, a direct object NP intervened between the verb and the reflexive (e.g., *the urban myth*). It is thus possible that the attraction profile that we observed for the emphatic reflexive in Experiment 2 and the attraction profile that King and colleagues observed for oblique reflexives reflect the same underlying phenomenon.

The goal of Experiment 3 was to assess whether the attraction profile that we observed for emphatic reflexives in Experiment 2 reflects the use of an error-prone memory access strategy. To test this possibility, we used a gender attraction paradigm similar to that used by Sturt (2003), in which the gender of the target and distractor NPs was manipulated, while holding animacy constant across conditions. This manipulation prevents the use of animacy to distinguish between the target and distractor NPs at the point of the critical retrieval for the null subject dependency in the adjunct clause (i.e., the gerundive verb). Specifically, the diagnosticity of the animacy cue will be reduced in memory retrieval for the missing subject, and mis-retrieval is less likely since the structural and semantic constraints on adjunct control will no longer be in competition with each other. In this test, the target should be reliably retrieved as the subject of the adjunct clause because it is a perfect match to the requirements of the gerundive verb. As a result, this manipulation allows us to target the impact of structurally irrelevant material on memory retrieval for the emphatic reflexive by isolating the gender feature-match, which is crucial only to the licensing of the emphatic reflexive dependency.

In this configuration, we did not expect to observe the effect of mis-retrieval at the gerundive site. However, at the reflexive site, there are two possibilities. If the reflexive provides a faithful reflection of what was retrieved at the gerundive, then we should fail to find evidence of attraction at the reflexive region. If, however, the reflexive engages a unconstrained retrieval, then gender-matching attractors should induce attraction at the reflexive region . As in the previous experiment, we used conditions with subject-verb agreement as a baseline measure for attraction effects, with the same predictions for attraction effects.

3.5.1 Participants

Thirty-two members of the University of Maryland community participated in Experiment 1. All participants gave informed consent, and received either course credit or payment of $10 for their participation. The experimental session lasted approximately 45 minutes.

3.5.2 Materials

Experiment 3 used sentences involving adjunct control configurations, and sentences involving subject-verb agreement to provide a baseline measure of attraction effects. Two experimental factors were manipulated for adjunct control and subject-verb agreement: grammaticality and attraction. Within the subject-verb agreement conditions, grammaticality was manipulated by varying the number of the agreeing verb, which was singular in the grammatical conditions and plural in the ungrammatical conditions. Attraction was manipulated by varying the number of the distractor NP. Within the adjunct control conditions, grammaticality was manipulated by varying the gender of the main clause subject. The main clause subject always matched the gender of the reflexive in the grammatical conditions and mismatched the gender in the ungrammatical conditions. Attraction was manipulated by varying the gender of the distractor NP. The main clause subject and distractor NPs were always semantically appropriate as potential subjects of the adjunct verb to prevent semantic or pragmatic biases in subject retrieval.

Forty-eight item sets of items like those in Table 3.2 were constructed. In all conditions, the main clause subject was modified by an object relative clause that contained the distractor NP in subject position. The relative clause verb never overtly expressed agreement to prevent biases in retrieval. A post-verbal adverbial was used in all conditions to signal the end of the relative clause. The main clause was passivized for all conditions for the same reasons described for Experiment 2. In the adjunct control conditions, an emphatic reflexive requiring an animate, gender-matching antecedent as the local subject, which in this case would be the null subject PRO, served as a critical probe to determine whether or not an animate NP was retrieved as the subject of the adjunct clause. Each sentence was followed by a comprehension question. Comprehension questions addressed various parts of the sentence to prevent the possibility that participants might develop superficial reading strategies whereby they extracted only the information necessary to answer the comprehension question without reading the entire sentence. The anomalies in the ungrammatical fillers involved several different types of grammatical errors, including unlicensed NPIs, subject-verb agreement mismatches, and unlicensed verbal morphology, e.g., *-ed* → *-ing*.

Table 3.2: Summary of experimental items for Experiment 3.

Subject-verb agreement conditions

Grammatical, no attraction
1. The harpist that the diva liked very much was congratulated after playing the beautiful song at the brand new recording studio.

Grammatical, attraction
2. The harpist that the divas liked very much were congratulated after playing the beautiful song at the brand new recording studio.

Ungrammatical, attraction
3. The harpist that the divas liked very much were congratulated after playing the beautiful song at the brand new recording studio.

Ungrammatical, no attraction
4. The harpist that the diva liked very much were congratulated after playing the beautiful song at the brand new recording studio.

Adjunct control conditions

Grammatical, no attraction
5. The harpist that the diva liked very much was congratulated after playing the beautiful song herself at the brand new recording studio.

Grammatical, attraction
6. The harpist that the guitarist liked very much was congratulated after playing the beautiful song herself at the brand new recording studio.

Ungrammatical, attraction
7. The drummer that the diva liked very much was congratulated after playing the beautiful song herself at the brand new recording studio.

Ungrammatical, no attraction
8. The drummer that the guitarist liked very much was congratulated after playing the beautiful song herself at the brand new recording studio.

3.5.4 Procedure

The same self-paced reading procedure was used as in Experiment 2.

3.5.5 Analysis

Statistical analysis followed the same steps as in Experiment 2. Data trimming affected less than 1% of the data. Model comparison to determine the appropriate random effects structure for linear mixed effects analysis revealed that a maximally-specified random effects structure did not provide a significantly better fit to the data than an intercept-only model (subject-verb agreement: $X^2_{(18)}$=11.16, *p*=0.88; adjunct control: $X^2_{(18)}$=14.53, *p*=0.69;).

3.5.6 Results

Comprehension question accuracy

Mean comprehension question accuracy for Experiment 3 was 82.3%. Accuracy across experimental conditions ranged from 81% to 86%.

Subject-verb agreement

The verb region showed neither significant effects nor an interaction. The immediate post-verbal region (Region 10) showed a main effect of grammaticality ($\hat{\beta}$=0.18, SE=0.03, *t*=-5.06) and a significant interaction between grammaticality and attraction ($\hat{\beta}$=-0.21, SE=0.07, *t*=-2.98). Mean reading times for the post-verbal region are shown in Figure 3.5. Planned pairwise comparisons revealed that the interaction was driven by attraction in the ungrammatical conditions ($\hat{\beta}$=-0.16, SE=0.05, *t*=-2.97), as reflected by facilitated reading times for ungrammatical sentences with a plural distractor compared to ungrammatical sentences without a plural distractor. No such difference was observed for the grammatical conditions.

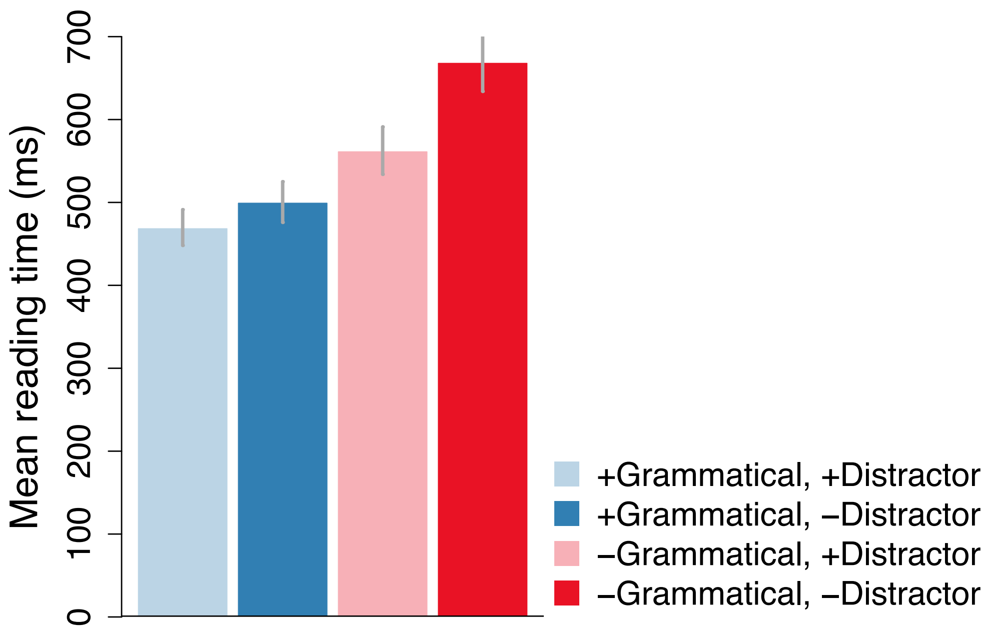


Figure 3.5: Self-paced reading times for Region 9 (first spill-over region following the agreeing verb). Error bars indicate standard error of the mean.

Adjunct control

The control dependency

The pre-critical adjunct subordinator region (e.g., *after*) showed no significant effects or an interaction. In addition, no evidence of attraction was observed at the critical adjunct verb region (Region 11), as the interaction between grammaticality and attraction was not significant. However, a main effect of distractor animacy was observed at this region, driven by a slow down in the grammatical condition with an animate distractor.[14] Mean reading times for the critical verb region are shown in Figure 3.6.

[14] This effect likely reflects a multiple-match inhibitory effect (a 'fan' effect) that is occasionally observed in on-line dependency formation in grammatical sentences (Anderson, 1974; Anderson & Reder, 1999; Chow, Lewis, & Phillips, 2014; Dillon, 2011; R. L. Lewis & Vasishth, 2005). In multiple match contexts, the presence of multiple, maximally similar items, such as two animate

The reflexive dependency

The reflexive region showed no significant effects or an interaction. The immediate post-reflexive region showed only a main effect of grammaticality ($\hat{\beta}$=-0.05, SE=0.02, *t*=-2.09), as reflected by slower reading times for ungrammatical sentences compared to grammatical sentences. No significant interactions were observed at this region or at the subsequent regions. Mean reading times for the post-reflexive region are shown in Figure 3.7.

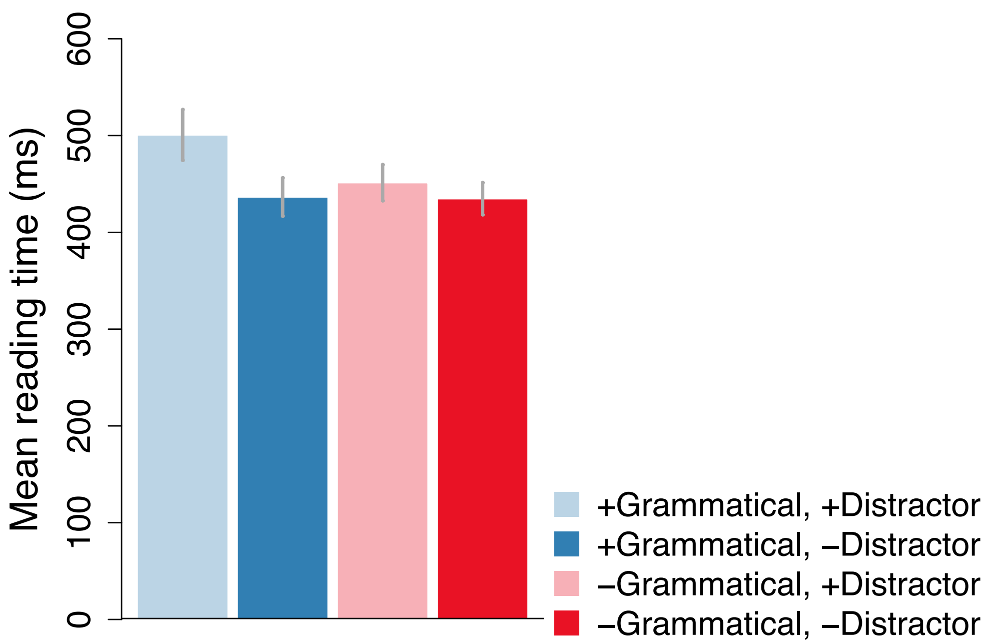


Figure 3.6: Self-paced reading times for Region 9 (first spill-over region following the agreeing verb). Error bars indicate standard error of the mean.

items, can retroactively impact the encoding of the target, making retrieval of the target memory more difficult. Crucially, however, inhibitory effects do not entail mis-retrieval of the structurally illicit distractor (see Dillon, 2011, for further discussion).

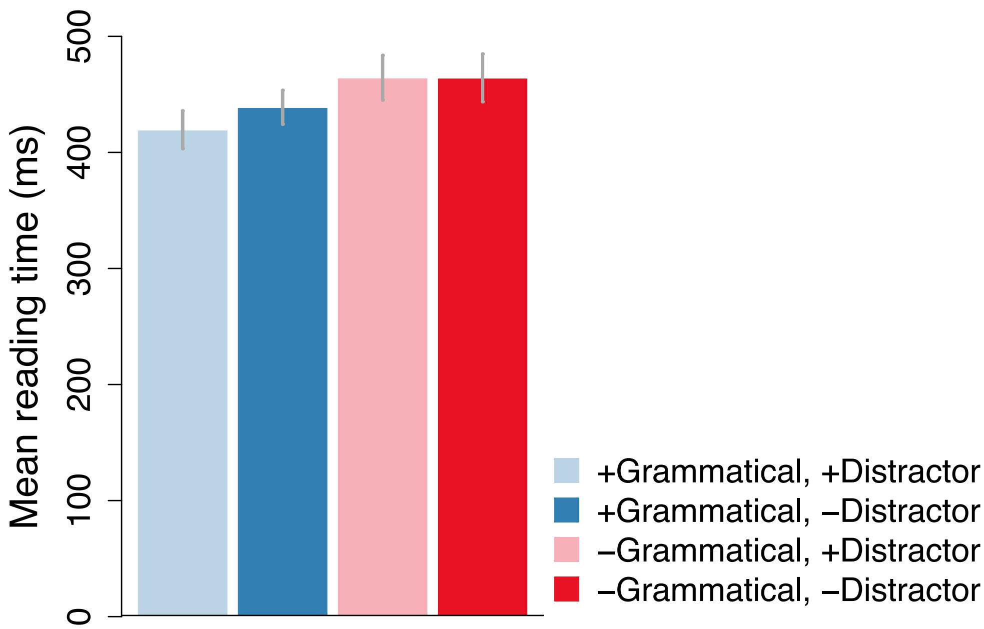


Figure 3.7: Self-paced reading times for Region 16 (first region following the reflexive). Error bars indicate standard error of the mean.

3.5.7 Discussion

The results from Experiment 3 revealed divergent profiles for subject-verb agreement and adjunct control dependencies. For subject-verb agreement, reading times showed reliable sensitivity to the structurally irrelevant but feature compatible distractor NP for subject-verb agreement, but no corresponding effect was observed for adjunct control. By contrast, no attraction effects were observed for reflexives in the adjunct control conditions. The attraction-free processing for adjunct control sentences suggests that the emphatic reflexive was not a source of attraction in Experiment 2. This finding supports the assumption from Experiment 2 that the emphatic reflexive was a passive reflection of the constituent that was retrieved as the subject of the adjunct clause.

The key difference between Experiments 2 and 3 is that the competition between the target and the distractor was eliminated in Experiment 3 for the gerundive verb. Although multiple nouns were present in the main clause, only the target NP satisfied both the structural and animacy requirements of the control dependency. As gender agreement is not a requirement for the control dependency, i.e. gender is not a feature required to resolve the null subject dependency, thus there was no reason to prefer the attractor over the perfectly matched target. The lack of attraction observed in Experiment 3 at the gerundive verb suggest that the target NP, i.e. the subject of the next higher clause, was reliably retrieved as the subject of the adjunct clause, with grammatical fidelity mirrored in the processing of the emphatic reflexive. These findings replicate previous findings that reflexives are not sensitive to the gender match of structurally irrelevant material (e.g., Sturt, 2003).[15]

3.6 General discussion

3.6.1 Summary of results

In the present study, we investigated the processing of null subject licensing in adjunct control structures and subject-verb agreement in English using untimed, off-line acceptability ratings and self-paced reading tasks. The acceptability study

[15] A puzzle that remains, however, is the contrasting profiles for emphatic reflexives and oblique reflexives. King and colleagues (2012) found reliable evidence that oblique reflexives are susceptible to attraction from structurally irrelevant, feature compatible material. Oblique reflexives and emphatic reflexives are superficially similar with respect to their hierarchical relation to the verb that they combine with as an argument. The results from Experiment 3 suggest that access to the contents of the focus of attention, in particular argument structure, is not required in order to ensure grammatically accurate parsing of a reflexive dependency. However, further research is needed to better understand why these two superficially similar reflexive anaphors show contrasting online profiles with respect to attraction effects.

played a crucial role in confirming that comprehenders prefer animate subjects for null subject licensing, as discussed in the formal literature (see Landau, 2001). This preference allowed us to exploit animacy in an interference paradigm to probe for effects of attraction and illusory licensing for null subject dependencies using reading time measures. The empirical question that the on-line, self-paced reading tasks addressed was whether the attraction effects commonly observed for subject-verb agreement extend to null subject licensing.

Null subject licensing showed an on-line attraction profile that was qualitatively similar to the attraction profile observed for subject-verb agreement, as reflected by the eased processing of ungrammatical sentences in the presence of a structurally irrelevant, semantically compatible distractor NP. This profile was observed initially in the processing of the null subject dependency and was latter mirrored in the processing of the emphatic reflexive. Experiment 3 supported the claim that the emphatic reflexive was a passive reflection of the retrieval for the missing subject, replicating previous findings that reflexives are not mislead by the gender match of structurally irrelevant material (e.g., Sturt, 2003).

3.6.2 The status of the interpretative hypothesis

The aim of the present study was to better understand what might be responsible for the contrasting illusory licensing profiles seen for subject-verb agreement and anaphora in previous studies. In particular, we asked which properties of the dependencies might cause the parser to differentially engage the access mechanisms. One possibility was that the contrasting profiles reflected differences in interpretative content between subject-verb agreement and reflexive dependencies. According to this hypothesis, comprehenders engaged structured access for interpreted, anaphoric dependencies, possibly in order to ensure proper interpretation of message intended by the utterance.

The similarity between null subject licensing and subject-verb agreement with respect to attraction effects from structurally irrelevant material suggests that structured access is not deployed for all interpreted, anaphoric dependencies. Null subjects and reflexive anaphors for example both participate in anaphoric dependencies and both contribute to the interpretation of the intended message. However, these two dependencies show qualitatively different access profiles, challenging the hypothesis that all interpreted, anaphoric dependencies engage a structured memory access mechanism. That is, interpretative content is not a sufficient condition to engage structured access.

3.6.3 Variability across dependencies

The findings from the present study challenge the recent claim that memory access mechanisms are tied to specific types of grammatical dependencies (e.g., Alcocer & Phillips, 2012 ms.; Dillon et al., 2013; Phillips et al., 2011), i.e., the grain-size of variability observed for subject-verb agreement and anaphora in previous studies is not dependency-type. Specifically, they suggest that structured-access is not directly linked to anaphoric dependencies.

I suggest, instead, that the current findings motivate a narrower focus on the computation of feature match in memory, specifically probe-to-target similarity. The current findings suggest that animacy features, but not other features like gender or number, induce attraction in interpretative anaphoric dependencies, such as null subject licensing. The next step would be to test whether animacy attraction extends to the processing of canonical reflexive anaphors, such as those previously tested by Sturt (2003) and others (e.g., Clifton et al., 1999; Dillon et al., 2013; Nicol & Swinney, 1989).[16]

If animacy attraction is observed for reflexive anaphors, there are two possibilities for why animacy features, but not gender or number features, would induce attraction for interpretive anaphoric dependencies. The first possibility is that animacy features induce attraction because anaphoric dependencies care more about animacy than other features, i.e., there is some primacy for animacy in retrieval. However, it is currently unclear why animacy might matter more than other more obvious features like gender or number for anaphoric dependencies. One reason that animacy might matter more is that since subjects are typically animate, animacy could be a reliable cue for locating a subject. In addition, recent research on the psychology of memory suggests that animacy information is one of the most important dimensions in controlling memory retention (e.g., Nairne, Van Arsdal., Pandeirada, Cogdill, & LeBreton, 2013). If memory retrieval is a skilled-based procedure, as suggested by Lewis and Vasishth (2005), then it could

[16] I test for effects of animacy attraction in reflexive processing in the next chapter.

be optimized as a function of language use to deploy only the most frequent and reliable cues like animacy to recover an antecedent.

Although it has recently been argued that all cues equally in memory retrieval (e.g., Lewis & Vasishth, 2005), there are psycholinguistic models in which different sources of information are differentially valued. In the Competition Model (MacWhinney, & Bates, 1989; MacWhinney, Bates, & Kliegl, 1984), the most frequent and reliable sources of information, such as word order or agreement, are valued more in interpretation. As suggested above, animacy might be a reliable source of information for memory retrieval, as it can be highly diagnostic of the relationship between a reflexive and its antecedent.

Alternatively, animacy features could induce attraction because an animacy mismatch is in some sense a larger mismatch than just a gender or number mismatch. In English, an animacy mismatch between a reflexive anaphor and its local subject entails a gender mismatch, for example. In an ACT-R style memory architecture (e.g., Lewis & Vasishth, 2005), where each cue contributes directly but independently in retrieval, a target that does not bear the required animacy feature would be doubly penalized due to the corresponding gender mismatch, increasing attraction to other animate NPs in structurally irrelevant positions. That is, the animacy attraction effect could be a consequence of low probe-to-target similarity.

In short, specific content features like animacy may differentially impact the strength of association between the retrieval probe and the target of memory retrieval (probe-to-target similarity), leading to strong attraction effects for anaphors. However, the present results do not decide between the implementations that there is an inherent primacy for animacy in anaphor resolution or the fact that since animacy match entails a (mis-)match to other content features, it contributes more to the probe-to-target similarity. While I tease these possibilities apart in the next chapter, it is worth mentioning that this hypothesis makes strong cross-linguistic predictions, since animacy and gender in English are conflated. In particular, it predicts non-attraction with reflexives in languages with syntactic gender, like Spanish or Polish, for example.

One issue that remains, however, is why subject-verb agreement dependencies fail so easily with respect to illusory licensing as compared to reflexive dependencies, even when an imperfect target provides a relatively strong fit to the retrieval probe. I also address this question in the next chapter.

3.7 Conclusion

In Chapter 3, we investigated the possibility that the contrasting illusory licensing profiles observed for subject-verb agreement anaphora in previous studies is a consequence of differences in interpretative content. We tested this possibility using a new type of dependency involving null subject licensing in adjunct control structures. Null subject licensing provided a good test to better understand what is responsible for the contrasting profiles because it is an interpreted anaphoric dependency that shares certain properties with subject-verb agreement. We found that null subject licensing showed effects of illusory licensing effects qualitatively similar to those observed for subject-verb agreement. These results suggest that the contrasting profiles are unlikely to reflect the interpretative status of subject-verb agreement vs. anaphora. Rather, we took these results to suggest that variability in susceptibility to illusory licensing is a consequence of the role of specific features in memory retrieval, or more generally the degree of match between the target of memory access and the retrieval, such that when the target is a poor fit, comprehenders become more sensitive to structurally irrelevant, but feature matched material in memory.

Chapter 4 Reflexive anaphors

4.1 Introduction

In Chapter 3, we found that memory retrieval for anaphoric dependencies involving null subject licensing can be mislead by structurally irrelevant animate distractors. We hypothesized that this effect could reflect either the role of specific content features in memory retrieval, or a more general property about the computation of feature similarity between the retrieval probe and the target, i.e., the degree of match between the target and retrieval probe or 'probe-to-target similarity'. The goal of Chapter 4 is to distinguish these alternatives using a more direct test of the impact of feature match on anaphoric dependencies involving reflexives in English. Results from three eye-tracking studies replicate the animacy attraction effect reported in Chapter 3, but a qualitatively similar attraction profile is also observed for contexts where the target memory mismatches multiple features beyond animacy. These results imply that attraction for anaphoric dependencies is a consequence of feature similarity between the target and retrieval probe (probe-to-target similarity), rather than the use of specific content features in retrieval. I also provide computational simulations that show that it is possible to derive both the presence and absence of attraction effects for reflexive anaphors from the same memory architecture. I conclude this chapter by proposing an account of why it is harder to obtain attraction effects in anaphor resolution than in some other dependencies, like subject-verb agreement, based on whether retrieval is triggered by error correction or normal resolution.

4.2 The present study

Previous studies of memory retrieval for reflexive licensing have relied on a narrow range of feature manipulations, making it difficult to assess the possibility that attraction effects are a consequence of the use of certain features or probe-to-target similarity. In particular, previous studies have tested contexts involving a single feature manipulation. For example, Sturt (2003) manipulated only gender match between the reflexive and potential antecedents, and Dillon et al. (2013) manipulated only number match. However, reflexive anaphors are subject to a number of other syntactic and semantic constraints beyond morphological feature match, such as animacy restrictions, that have not yet been investigated in studies of attraction.

In the present study, I test a broader range of features in different syntactic configurations to better understand the scope of (non-)attraction effects in sentence comprehension. I also manipulate the degree of feature similarity between the reflexive and potential antecedents to maximize the chances of observing an attraction effect for reflexive dependencies. Results from three eye-tracking studies show that reflexive anaphors are indeed susceptible to attraction effects, but only selectively. When the reflexive mismatches the local subject in just one feature, e.g., gender or number, there are no attraction effects, replicating previous findings (e.g., Sturt, 2003; Dillon et al., 2013). However, when the local subject mismatches in two features, strong attraction effects are found, comparable in size to those observed for subject-verb agreement.

4.3 Experiment 1

The goal of Experiment 1 was to extend the test of animacy attraction that we used for null subject licensing in Chapter 3 to reflexive licensing. As noted above, previous tests of reflexive licensing have been restricted to contexts involving manipulations of only gender or number. However, reflexive licensing is subject to a range of requirements beyond morphological agreement. For example, the relationship between a reflexive anaphor and its antecedent is also subject to an animacy requirement, as illustrated in (1).

(1) The {man | *movie} injured himself.

In Experiment 1, I test the impact of structurally irrelevant animate distractors on antecedent retrieval for reflexive licensing. Given the contrast in (1), taken together with the animacy attraction effects observed for null subject licensing in Chapter 3, animacy information might be used to recover an antecedent in memory, as expected in a cue-based, content-addressable memory architecture that uses every piece of available information to guide retrieval. If a uniform, content-addressable retrieval mechanism mediates reflexive licensing, I expect to find evidence of attraction effects from animate attractions, consistent with the findings for null subject licensing. If, on the other hand, a structure-guided access mechanism mediates reflexive licensing, as has been suggested in previous studies, I expect the narrow use of syntactic cues in memory retrieval to block access to structurally irrelevant animate distractors, just as it blocks access to structurally irrelevant but gender or number matched distractors.

4.3.1 Participants

Thirty members of the University of Maryland community participated in Experiment 1. All participants gave informed consent, and received either course credit or payment of $10 for their participation. The experimental session lasted approximately 45 minutes.

4.3.2 Materials

Thirty-six item sets of the form shown in Table 4.1 were constructed. The subject-verb agreement conditions in (1-4) provided a baseline measure of attraction for the reflexive conditions in (5-6). I initially tested only ungrammatical sentences for the reflexive conditions because attraction usually only appears in ungrammatical configurations where neither the target nor the distractor is a perfect match to the retrieval cues (hence the term 'illusion of grammaticality').[17] Two experimental factors were manipulated: grammaticality and attraction. In the agreement conditions, grammaticality was manipulated by varying the number feature of the agreeing verb. The agreeing verb was always a present tense agreeing form of *be* (*was* in the grammatical conditions and *were* in the ungrammatical conditions). In the reflexive conditions, the reflexive was always plural (*themselves*). For both dependencies, attraction was manipulated by varying

[17] In Experiment 2 I test a full paradigm involving both grammatical and ungrammatical sentences for reflexive dependencies.

the number of the distractor noun. The head noun was always singular and inanimate for both dependencies, and the distractor noun was always animate. Across all conditions, the subject head noun (i.e. the local subject) was held constant. The subject head noun was always modified by an object relative clause that contained the distractor noun. The relative clause verb never expressed agreement to prevent biases in retrieval. In the reflexive conditions, the main verb was always a non-agreeing past tense verb that was immediately followed by a direct object reflexive. The dependent element (the agreeing verb or the reflexive) was always followed by a three to six word spill-over region.

The 36 item sets were mixed with 72 grammatical filler sentences for a total of 108 sentences, yielding a grammatical-to-ungrammatical ratio of 3:1. Half of the target items and half of the filler sentences were followed by a comprehension question. The comprehension questions addressed various parts of the sentence in order to prevent participants from developing superficial reading strategies that would extract the information necessary to answer the question without fully comprehending the sentence.

Table 4.1: Summary of subject-verb agreement and reflexive conditions for Experiment 1. Pre-critical, critical, and spill-over regions included in the analysis are indicated by brackets.

Subject-verb agreement conditions for Experiment 1

Grammatical, no attraction
1. The soothing tea that [the nervous student drank] [was imported] [from India.]

Grammatical, attraction
2. The soothing tea that [the nervous students drank] [was imported] [from India.]

Ungrammatical, attraction
3. The soothing tea that [the nervous students drank] [were imported] [from India.]

Ungrammatical, no attraction
4. The soothing tea that [the nervous student drank] [were imported] [from India.]

Reflexive conditions for Experiment 1

Ungrammatical, attraction
5. The soothing tea that [the nervous students drank calmed] [themselves] [down after the test.]

Ungrammatical, no attraction
6. The soothing tea that [the nervous student drank calmed] [themselves] [down after the test.]

4.3.5 Procedure

The 36 experimental items were distributed into 36 into 6 lists in a Latin Square design. The order of each list was pseudo-randomized for each participant such that no two experimental sentences were presented in succession. All sentences were allowed on a single line on the visual display (142 character limit per line). Sentences were presented using a 12-point fixed-width font (Courier). All characters had a size of 9x16 pixels. The resolution of the LCD screen screen was 1280x720 pixels. Eye-movements were recorded using an Eyelink 1000 tower mount eye-tracker, which sampled eye movements at 1000 Hz. The tower was 32 in. from the visual display, which gave participants approximately 4-5 characters per degree of visual angle. Participants had binocular vision while eye movements were measured, but only the right eye was tracked.

Before the beginning of the experiment, participants were familiarized with the apparatus and given eight practice trials. While seated, participants' heads were immobilized using a chin rest and forehead restraint that was adjusted for each participant. At the start of the experiment and whenever necessary throughout the experiment, the experimenter calibrated the eye-tracker using a 9-point display to ensure accurate measurement of eye-movements across the screen. Participants began each trial by fixating on a square marker at the beginning of the sentence on the display. Once a fixation on the target was recognized by the experimental software, the trial sentence was displayed all at once. Participants indicated completion of reading using a response pad. On trials that had a comprehension question, the question was presented immediately after the test sentence. Participants indicated their response on the response pad. Participants were allowed to take breaks throughout the experimental session at their discretion. After each break, participants were recalibrated. The experimental session lasted approximately 35 minutes.

4.3.6 Data analysis

Fixations of less than 80 ms in duration and within one character of the previous or following fixation were incorporated into the neighboring fixation. All remaining fixations shorter than 50 ms were excluded, since readers do not extract much information from such short fixations (Rayner & Pollatsek, 1989). I report means and statistical analysis from three regions of interest, corresponding to the pre-critical, critical, and spill-over regions, as indicated in Table 4.1. For the subject-verb agreement conditions, the pre-critical region consisted of the material between the relative clause complementizer and the agreeing verb (exclusive), the critical region consisted of the agreeing verb and the following word, and the spill-over region consisted of the remaining words of the sentence. An extended window for the critical region in the agreement conditions was adopted to maximize similarity with the critical region in the reflexive conditions. Analyses for the subject-verb agreement conditions and the reflexive conditions were conducted independently.

I report four measures for each region of interest, which can be divided into 'early' and 'late' measures. The early measures reported here include first-pass reading time and right-bound reading time. First-pass reading is calculated by summing all fixations in a region after eye-gaze first enters the region until the first saccade out of that region (either to the right or left). Right-bound reading time is the sum of all fixations in a region beginning when eye-gaze first enters that region from the left until that region is first excited to the right. Right-bound reading time includes fixations that occur after leftward regressive eye movements from the region, subsuming first-pass reading times. The late measures reported here include regression path duration and re-read (second-pass) time. Regression path duration is the sum of all fixations from the time when eye gaze first enters the region from the left to the time when eye gaze is first exited to the right. Re-read time is the sum of all fixations in a region after that region has been exited (either to the right or left) for the first time.

I used the Box-Cox procedure (Box & Cox, 1964) to determine that a natural log would be the appropriate transformation to obtain normally distributed residuals (average across critical regions of interest: λ=-0.2) (see Vasishth, Chen, Li, & Guo, 2013 for discussion about the importance of appropriately transforming reading time data in psycholinguistic research) . Once this transformation was carried out, reading time data were analyzed using linear mixed-effects models. I estimated models using the *lme4* package (Bates et al., 2011) in the R software environment (R Development Core Team, 2014). Each model included fixed effects for the experimental factors *grammaticality* and *attraction*, and their interaction for the subject-verb agreement conditions. The random effect structure included random intercepts for item and participant, as well as random slopes for the fixed effects and their interaction (Barr et al., 2013)**.** The fixed effects were coded using sum contrasts. Grammaticality was coded as -0.5 for grammatical conditions and 0.5 for ungrammatical conditions. Attraction was coded as -0.5 for attractor conditions and 0.5 for no attractor conditions. Thus, a negative sign on the coefficient indicates a grammatical advantage in the case of grammaticality or an attractor advantage in the case of attraction. All models reported here are the maximal models that have converged (Barr et al. 2013). Analyses were also carried out on data in which missing observations contributed a value of 0 ms to the cell mean. Similar patterns of results obtain when missing observations were omitted from the analysis. A fixed effect or an interaction was considered significant if its *t*-value was greater than 2, indicating that the effect's 95% confidence interval does not include 0 (Gelman & Hill, 2007). Thus, *t*-values with an absolute value greater than 2 are significant with an alpha value at 0.05.

4.3.7 Results

Comprehension question accuracy

The comprehension question accuracy for Experiment 1 was 95%, indicating that participants successfully comprehended the experimental stimuli.

Reading times

Raw mean reading times and the statistical analyses by measure and by region for Experiment 1 are presented in Tables 4.2 and 4.3. Before we discuss the results in detail, I will first provide a brief overview of the main findings. The results from Experiment 1 show that reflexive anaphors are susceptible to attraction effects from structurally irrelevant material, revealing an attraction profile that is qualitatively similar to subject-verb agreement. Specifically, we observe attraction from structurally irrelevant, animate distractors, replicating the animacy attraction effect reported in Chapter 3.

In the subject-verb agreement sentences, reading times for ungrammatical sentences were initially disrupted by agreement errors in early measures (first pass), but this disruption was reduced in later measures (re-read) for ungrammatical sentences with a plural distractor, as compared to ungrammatical sentences with a singular distractor. These findings replicate previous findings for facilitation effects due to the presence of a morphological compatible but structurally irrelevant distractor in contexts involving illicit subject-verb agreement (e.g., Dillon et al., 2013; Wagers et al., 2009). Sentences with a reflexive anaphor showed a similar facilitatory effect beginning in early measures (first pass) extending through to later measures (re-read), as the reading time disruption for an illicit reflexive was reduced due to the presence of a structurally irrelevant but animate, number matching distractor.

Table 4.2: Table of means for subject-verb agreement conditions and reflexive conditions by measure and by region for Experiment 1. Standard error by participants is shown in parentheses.

		Region		
		Pre-critical	**Critical**	**Spill-over**
	First pass reading time			
Agreement	Grammatical, attractor	1119 (41)	333 (14)	647 (35)
	Grammatical, no attractor	1098 (42)	338 (14)	669 (31)
	Ungrammatical, attractor	1111 (40)	394 (20)	662 (35)
	Ungrammatical, no attractor	1072 (46)	415 (18)	671 (37)
Reflexives	Ungrammatical, attractor	1470 (59)	228 (8)	765 (35)
	Ungrammatical, no attractor	1454 (58)	275 (10)	71 (33)
	Right-bound reading time			
Agreement	Grammatical, attractor	1268 (44)	391 (16)	950 (44)
	Grammatical, no attractor	1217 (45)	394 (16)	965 (43)
	Ungrammatical, attractor	1279 (41)	489 (25)	1041 (52)
	Ungrammatical, no attractor	1207 (43)	511 (23)	1057 (53)
Reflexives	Ungrammatical, attractor	1747 (77)	268 (11)	1085 (44)
	Ungrammatical, no attractor	1735 (71)	310 (12)	1179 (48)
	Regression path duration			
Agreement	Grammatical, attractor	1333 (57)	469 (39)	1962 (125)
	Grammatical, no attractor	1294 (56)	424 (22)	1928 (151)
	Ungrammatical, attractor	1332 (47)	602 (49)	2279 (150)
	Ungrammatical, no attractor	1273 (47)	735 (57)	2396 (164)
Reflexives	Ungrammatical, attractor	1869 (106)	349 (32)	2320 (149)
	Ungrammatical, no attractor	1859 (97)	400 (35)	2796 (162)
	Re-read time			
Agreement	Grammatical, attractor	675 (74)	265 (25)	296 (35)
	Grammatical, no attractor	588 (68)	220 (26)	285 (33)
	Ungrammatical, attractor	792 (73)	340 (35)	360 (45)
	Ungrammatical, no attractor	868 (81)	409 (35)	379 (43)
Reflexives	Ungrammatical, attractor	1011 (109)	147 (19)	313 (39)
	Ungrammatical, no attractor	1205 (112)	273 (22)	458 (44)

Table 4.3: Summary of the statistical analysis by measure and by region for Experiment 1. All *t*-values with an absolute value greater than 2 are significant with an alpha value at 0.05, indicated in bold.

	REGIONS								
	PRE-CRITICAL			CRITICAL			SPILL-OVER		
	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t
First pass reading time									
grammaticality	-0.06	0.04	-1.29	**0.29**	**0.09**	**3.09**	-0.02	0.06	-0.39
agreement attraction	0.03	0.06	0.57	0.10	0.10	-1.04	-0.05	0.05	-0.88
grammaticality×attraction	0.02	0.11	0.20	-0.16	0.17	-0.95	-0.08	0.15	-0.55
animacy attraction	0.02	0.08	0.33	**-0.29**	**0.14**	**-2.03**	0.05	0.08	0.68
Right-bound reading time									
grammaticality	-0.01	0.03	-0.34	**0.37**	**0.09**	**3.94**	0.04	0.06	0.64
agreement attraction	0.04	0.03	1.11	-0.12	0.11	-1.04	-0.05	0.05	-1.02
grammaticality×attraction	-0.00	0.08	-0.11	-0.19	0.17	-1.12	-0.14	0.12	-1.17
animacy attraction	-0.00	0.03	-0.02	**-0.31**	**0.14**	**-2.11**	-0.08	0.08	-1.00
Regression path duration									
grammaticality	-0.00	0.03	-0.15	**0.43**	**0.09**	**4.42**	0.13	0.08	1.48
agreement attraction	0.02	0.03	0.66	-0.14	0.11	-1.22	-0.01	0.06	-0.28
grammaticality×attraction	-0.01	0.08	-0.23	-0.30	0.18	-1.62	-0.20	0.15	-1.31
animacy attraction	0.00	0.03	0.08	-0.28	0.15	-1.86	-0.19	0.11	-1.62
Re-read time									
grammaticality	**0.74**	**0.22**	**3.38**	**0.57**	**0.20**	**2.78**	0.25	0.21	1.18
agreement attraction	-0.03	0.24	-0.15	-0.11	0.23	-0.49	-0.14	0.22	-0.65
grammaticality×attraction	-0.64	0.42	-1.51	**-1.15**	**0.43**	**-2.65**	-0.23	0.44	-0.53
animacy attraction	-0.50	0.41	-1.20	**-1.47**	**0.34**	**-4.23**	**-1.22**	**0.44**	**-2.71**

Below, I divide the detailed discussion of the results into two sections for subject-verb agreement and reflexives. In each section, I discuss first the results for early processing (first-pass reading time and right-bound reading time) and then discuss the results for late processing (regression path duration and re-read time).

Subject-verb agreement

Early processing (first-pass reading time and right bound reading time)

No significant effects or interactions were observed in the pre-critical region for first-pass or right-bound reading times, providing a appropriate baseline for interpreting the results at the critical verb region. At the critical verb region, there was a main effect of grammaticality in first-pass reading times, reflected by increased reading times for ungrammatical sentences relative to grammatical sentences. A similar main effect of grammaticality was observed at the reflexive region in right-bound reading times. No other effects were observed at the critical verb region. No significant effects or interactions were observed in the spill-over region in first-pass reading times or right-bound reading times.

Late processing (regression path duration and re-read time)

Regression path times at the critical verb showed a main effect of grammaticality, reflected by increased reading times for ungrammatical sentences relative to grammatical sentences. Re-read times at this region showed a similar main effect of grammaticality as well as an interaction between grammaticality and attraction. Planned pairwise comparisons revealed that this interaction was driven by a significant effect of attraction in the ungrammatical conditions, reflected by facilitated reading times for ungrammatical sentences with an attractor relative to ungrammatical sentences without an attractor. No such effect was observed in the grammatical conditions. A main effect of grammaticality was also observed in re-read times at the pre-critical regions, indicating that the processing disruption associated with the illicit verb impacted re-reading both of the critical verb and the preceding material. No other effects were observed in late measures.

Reflexive anaphors

Early processing (first-pass reading time and right bound reading time)

No significant effects were observed in the pre-critical region for first-pass reading times or right-bound reading times, providing an appropriate baseline for interpreting the results at the reflexive region. At the reflexive region, there was a main effect of attraction both in first-pass and right-bound measures, reflected by facilitated reading times for ungrammatical sentences with an attractor relative to ungrammatical sentences without an attractor. No other effects were observed in early measures.

Late processing (regression path duration and re-read time)

Both regression path and re-read times showed a main effect of attraction at the reflexive region, reflected by facilitated reading times for ungrammatical sentences with an attractor relative to ungrammatical sentences without an attractor. A similar effect was observed at the spill-over region in re-read times

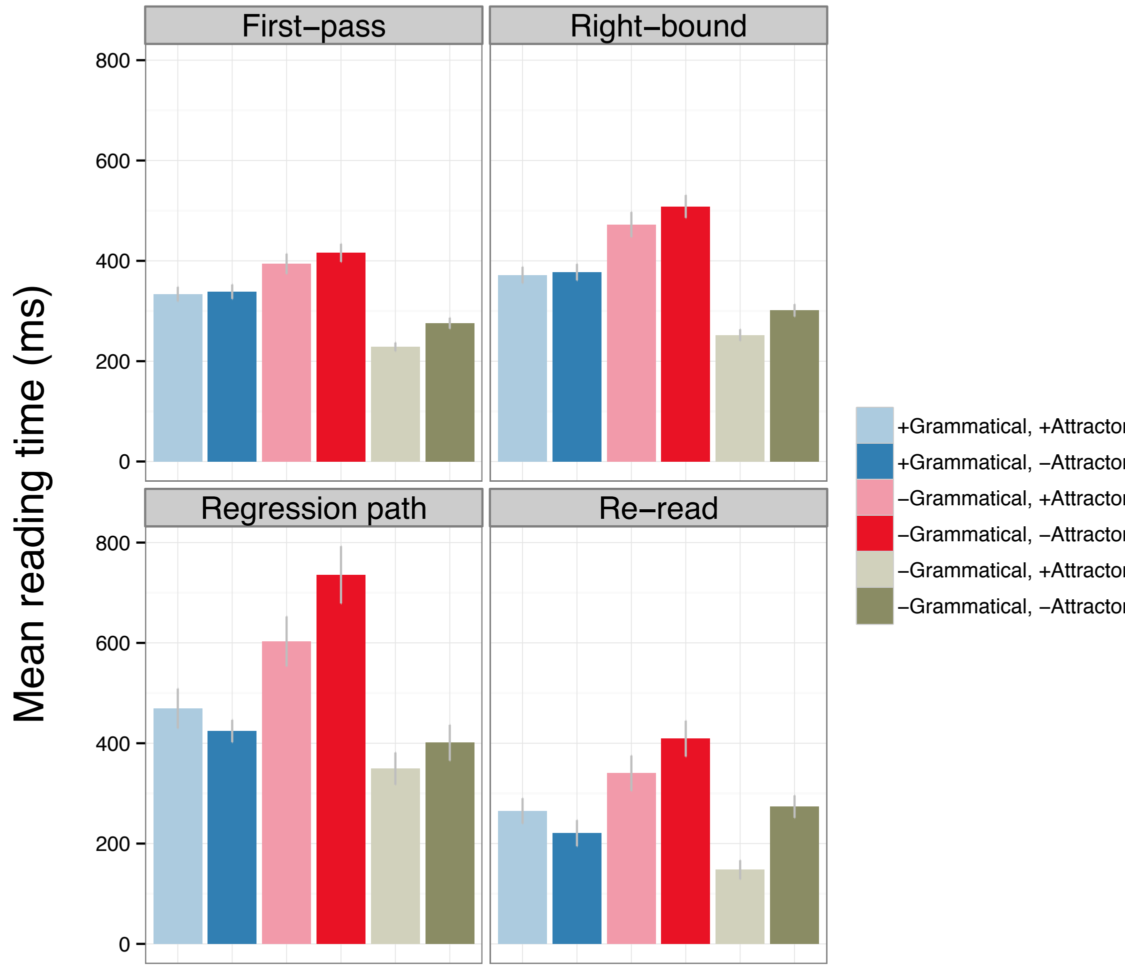


Figure 4.1: Mean reading times by measure by condition Experiment 1). Error bars indicate standard error by participant.

4.3.7 Discussion

The results of Experiment 1 demonstrate that the animacy attraction effects observed for null subject licensing in Chapter 3 extend to reflexive anaphors, which show a qualitatively similar attraction profile in the presence of a structurally irrelevant but animate, feature compatible distractor. However, these findings are inconsistent with previous studies of reflexive anaphors, which have consistently failed to find reliable evidence of facilitatory attraction effects in online processing (Clifton et al., 1999; Cunnings & Sturt, 2012; Dillon, 2011; Dillon et al., submitted; Nicol, 1988; Nicol & Swinney, 1989; Xiang et al., 2009). In our study, we observed clear effects of facilitatory attraction effects across multiple measures at the reflexive region, starting in early measures. In particular, the finding of attraction effects in first pass reading times suggests that the structurally irrelevant, animacy matched antecedent was considered in the earliest stages of processing. These results imply that syntactic structure may not provide a hard constraint on the items that are considered for participation in a reflexive dependency, as previously assumed.

Instead, sensitivity to structurally irrelevant but partially-matched material is expected in a direct-access, content-addressable memory architecture, in which all available information is used to probe the contents of memory in parallel, leading to mis-retrieval of partially activated material regardless of grammaticality. More generally, the findings from Experiment 1 suggest that animacy is a feature that is actively used to guide memory retrieval for anaphoric dependencies, as qualitatively similar animacy attraction effects have been observed for both reflexive and null subject licensing. These findings challenge accounts that assume that different dependencies are directly linked to specific retrieval mechanisms (e.g., Alcocer & Phillips, 2012 ms.; Dillon et al., 2013; Phillips et al., 2011). Instead, these findings are compatible with a feature-based account of the contrast between reflexive anaphors and subject verb agreement. Under a feature-based account, susceptibility to attraction effects could be directly linked to specific features, like animacy, or it could reflect differences in the degree of feature similarity between the target memory and the retrieval probe. I distinguish these alternatives in Experiment 3.

One interesting point about Experiment 1 is that subject-verb agreement and reflexives show different time profiles with respect to attraction effects. In particular attraction was observed in both early and late measures for reflexives, but only as a late effect for subject-verb agreement. The timing profile for subject-verb agreement is consistent with the hypothesis that memory retrieval for subject-verb agreement is an error-driven process, as proposed by Wagers et al., (2009). According to this hypothesis, a noisy, error-prone memory retrieval is engaged when the expected features of the verb do not align with the bottom up input. This hypothesis explains why agreement attraction occurs only in ungrammatical contexts, where the verb does not bear the expected number features. In our experiment, the timing profile for subject-verb agreement could reflect two effects, such as an initial error detection, which would give rise to the main effect of grammaticality with an absence of attraction observed in early measures, followed by a error-driven, noisy memory retrieval, which would give rise to the attraction effects observed in later measures. I return to a fuller discussion of this possibility and its implications in the general discussion.

4.4 Experiment 2

Experiment 2 was designed to test an alternative explanation for the animacy attraction effect observed for reflexive anaphors. It is possible, for example, that animacy attraction for reflexive anaphors reflects an effect of proximity concord (e.g., Quirk, Greenbaum, Leech, & Svartvik, 1985) . Specifically, in an attraction configuration, the reflexive could have been incorrectly licensed by the structurally irrelevant, but linearly proximate antecedent. Although attraction effects have been shown to be robust against effects of linear proximity for subject-verb agreement (see Phillips et al, 2011, for discussion), the same may not be true for reflexive anaphors. In particular, the reflexive attraction effect (including the animacy attraction effect reported for null subject licensing in Chapter 3) may be consistent with models of sentence comprehension in which online comprehension processes are impacted by local surface relations (e.g., Tabor et al., 2004).

The first goal of Experiment 2 was to determine whether the animacy attraction effect observed for reflexive anaphors extends beyond configurations with a linearly proximate attractor. The second goal of Experiment 2 was to directly compare susceptibility to attraction effects within reflexive dependencies to further test the hypothesis proposed in Chapter 3 that attraction effects are not directly linked to specific dependencies, but rather specific features. To this end, I compared reflexive dependencies involving an animacy manipulation to reflexive dependencies involving a gender manipulation. Importantly, reflexive dependencies involving a gender manipulation have been shown in previous studies to strongly resist attraction effects (e.g., Sturt, 2003), providing an appropriate baseline for comparison. Evidence of contrasting attraction profiles for these two feature manipulations within a single reflexive dependency would challenge the hypotheses that attraction effects are directly linked to specific dependencies or that a structured access mechanism is directly linked to anaphoric dependencies. Rather, such evidence would imply that attraction effects are directly linked to specific features in memory retrieval.

4.4.1 Participants

Thirty members of the University of Maryland community participated in Experiment 1. All participants gave informed consent, and received either course credit or payment of $10 for their participation. The experimental session lasted approximately 45 minutes.

4.4.2 Materials

Thirty-six item sets of the form shown in Table 4.4 were constructed to directly compare the impact of the distractor's gender and animacy features on the processing of reflexive anaphors. Two experimental factors were manipulated: grammaticality and attraction. Grammaticality was manipulated by varying either the (stereotypical) gender or the animacy of the local subject noun. Likewise, attraction was manipulated by varying the (stereotypical gender) or the animacy of the distractor noun. The stereotyped nouns were selected partly from previous studies using stereotyped nouns (e.g., Sturt, 2003; Dillon et al., 2013) and on the basis of intuition. Across all conditions, the distractor was the subject of the main clause and the target was always the subject a subordinate clause that contained the reflexive. This configuration was used in order to prevent effects of proximity concord. In order to prevent any biases in retrieval, neither the main clause verb nor the embedded clause verb expressed agreement, and both the target and the distractor nouns were semantically appropriate antecedents for the reflexive.

Table 4.4: Summary of reflexive gender and animacy manipulation for Experiment 2. Pre-critical, critical, and spill-over regions included in the analysis are indicated by brackets.

Reflexives gender manipulation for Experiment 2

<u>Grammatical, no attraction</u>
1. The strict librarian said that [the studious schoolgirl reminded] [herself] [about the] overdue book.

<u>Grammatical, attraction</u>
2. The strict father said that [the studious schoolgirl reminded] [herself] [about the] overdue book.

<u>Ungrammatical, attraction</u>
3. The strict librarian said that [the studious schoolboy reminded] [herself] [about the] overdue book.

<u>Ungrammatical, no attraction</u>
4. The strict father said that [the studious schoolboy reminded] [herself] [about the] overdue book.

Reflexives animacy manipulation for Experiment 2

<u>Ungrammatical, attraction</u>
5. The strict librarian said that [the brief memo reminded] [herself] [about the] overdue book.

<u>Ungrammatical, no attraction</u>
6. The strict father said that [the brief memo reminded] [herself] [about the] overdue book.

4.4.5 Procedure and data analysis

The same procedure was used as in Experiment 1 and the analysis followed the same steps, with changes made to the fixed effects to reflect the experimental manipulation. Each model included a fully specified random effects structure. The box-cox procedure determined that a natural log was the appropriate transformation to obtain normally distributed residuals (average across critical region of interest: λ=-0.3).

4.4.6 Results

Comprehension question accuracy

The comprehension question accuracy for Experiment 1 was 87%, indicating that participants successfully comprehended the experimental stimuli.

Reading times

Raw mean reading times and statistical analyses by measure and by region for Experiment 2 are presented in Tables 4.5 and 4.6. Experiment 2 replicated the animacy attraction effect for reflexives in a non-local attraction configuration, suggesting that animacy attraction for reflexives is not a consequence of proximity concord. Specifically, results showed that reflexives are indeed susceptible to attraction, but only selectively: when the reflexive mismatched the target memory (the local subject) in gender, there were no attraction effects from structurally irrelevant material, but when the target mismatched in animacy, there were strong attraction effects, qualitatively similar to the subject-verb agreement attraction effects observed in Experiment 1. In the gender manipulation conditions, only the gender (mis-)match of the structurally appropriate antecedent impacted reading times. Ungrammatical conditions where the target did not match in gender with the reflexive showed longer reading times, but crucially, there were no significant effects of attraction or an interaction in the gender manipulation conditions at any region or in any measure, replicating previous findings for reflexives (e.g., Sturt, 2003; Dillon et al., 2013). However, a different profile is observed for the conditions involving an animacy manipulation. Results for the animacy manipulation conditions showed a significant effect of attraction from structurally irrelevant animacy matched antecedents. Importantly, the results of Experiment 2 suggest that animacy attraction for reflexive dependencies is not a consequence of proximity concord, as non-local attraction effects are observed in configurations where the distractor appears before the target.

Table 4.5: Table of means for reflexive gender and animacy manipulations, by measure and by region for Experiment 2. Standard error by participants is shown in parentheses.

		Region		
		Pre-critical	**Critical**	**Spill-over**
	First pass reading time			
Agreement	Grammatical, attractor	1052 (47)	215 (12)	227 (17)
	Grammatical, no attractor	1020 (48)	205 (10)	235 (16)
	Ungrammatical, attractor	924 (39)	229 (13)	261 (20)
	Ungrammatical, no attractor	960 (37)	243 (14)	274 (23)
Reflexives	Ungrammatical, attractor	873 (37)	216 (11)	202 (15)
	Ungrammatical, no attractor	934 (48)	226 (12)	221 (19)
	Right-bound reading time			
Agreement	Grammatical, attractor	1207 (46)	225 (13)	249 (21)
	Grammatical, no attractor	1180 (47)	218 (11)	262 (19)
	Ungrammatical, attractor	1132 (37)	256 (15)	327 (28)
	Ungrammatical, no attractor	1088 (35)	276 (17)	337 (30)
Reflexives	Ungrammatical, attractor	1020 (38)	228 (12)	242 (18)
	Ungrammatical, no attractor	1091 (44)	258 (16)	284 (26)
	Regression path duration			
Agreement	Grammatical, attractor	1265 (50)	274 (21)	314 (38)
	Grammatical, no attractor	1250 (53)	260 (17)	403 (55)
	Ungrammatical, attractor	1211 (43)	318 (27)	486 (59)
	Ungrammatical, no attractor	1139 (39)	373 (37)	539 (65)
Reflexives	Ungrammatical, attractor	1081 (42)	273 (18)	329 (34)
	Ungrammatical, no attractor	1189 (51)	365 (44)	453 (60)
	Re-read time			
Agreement	Grammatical, attractor	761 (78)	193 (22)	223 (26)
	Grammatical, no attractor	750 (66)	196 (21)	254 (25)
	Ungrammatical, attractor	1027 (93)	338 (30)	368 (33)
	Ungrammatical, no attractor	1005 (80)	314 (27)	336 (34)
Reflexives	Ungrammatical, attractor	561 (54)	169 (18)	248 (24)
	Ungrammatical, no attractor	749 (65)	350 (34)	352 (38)

Table 4.6: Summary of the statistical analysis by measure and by region for Experiment 2. All *t*-values with an absolute value greater than 2 are significant with an alpha value at 0.05, indicated in bold.

	REGIONS								
	PRE-CRITICAL			CRITICAL			SPILL-OVER		
	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t
First pass reading time									
grammaticality	-0.02	0.07	-0.35	0.21	0.15	1.32	0.14	0.18	0.78
agreement attraction	-0.03	0.07	-0.45	-0.02	0.15	-0.14	-0.03	0.16	-0.20
grammaticality×attraction	-0.03	0.14	-0.28	-0.10	0.31	-0.32	0.20	0.35	0.56
animacy attraction	0.03	0.10	0.34	-0.04	0.20	-0.24	0.15	0.29	0.52
Right-bound reading time									
grammaticality	-0.00	0.06	-0.13	0.25	0.15	1.60	0.21	0.18	1.20
agreement attraction	0.00	0.05	0.13	-0.03	0.15	-0.21	-0.04	0.17	-0.26
grammaticality×attraction	0.09	0.11	0.82	-0.09	0.30	-0.29	0.22	0.36	0.60
animacy attraction	0.00	0.07	0.06	-0.08	0.20	-0.42	0.11	0.29	0.39
Regression path duration									
grammaticality	-0.00	0.06	-0.12	0.27	0.16	1.67	0.29	0.19	1.53
agreement attraction	0.01	0.06	0.28	-0.05	0.16	-0.33	-0.06	0.17	-0.36
grammaticality×attraction	0.11	0.12	0.90	-0.13	0.31	-0.42	0.24	0.37	0.65
animacy attraction	-0.01	0.08	-0.16	-0.11	0.21	-0.54	0.09	0.30	0.29
Re-read time									
grammaticality	**0.65**	**0.19**	**3.37**	**1.03**	**0.25**	**4.00**	**0.61**	**0.23**	**2.57**
agreement attraction	-0.23	0.39	-1.09	-0.02	0.21	-0.13	0.12	0.24	0.52
grammaticality×attraction	0.11	0.39	0.27	-0.02	0.41	-0.05	0.35	0.40	0.87
animacy attraction	-0.68	0.39	-1.74	**-0.92**	**0.28**	**-3.21**	0.03	0.33	0.11

Reflexives: Gender manipulation

Early processing (first-pass reading time and right bound reading time)

No significant effects or interactions were observed in first-pass reading times or right-bound reading times at the pre-critical, critical or spill-over regions.

Late processing (regression path duration and re-read time)

A main effect of grammaticality was observed in re-read times at the critical reflexive region for the gender manipulation, reflected by a slow down in the ungrammatical conditions relative to the grammatical conditions. A similar effect was observed in re-read times at the spill-over region, as well as at the pre-critical region, indicating that the processing disruption impacted re-reading of the preceding region. Importantly for the gender manipulation, we failed to find significant effects of attraction or an interaction of grammaticality with attraction at any region or measure, replicating previous findings for reflexives (e.g., Sturt, 2003).

Reflexives: Animacy manipulation

Early processing (first-pass reading time and right bound reading time)

No significant effects or interactions were observed in first-pass reading times or right-bound reading times at the pre-critical, critical or spill-over regions.

Late processing (regression path duration and re-read time)

A main effect of attraction was observed in re-read times at the critical reflexive region, reflected by facilitated reading times for ungrammatical sentences with an attractor relative to ungrammatical sentences without an attractor. A similar effect was observed in the pre-critical region, indicating that the attraction effect impacted re-reading of the preceding region.

Direct comparison of attraction effects

The results from Experiment 2 showed a reliable effect of attraction for the reflexive animacy manipulation, and no corresponding effect for the reflexive gender manipulation. Because the contrast between previous tests of reflexives using a gender manipulation and the current manipulation of animacy is of primary theoretical interest, I directly tested the attraction effects between the gender manipulation and the animacy manipulation. Importantly, contrasting attraction profiles were observed for these manipulations, as reflected by an interaction between attraction and feature-type (*gender* vs. *animacy*) ($\hat{\beta}$=0.89, SE=0.44; t=2.02) in re-read times at the critical reflexive region. Pairwise comparisons revealed that this interaction was driven by the significant attraction effect for the animacy manipulation reported above.

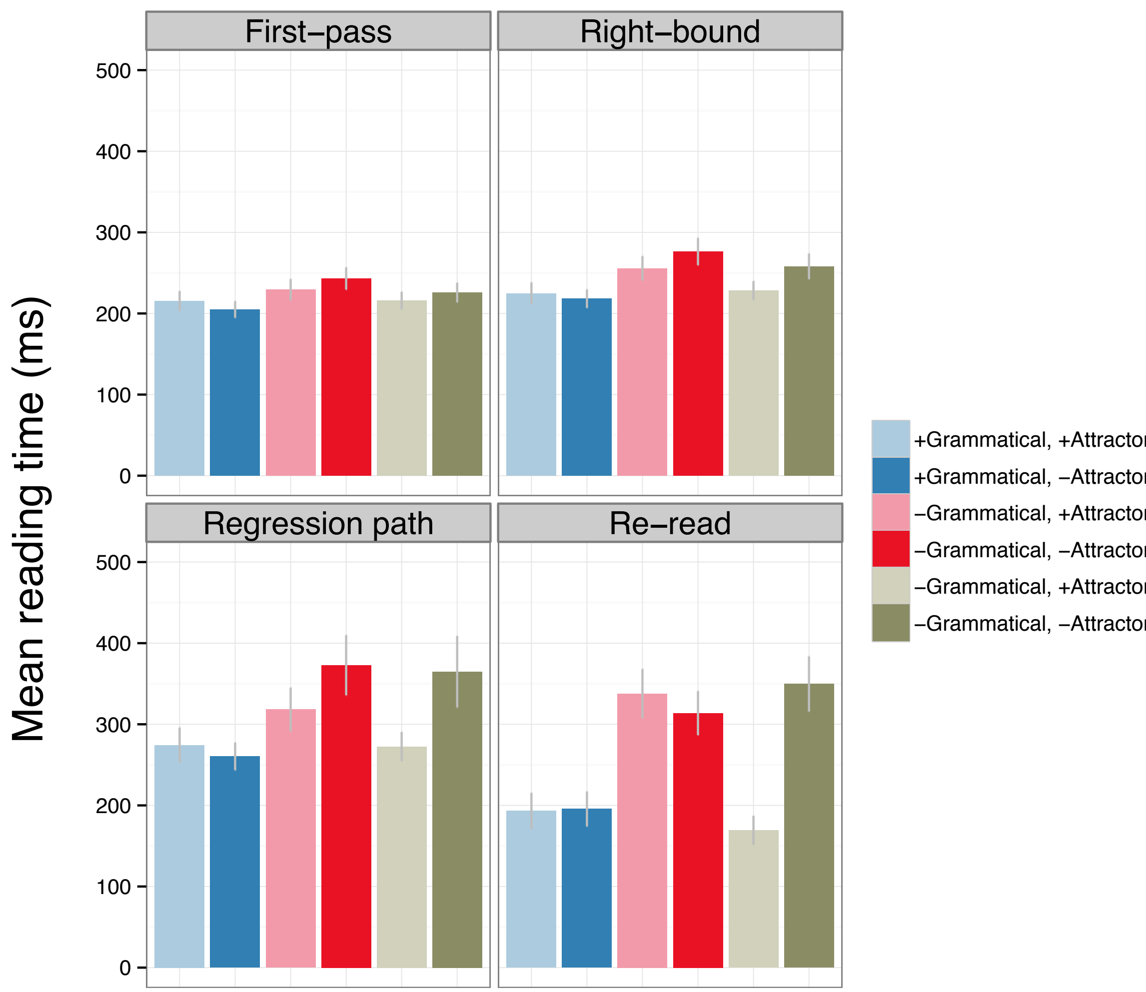


Figure 4.2: Mean reading times by measure by condition (Experiment 2). Error bars indicate standard error by participant.

4.4.7 Discussion

The results of Experiment 2 provide further evidence that reflexive anaphors are indeed susceptible to attraction from structurally irrelevant material, as they replicated the animacy attraction effect in a non-local attraction configuration. However, perhaps the most interesting finding from Experiment 2 was the contrasting attraction profiles observed within a single dependency. In particular, the attraction-free profile observed for the gender manipulation replicates previous findings for reflexive processing (e.g., Sturt, 2003; Dillon et al., 2013) and suggests the use of a structure-guided memory access mechanism in which structural information has a privileged status in antecedent retrieval, blocking access to feature compatible material in structurally irrelevant positions. However, the attraction profile observed for the animacy manipulation is inconsistent with previous findings for reflexive processing and suggests that structural priority is much more limited.

These findings inform our understanding about the source and scope of attraction effects in comprehension in several ways. First, the attraction effects observed for the animacy manipulation suggest that the parser is in fact able to engage a direct-access, content-addressable retrieval mechanism for reflexive dependencies. In particular, these results suggest that retrieval for reflexive processing can utilize both structural and non-structural cues. Second, these results from Experiments 1 and 2 (taken together with the results from Chapter 3) suggest that attraction effects are not directly linked to specific dependencies, and that there is not a direct link between the type of grammatical dependency and retrieval mechanisms or retrieval strategies.

As suggested earlier in Chapter 3, there are two possibilities for why animacy, but not gender or number, induces attraction for anaphoric dependencies. The first possibility is that there is an inherent primacy for animacy in retrieval. For example, animacy could be a particularly reliable cue to the subject, since subjects are typically animate. As such, the parser could prioritize animacy in memory retrieval for an antecedent. The second possibility is that an animacy match contributes more to the probe-to-target similarity, since it entails a gender match in English. For example, a retrieval target that does not carry the required animacy feature, e.g., [+animate], could be doubly penalized because it also does not bear the required gender feature, increasing sensitivity to other animate NPs in structurally irrelevant positions. I distinguish these alternatives in Experiment 3.

4.5 Experiment 3a

Experiments 3a-b were designed to distinguish the possibilities that attraction effects within reflexive dependencies are a consequence of specific features in memory retrieval on the one hand, or probe-to-target similarity on the other hand. To achieve this, I manipulated the degree of probe-to-target similarity, comparing contexts that involved a single feature manipulation, e.g., gender, as done in previous studies of reflexives to contexts that involved multiple feature manipulations, e.g., gender + number. Crucially, both gender and number manipulations alone have been independently shown to resist attraction effects. However, if the animacy attraction effects observed for reflexives are a consequence of decreased probe-to-target similarity, then we might expect to find effects of attraction in a combined gender and number manipulation, i.e. where the target mismatches the reflexive in both gender and number, due to the decreased probe-to-item similarity.

4.5.1 Computational model of associative retrieval

To better understand how probe-to-target similarity might impact attraction effects in memory retrieval, I first conducted simulations using an explicit computational model of memory retrieval. I used a computational model that is based on an implementation in the R software environment of the core ACT-R (Adaptive Character of Thought – Rational) equations described in Lewis and Vasishth (2005).

ACT-R is a cognitive architecture that is based on empirically motivated principles of working memory. It has served as the basis for a computational model of sentence processing. In this model, the hierarchical structure of a sentence is represented as a set of disconnected chunks in content-addressable memory. Linguistic dependencies, such as those involving the relationship between a reflexive and its antecedent are formed using retrieval cues that target specific licensing features of individual linguistic memory chunks. Specifically, retrieval is used to access chunks for completing dependencies, and the hierarchical structure arises as a consequence of the pointer mechanism inspired from HPSG attribute-value matrices (Pollard & Sag, 1994).

The defining property of CAM is that lexical and grammatical constraints provide the retrieval cues that allow direct access to the necessary licensing information in memory, obviating the need to search through extraneous representations. Chunks are encoded as a bundle of feature-value pairs called ‘content features’. Features that are used for encoding and retrieval include lexical content (e.g., category information, morphological features, etc.), structural features (e.g., Case, grammatical role, etc.), and local hierarchical relations (e.g., sisterhood/complementation). Values for feature may include symbols (e.g., ±singular, ±nominative, etc.) or pointers to other chunks in memory (e.g., NP_1, IP_2, etc.). A linguistic chunk may be retrieved from CAM for further processing or to build a linguistic dependency if the stored memory representation contains some features that overlap with the retrieval cues, i.e., linguistic information is retrieved based upon the content of the representations rather than their location in the hierarchical representation of the sentence in memory. Chunks are differentially activated based on the degree of similarity between their content features and the retrievals, and the total activation level of an individual chunk determines both the probability of retrieval and the retrieval latency. The activation of a chunk A_i is defined by Equation 1.

$$A_i = B_i + \sum_{j=1}^{m} W_j S_{ji} + \sum_{k=1}^{p} P\, M_{ki} + \epsilon \qquad (1)$$

Before I describe each term of Equation 1 in detail, it is important to emphasize that Equation 1 reflects fairly neutral assumptions about human memory. In particular, each term corresponds to a basic property of human memory. For example, consider the task of remembering a phone number. There are several factors that will impact your ability to successfully remember that number, including the time since you last used that phone number (=baseline activation, term 1), how many other phone numbers you have to remember (=similarity-based interference, term 2), how similar those other phone numbers are to the one that you are trying to remember (=partial-match interference, term 3), and random distractions (=noise, term 4). As such, ACT-R is based on relatively neutral assumptions about the operation of human memory.

In more detail, the first term corresponds to the baseline activation of a chunk i, which reflects the frequency and recency of use and time-based decay. The baseline activation is defined by Equation 2, where t_j reflects the time since the j^{th} retrieval of chunk i, summing over all n retrieval results.

$$B_i = \ln\left(\sum_{j=1}^{n} t_j^{-d}\right) \quad (2)$$

The second term corresponds to the associative activation of a chunk i, which is the summation of the weighted strength of association $W_j S_{ji}$ between each probe cue (Q_j ..., Q_n) in the retrieval probe and the memory item I_i, where W_j is the weight associated with each cue j. This weight determines the contribution of a cue to the overall activation of an item. Cue weighting W_j is calculated by

G/n, where G is the total amount of activation available and n is the total number of cues in the retrieval probe. The strength of association from cue j to item i, represented as S_{ji}, is calculated by Equation 3, where S reflects the total associative strength between a cue j and item i. The strength of association is reduced as a function of the number of items in memory that are associated with cue j by the log of the "fan" of an item.

$$S_{ji} = S - \ln(\text{fan}_j) \quad (3)$$

The third term corresponds to the partial-match mismatch penalty for chunk i, which is the summation of the partial cue-matches with the retrieval probe. The negatively-valued parameter P corresponds to the match scale, i.e., the amount of weighting given to the similarity to cue k in the retrieval probe. The term M_{ki} corresponds to the similarity between the cue k in retrieval probe and the corresponding feature in chunk i, which is expressed by a maximum similarity (e.g., 1) and a maximum difference (e.g., 0). Thus, the total partial-match penalty is the number of cues that are not matched by chunk i. This component is crucial in allowing partial-match retrievals, as it provides an activation boost for partial-matching chunks. When combined with stochastic activation noise, as defined by Equation (4), a grammatically illicit item may be incorrectly retrieved from memory, giving rise to attraction effects. Stochastic noise is distributed according to a logarithmic distribution, with a mean of 0 and a variance that is a function of the noise parameter s.

$$\sigma^2 = \frac{\pi^2}{3} s^2 \tag{4}$$

At the point of retrieval, the ACT-R model computes the activation for each chunk i in memory, with respect to the retrieval probe. Higher activations map to faster retrieval latencies, as defined by Equation (5), where T_i corresponds to the retrieval latency for chunk *i.* The variable *Fi* is a scaling parameter that is adjusted in order to fit the dependent measure. Thus, the chunk with the highest activation will be the memory that is retrieved. In this respect, ACT-R is similar to other race models of syntactic comprehension (e.g., McRoy & Hirst, 1990; Traxler, Pickering, & Clifton, 1998; Van Gompel, Traxler, & Pickering, 2001) .

$$T_i = Fe_i^{-A} \tag{5}$$

The ACT-R model provides several dependent measures, including the predicted error rate by condition, the predicted retrieval latency by condition, and the predicted attraction effect. The predicted error rate corresponds to the percentage of the model runs in which the distractor chunk was retrieved from memory for the critical retrieval (e.g., at the reflexive or agreeing verb). I assume with others (e.g., Vasishth et al. 2008; Dillon 2011) that predicted retrieval error is the index of attraction in comprehension, if attraction results from incorrect retrieval of the distractor. The predicted latency indexes the time for the winning retrieval, computed based on Equation (5). Following others, I adopt the simplifying assumption that there is a monotonic relation between the reading times in empirical measures (e.g., total time reading measures in eye-tracking or self-paced reading times) that are thought to partly reflect the timing of retrieval operations in real-time comprehension and the predicted retrieval latencies that are generated by the model. That is, longer retrieval latencies correspond to longer overall reading times. Lastly, the attraction effect is calculated as the difference in predicted error rates or retrieval latencies between two conditions, e.g., the difference between an ungrammatical sentence with a feature-matched attractor and its ungrammatical counterpart without a feature-matched attractor. This measure estimates the magnitude and direction of the attraction effect from.

There are two key assumptions about the ACT-R model that are worth pointing out before we continue. First, it is assumed in the ACT-R framework that that all retrospective linguistic dependencies are implemented using a single uniform retrieval mechanism that deploys all available information to guide retrieval (see the right side of the decision tree in Chapter 2). For example, reflexive dependencies, just like subject-verb agreement, deploy a combination of structural and non-structural cues. This assumption differs from existing accounts of reflexive processing, which suggest that antecedent retrieval relies on a narrow set of syntactic cues (e.g., Dillon et al., 2013). Importantly, a model that uses a narrow set of syntactic cues has been shown to capture the effects of non-attraction reported in previous studies of reflexive anaphors. In contrast, I implement a model that utilizes all available information to examine the impact of multiple feature combinations on attraction profiles. Second, it is generally assumed that retrieval involves the head of the retrieval target, rather than the entire phrase, e.g. a complex NP for licensing a reflexive. This is neither an obvious nor an innocent assumption, as it impacts the relative activation of the target and distractor positions.

4.5.2 Model parameters

We used the default ACT-R parameter values shown in Table 4.7. The only exception to the default parameters is the scaling parameter *F*, which was adjusted to fit the appropriate time scale. 5,000 Monte Carlo simulations were run for each model. Each trial included the full series of hypothesized retrievals, and each trial generated a predicted retrieval latency and predicted retrieval error for the chunk with the highest probability.

Table 4.7: Summary of ACT-R model parameters (Experiment 3a).

Parameter	**Value**
Latency factor (*F*)	1.5
Total goal activation (*G*)	1.0
Noise (*ans*)	0.4
Maximum associative strength (fan)	1.5
Decay (*d*)	0.5
Maximum difference (*P*)	-1.0

4.6.3 Materials and retrieval schedules

I simulated reflexives dependencies using a retroactive interference paradigm based on the constituent creation times described in Tables 4.8 (retroactive interference paradigm, Experiment 1) and 4.10 (proactive interference paradigm, Experiment 2) and the respective retrieval schedules described in Tables 4.9 and

4.11. Creation times were estimated from the empirical reading times (re-read times) from Experiments 1 and 2. Retrievals are linked to the processing of a given constituent, e.g., a main clause VP, a relative clause VP, or a reflexive. The structural constraints on reflexive anaphors, which require a local, c-commanding subject, were implemented as a combination of structural location features, e.g. [Spec, IP1], and features indicating the level of embedding, e.g., IP1, following Dillon (2011). To test the impact of probe-to-target similarity on attraction, I manipulated the degree of match between the target (the local subject) and retrieval cues, holding the feature composition of the structurally irrelevant but feature matching distractor constant. I varied the degree of probe-to-target similarity along a scale, ranging from a perfect match to mismatching on three cues, e.g., number, gender, and animacy.

Table 4.8: Constituent creation times and feature composition for the reflexives retroactive interference paradigm from Experiment 1 (Experiment 3a).

	NP-Target	**NP-Distractor**	**VP1**	**VP2**
Time	550	1150	1700	2300
Category	NP	NP	VP	VP
Person	3	3	3	3
Number	sing	sing	sing	sing
Gender	fem	fem	-	-
Animacy	animate	animate	animate	animate
Role	[Spec, IP1]	[Spec, IP2]	[head, V1]	[head, V2]
Embedding	IP1	IP2	IP2	IP1

Table 4.9: Schedule of retrievals and cue sets for the reflexives retroactive interference paradigm from Experiment 1 (Experiment 3a). VP1 = object retrieval of the main clause subject inside the relative clause. VP2 = thematic binding of the main clause subject for the main clause verb. Reflexive = antecedent retrieval for the reflexive.

	VP1	**VP2**	**Reflexive**
Time	1700	2300	2700
Category	NP	NP	NP
Person	-	-	3
Number	-	-	sing
Gender	-	-	fem
Animacy	-	-	animate
Role	[Spec, IP1]	[Spec, IP1]	[Spec, IP1]
Embedding	IP1	IP1	IP1

Table 4.10: Constituent creation times and feature composition for the reflexives proactive interference paradigm from Experiment 1 (Experiment 3a).

	NP-Distractor	**VP1**	**NP-Target**	**VP**
Time	550	1150	1700	2300
Category	NP	VP	NP	VP
Person	3	3	3	3
Number	sing	sing	sing	sing
Gender	fem	-	fem	-
Animacy	animate	animate	animate	animate
Role	[Spec, IP1]	[head, V1]	[Spec, IP2]	[head, V2]
Embedding	IP1	IP1	IP2	IP2

Table 4.11: Schedule of retrievals and cue sets for the reflexives proactive interference paradigm from Experiment 1 (Experiment 3a). VP1 = thematic binding of main clause subject for the main clause verb. VP2 = thematic binding of the embedded clause subject for the embedded clause verb. Reflexive = antecedent retrieval for the reflexive.

	VP1	**VP2**	**Reflexive**
Time	1150	2300	2700
Category	NP	NP	NP
Person	-	-	3
Number	-	-	sing
Gender	-	-	fem
Animacy	-	-	animate
Role	[Spec, IP1]	[Spec, IP2]	[Spec, IP2]
Embedding	IP1	IP2	IP2

4.5.4 Modeling results

The results from the ACT-R computational simulations are shown in Figure 4.3 (retroactive interference paradigm) and Figure 4.4 (proactive interference paradigm. Simulations revealed an exponential increase in the presence of an illusion as a function of the probe-to-target similarity across both interference configurations. Further investigation revealed that this modulation of the attraction effect is a result of the non-linearity of Equation 3, which is graphically represented in Figure 4.5.

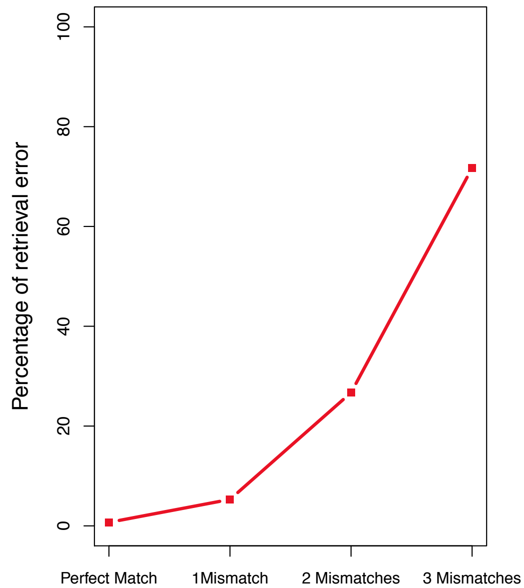


Figure 4.3: Computational simulations of probe-to-target similarity for a retroactive interference paradigm, e.g., Experiment 1 (Experiment 3a). Predicted percentage of retrieval error is plotted against varying degrees of probe-to-target similarity range from a perfect match to mismatching on three retrieval cues.

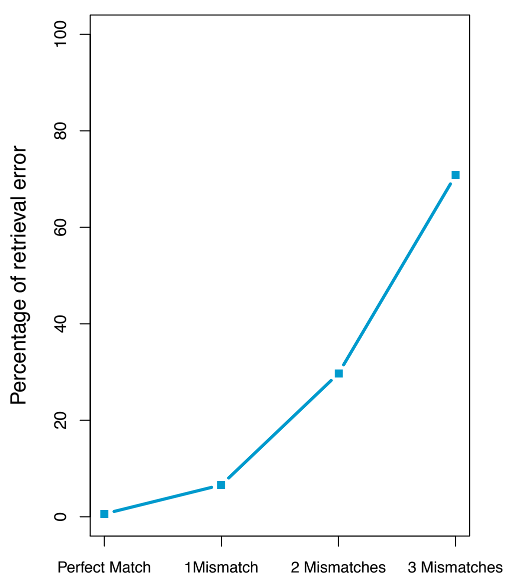


Figure 4.4: Computational simulations of probe-to-target similarity for a proactive interference paradigm, e.g., Experiment 2 (Experiment 3a). Predicted percentage of retrieval error is plotted against varying degrees of probe-to-target similarity range from a perfect match to mismatching on three retrieval cues.

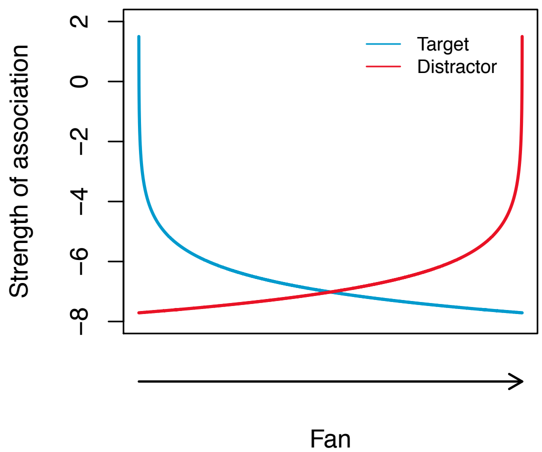


Figure 4.5: A graphical representation of the non-linearity of the ACT-R fan parameter, which gives rise to the exponential increase in attraction effects shown in Figures 4.2 and 4.3. This figure shows that as probe-to-target similarity decreases (described as the fan parameter), the strength of association to the target decreases exponentially, and sensitivity to a feature-compatible attractor increases exponentially.

4.5.5 Modeling discussion

The ACT-R model of memory retrieval revealed a surprising impact of probe-to-target similarity, as we observed a substantial, non-linear growth in the presence of an illusion as the probe-to-target similarity decreased. These results suggest that the degree of similarity between the target memory and the retrieval probe may have a stronger impact on susceptibility to attraction effects than we previously assumed. These findings might have important consequences for our understanding of how attraction effects arise in reflexive dependencies. In particular, the modeling results imply that increased susceptibility to attraction effects, as observed in our comparison between gender and animacy match in Experiment 2 is not a consequence of the use of specific content features like animacy in memory retrieval. Rather, the modeling results suggest that increased susceptibility to attraction is a consequence of probe-to-target similarity, such that when the memory target is an extremely poor fit to the retrieval probe, as in the case of an animacy mismatch, comprehenders are more sensitive to the feature match of structurally irrelevant distractors, leading to an increased likelihood of mis-retrieval.

According to this account, the issues regarding goodness of fit for animacy (mis-)matches could be made to follow from the computation of feature match in a content-addressable memory architecture. As suggested above, for example, a retrieval target that does not bear the required animacy feature could be doubly penalized in accordance with Equation 1 because the target does not also bear the required gender feature, increasing the likelihood of mis-retrieval, as shown in Figures 4.2 and 4.3. In short, the modeling results quantify the hypothesis that when the target is a poor fit to the retrieval cues, comprehenders are more susceptible to attraction.

4.6 Experiment 3b

Experiment 3b tested the predictions of the ACT-R model using eye-tracking, comparing contexts that involved a single feature manipulation, e.g., gender, as done in previous studies of reflexives to contexts that involved multiple feature manipulations, e.g., gender + number.

4.6.1 Participants

Twenty-four members of the University of Maryland community participated in Experiment 1.[18] All participants gave informed consent, and received either course credit or payment of $10 for their participation. The experimental session lasted approximately 45 minutes.

4.6.2 Materials

Thirty-six item sets of the form shown in Table 4.12 were constructed to directly test the effect of probe-to-item similarity. Two experimental factors were manipulated: grammaticality and attraction. Grammaticality was manipulated by varying the number of feature mismatches between the reflexive and the target, such that it either mismatched one feature, e.g., gender, or two features, e.g., gender and number. Likewise, attraction was manipulated by varying the gender of the distractor noun. Across all conditions, the distractor was the subject of the main clause and the target was always the subject a subordinate clause that contained the reflexive. This configuration was used in order to prevent effects of proximity concord. In order to prevent any biases in retrieval, neither the main clause verb nor the embedded clause verb expressed agreement, and both the target and the distractor nouns were semantically appropriate antecedents for the reflexive.

[18] Only 24 participants were tested, rather than 30, as in the previous experiments, due to time constraints.

Table 4.12: Summary of reflexive 1- vs. 2-feature mismatch manipulation for Experiment 3. Pre-critical, critical, and spill-over regions included in the analysis are indicated by brackets.

Reflexives 1-feature mismatch manipulation for Experiment 3
Grammatical, no attraction
1. The talented actor mentioned that [the attractive spokesman praised] [himself] [for a] great job.
Grammatical, attraction
2. The talented actress mentioned that [the attractive spokesman praised] [himself] [for a] great job.
Ungrammatical, attraction
3. The talented actor mentioned that [the attractive spokeswoman praised] [himself] [for a] great job.
Ungrammatical, no attraction
4. The talented actress mentioned that [the attractive spokeswoman praised] [himself] [for a] great job.
Reflexives 2-feature mismatch manipulation for Experiment 3
Ungrammatical, attraction
5. The talented actor mentioned that [the attractive spokeswomen praised] [himself] [for a] great job.
Ungrammatical, no attraction
6. The talented actress mentioned that [the attractive spokeswomen praised] [himself] [for a] great job.

4.6.5 Procedure and data analysis

The same procedure was used as in Experiment 1 and the analysis followed the same steps. The box-cox procedure determined that a natural log was the appropriate transformation to obtain normally distributed residuals (average across critical region of interest: λ=-0.2).

4.6.6 Results

Comprehension question accuracy

The comprehension question accuracy for Experiment 1 was 87%, indicating that participants successfully comprehended the experimental stimuli.

Reading times

Raw mean reading times and statistical analyses by measure and by region for Experiment 2 are presented in Tables 4.13 and 4.14. The results of Experiment 3b replicated the attraction effects observed in Experiments 1 and 2, further demonstrating that reflexives are in fact susceptible to attraction effects from structurally irrelevant but feature compatible material in memory. In addition, Experiment 2 replicated the selectivity of the attraction effects within reflexive dependencies as observed in Experiment 2. In particular, when the reflexive mismatched the target (the local subject) in only one feature, gender, there were no attraction effects. But, when the target mismatched in two features, gender and number, there were strong attraction effects, qualitatively similar to subject-verb agreement. In the single feature manipulation conditions, only the (mis-)match of the structurally defined target impact reading times at the reflexive. Ungrammatical conditions where the target mismatched one feature showed longer reading times, and crucially, there were no significant effects of attraction or an interaction between grammaticality and attraction at any region or in any measure, replicating previous findings for reflexives (e.g., Sturt, and Dillon). However, we observed a different profile for the conditions involving multiple feature mismatches, which showed a significant effect of attraction from structurally irrelevant material.

Table 4.13: Table of means for reflexive 1-feature and 2-feature mismatch manipulations, by measure and by region for Experiment 2. Standard error by participants is shown in parentheses.

		Region		
		Pre-critical	**Critical**	**Spill-over**
	First pass reading time			
1-feature	Grammatical, attractor	947 (50)	197 (11)	166 (13)
	Grammatical, no attractor	928 (43)	223 (12)	155 (15)
	Ungrammatical, attractor	805 (34)	225 (13)	166 (23)
	Ungrammatical, no attractor	903 (40)	223 (12)	165 (16)
2-features	Ungrammatical, attractor	913 (36)	185 (12)	156 (14)
	Ungrammatical, no attractor	882 (37)	290 (18)	129 (14)
	Right-bound reading time			
1-feature	Grammatical, attractor	1162 (54)	200 (11)	196 (17)
	Grammatical, no attractor	1108 (41)	228 (12)	185 (23)
	Ungrammatical, attractor	1025 (36)	245 (14)	215 (31)
	Ungrammatical, no attractor	1088 (39)	234 (13)	195 (20)
2-feature	Ungrammatical, attractor	1117 (39)	191 (13)	191 (20)
	Ungrammatical, no attractor	1062 (40)	329 (21)	178 (23)
	Regression path duration			
1-feature	Grammatical, attractor	1295 (79)	221 (16)	312 (43)
	Grammatical, no attractor	1203 (57)	251 (14)	282 (44)
	Ungrammatical, attractor	1199 (65)	304 (29)	339 (54)
	Ungrammatical, no attractor	1206 (58)	317 (40)	342 (67)
2-feature	Ungrammatical, attractor	1188 (47)	229 (24)	318 (53)
	Ungrammatical, no attractor	1199 (64)	456 (65)	405 (91)
	Re-read time			
1-feature	Grammatical, attractor	873 (97)	223 (30)	265 (36)
	Grammatical, no attractor	799 (93)	224 (27)	253 (36)
	Ungrammatical, attractor	881 (98)	318 (34)	266 (38)
	Ungrammatical, no attractor	1044 (106)	339 (34)	234 (26)
2-feature	Ungrammatical, attractor	925 (99)	212 (25)	287 (36)
	Ungrammatical, no attractor	1210 (119)	445 (55)	288 (37)

Table 4.14: Summary of the statistical analysis by measure and by region for Experiment 3b. All *t*-values with an absolute value greater than 2 are significant with an alpha value at 0.05, indicated in bold.

	REGIONS								
	PRE-CRITICAL			CRITICAL			SPILL-OVER		
	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t
First pass reading time									
grammaticality	-0.06	0.05	-1.29	0.13	0.20	0.65	0.02	0.23	0.11
agreement attraction	-0.05	0.05	-1.03	-0.16	0.18	-0.89	0.14	0.22	0.65
grammaticality×attraction	-0.11	0.09	-1.16	0.45	0.33	1.36	-0.53	0.43	-1.22
animacy attraction	0.03	0.06	0.53	**-0.97**	**0.28**	**-3.40**	0.53	0.33	1.61
Right-bound reading time									
grammaticality	-0.05	0.03	-1.57	0.16	0.20	0.79	0.05	0.23	0.22
agreement attraction	-0.01	0.02	-0.61	-0.15	0.18	-0.84	0.17	0.23	0.74
grammaticality×attraction	-0.06	0.05	-1.09	0.48	0.33	1.46	-0.51	0.44	-1.17
animacy attraction	0.05	0.04	1.36	**-1.06**	**0.29**	**-3.62**	0.51	0.34	1.50
Regression path duration									
grammaticality	-0.03	0.03	-0.84	0.20	0.21	0.94	0.06	0.24	0.24
agreement attraction	0.00	0.03	0.22	-0.17	0.18	-0.93	0.19	0.25	0.79
grammaticality×attraction	-0.04	0.06	-0.66	0.49	0.34	1.43	-0.49	0.46	-1.07
animacy attraction	0.02	0.05	0.49	**-1.12**	**0.30**	**-3.60**	0.52	0.37	1.39
Re-read time									
grammaticality	0.29	0.23	1.24	**0.93**	**0.23**	**3.89**	-0.20	0.27	-0.74
agreement attraction	-0.02	0.29	-0.06	-0.17	0.21	-0.81	0.02	0.27	0.09
grammaticality×attraction	-0.45	0.50	-0.90	0.03	0.42	0.09	-0.64	0.51	-1.25
animacy attraction	**-0.79**	**0.37**	**-2.12**	**-1.08**	**0.31**	**-3.48**	0.32	0.36	0.8

Reflexives: 1-feature mismatch manipulation

Early processing (first-pass reading time and right bound reading time)

No significant effects or interactions were observed in first-pass reading times or right-bound reading times at the pre-critical, critical or spill-over regions.

Late processing (regression path duration and re-read time)

A significant main effect of grammaticality was observed in re-read times at the critical reflexive region, as reflected by a slow down in reading times for ungrammatical sentences relative to grammatical sentences. Importantly for the 1-feature mismatch manipulation, no significant effects of attraction and no interaction of grammaticality with attraction was observed at any region or in any measure, replicating previous findings for reflexives (e.g., Sturt, 2003; Dillon et al., 2013).

Reflexives: 2-feature mismatch manipulation

Early processing (first-pass reading time and right bound reading time)

No significant effects were observed at the pre-critical region. A significant main effect of attraction was observed in both first-pass and right-bound times at the critical reflexive region, as reflected by facilitated reading times for ungrammatical sentences with an attractor relative to ungrammatical sentences without an attractor. No other effects were observed in early measures.

Late processing (regression path duration and re-read time)

A significant main effect of attraction was observed in regression path duration and re-read times at the critical reflexive region, as reflected by facilitated reading times for ungrammatical sentences with an attractor relative to ungrammatical sentences without an attractor. A similar effect was observed in the pre-critical region, indicating that the attraction effect impact re-reading of the preceding region. No other effects were observed in late measures.

Direct comparison of attraction effects

The results from Experiment 3 showed a reliable effect of attraction when the reflexive target mismatched 2-features, and no corresponding effect when the target mismatched 1-feature. Because the contrast between the 1- and 2-feature mismatch manipulation is of primary theoretical interest, I directly tested the attraction effects between the two feature manipulations. Importantly, contrasting attraction profiles were observed for these manipulations, as reflected by an interaction between attraction and probe-to-target similarity (1- vs. 2-feature mismatch) ($\hat{\beta}$=-0.92, SE=0.42, t=2.16) in re-read times at the critical reflexive region. Pairwise comparisons revealed that this interaction was driven by the significant attraction effect for the 2-feature mismatch manipulation reported above.

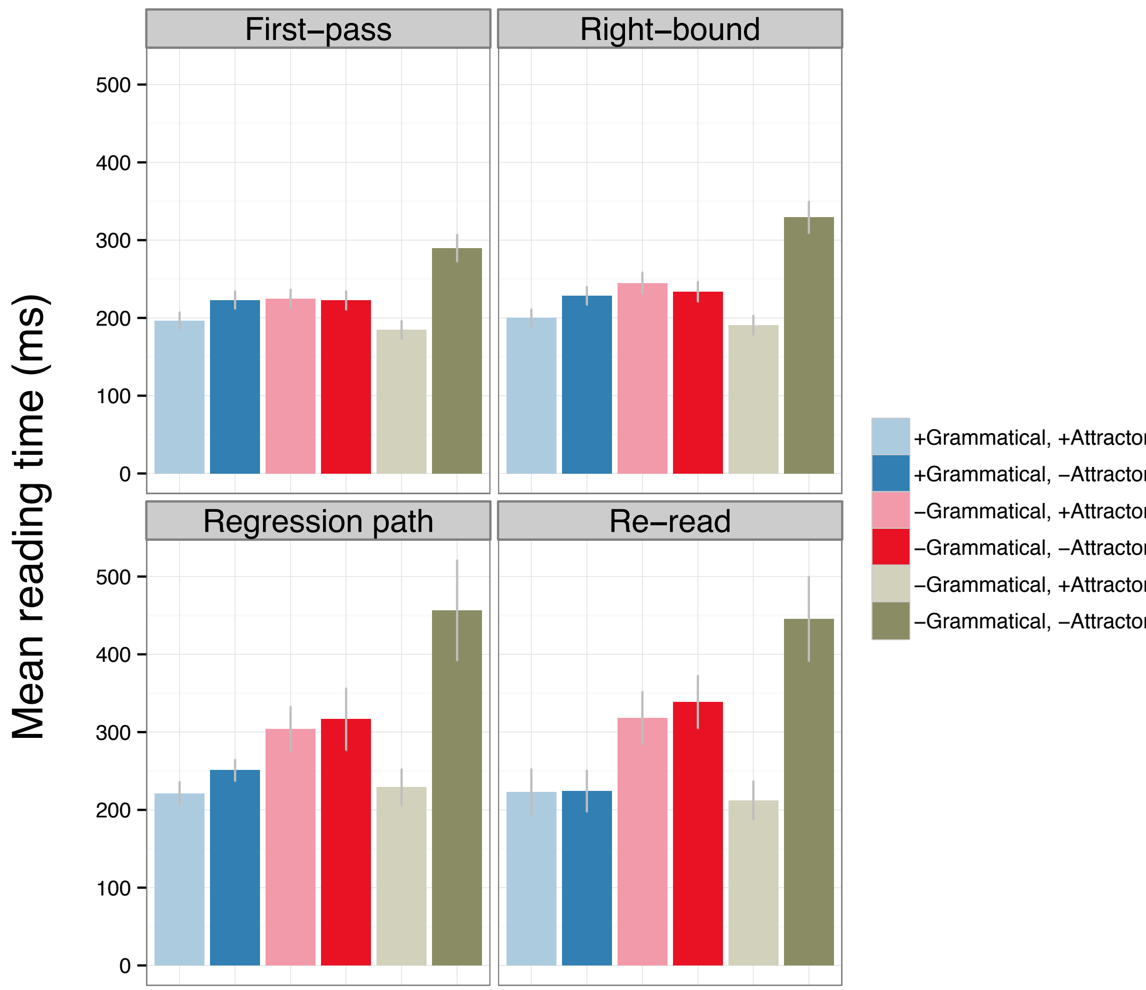


Figure 4.6: Mean reading times by measure by condition (Experiment 3b). Error bars indicate standard error by participant.

4.6.7 Discussion

The results of Experiment 3b are consistent with the predictions of the computational simulations in Experiment 3a. In particular, the results of Experiment 3b suggest that susceptibility to attraction effects within reflexive dependencies is not a consequence of the use of specific features like animacy in retrieval, as we observed strong attraction effects using a different feature combination that did not rely on animacy to probe for attraction effects. Rather, the results of Experiment 3b suggest that susceptibility to attraction effects within reflexive dependencies is likely to be a consequence of the probe-to-target similarity, such that when the target is a poor fit to the retrieval cues, as measured by the number of matching cues (e.g., Equation 1), we see strong attraction effects. The generalization that emerges from Experiments 1-3 is that when the target memory is a poor match to the retrieval cues, comprehenders notice structurally irrelevant but feature matched lures.

4.7 Experiment 4

One issue that the results from Experiment 3 raises is how a quantitative difference in probe-to-target similarity (1 vs. 2 feature mismatch) can yield a qualitative difference in attraction effects, with non-attraction effects in a 1-feature mismatch context and strong attraction effects in 2-feature mismatch contexts. In particular, the empirical results from Experiment 3b differ from the predictions of the computational simulations reported in Experiment 3a, in that the computational simulations predicted a gradient effect of probe-to-target similarity on the attraction effect rather than the categorical differences observed in the empirical data. Furthermore, the results from Experiments 1-3 do not explain why subject-verb agreement is much more susceptible to attraction effects than reflexives, even in 1-feature mismatch contexts, where the target mismatches only on number.

The differences between the computational simulations and empirical results and the differences between reflexives and subject verb agreement suggest that additional factors beyond simple probe-to-target similarity must impact susceptibility to attraction effects. In particular, the categorical differences in attraction effects observed in Experiment 3b suggest two things that are in apparent opposition: On the one hand, the non-attraction effects for reflexives suggest that structural information has a privileged status in antecedent retrieval. On the other hand, evidence of attractions effects for reflexive dependencies suggests that structural priority is much more limited.

The goal of Experiment 4 was to investigate how priority for structural information in memory retrieval can impact attraction effects using computational simulations. To achieve this, I manipulated the strength of structural retrieval cues relative to non-structural cues. In particular I asked: is there a weighting for structural information that can capture both the presence and absence of reflexive attraction within the same memory architecture?

4.7.1 Cue-combinatorics schemes

A computationally complete model of sentence comprehension (in the sense of Lewis and Vasishth, 2005) must explicitly describe (i) the cues that are used in memory retrieval, and (ii) the way in which those cues combine at the retrieval site to generate a retrieval probe, i.e. the cue-combinatorics scheme. However, the type of cue-combinatorics scheme that is used in memory retrieval for dependency resolution remains unresolved.

Research on cue-combinatorics has identified several schemes that could be used in sentence comprehension. One possibility is that cues are combined at the retrieval site using a linear/additive scheme, as defined by Equation (6), where each cue in the retrieval probe contributes independently to the activation of a memory item.[19] In a linear scheme, the probability P of retrieving a memory item A_i ($P(A_i)$) is calculating by the summation of the weighted strength of association ($W_jS(Q_j,I_i)$) between each probe cue (Q_j ..., Q_m) in the retrieval probe and the memory item I_i, where W_j is the weight associated with each cue j. This weight determines the contribution of a cue to the overall activation of an item. Cue weighting W_j is calculated by G/n, where G is the total amount of activation available and n is the total number of cues in the retrieval probe. The strength of association from cue j to item i, represented as $S(Q_j,I_i)$, is calculated by Equation 3, where S reflects the total associative strength between a cue j and item i. The strength of association is reduced as a function of the number of items in memory that are associated with cue j by the log of the "fan" of an item (see Equation 3).

---

[19] Equation 6 is a variant of the second term of Equation 1 .

$$P(A_i|Q_1, \dots, Q_m) = \sum_{j=1}^{m} W_j \, S(Q_j, I_i) \tag{6}$$

Retrieval cues may also be combined in a non-linear/multiplicative fashion, as defined by Equation (7) (from Van Dyke & McElree, 2006) , where the contribution of each cue is dependent on the other cues in the retrieval probe. In a non-linear scheme, the probability *P* of retrieving a memory item I_i ($P(I_i)$) is calculating by the multiplication of the strength of association *S* between each probe cue (Q_j …, Q_m) and the memory item I_j, where w_j is the weighting factor.

$$P(A_i|Q_1, \dots, Q_m) = \frac{\prod_{j=1}^{m} S(Q_j, I_j)^{w_j}}{\sum_{k=1}^{N} \prod_{j=1}^{m} S(Q_j, I_j)^{w_j}} \tag{7}$$

In both a linear and non-linear scheme, the probability of retrieving an item from memory is an increasing function of the degree to which the retrieval cues match the memory item, providing a probe-to-item similarity strength described in terms of activation.[20] Crucially, however, because retrieval in both schemes is driven by probe-to-item similarity, the probability of retrieving a particular memory item will be reduced as a function of the degree to which the retrieval cues match other items in memory (e.g., the “fan”). In a linear scheme, where each cue contributes independently to the activation of an item, multiple items may be differentially activated based on their degree of match to the retrieval cues, items that partially match the retrieval cues may be activated. One

[20] The term “activation” refers to an abstract activation quantity, rather than actual neural activity. However, Anderson (2007) describes activation in ACT-R as the input that drives the actual rate of firing of neurons.

consequence of a linear scheme, then, is that items that partially match the retrieval cues may be incorrectly retrieved, especially when the probe-to-distractor strength is similar to the probe-to target strength. In the domain of sentence processing, partial-match activation can give rise to attraction effects from morphologically feature-matched items that are in structurally irrelevant positions in the hierarchical representation of the sentence in memory. That is, the retrieval cues in a linear cue-combinatorics scheme impose only a soft constraint on retrieval. In a non-linear, multiplicative scheme, by contrast, mis-retrieval due to partial-matching is less likely since the contribution that each cue makes to the probe-to-item strength of association is dependent on the contribution of the other cues in the retrieval probe.

Current leading models of memory retrieval in sentence processing, such as ACT-R (e.g., Lewis and Vasisth, 2005), assume that retrieval cues combine in a linear fashion. In ACT-R, for example, cues are implemented with equal weights, which amounts to the implicit claim that no one type of information, e.g. structural or non-structural, takes priority in retrieval. The use of a linear cue-combinatorics scheme in sentence comprehension is attractive for several reasons. First, the notion of partial-match activation provides a relatively straightforward explanation for the presence of attraction effects from items that only partially match the retrieval cues. Second, there is independent evidence that cues are combined in a linear fashion across other cognitive and perceptual domains (e.g., Trommershäuser et al., 2011). However, it has been argued that a weighted, linear cue-combinatorics scheme might be necessary in sentence comprehension. For example, in a recent study that used SAT measures, Van Dyke and McElree (2011) found that the structural position of a semantically appropriate distractor determined the presence or absence of interference effects in memory retrieval, as measured by the difference in SAT asymptotes. Based on these findings, Van Dyke and McElree argued that structural cues are prioritized over non-structural cues in calculating probe-to-item strength, such that structural cues effectively "gate" access to non-structural cues in memory retrieval.

4.7.2 Simulating structural priority in memory retrieval

To better understand how priority for structure impacts attraction effects in memory retrieval, I systematically varied the strength of structural cues relative to non-structural cues. I achieved this by varying the weight of structural cues (W_j) in the ACT-R model described in Experiment 3a, ranging from equal weighting, 1:1, to a weighting of 6:1. Results from our simulations are shown in Figure 4.6.

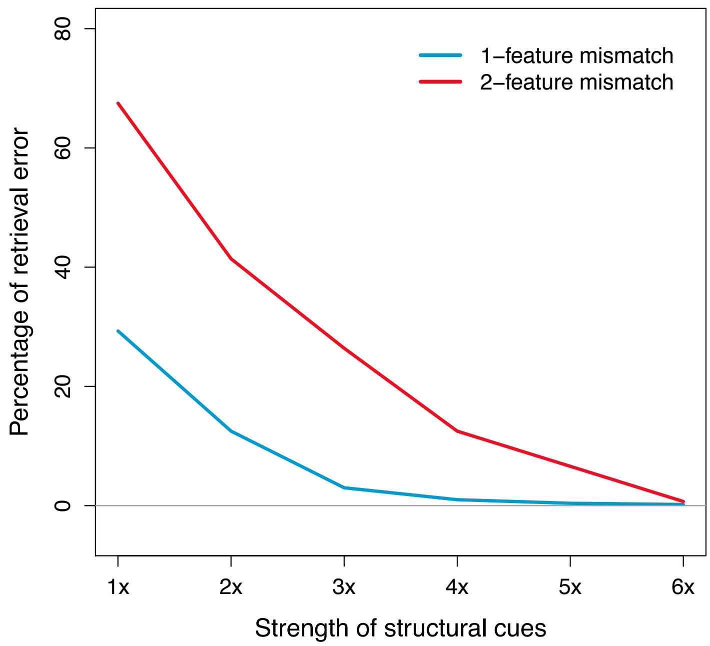


Figure 4.7: Computational simulations systematically varying the priority for structural information in memory retrieval relative to non-structural, morphological information.

Computational simulations revealed that when structural and non-structural cues were weighted evenly, e.g., 1:1, rampant illusions were incorreclty predicted for both 1- and 2-feature mismatch contexts. Conversely, if structural cues are weighted strongly relative to non-structural cues, e.g., 6:1, then structural cues effectively "gate" memory access such that only candidate memory items that match the desired structural properties are considered, completely eliminating illusions for both 1- and 2-feature mismatch contexts. Interestingly, however, there is a range of weights between these extremes that predicts both the presence of an illusion in the 2-feature mismatch context and the absence of an illusion in the 1-feature mismatch context.

4.7.3 Discussion

ACT-R simulations suggest that the contrasting attraction effects for reflexive dependencies may reflect a linear, weighted cue combinatorics scheme in which structural cues are weighted more strongly in retrieval than non-structural, morphological cues. Crucially, the simulations revealed a (narrow) range of weights that can capture the categorical presence and absence of attraction effects for reflexive dependencies involving 1 and 2 feature mismatches from within the same direct-access, content-addressable memory architecture.

The results of the computational simulations in Experiment 4 as well as Experiment 3a are consistent with the findings from Van Dyke and McElree (2011), which suggest that syntactic constraints restrict interference from semantically appropriate distractors for thematic binding. Van Dyke and McElree argued that these data support the use of a linear, weighted cue-combinatoric scheme in which syntactic information effectively gates access to other features in memory retrieval. However, the link between their study and ours is not yet completely clear, as different methodologies were used (e.g., SAT vs. eye-tracking) for different types of interference configurations (e.g., multiple match vs. partial match). Importantly, as discussed in Chapter 2, interference effects in these different configurations license different conclusions about the items that are considered in memory retrieval. In particular, in the case of multiple match configurations, interference effects do not always license the conclusion that the distractor was incorrectly retrieved, whereas this conclusion is licensed in partial match configurations (see Dillon, 2011, for further discussion).

One remaining question that the current results raise is why subject-verb agreement shows strong attraction effects even in a 1-feature mismatch context. The fact that subject-verb agreement dependencies are so easily susceptible to attraction effects from structurally irrelevant material implies that structural cues are not weighted as strongly for subject-verb agreement as they are for reflexive dependencies. But why should structural cues be prioritized for anaphora (including the reflexive anaphor licensing and null subject licensing), but not for subject-verb agreement?

I argue that the contrasting sensitivity to attraction effects observed for subject-verb agreement and reflexive dependencies reflects whether retrieval is triggered by error correction, as in the case of subject-verb agreement, or normal resolution. Specifically, I suggest that whereas structural cues are prioritized (e.g., via weighting) in retrieval in the normal course of processing unpredictable dependencies such as anaphora and thematic binding (e.g., Van Dyke & McElree, 2011), they are not prioritized when retrieval is triggered by an unfulfilled prediction, as in the case of subject-verb agreement. For example, current leading accounts of agreement attraction effects in comprehension (e.g., Wagers et al., 2009) claim that retrieval for subject-verb agreement is an error-driven process that is triggered only when the anticipated features of the verb do not match the actual form of the verb as it appears in the bottom up input. I argue that the error signal that arises from the violated agreement prediction impacts the cue-combinatorics scheme by limiting the priority for structural information in the retrieval probe. For example, it is possible that the error signal from the violated agreement prediction leads the parser to be uncertain about the accuracy of its existing structural encoding of the complex subject NP. If the parser is uncertain about the previous structural encoding, the memory retrieval mechanism may not prioritize structural cues in retrieval, thereby increasing sensitivity to structurally irrelevant morphological features. Specifically, when the parser is uncertain about the encoding of the previous structural analysis, structural cues are not prioritized or weighted more strongly relative to non-structural morphological cues. Anaphoric elements like null subjects or reflexives, by contrast, are not predictable like subject-verb agreement morphology, and retrieval for an antecedent is triggered by normal resolution. Since the features of the anaphoric element cannot be as reliably anticipated as subject-verb agreement morphology, there is no expectation for the actual features of the anaphor to violate. Thus, the parser should be confident about the accuracy of its existing structural encoding and prioritize structural cues in retrieval. That is, structural information should always be prioritized in memory retrieval when the parser is confident about the encoding of the previous structural analysis.

Empirical support for this account comes from our comparison of subject-verb agreement reflexive dependencies. In particular, the eye-tracking data show different timing profiles for subject-verb agreement and reflexive dependencies. In the case of reflexive dependencies, we observed in two of the three studies consistent attraction effects beginning in early measures (first pass) persisting to later measures (e.g., re-read). In the case of subject-verb agreement, by contrast, reading times were initially disrupted in the ungrammatical conditions by the unfulfilled prediction for agreement in early measures (first pass), followed by an attraction effect in later measures (re-read). By hypothesis, initial detection of the violated agreement prediction reduced certainty about the existing structural encoding, leading the parser to reduce the priority for structural information in retrieval for an agreement licensor. In turn, decreased priority for structural information differentially increased sensitivity to non-structural feature matches.

This account is also supported by recent empirical findings that suggest that comprehenders actively maintain and rationally update uncertainty about the prior structure throughout the course of incremental comprehension (e.g., Bicknell, Tanenhaus, & Jaeger, 2014; Levy, Bicknell, Slattery, & Rayner, 2009; Levy, 2008). My account is also consistent with the recent argument that structural cues might not even be used in retrieval for subject-verb agreement (e.g., Dillon, Levy, Staub, & Clifton, 2014). However, further research is needed to better understand whether structural information is used, but with a malleable weighting, as I have suggested, or not at all, as Dillon et al. (2014) have suggested.

4.8 General discussion

4.8.1 Summary of results

The role of structural information in guiding retrieval operations in sentence comprehension is a controversial topic. Evidence from multiple tasks and phenomena has been used to argue that structural information has a privileged status in retrieving the heads of dependencies. However widespread evidence from non-structural attraction effects suggests that structural priority is more limited. This contrast has motivated various proposals regarding the memory architecture for the parser, including the use of qualitatively different memory retrieval mechanisms or the selective use of non-structural cues for different dependencies.

In the present study, I showed that for reflexive licensing, a phenomenon that has resisted non-structural attraction effects in past studies, we can systematically control where attraction effects do and do not occur. Attraction effects manifest as eased processing of illicit dependent elements due to the mis-retrieval of a structurally irrelevant licensor, giving rise to an illusion of grammaticality. Previous studies on reflexive anaphors have failed to consistently find such effects, but they tested only contexts in which the illicit dependency involved one feature mismatch, such as only gender (e.g., Sturt, 2003), or number (e.g., Dillon et al., 2013). Instead, I compared contexts involving a 1-feature mismatch and a 2-feature mismatch. The results from our study show that reflexives are indeed susceptible to attraction effects like those commonly observed for subject-verb agreement, but only selectively. In particular, when the reflexive mismatches the true subject in just one feature, there is attraction, but when it mismatches in two features, strong attraction effects are observed. These findings suggest that reflexive attraction is not a consequence of the use of specific features, like animacy, but rather the degree of probe-to-target similarity.

Furthermore, I showed based on computational simulations that is possible to derive both the presence and absence of the reflexive attraction from within the same memory architecture. In particular, ACT-R simulations suggest that the contrasting attraction effects for reflexive dependencies may reflect a weighted, linear cue-combinatorics scheme in which structural cues are weighted more strongly in antecedent retrieval than non-structural, morphological cues.

Lastly, I proposed an account of why it is harder to obtain attraction effects in anaphor resolution than in some other dependencies, based on whether retrieval is triggered by error correction or normal resolution. Specifically, I suggested that whereas structural cues are prioritized in retrieval in the normal course of processing unpredictable dependencies, such as anaphora or thematic binding, they are not prioritized when retrieval is triggered by an unfulfilled prediction, as in the case of subject-verb agreement violations, leading the parser to be uncertain about the accuracy of its existing structural encoding. I provided supporting timing evidence from Experiments 2 and 3, which show different timing profiles for attraction effects, which are observed early (first pass) for reflexive dependencies, but only as a late effect (re-read) for subject-verb agreement.

4.8.2 Revised generalizations

Recent research on the processing of subject-verb agreement and anaphora have converged on the generalization that the selective nature of attraction effects is a consequence of how the two dependencies engage the candidate memory mechanisms, including the use of qualitatively different retrieval mechanisms or the selective use of non-structural cues for the two dependencies (left vs. right side of the decision tree in Chapter 2). In particular, it has been argued that the candidate memory access mechanisms or retrieval strategies may be directly linked to specific types of grammatical dependencies. For example, Dillon and colleagues (2013) suggest that a structure memory access strategy is used for anaphoric dependencies, but not for subject verb agreement, which they implemented as a narrow set of syntactic retrieval cues for reflexive dependencies. They acknowledge that the empirical results from their study are also consistent with models that assume different retrieval mechanisms for different dependencies. In their simulations, the use of both structural and non-structural cues for subject-verb agreement results in mis-retrieval of a morphologically compatible but structurally irrelevant distractor, but the use of only structural cues for reflexive dependencies guarantees retrieval of the structurally appropriate antecedent. These accounts predict that structured access should always be observed for anaphoric dependencies.

However, the findings from the present study, taken together with the results for null subject licensing reported in Chapter 3, demonstrate that anaphoric dependencies are indeed susceptible to attraction effects from structurally irrelevant material, leading to illusions of grammaticality. These findings suggest that structured memory access is not directly linked to anaphoric dependencies. Rather, the results from the present study point to a new generalization, namely that susceptibility to attraction effects in sentence comprehension is a consequence of the degree of similarity between the target memory and the retrieval probe, such that when probe-to-target similarity is extremely poor, as in the case of an animacy mismatch, or mismatching on multiple features, comprehenders notice the feature match of structurally irrelevant distractors.

In contrast to previous proposals, the finding that subject-verb agreement and reflexives are both susceptible to attraction, albeit differentially susceptible, suggests that they deploy the same memory retrieval mechanism. In particular, sensitivity to structurally irrelevant, but morphologically or semantically compatible material is expected in a direct-access, content-addressable memory architecture that deploys all available information, including both structural and non-structural retrieval cues using a weighted, linear cue-combinatorics scheme (e.g., Van Dyke and McElree, 2011). More generally, the current findings are consistent with accounts which claim that memory retrieval in sentence comprehension relies on a single, uniform retrieval mechanism (see the right side of the decision tree in Chapter 2).

The current findings are also compatible with previous accounts that claim that syntactic structure has a privileged status in antecedent retrieval for reflexive licensing (e.g., Clifton et al., 1999; Cunnings & Sturt, 2012; Dillon et al., 2013; Nicol & Swinney, 1989; Sturt, 2003; Xiang et al., 2009). However, sensitivity to structurally irrelevant feature matches during antecedent retrieval suggests that syntactic structure does not provide a hard constraint on the antecedents that are initially considered for reflexive resolution. In particular, the differential sensitivity to non-structural information observed in the present study, taken together with the results of the computational simulations, suggests that structural information is simply valued more strongly in retrieval than morphological information, but not implementation as a constant hard constraint.

A weighted cue-combinatorics scheme for reflexive dependencies is also compatible with earlier claims that morphological information is not as critical as structural information in antecedent retrieval. For example Brysbaert and Mitchell (2000:465) note that “For some as yet unexplained reason, grammatical gender information does not appear to play as rapid and efficient a role in guiding syntactic processing as might have been expected from the formal constraints such cues place on the structures of sentences.” The profile described by Brysbaert and Mitchell is precisely the type of effect expected using a weighted cue-combinatorics scheme in which structural cues are valued more strongly that non-structural morphological cues.

Lastly, my proposal for dynamic structural weighting is strengthened by the recent argument that comprehenders maintain and rationally update (un)certainty about the prior structure (e.g., Bicknell et al., 2014; Levy et al., 2009; Levy, 2008). For example, the comprehender may experience decreased uncertainty about the existing structural encoding in the case of an unfulfilled prediction, as in the case of an subject-verb agreement violations. This shift in certainty about the prior structure could cause the parser to shift priority from structural information to non-structural information in memory retrieval, resulting in increased susceptibility to attraction effects for subject-verb agreement.

4.9 Conclusion

The goal of Chapter 4 was to better understand the source of the contrasting attraction profiles observed in sentence comprehension. Based on the findings from Chapter 3, I explored two possibilities. The first possibility was that susceptibility to attraction effects is directly linked to specific features. That is, the contrasting attraction profiles are a consequence of the use of specific content features, like animacy, in memory retrieval. However, this proposal was puzzling, since certain features like number show contrasting profiles for subject-verb agreement and reflexive (see Dillon et al., 2013). A second possibility, which is really a variant of the first, suggests that probe-to-target similarity is the key determinant of attraction effects in comprehension.

I tested these possibilities focusing on the processing of reflexive anaphors using eye-tracking while reading and computational simulations. The data from reported in Chapter 4 revealed several key insights about the source and scope of attraction effects in comprehension. First, I extended previous investigations of reflexive anaphors to test a wider ranger of features (e.g., animacy) and feature combinations (e.g., 2-feature mismatch contexts) and a broader range of syntactic contexts (e.g., proactive and retroactive interference paradigms). I found that for reflexive anaphor resolution, a phenomenon that has resisted non-structural attraction effects in previous studies, we can systematically control where attraction effects do and do not occur. Specifically, my results show that reflexives are selectively susceptible to attraction effects: when the reflexive mismatches the true subject in just one feature (e.g., number or gender), there are no attraction effects, but when it mismatches in two features (e.g., gender + number), strong attraction effects are found, comparable in size to agreement attraction. Furthermore, I showed based on computational simulations that is possible to derive both the presence and absence of the selective attraction effects from within the same memory architecture.

Specifically, ACT-R simulations suggest that the selective attraction effects for reflexive anaphors might reflect a weighted cue-combinatorics scheme in which structural cues are weighted more strongly in retrieval than morphological cues. Lastly, I proposed an account of why it is harder to obtain attraction effects in anaphor resolution than in subject-verb agreement, based on whether memory retrieval for a licensor is triggered by error-correction or normal resolution, and I provided supporting reading time evidence from Experiment 2 and 3. In particular, I argued that whereas structural cues are prioritized in retrieval in the normal course of processing unpredictable dependencies, like anaphora and thematic binding, they are not prioritized when retrieval is triggered by an unfulfilled prediction, as in the case of agreement violations, leading the parser to be uncertain about the accuracy of its existing structural encoding.

In sum, the findings from Chapter 4, taken together with recent findings from studies by Van Dyke and McElree (2011) and Kush and colleagues (2013; submitted), motivate a shift from dependency-wise differences in attraction effects to the micro-structure of how different sources of information combine to access information in structured representations.

Chapter 5 Negative Polarity Items (NPIs)

5.1 Introduction

In Chapters 3 and 4, I showed that we can systematically "turn on" linguistic illusions for anaphor resolution, a phenomenon that has resisted illusions in the past, revealing valuable insights into the memory architecture for the parser. In this chapter, I show the opposite. In particular, I show that we can "turn off" linguistic illusions for negative polarity item (NPI) licensing, a phenomenon that is highly susceptible to illusions, revealing equally important insights into how we encode and access linguistic information in memory.

Specifically, I present the results from eight experiments involving untimed, off-line judgments, speeded-acceptability judgments, self-paced reading, and computational simulations which show that illusory licensing effects for NPIs show a fleeting time profile, such that it is categorically present or absent depending on the timing of when the NPI is introduced in the sentence, relative to a licensor. I also tested whether the contrasts observed for NPIs, extends to subject-verb agreement. Subject-verb agreement and NPIs depend on different licensing mechanisms, but show qualitatively similar profiles with respect to illusory licensing (this difference was discussed in detail in Chapter 2). Interestingly, the same timing profile observed for NPI illusions does not extend to agreement attraction, shedding new light on source and scope of illusory licensing effects in comprehension.

I argue that the contrast in illusory NPI licensing effect is a consequence of changes over time in the encoding or interpretation of the licensing context. Specifically, I suggest two possibilities: (i) the initial interpretation of the licensing context is accurately and rapidly built and remains constant over time, but the format of its encoding changes over time, or (ii) the interpretation of the licensing context changes over time. According to this proposal, the fleeting NPI illusion reflects access to the internal stages of encoding and interpretative processes.

Importantly, these findings provide additional evidence that the scope of linguistic illusions in sentence comprehension is more limited than previously assumed and the profile of selective fallibility is not a consequence of dependency-wise differences in memory access mechanisms. That is, attraction effects are not directly linked to specific dependencies or rather specific access mechanisms are not directly linked to specific dependencies. Furthermore, the finding that NPI and subject-verb agreement show different time profiles suggests that there is not a homogenous cause for linguistic illusions. Based on their alignment with distinct formal licensing mechanisms, I argue that variability in the ability of the parser to accurately implement linguistic constraints in real-time processing is, in part, a consequence of general cognitive mechanisms that distinguish the encoding and interpretation of emerging syntactic and semantic representations.

5.2 The present study

Previous research on linguistics illusions has highlighted the impact of the interval after the appearance of the dependent element (e.g., the NPI or agreeing verb), and the resulting theories such as those proposed by Vasishth et al. (2008), Xiang et al. (2009), Wagers et al., (2009), and Phillips et al., 2009 have emphasized that the sensitivity of the access mechanisms to the structural properties of a potential licensor varies as a function of the amount of time that a comprehender has had to process the dependent element, with more time yielding greater grammatical accuracy. Lewis and Phillips (in press) have recently suggested that this difference between fast (timed, online) and slow (untimed, offline) responses could reflect an improvement over time in the signal-to-noise ratio in the responses: by holding all components of the NPI licensing mechanism (i.e., the licensing conditions, the encoding of the context, and the access mechanisms) constant, while assuming that the cognitive architecture is noisy, the signal-to-noise ratio can improve over time with repeated access attempts. For example, an outcome of intrusive licensing that has a 25% probability of occurrence on a single access trial will have a substantially reduced probability of being the dominant outcome over the course of multiple access trials, leading to greater grammatical accuracy.[21]

These theories predict that variation in the time or position where the dependent element appears in the sentence should not strongly impact the illusion,

[21] In support of this proposal, I provide computational simulations in Appendix A for an uncontroversial case involving subject-verb agreement attraction. One notable achievement of these simulations is that it predicts improvement over time with iterative retrievals, all while maintaining a constant representation of the sentence.

given the assumption that the components of the licensing mechanism (i.e., the licensing conditions, the encoding of the context, and the access mechanisms) are stable over time. In order to better understand the source and scope of illusions in comprehension, I test this prediction in the present study. Specifically, I focus on a different part of the time course of dependency resolution for NPIs and subject-verb agreement, namely the time that elapses between the potential licensors and the introduction of the dependent element.

5.3 The limitations of previous studies

Linguistic illusions have provided valuable insights into how linguistic representations are encoded and navigated mentally. However, rather general conclusions about the encoding and access mechanisms have been drawn from a narrow range of findings. For example, while existing evidence shows that illusory licensing is robust across multiple tasks and languages, all previous demonstrations of illusory NPI licensing involve the NPI *ever* and a configuration where the intrusive licensor is the subject or object of a subject-modifying relative clause (e.g., (1c) and (1c)). Existing accounts such as those proposed by Vasishth et al. (2008) and Xiang et al. (2009) predict that all other NPIs should behave similarly with respect to illusory licensing. In the first set of experiments, I directly compared the NPIs *ever* and *any*. The prototypical NPI *any* is similar to the NPI *ever*, in that it may be directly licensed in DE contexts, as shown in (1).

(1) a. Nobody has eaten any of the cake that Mary made.
b. No student wants any homework over the holiday.

The NPI *any* has not featured prominently in previous research on the processing of NPIs due to the confound of the so-called *free-choice* interpretation. Free-choice *any* is licensed in non-DE contexts that invite a choice among a set of alternatives. However, unlike its NPI counterpart *ever*, free-choice *any* is not polarity sensitive and it is acceptable without negation as shown in (2) (see Giannakidou, 2001, for discussion). Consequentially, *any* is less amenable to tests of illusory NPI licensing that rely on a negative intrusive licensor.

(2) a. Sally will marry any doctor.
b. Pick a card, any card.
c. Students can purchase any book online.

It may be possible, however, to force polarity sensitivity for *any* in contexts that do not provide the kind of alternatives needed for the free-choice interpretation. For example, the lack of a clear set of alternatives for the abstract mass noun *satisfaction* in (3a) together with the episodocity of the entire proposition renders the free-choice interpretation of *any* infelicitous.[22] The negative polarity interpretation, by contrast, is readily accessible when *any* appears in the scope of negation, as in (3b). If comprehenders reliably interpret *any* as a negative polarity item in contexts like (3b), then it may provide a new context to test the scope of illusory NPI licensing.

[22] If there were a modal verb instead of a simple preterite, the free-choice interpretation of *any* would be acceptable.

(3) a. ??The criminals felt any satisfaction from the crime.
b. No criminals felt any satisfaction from the crime.

Although semantic analyses of NPIs (e.g., Chierchia, 2006; Giannakidou, 2011; Kadmon & Landman, 1993) treat *any* and *ever* similarly in licensing environments like (8a), these two items could nevertheless show different profiles in on-line comprehension. For instance, *any* and *ever* must appear in different positions of the sentence due to differences in syntactic category (e.g., *any* is a determiner whereas *ever* is an adverb). This unavoidable difference consequently varies the distance from the licensor, and it is possible that well-known effects of distance, such as time or decay could contribute to more or fewer errors of illusory licensing for *any* compared to *ever* (e.g., Altmann & Gray, 2002; Anderson & Schooler, 1991; Anderson & Lebiere, 1998; Anderson et al., 2004; Anderson, 2007; Berman, Jonides, & Lewis, 2009; Gibson, 2000; Jonides et al., 2008; R. L. Lewis & Vasishth, 2005; McKone, 1995; McKone, 1998; Van Dyke & Lewis, 2003; but cf. Lewandowsky & Oberauer, 2008; R. L. Lewis & Badecker, 2010)

5.2 Experiment 1

In Experiment 1, I compared sentences with the items *ever* and *any* using untimed, off-line acceptability ratings. The goal of this experiment was to determine whether comprehenders would reliably interpret *any* as a negative polarity item, rather than a free-choice item, when it occurred with an abstract mass term that does not support the kind of alternatives needed for the free-choice interpretation. If comprehenders interpret *any* as a negative polarity item, then we should expect to see rejection of sentences that lack a licensor. By contrast, if comprehenders interpret *any* as a free-choice item, then we should expect to see either no difference between sentences with *any*, or at least increased acceptability for sentences with an intrusive licensor and no licensor relative to their *ever* counterparts. Sentences with the NPI *ever* provided baseline conditions against which comprehenders' sensitivity to the negative polarity interpretation was measured.

5.2.1 Participants

Participants were 24 native speakers of English who were recruited using Amazon's Mechanical Turk crowdsourcing web-service (https://aws.amazon.com/mturk). All participants in this and the following experiments provided informed consent, and they were required to pass a short English proficiency test in order to participate in the experiment for payment. The proficiency test probed various constraints on tense, aspects, modality, morphology, ellipsis, and syntactic islands. Participants in Experiment 1 were compensated $2.50. The experiment lasted approximately 15 minutes.

5.2.2 Materials

Experimental materials consisted of 36 sets of 6 items, which varied in terms of the presence and structural location of an NPI licensor (grammatical licensor /

intrusive licensor / no licensor), and the type of NPI (*ever* / *any*). All items contained a subject NP that was modified by an object relative clause. This subject NP was always followed by a main clause predicate that contained the NPI. The grammatical licensor condition had the NPI licensor *no* as the determiner of the main clause subject NP. The intrusive licensor condition had the same licensor as the determiner of the relative clause subject NP. In the no licensor condition, the NPI licensor was replaced with the definite determiner *the*, which fails to license NPIs. The relative clause was always followed by the auxiliary *have*, which served to clearly mark the right edge of the relative clause. This demarcation was included to ensure that participants would correctly construct a parse in which the intrusive licensor did not c-command the NPI. The NPI *ever* always appeared immediately following the auxiliary. The NPI *any* was always interpolated between the main clause verb to its left and an abstract mass noun to its right. An example set of items is given in Table 5.1, showing all 6 conditions.

Table 5.1: Sample set of items for Experiment 1. NPIs and NPI licensors are in bold.

ever	GRAMMATICAL LICENSOR	No authors [that the critics recommended] have ever received acknowledgment for a best-selling novel.
	INTRUSIVE LICENSOR	The authors [that **no** critics recommended] have **ever** received acknowledgment for a best-selling novel.
	NO LICENSOR	The authors [that the critics recommended] have **ever** received acknowledgment for a best-selling novel.
any	GRAMMATICAL LICENSOR	**No** authors [that the critics recommended] have received **any** acknowledgment for a best-selling novel.
	INTRUSIVE LICENSOR	The authors [that **no** critics recommended] have received **any** acknowledgment for a best-selling novel.
	NO LICENSOR	The authors [that the critics recommended] have received **any** acknowledgment for a best-selling novel.

Each participant rated 108 sentences, consisting of 36 NPI sentences and 72 filler sentences. The 36 sets of NPI items were distributed across 6 lists in a Latin Square design. The filler sentences were of similar length and complexity to the NPI sentences, but lacked an NPI. Roughly half of the filler sentences used determiners similar to those used in the NPI sentences in similar positions to prevent the possibility that participants might develop superficial reading strategies based on the distribution of the determiners in the NPI sentences. Materials were balanced so that across the experiment half of the sentences were ungrammatical. The anomalies in the filler sentences comprised a variety of grammatical violations, including agreement errors, pronoun gender violations, and unlicensed verbal morphology.

5.2.3 Procedure

Sentences were presented using the web-based Ibex presentation software (Alex Drummond, http://spellout.net/ibexfarm/). Participants were instructed to rate the acceptability of each sentence using a 7-point scale, with '7' being the most acceptable, and '1' the least acceptable. Participants could take as much time as needed to rate each sentence, so long as they finished the experiment within the 30 minute time restriction imposed by the Mechanical Turk session. Each sentence was displayed in its entirety on the screen along with the rating scale. Participants could click boxes to enter their rating, or use a numerical keypad. The order of presentation was randomized for each participant. Eight practice items were presented before the beginning of the experiment.

5.2.4 Data analysis

Data were analyzed using linear mixed-effects models, with fixed factors for experimental manipulations and random effects for participant and item. I estimated models using the *lme4* package (Bates et al., 2011) in the R software environment (R Development Core Team, 2014). I used helmert contrast coding for experimental fixed effects, as shown in Table 5.2, and a fully-specified random effects structure, which included random intercepts and slopes for all fixed effects by participants and by items (Baayen et al., 2008; Barr et al., 2013). If there was a convergence failure, or if the model converged but the correlation estimates were high (e.g., > .9), the random effects structure was simplified following Baayen et al. (2008). An effect was considered significant if the absolute *t*/*z*-value was greater than 2 (Gelman & Hill, 2007).

Table 5.2: Helmert contrast coding for experimental factors.

	ever			*any*		
	GRAMMATICAL LICENSOR	INTRUSIVE LICENSOR	NO LICENSOR	GRAMMATICAL LICENSOR	INTRUSIVE LICENSOR	NO LICENSOR
C1	-2	1	1	-	-	-
C2	-	-1	1	-	-	-
C3	-	-	-	-2	1	1
C4	-	-	-	-	-1	1

5.2.5 Results

The raw results of the off-line acceptability judgment study are presented in Table 5.3 and the statistical analysis is presented in Table 5.4. A significant main effect of grammaticality was observed for both NPIs, as ratings were significantly higher for sentences with a grammatical licensor relative to sentences with an intrusive licensor or no licensor (contrasts C1 and C3). No significant differences were found within the ungrammatical conditions for either NPI (contrasts C2 and C4).

Table 5.3: Mean acceptability ratings and standard error by participant for Experiment 1. Values are on a 7-point scale, with '7' being the most acceptable, and '1' the least acceptable.

NPI	GRAMMATICAL LICENSOR	INTRUSIVE LICENSOR	NO LICENSOR
ever	4.68 (±0.20)	2.02 (±0.13)	1.93 (±0.13)
any	4.46 (±0.19)	1.83 (±0.11)	1.74 (±0.10)

Table 5.4: The four orthogonal contrasts for experimental factors (Experiment 1). Contrast 1 (C1) is the effect of grammaticality for the NPI *ever*. Contrast 2 (C2) is the effect of attraction for the NPI *ever*. Contrast 3 (C3) is the effect of grammaticality for the NPI *any*. Contrast 4 (C4) is the effect of attraction for the NPI *any*. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

COMPARISON	$\hat{\beta}$	SE	t
C1	**2.64**	**0.30**	**8.73**
C2	0.09	0.20	0.47
C3	**2.65**	**0.28**	**9.18**
C4	0.08	0.17	0.48

5.2.6 Discussion

Untimed acceptability ratings from Experiment 1 revealed similar profiles for *ever* and *any*. Sensitivity to a licensor in a grammatically appropriate position, taken together with the failure to find a difference between the intrusive licensor and no licensor sentences for *any* is consistent with an NPI interpretation for *any*. A free-choice interpretation, by contrast, would predict either no differences between the sentences with *any*. However, this was not the case, as we observed a clear main effect of grammaticality expected under an NPI interpretation. For both items, participants showed robust sensitivity to the structural licensing conditions on NPIs, as they reliably rated sentences with a grammatical licensor higher than sentences with an intrusive licensor and no licensor. Furthermore, we failed to find evidence that the presence of a semantically compatible, but structurally illicit intrusive licensor improved the off-line acceptability of an illicit NPI.

5.3 Experiment 2

Experiment 2 used speeded-acceptability judgments to directly compare the NPIs *ever* and *any*. The main aim of this experiment was to test whether the illusory licensing effect that is commonly observed for the NPI *ever* would extend to the NPI *any*. Previous studies on the processing of NPIs have shown that comprehenders are most susceptible to illusory NPI licensing in on-line, time sensitive measures. The speeded-acceptability judgment task used in Experiment 2 has been shown to reliably elicit the illusory NPI licensing effect by restricting the amount of time that comprehenders have to reflect on grammatical intuitions (e.g., Drenhaus et al., 2005; Xiang et al., 2006).

5.3.1 Participants

Participants were 18 native speakers of English from the University of Maryland community. Participants were either compensated $10 or received credit in an introductory linguistics course. All participants were naïve to the purpose of the experiment. The speeded-acceptability task lasted approximately 20 minutes, and was administered as a part of a one-hour session involving unrelated experiments.

5.3.2 Materials

Experiment materials consisted of the same 36 sets of 6 items as in Experiment 1, with the same filler sentences

5.3.3 Procedure

Sentences were presented on a desktop PC using the Ibex presentation software. Sentences were presented one word at a time in the center of the screen in a rapid serial visual presentation (RSVP) paradigm (Potter, 1988) at a rate of 300 milliseconds per word. A response screen appeared for 3 seconds at the end of each sentence during which participants made a 'yes/no' response by button press. Participants were instructed to read each sentence carefully, and to judge each sentence according to its acceptability in colloquial speech. If participants waited longer than 3 seconds to respond, they were given feedback that their response was too slow. The order of presentation was randomized for each participant. Eight practice items were presented before the beginning of the experiment.

5.3.4 Data analysis

Data were analyzed using logistic mixed-effects models, since the dependent measure was categorical (i.e., 'yes' or 'no'). I used the helmert contrast coding in Table 5.2 for experimental fixed effects, and a fully-specified random effects structure, which included random intercepts and slopes for all fixed effects by participants and by items (Baayen et al., 2008; Barr et al., 2013). If there was a convergence failure, or if the model converged but the correlation estimates were high (e.g., > .9), the random effects structure was simplified following Baayen et al. (2008).

5.3.5 Results

Figure 5.1 shows the proportion of 'yes' responses for the six experimental conditions in Experiment 2. The results of the statistical analysis are presented in Table 5.5. Results showed reliable detection of an illicit NPI, as grammatical sentences were more likely to be accepted than their grammatical counterparts (contrasts C1 and C3). But results for the ungrammatical conditions sharply diverged. Contrasting illusory licensing profiles were observed for *ever* and *any*, as reflected by an interaction between attraction and NPI (contrast C2 vs. C4). This interaction was driven by a significant illusory licensing effect for *ever* (contrast C2), as participants were more likely to accept ungrammatical sentences with an intrusive licensor than ungrammatical sentences with no licensor. We failed to find a corresponding effect for *any*. Figure 5.2 presents a by-participant plot of illusory licensing effects for *ever* and *any*, illustrating a clear difference in judgments for the two NPIs.

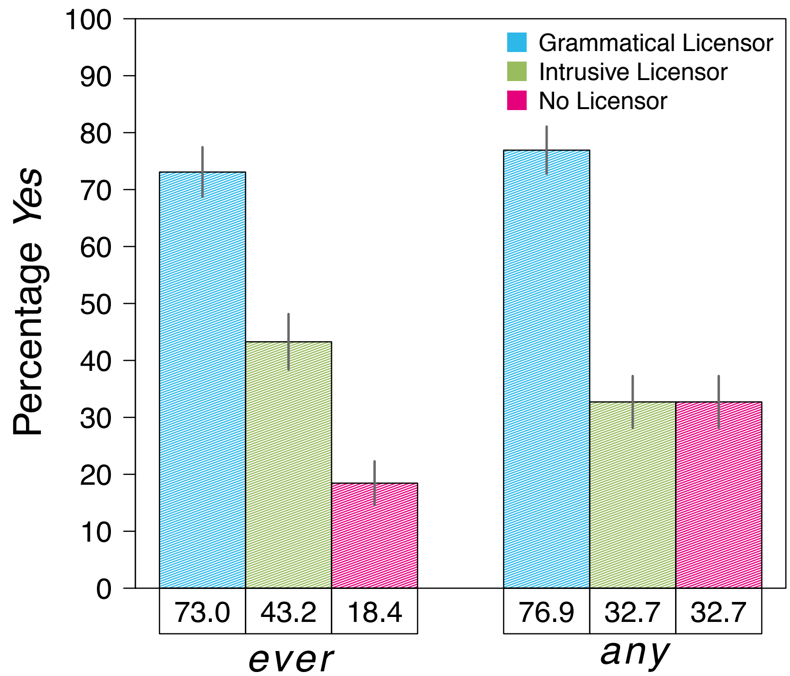


Figure 5.1: Speeded-acceptability judgments (Experiment 2). Mean percentage 'yes' responses for sentences with the NPIs *ever* (left) and *any* (right). Error bars indicate standard error of the mean.

Table 5.5: The four orthogonal contrasts for experimental factors (Experiment 2). Contrast 1 (C1) is the effect of grammaticality for the NPI *ever*. Contrast 2 (C2) is the effect of attraction for the NPI *ever*. Contrast 3 (C3) is the effect of grammaticality for the NPI *any*. Contrast 4 (C4) is the effect of attraction for the NPI *any*. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

COMPARISON	$\hat{\beta}$	SE	z
C1	**2.69**	**0.62**	**4.3**
C2	**1.57**	**0.39**	**4.01**
C3	**3.04**	**0.72**	**4.18**
C4	0.02	0.37	0.06
C2 vs. C4	**-1.83**	**0.62**	**-2.91**

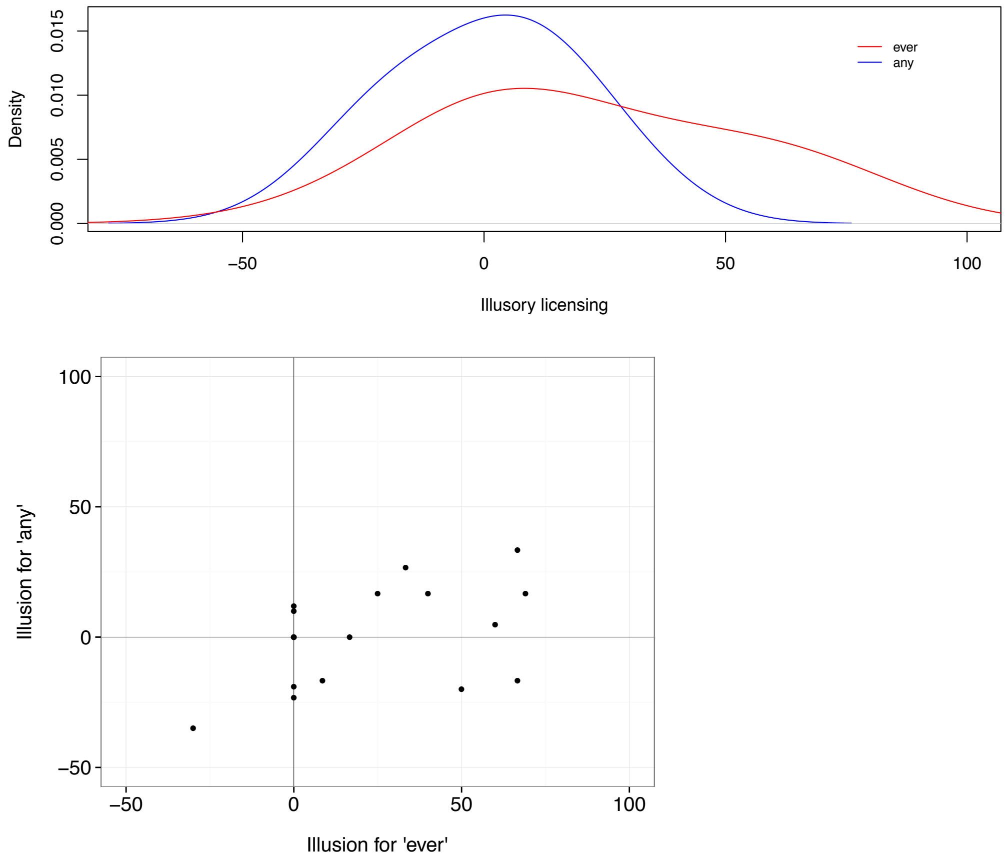


Figure 5.2: Top: Distributions of attraction effects by participant for *ever* (red) and *any* (blue) (Experiment 2). Bottom: Attraction effects by participant for *ever* and *any* by participant. Attraction is positive-going, reflecting increased acceptability.

5.3.6 Discussion

In Experiment 2, I investigated the scope of the illusory NPI licensing effect in a direct comparison of the two NPIs *ever* and *any*. Speeded-acceptability judgments revealed contrasting illusory licensing profiles. The NPI *ever* showed a clear illusory licensing effect, as comprehenders were more likely to accept ungrammatical sentences with *ever* when an intrusive licensor was present. The profile observed for *ever* replicates previous findings of illusory NPI licensed noted in Drenhaus et al. (2005) and Xiang et al. (2006, 2009). By contrast, I failed to an attraction effect for the NPI *any*, as the presence of an intrusive licensor did not increase rates of acceptance of sentences that contained an illicit NPI.

However, one potential concern with the results of Experiment 2 is the high rate of acceptance of ungrammatical sentences with *any*. It is possible, for example, that an illusory licensing effect for sentences involving the NPI *any* was masked by the increased rate of acceptance of sentences with no licensor. One factor that may differentially impact judgments is the free-choice interpretation of *any*. Although the off-line judgments from Experiment 1 suggests that comprehenders ultimately favor the negative polarity interpretation of *any* in configurations such as those used in Experiment 2, the mechanisms that are recruited for incremental processing may have nevertheless accessed the free-choice interpretation before the restricting abstract mass noun was encountered in the input sequence. It is possible, then, that in some portion of trials, some residue of this temporary free-choice interpretation could have influenced end-of-sentence judgments in an attempt to make sense of an ungrammatical sentence. In Experiment 3, I investigated this possibility by examining the time course of the contrast in profiles observed for *ever* and *any*.

5.4 Experiment 3

To investigate the time-course of the contrasting profiles observed for *ever* and *any*, I conducted a self-paced reading experiment using the same items from Experiments 1 and 2. In self-paced reading measures, illusory licensing effects are predicted at the NPI as facilitated reading times for ungrammatical sentences with an intrusive licensor, relative to ungrammatical sentences with no licensor. The absence of an illusion, by contrast, is predicted as only a main effect of grammaticality, with no divergence in reading times for the two types of ungrammatical sentences.

5.4.1 Participants

Participants were 24 native speakers of English from the University of Maryland community. Participants were either compensated $10 or received credit in an introductory Linguistics course. All participants were naïve to the purpose of the experiment. The self-paced reading task lasted approximately 35 minutes, and was administered as a part of a one-hour session involving unrelated experiments.

5.4.2 Materials

Experimental materials consisted of the same 36 sets of 6 items as in Experiments 1 and 2. 72 grammatical fillers were also included, such that each participant read a total of 108 sentences. The filler sentences were of similar length and complexity to the NPI sentences, but lacked an NPI. Roughly half of the filler sentences used determiners similar to those in the NPI conditions in similar positions to prevent the possibility that participants might develop superficial reading strategies based on the distribution of the determiners in the NPI sentences. Each sentence was followed by a comprehension question. Comprehension questions addressed various parts of the sentence to prevent the possibility that participants might develop superficial reading strategies whereby they extracted only the information necessary to answer the comprehension question without fully interpreting the sentence.

5.4.3 Procedure

Sentences were presented on a desktop PC in a moving-window self-paced reading display using the Linger software package (Doug Rohde, MIT). The increased likelihood of mis-retrieval, MIT). Sentences were initially masked by dashes, with white spaces and punctuation intact. Participants pushed the space bar to reveal each word. Presentation was non-cumulative, such that the previous word was replaced with a dash when the next word appeared on the screen. Each sentence was followed by a 'yes/no' comprehension question. Comprehension questions addressed various parts of the sentence in order to prevent participants from developing superficial reading strategies. For ungrammatical sentences, comprehension questions addressed only content prior to the NPI. Onscreen feedback was provided for incorrect answers. Participants were told to read more carefully if they answered multiple questions incorrectly. The order of presentation was randomized for each participant. Eight practice items were presented before the beginning of the experiment.

5.4.4 Data analysis

Self-paced reading times for experimental sentences were examined region-by-region. The regions used for analysis consisted of single words. I report statistical analyses for three regions of interest, including the region immediately before the

NPI (pre-critical region), the NPI region (critical region), and the region immediately after the NPI (spill-over region). I used the Box-Cox procedure (Box & Cox, 1964) to determine that a natural log would be the appropriate transformation to obtain normally distributed residuals (box-cox results for the critical regions: λ=-0.3) (see also Vasishth et al., 2013, for further discussion about the importance of appropriately transforming reading time data in psycholinguistic research). Following the box-cox procedure, statistical analyses were carried out with the untrimmed, log-transformed reading time data using linear mixed-effects models. I used helmert contrast coding for experimental fixed effects, as shown in Table 5.2, and a fully-specified random effects structure, which included random intercepts and slopes for all fixed effects by participants and by items (Baayen et al, 2008; Barr et al., 2013 If there was a convergence failure, or if the model converged but the correlation estimates were high (e.g., > .9), the random effects structure was simplified following Baayen et al. (2008). One participant was excluded from analysis due to accuracy below 80%.

5.4.5 Results

Comprehension question accuracy

The mean comprehension question accuracy was 94%, indicating that participants successfully comprehended the experimental materials.

Self-paced reading times

Figures 5.3 and 5.4 show the region-by-region condition means for sentences with the NPIs *ever* (Figure 5.3) and *any* (Figure 5.4). Results of the statistical analysis are presented in Tables 5.6 and 5.7.

Self-paced reading times: ever

No significant effects for any of the comparisons were observed in the pre-critical region. At the critical NPI region, there was a significant main effect of grammaticality, as reflected by a slowdown in ungrammatical conditions relative to the grammatical condition. This grammaticality effect persisted into the first and second spill-over regions. The critical NPI region and spill-over region 1 also showed a significant illusory licensing effect, due to facilitated reading times for ungrammatical sentences with an intrusive licensor relative to ungrammatical sentences with no licensor.

Self-paced reading times: any

No significant effects for any of the comparisons were observed in the pre-critical region. At the critical region, there was a significant main effect of grammaticality, as reflected by a slowdown in ungrammatical conditions relative to the grammatical condition. This effect persisted into the spill-over regions. Importantly, there was no reliable evidence for illusory licensing at any region.

Direct comparison of ever and any

A direct comparison of *ever* and *any* (contrast C2 vs. C4) revealed contrasting illusory licensing profiles for *ever* and *any*, as reflected by a significant interaction between NPI type and illusory licensing. This interaction was driven by a significant illusory licensing effect for *ever* (contrast C2). Figure 5.5 presents a by-participant plot of illusory licensing effects for *ever* and *any*, illustrating a clear difference in illusory licensing profiles for the two NPIs.

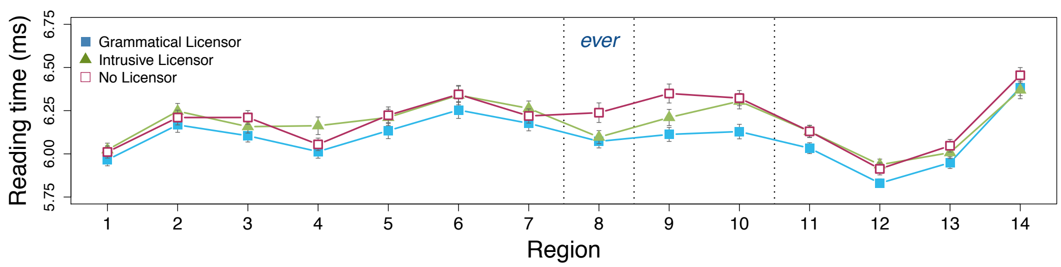


Figure 5.3: Self-paced reading results for sentences with the NPI *ever* (Experiment 3). Region-by-region means separated by the presence and location of a potential licensor. Error bars indicate standard error of the mean. Sample sentence: *No|The*$_1$ *authors*$_2$ *that*$_3$ *the|no*$_4$ *critics*$_5$ *recommended*$_6$ *have*$_7$ *ever*$_8$ *received*$_9$ *acknowledgment*$_{10}$ *for*$_{11}$ *a*$_{12}$ *best-selling*$_{13}$ *novel*$_{14}$.

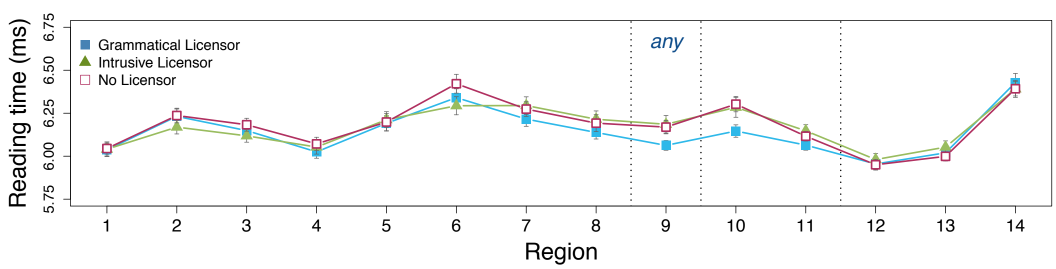


Figure 5.4: Self-paced reading results for sentences with the NPI *any* (Experiment 3). Region-by-region means separated by the presence and location of a potential licensor. Error bars indicate standard error of the mean. Sample sentence: *No|The*$_1$ *authors*$_2$ *that*$_3$ *the|no*$_4$ *critics*$_5$ *recommended*$_6$ *have*$_7$ *received*$_8$ *any*$_9$ *acknowledgment*$_{10}$ *for*$_{11}$ *a*$_{12}$ *best-selling*$_{13}$ *novel*$_{14}$.

Table 5.6: The four orthogonal contrasts for experimental factors (Experiment 3). Contrast 1 (C1) is the effect of grammaticality for the NPI *ever*. Contrast 2 (C2) is the effect of attraction for the NPI *ever*. Contrast 3 (C3) is the effect of grammaticality for the NPI *any*. Contrast 4 (C4) is the effect of attraction for the NPI *any*. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

	REGIONS											
	PRE-CRITICAL			CRITICAL			SPILL-OVER 1			SPILL-OVER 2		
	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t
C1	-0.06	0.04	-1.49	**-0.09**	**0.04**	**-2.07**	**-0.16**	**0.04**	**-3.40**	**-0.16**	**0.04**	**-3.57**
C2	0.03	0.05	0.69	**-0.14**	**0.05**	**-2.37**	**-0.13**	**0.06**	**-2.00**	-0.03	0.04	-0.75
C3	-0.06	0.05	-1.25	**-0.11**	**0.04**	**-2.81**	**-0.14**	**0.04**	**-3.10**	**-0.06**	**0.04**	**-2.04**
C4	0.02	0.05	0.44	-0.01	0.04	-0.35	-0.01	0.07	-0.25	0.03	0.03	0.81

Table 5.7: Interaction of NPI type × attraction (C2 vs. C4; NPI region) (Experiment 3).

	INTERACTION		
	$\hat{\beta}$	SE	*t*
C2 vs. C3	**-0.15**	**0.07**	**2.04**

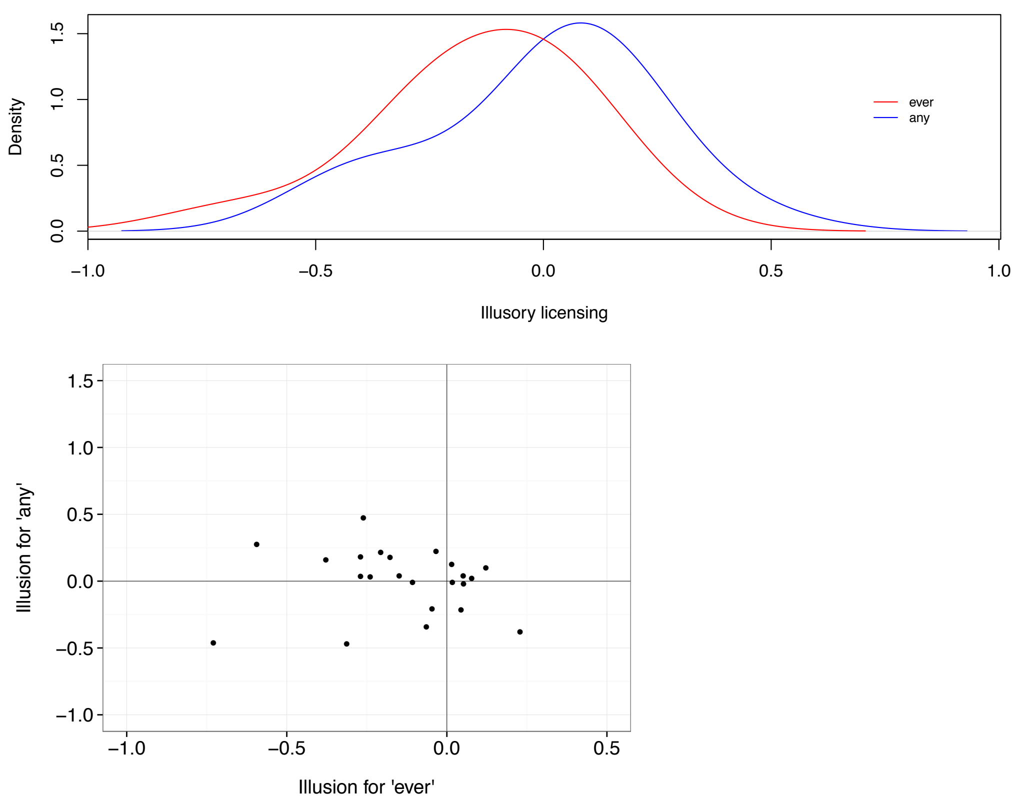


Figure 5.5: Distributions of attraction effects by participant for *ever* (red) and *any* (blue) (Experiment 3) over the NPI and first spill-over regions. Bottom: Attraction effects by participant for *ever* and *any*. Attraction is negative-going, reflecting facilitated processing.

5.4.6 Discussion

The most important finding from Experiment 3 is the replication of the contrasting illusory licensing profiles for *ever* and *any* that was observed in Experiment 2. This replication demonstrates that the contrast is robust across different measures. As in Experiment 2, the NPI *ever* showed a reliable illusory licensing effect, reflected in self-paced reading times as facilitated processing of an illicit NPI in the presence of an intrusive licensor. The NPI *any*, by contrast, failed to show an illusory licensing effect, as the presence of an intrusive licensor did not ease the processing of an illicit NPI.

Importantly, sentences with the item *any* showed rapid detection of an illicit NPI as reflected by increased reading times for ungrammatical sentence (both with and without an intrusive licensor) beginning at the NPI region, with the difference reaching significance at the immediate following region. This profile is not expected if comprehenders pursued a free-choice interpretation of *any*, and it suggests that the structural constraints on negative polarity licensing can impact real-time comprehension, at least selectively.

There are several possible explanations for the contrasting illusory licensing profiles observed for *ever* and *any*. One possibility is that the contrast in profiles reflect inherent properties of the two NPIs. The NPIs *ever* and *any* differ along several dimensions, including syntactic category and quantificational status. For example, *any* is a quantificational determiner whereas *ever* is an adverb. Although current leading theories of polarity phenomenon treat *ever* and *any* similarly in terms of their semantics and compositional properties (e.g., Chierchia, 2006; Giannakidou, 2011), it is possible that differences in syntactic category for example could contribute to differences in how the parser integrates these items. For example, while the attachment of a determiner like *any* is relatively straightforward, there could be some uncertainty about the attachment position of the adverb *ever* in the hierarchical representation of the sentence.

Another possibility is that the contrasting profiles reflect selective mis-identification or repair of *ever*, but not *any*. The NPI *ever* has a phonological and orthographic near neighbor, *never*. This near neighbor is semantically compatible with the sentence-initial context, and the parser could have accessed this near neighbor to generate a perfectly grammatical representation of an otherwise ungrammatical sentence. For instance, assuming a noisy-channel model of sentence comprehension (Gibson et al., 2013; Levy, 2008), comprehenders may have maintained some uncertainty about whether the input contained the NPI *ever*, or its near-neighbor *never*. It is possible, then, that in some portion of trials, comprehenders mis-identified *ever* as *never* in contexts that could not support a grammatical interpretation of the NPI. The NPI *any* was less likely to be mis-identified as one its near-neighbors, e.g., *many*, as this would result in semantic incompatibility with the abstract mass term, e.g., ?*many satisfaction*. However, I suggest that this account cannot capture the contrasting profiles for several reasons. First, this account does not predict that mis-identification should occur more in sentences with intrusive negation, which show more errors of illusory licensing relative to sentences that lack negation. Second, this account incorrectly predicts similar profiles for *ever* and *any,* and it incorrectly predicts immediate susceptibility to illusory licensing for *any*, since the restricting abstract mass term that would prevent re-analysis or phonological repair was not encountered until after the NPI. Self-paced reading times for *any*, however, showed resistance to illusory licensing immediately at the NPI.

A third possibility is that the contrasting profiles reflect differences in time or distance between *ever* and *any* with respect to their position and the parsing operations that precede these items. In the items that I tested, *ever* appeared pre-verbally, whereas *any* appeared post-verbally. Routine parsing operations associated with thematic binding at the main clause verb, for example, may have selectively reactivated the target main clause subject NP immediately prior to *any*, giving this NP an activation advantage that could reduce or eliminate attraction from an intrusive licensor NP. No such activation bias would be available for *ever*, since it appeared before the main clause verb. More generally, the contrasting profiles for *ever* and *any* could reflect differences in time or distance between the two items, as the NPI *ever* appeared earlier in the sentence than *any*. It is possible, for example, that well-known effects of time-based decay (Altmann & Gray, 2002; Anderson & Schooler, 1991; Anderson & Lebiere, 1998; Anderson et al., 2004; Anderson, 2007; Berman et al., 2009; Gibson, 2000; Jonides et al., 2008; R. L. Lewis & Vasishth, 2005; McKone, 1995; McKone, 1998; Van Dyke & Lewis, 2003; but cf. Lewandowsky & Oberauer, 2008; R. L. Lewis & Badecker, 2010), could have dampened the impact of the intrusive licensor for the NPI *any* that appeared later in time.

Alternatively, the position difference could have given rise to a kind of anti-locality effect (Häussler & Bader, 2012; Konieczny, 2000). For example, the increased distance and additional intervening material between the NPI and the main subject NP for sentences involving *any* could have narrowed down the range of possible continuations or it could have changed some component of the NPI licensing process, such as the licensing conditions, the encoding of the context, or the access mechanisms, leading to a heightened sensitivity to an unlicensed NPI. No such anti-locality advantage would be available for *ever* because it appeared in an earlier position relative to *any*. More specifically, it is possible that the NPI *any* in a later position could be more predictable given the left prefix context of the sentence. To address this possibility, I conducted a follow-up corpus study that investigated the conditional probabilities of *ever* vs. *any* based on the left prefix context of the experimental items from Experiments 1-3. However, this study was inconclusive because the data are too sparse.

5.5 Experiment 4

The next set of experiments were designed to distinguish between accounts of the contrasting illusory licensing profiles for *ever* and *any* that appeal to lexical differences and selective repair on the one hand, and an account that appeals to positional differences on the other hand. I achieved this by holding the NPI constant, testing only *ever*, and by manipulating whether the NPI appeared pre-verbally or post-verbally. If the contrasting illusory licensing profiles observed in Experiments 2 and 3 reflect inherent differences between *ever* and *any* or selective repair of *ever*, then we expect to see similar illusory licensing effects across all positions. If, however, the contrasting profiles reflect positional differences between pre-verbal *ever* and post-verbal *any,* then we expect to see selective susceptibility to illusory licensing, with illusory licensing effects only in pre-verbal positions. I used both speeded-acceptability judgments (Experiment 4) and self-paced reading measures (Experiment 5) to directly compare the processing of *ever* in pre-verbal and post-verbal positions.

5.5.1 Participants

Participants were 18 native speakers of English from the University of Maryland. Participants were either compensated $10 or received credit in an introductory Linguistics course. All participants were naïve to the purpose of the experiment. The task lasted approximately 20 minutes, and was administered as a part of a one-hour session involving unrelated experiments.

5.5.2 Materials

Experimental materials consisted of 36 sets of 6 items, which varied in terms of the presence and structural location of an NPI licensor (grammatical licensor / intrusive licensor / no licensor), and the position of the NPI (*pre-verbal* / *post-verbal*). As in Experiments 2 and 3, the NPI licensor *no* appeared either as the determiner of the main subject (grammatical licensor) or as the determiner of the relative clause subject (intrusive licensor), or was replaced with the definite determiner *the* (no licensor). The NPI *ever* appeared either immediately before the main clause verb (pre-verbal conditions), as in Experiments 2 and 3 or in a later position immediately after the auxiliary of an embedded sentential complement clause (post-verbal conditions). An example set of items is given in Table 5.8.

Table 5.8: Sample set of items for Experiment 3. NPIs and NPI licensors are in bold.

Pre-verbal ever	GRAMMATICAL LICENSOR	No journalists [that the editors recommended for the assignment] ever thought that the readers would understand the complicated situation.
	INTRUSIVE LICENSOR	The journalists [that **no** editors recommended for the assignment] **ever** thought that the readers would understand the complicated situation.
	NO LICENSOR	The journalists [that the editors recommended for the assignment] **ever** thought that the readers would understand the complicated situation.
Post-verbal ever	GRAMMATICAL LICENSOR	No journalists [that the editors recommended for the assignment] thought that the readers would **ever** understand the complicated situation.
	INTRUSIVE LICENSOR	The journalists [that **no** editors recommended for the assignment] thought that the readers would **ever** understand the complicated situation.
	NO LICENSOR	The journalists [that the editors recommended for the assignment] thought that the readers would **ever** understand the complicated situation.

Each participant read 108 sentences, consisting of 36 NPI sentences and 72 filler sentences. The 36 sets of NPI items were distributed across 6 lists in a Latin Square design. The filler sentences were of similar length and complexity to the NPI sentences, but lacked an NPI. Roughly half of the filler sentences used determiners similar to those used in the NPI sentences in similar positions to prevent the possibility that participants might develop superficial reading strategies based on the distribution of the determiners in the NPI sentences. Materials were balanced so that across the experiment half of the sentences were ungrammatical. The anomalies in the filler sentences comprised a variety of grammatical violations, including agreement errors, pronoun gender violations, and unlicensed verbal morphology.

5.5.3 Procedure

The same procedure was used as in Experiment 2.

5.5.4 Data analysis

Data analysis followed the same steps as in Experiment 2.

5.5.5 Results

Figure 5.6 shows the proportion of 'yes' responses for the six experimental conditions in Experiment 2. The results of the statistical analysis are presented in Table 5.9. Results showed reliable detection of an unlicensed NPI, as grammatical sentences were more likely to be accepted than their ungrammatical counterparts. But judgments for the ungrammatical sentences sharply diverged. Contrasting illusory licensing profiles were observed for pre-verbal and post-verbal *ever*, as reflected in an interaction between illusory licensing and NPI position within the ungrammatical conditions. Planned pairwise comparisons revealed that this interaction was driven by a significant illusory licensing effect for pre-verbal *ever*, as participants were more likely to accept ungrammatical sentences with an intrusive licensor than ungrammatical sentences with no licensor. No such effect was observed for post-verbal *ever*. Figure 5.7 presents a by-participant plot of illusory licensing effects for pre-verbal and post-verbal *ever*, illustrating a clear difference in illusory licensing profiles for the two positions.

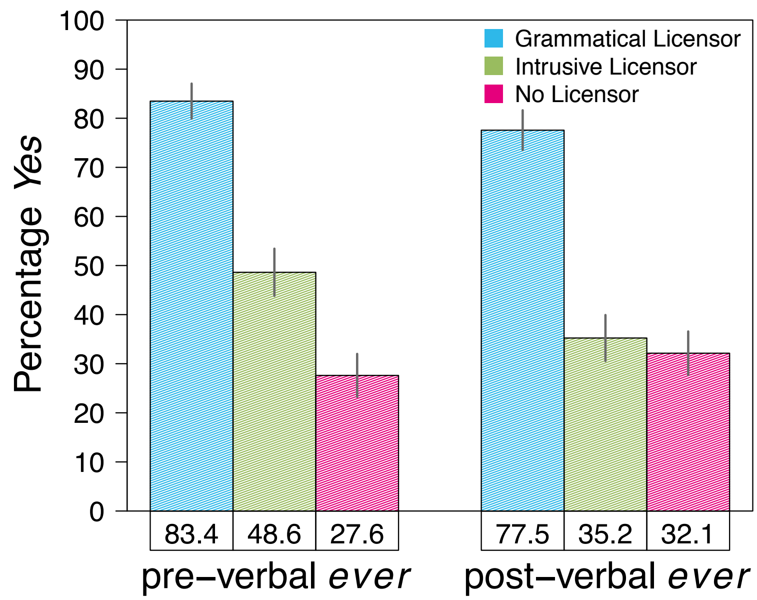


Figure 5.6: Speeded-acceptability judgments (Experiment 4). Mean percentage 'yes' responses for sentences with the NPI *ever* in a pre-verbal position (left) and a post-verbal position (right). Error bars indicate standard error of the mean.

Table 5.9: The four orthogonal contrasts for experimental factors (Experiment 2). Contrast 1 (C1) is the effect of grammaticality for sentences with a pre-verbal NPI. Contrast 2 (C2) is the effect of attraction for sentences with a pre-verbal NPI. Contrast 3 (C3) is the effect of grammaticality for sentences with post-verbal NPI. Contrast 4 (C4) is the effect of attraction for sentences with post-verbal NPI. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

COMPARISON	$\hat{\beta}$	SE	z
C1	**2.73**	**0.56**	**4.83**
C2	**1.14**	**0.32**	**3.58**
C3	**2.28**	**0.50**	**4.51**
C4	0.14	0.32	0.46
C2 vs. C4	**1.05**	**0.45**	**2.30**

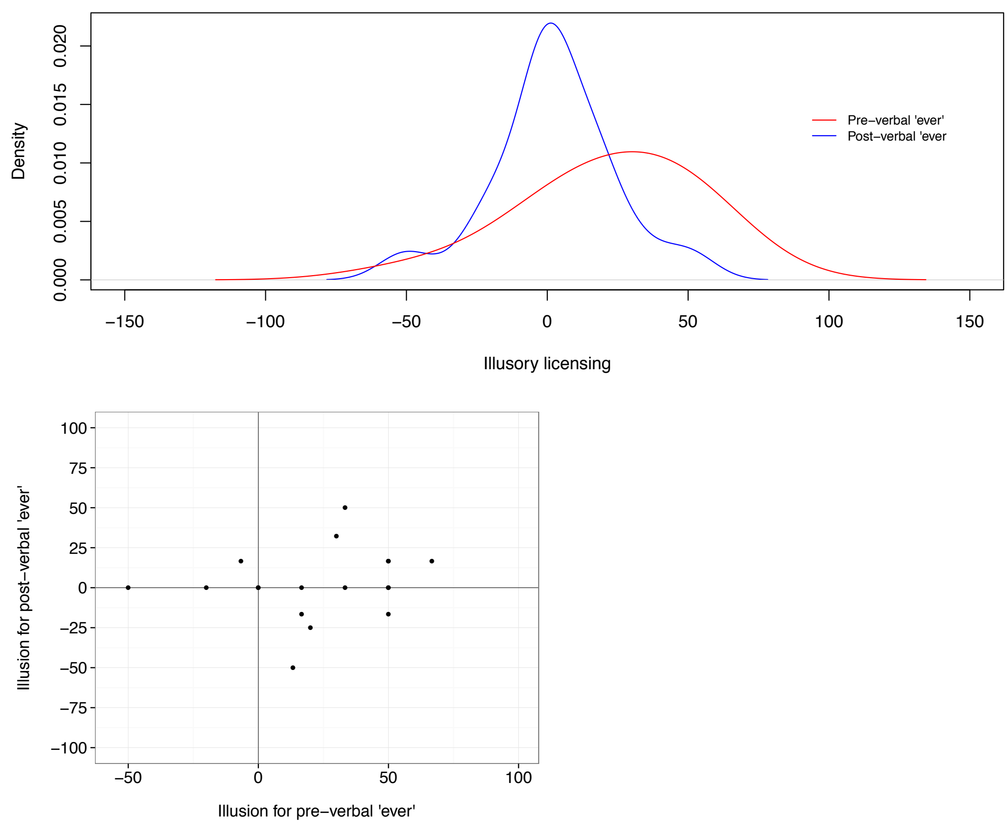


Figure 5.7: Distributions of attraction effects by participant for pre-verbal *ever* (red) and post-verbal *ever* (blue) (Experiment 4). Bottom: Attraction effects by participant for pre-verbal *ever* and post-verbal *ever* by participant. Attraction is positive-going, reflecting increased acceptability.

5.5.6 Discussion

Speeded-acceptability judgment results revealed the same modulation of illusions seen in the *ever*/*any* comparison, with contrasting illusory profiles for a single NPI *ever* in pre-verbal and post-verbal positions. Pre-verbal *ever* showed a reliable illusory licensing effect, as comprehenders were more likely to accept ungrammatical sentences when an intrusive licensor was present. This pattern replicates the illusory licensing effect observed for *ever* in Experiment 2, further demonstrating the robustness of the basic illusory licensing effect. Post-verbal *ever*, by contrast, showed no such illusory licensing effect, as the presence of an intrusive licensor did not reliably increase the rate of acceptance of sentences that contained an illicit post-verbal NPI.

The contrasting illusory licensing profiles observed for pre-verbal and post-verbal *ever* are consistent with the hypothesis that the contrasting illusory licensing profiles seen for the *ever*/*any* comparison reflect positional differences between the two NPIs, relative to the potential licensors. However, this requires further confirmation, and I return to a fuller discussion after Experiment 5.

5.6 Experiment 5

Experiment 5 used self-paced reading to examine the time-course of the contrasting illusory licensing profiles for pre-verbal and post-verbal *ever*. As in Experiment 3, illusory licensing effects would manifest as a main effect of grammaticality, accompanied by facilitated readings for ungrammatical sentences with an intrusive licensor relative to ungrammatical sentences with no licensor.

The absence of an illusion, by contrast, is predicted as only an effect of grammaticality, with no divergence in reading times for the ungrammatical sentences.

5.6.1 Participants

Participants were 30 native speakers of English from the University of Maryland. Participants were either compensated $10 or received credit in an introductory Linguistics course. All participants were naïve to the purpose of the experiment. The task lasted approximately 35 minutes, and was administered as a part of a one-hour session involving unrelated experiments.

5.6.2 Materials

Experimental materials consisted of the same 36 sets of 6 items as in Experiment 4. 72 grammatical fillers were also included, such that each participant read a total of 108 sentences. The filler sentences were of similar length and complexity to the NPI sentences, but lacked an NPI. Roughly half of the filler sentences used determiners similar to those in the NPI conditions, in similar positions, to prevent the possibility that participants might develop superficial reading strategies based on the distribution of determiners in the NPI sentences. Each sentence was followed by a 'yes/no' comprehension question. Comprehension questions addressed various parts of the sentence to prevent the possibility that participants might develop superficial reading strategies whereby they extract only the information necessary to answer the comprehension question without fully interpreting the sentence.

5.6.3 Procedure

The same procedure was used as in Experiment 3.

5.6.4 Data analysis

Data analysis followed the same steps as in Experiment 3. The box-cox procedure determined that a natural log was the appropriate transformation to obtain normally distributed residuals (average across critical region of interest: λ=-0.4).

5.6.6 Results

Comprehension question accuracy

The mean comprehension question accuracy was 91%, indicating that participants successfully comprehended the experimental materials.

Self-paced reading times

Figures 5.8 and 5.9 show the region-by-region condition means for sentences with a pre-verbal NPI *ever* (Figure 5.8) and a post-verbal NPI *ever* (Figure 5.9). Results of the statistical analysis are presented in Tables 5.10 and 5.11.

Self-paced reading times: pre-verbal ever

No significant effects for any of the comparisons were observed in the pre-critical and critical regions. At the first spill-over region, there was a significant main effect of grammaticality, as reflected by a slowdown in ungrammatical conditions relative to the grammatical condition. Spill-over region 1 also showed a significant illusory licensing effect, due to facilitated reading times for ungrammatical sentences with an intrusive licensor relative to ungrammatical sentences with no licensor. This illusory licensing effect persisted into the second spill-over region. No other significant effects were observed in spill-over region 2.

Self-paced reading times: post-verbal ever

No significant effects for any of the comparisons were observed in the pre-critical and critical regions. At the first spill-over region, there was a significant main effect of grammaticality, as reflected by a slowdown in ungrammatical conditions relative to the grammatical condition. This effect of grammaticality persisted into the second spill-over region. Importantly, there was no reliable evidence for illusory licensing at any region.

Direct comparison of ever and any

A direct comparison of *ever* and *any* (contrast C2 vs. C4) revealed contrasting illusory licensing profiles for pre-verbal and post-verbal *ever*, as reflected by a significant interaction between NPI type and illusory licensing. This interaction was driven by a significant illusory licensing effect for *ever* (contrast C2). Figure 5.10 presents a by-participant plot of illusory licensing effects for pre-verbal and post-verbal *ever*, illustrating a clear difference in illusory licensing profiles for the two NPIs.

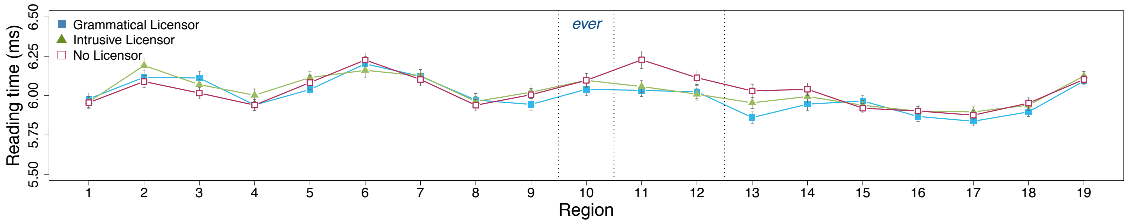


Figure 5.8: Self-paced reading results for sentences with pre-verbal *ever* (Experiment 5). Region-by-region means separated by the presence and location of a potential licensor. Error bars indicate standard error of the mean. Sample sentence: *No|The$_{1}$ authors$_{2}$ that$_{3}$ the|no$_{4}$ editors$_{5}$ recommended$_{6}$ for$_{7}$ the$_{8}$ assignment$_{9}$ ever$_{10}$ thought$_{11}$ that$_{12}$ the$_{13}$ readers$_{14}$ would$_{15}$ understand$_{16}$ the$_{17}$ complicated$_{18}$ situation$_{19}$.*

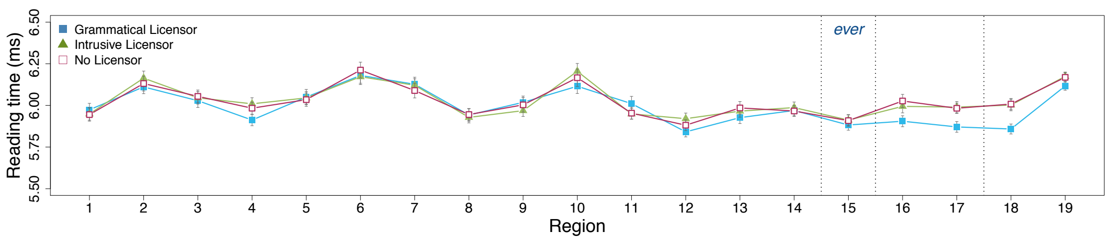


Figure 5.9: Self-paced reading results for sentences with post-verbal *ever* (Experiment 5). Region-by-region means separated by the presence and location of a potential licensor. Error bars indicate standard error of the mean. Sample sentence: *No|The$_{1}$ authors$_{2}$ that$_{3}$ the|no$_{4}$ editors$_{5}$ recommended$_{6}$ for$_{7}$ the$_{8}$ assignment$_{9}$ thought$_{10}$ that$_{11}$ the$_{12}$ readers$_{13}$ would$_{14}$ ever$_{15}$ understand$_{16}$ the$_{17}$ complicated$_{18}$ situation$_{19}$.*

Table 5.10: The four orthogonal contrasts for experimental factors (Experiment 2). Contrast 1 (C1) is the effect of grammaticality for sentences with a pre-verbal NPI. Contrast 2 (C2) is the effect of attraction for sentences with a pre-verbal NPI. Contrast 3 (C3) is the effect of grammaticality for sentences with post-verbal NPI. Contrast 4 (C4) is the effect of attraction for sentences with post-verbal NPI. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

	REGIONS											
	PRE-CRITICAL			CRITICAL			SPILL-OVER 1			SPILL-OVER 2		
	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t	$\hat{\beta}$	SE	t
C1	-0.06	0.04	-1.32	-0.05	0.04	-1.17	**-0.10**	**0.04**	**-2.50**	-0.03	0.05	-0.74
C2	0.00	0.05	-0.08	0.00	0.05	-0.08	**-0.16**	**0.04**	**-3.41**	**-0.10**	**0.04**	**-2.26**
C3	0.00	0.03	-0.17	-0.02	0.02	-0.94	**-0.10**	**0.03**	**-3.01**	**-0.11**	**0.03**	**-3.40**
C4	0.02	0.04	0.47	0.00	0.03	0.09	-0.03	0.04	-0.81	0.00	0.03	0.21

Table 5.11: Interaction of NPI type × attraction (C2 vs. C3; Regions = spill-over 1 and 2) (Experiment 5).

	INTERACTION		
	$\hat{\beta}$	SE	t
C2 vs. C3	**-0.12**	**0.05**	**-2.14**

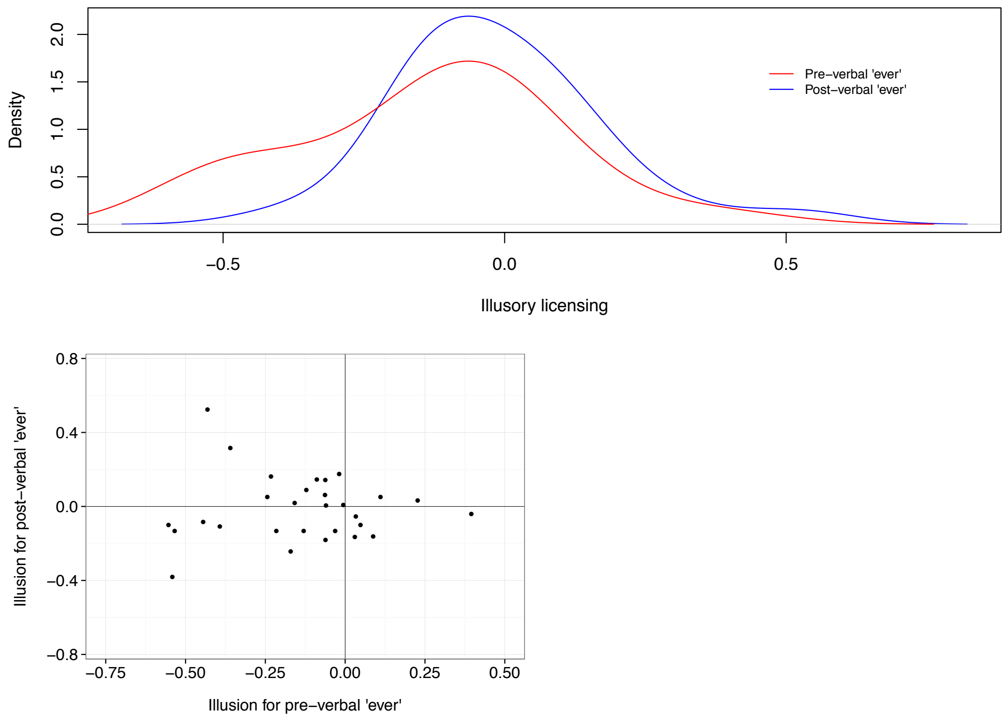


Figure 5.10: Distributions of attraction effects by participant for pre-verbal *ever* (red) and post-verbal *ever* (blue) (Experiment 4). Bottom: Attraction effects by participant for pre-verbal *ever* and post-verbal *ever* by participant. Attraction is negative-going, reflecting facilitated processing.

5.6.7 Discussion

As in Experiment 4, pre-verbal *ever* showed a reliable illusory licensing effect, reflected in self-paced reading times as facilitated processing of an illicit NPI in the presence of an intrusive licensor. Post-verbal *ever*, by contrast, showed no such illusory licensing effect, as the presence of an intrusive licensor did not ease the processing of an illicit NPI. These results replicate the contrasting illusory licensing profiles for pre-verbal and post-verbal *ever* observed in Experiment 5, demonstrating that the contrast is robust across methodologies. More generally, these results provide further evidence that the illusory licensing effect is much more limited than expected under current theories of linguistic illusions.

One aim of Experiments 4 and 5 was to better understand the source of the contrasting illusory licensing profiles observed for *ever* and *any* in Experiments 2 and 3. I suggested several possible accounts for this contrast. Experiments 4 and 5 were designed to distinguish between accounts of the contrasting illusory licensing profiles that appeal to inherent lexical differences between *ever* and *any* and selective repair of *ever* on the one hand, and an account that appeals to positional differences between the two NPIs on the other hand. The data from Experiments 4 and 5 show that even when holding the lexical NPI *ever* constant, we see a qualitatively similar contrast in illusory licensing effects. These data challenge the hypotheses that the presence or absence of the NPI illusion depends solely on the lexical status of the NPI, or the selective repair of *ever*. The finding that a single lexical NPI can exhibit different behaviors with regard to illusory licensing effects suggest that additional factors above and beyond lexical differences can cause a change in susceptibility to the illusion. I note that it is still possible that the contrast observed for *ever* and *any* may be due to lexical differences. However, the finding that a single lexical NPI can disassociate the behavior suggests that lexical status is not the only relevant dimension for determining susceptibility to the illusion.

In short, the findings from Experiments 4 and 5 suggest that there might be something else that drives the differences in behavior, such as the position of the NPI relative to the potential licensors. However, the notion of position could refer to several different possibilities, including differences in structural position (e.g., pre- vs. post-verbal positions, or main vs. embedded clause positions), or differences in time/distance (e.g., the time/distance between the potential licensors and the NPI, or simply earlier vs. later). I test these alternatives in Experiment 6.

5.7 Experiment 6

The aim of Experiment 6 was to better understand how the position of the NPI in the sentence might impact the illusory licensing effect. As noted above, position could describe structural position or position relative to a point in time/distance relative to the potential licensors. To distinguish these alternatives, I held constant the linear and structural position of the NPI *ever*, and manipulated the position of a parenthetical phrase to vary the time between the context containing the potential licensors and the NPI, as illustrated in (9).

(4) (As the editors mentioned) **no|the** authors [that **the|no** critics recommended for the assignment] have (as the editors mentioned) ***ever*** received a pay raise.

The parenthetical phrase in (8) extends the meaning of the main predicate, but it is not a primary constituent of the sentence. If the structural position of the NPI is critical for eliminating the illusion, then we should expect to see similar profiles for sentences with an intervening parenthetical phrase and sentences with a non-intervening parenthetical phrase, since the linear and structural position of the NPI is held constant. However, if elimination of the illusion is a consequence of extended time/distance between the intrusive licensor and the NPI, then we should expect to see the same modulation of the illusion seen for the previous comparisons, with a disappearance of the illusion for sentences involving an intervening parenthetical phrase.

5.7.1 Participants

Participants were 18 native speakers of English who were recruited using Amazon's Mechanical Turk web-service. Participants were compensated $2.50. The experiment lasted approximately 15 minutes.

5.7.2 Materials

Experimental materials consisted of 36 sets of 6 items, which varied in terms of the presence and structural location of an NPI licensor (grammatical licensor / intrusive licensor / no licensor), and the position of a parenthetical phrase (non-intervening / intervening). All items contained a main subject NP that was modified by an object relative clause. As in Experiments 2-5, the NPI licensor *no* appeared either as the determiner of the main subject NP (grammatical licensor) or as the determiner of the relative clause subject (intrusive licensor), or was replaced with the definite determiner *the* (no licensor). The relative clause was always followed by the auxiliary *have*, which served to clearly mark the right edge of the relative clause, ensuring that participants would correctly construct a parse in which the intrusive licensor did not c-command the NPI. A four to seven word parenthetical phrase appeared either at the beginning of the sentence (non-intervening) or between the auxiliary and the NPI (intervening). The intervening position was chosen to ensure the intended main clause predicate attachment and interpretation of the parenthetical clause. In order to ensure that the parenthetical phrase did not specifically highlight either of the NP positions where the potential NPI licensors appeared, the parenthetical phrase never directly engaged or referred back to any component of the complex subject NP, and no component of the parenthetical phrase required access to this NP. Across all conditions the NPI *ever* appeared in the same position immediately before the main verb. An example set of the items is given in Table 12.

Table 5.12: Sample set of items for Experiment 6. NPIs and licensors are in bold.

NON-INTERVENING PARENTHETICAL	GRAMMATICAL LICENSOR	As the editors mentioned, **no** authors [that the critics recommended for the assignment] have **ever** received a pay raise.
	INTRUSIVE LICENSOR	As the editors mentioned, the authors [that **no** critics recommended for the assignment] have **ever** received a pay raise.
	NO LICENSOR	As the editors mentioned, the authors [that the critics recommended for the assignment] have **ever** received a pay raise.
INTERVENING PARENTHETICAL	GRAMMATICAL LICENSOR	**No** authors [that the critics recommended for the assignment] have, as the editor mentioned, **ever** received a pay raise.
	INTRUSIVE LICENSOR	The authors [that **no** critics recommended for the assignment] have, as the editor mentioned, **ever** received a pay raise.
	NO LICENSOR	The authors [that the critics recommended for the assignment] have, as the editor mentioned, **ever** received a pay raise.

Each participant read 108 sentences, consisting of 36 NPI sentences and 72 filler sentences. The 36 sets of NPI items were distributed across 6 lists in a Latin Square design. The filler sentences were of similar length and complexity to the NPI sentences, but lacked an NPI. Roughly half of the filler sentences used determiners similar to those used in the NPI sentences in similar positions to prevent the possibility that participants might develop superficial reading strategies based on the distribution of the determiners in the NPI sentences. Materials were balanced so that across the experiment half of the sentences were ungrammatical. The anomalies in the filler sentences comprised a variety of grammatical violations, including agreement errors, pronoun gender violations, and unlicensed verbal morphology.

5.7.3 Procedure

The same procedure was used as in Experiments 2 and 4. [23]

5.7.4 Data Analysis

Data analysis followed the same steps as in Experiments 2 and 4.

5.7.5 Results

Figure 5.11 shows the proportion of 'yes' responses for the six experimental conditions in Experiment 2. The results of the statistical analysis are presented in

---

[23] Experiment 6 is not accompanied by a self-paced reading task. Based on the previous comparisons (e.g., Experiments 2-5), I did not anticipate a difference across methodologies, and a follow up study was not conducted because of time constraints.

Table 5.13. Results showed reliable detection of an unlicensed NPI, as grammatical sentences were more likely to be accepted than their ungrammatical counterparts. But judgments for the ungrammatical sentences sharply diverged. Contrasting illusory licensing profiles were observed for sentences with a non-intervening parenthetical phrase and sentences with an intervening parenthetical phrase, as reflected by increased acceptability for ungrammatical sentences with an intrusive licensor. No such difference was observed for the corresponding sentences with an intervening parenthetical. However, the data failed to show an interaction in the direct comparison of illusory licensing effects (t=1.42). The results of the statistical analysis suggest that there is in fact a qualitative difference between sentences with a non-intervening parenthetical and those with an intervening parenthetical, but the contrast is much more fragile than those observed in the previous experiments. The lack of an interaction between the two illusory licensing profile could be due to modest size of the experiment. The individual illusory licensing profiles by participant are shown in Figure 5.12.

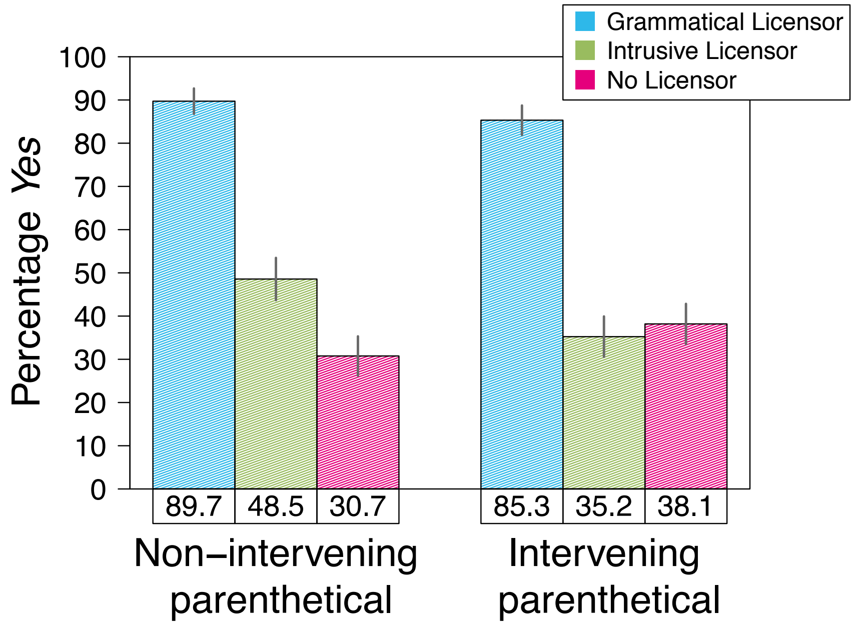
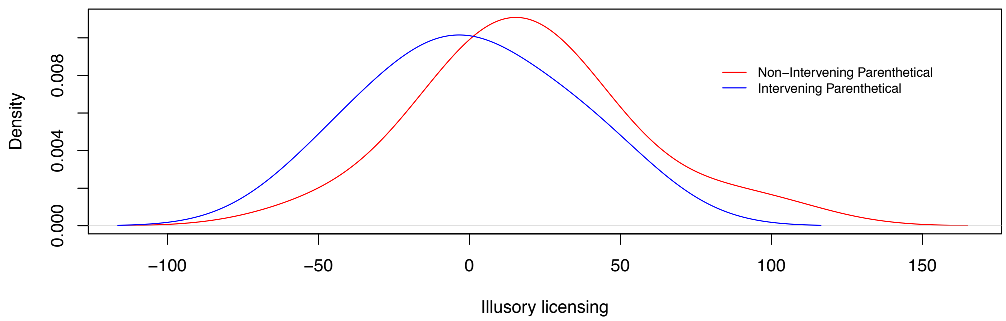


Figure 5.11: Speeded-acceptability judgments (Experiment 5). Mean percentage 'yes' responses for sentences with a non-intervening parenthetical phrase (left) and an intervening parenthetical phrase (right). Error bars indicate standard error of the mean.

Table 5.13: The four orthogonal contrasts for experimental factors (Experiment 2). Contrast 1 (C1) is the effect of grammaticality for sentences with a non-intervening parenthetical phrase. Contrast 2 (C2) is the effect of attraction for sentences with a non-intervening parenthetical phrase. Contrast 3 (C3) is the effect of grammaticality for sentences with an intervening parenthetical phrase. Contrast is the effect of attraction for sentences with an intervening parenthetical phrase. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

COMPARISON	$\hat{\beta}$	SE	z
C1	**2.96**	**0.44**	**6.71**
C2	**0.82**	**0.30**	**2.66**
C3	**3.97**	**1.31**	**3.02**
C4	-0.01	0.30	-0.05
C2 vs. C4	**0.95**	**0.45**	**2.07**



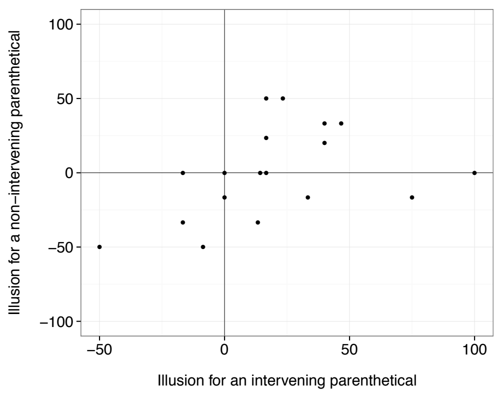


Figure 5.12: Distributions of attraction effects by participant for sentences with an intervening parenthetical phrase (red) and a non-intervening parenthetical phrase (Experiment 6). Bottom: Attraction effects by participant for sentences with an intervening parenthetical phrase and a non-intervening parenthetical phrase. Attraction is positive-going, reflecting increased acceptability.

5.7.6 Discussion

In Experiment 6, I held constant the linear and structural position of the NPI *ever*, and manipulated the position of a parenthetical phrase (non-intervening vs. intervening) to vary the time between the context containing the potential licensors and the NPI. Sentences with a non-intervening parenthetical phrase showed a reliable illusory licensing effect, as comprehenders were more likely to accept ungrammatical sentences when an intrusive licensor was present. By contrast, sentences with an intervening parenthetical phrase failed to show a corresponding illusory licensing effect, as the presence of an intrusive licensor did not increase the rate of acceptance of sentences that contained an illicit NPI. These results suggest that the presence or absence of the illusion is strongly impacted by the time or distance between the potential licensors and the NPI, rather than the structural or linear position of the NPI in the sentence.

5.8 Experiment 7

The results of the parenthetical phrase comparison revealed that extending the time between the intrusive licensor and the NPI can eliminate the illusion. However, it is possible that other factors, such as differences in passive memory dynamics that are unrelated to NPI licensing, could also make an important contribution to the illusory licensing effect. In Experiment 7 I used an explicit computational model of memory access to generate predictions about how these factors might impact NPI licensing and compared them to the behavioral data. To generate these predictions, I used a variant of the ACT-R (Adaptive Character of Thought – Rational) computational model of memory access described in Lewis and Vasishth (2005), Lewis et al. (2006), and Vasishth et al. (2008). Specifically, I used the same model described in Chapter 4 for our investigation of reflexive anaphors.

5.8.2 Model parameters

We tested a range of ACT-R parameter settings used in previous ACT-R research to ensure that the modeling results would be robust against variation in these parameters. The only exception to this approach was the scaling parameter *F*, which was set at 1.4 across all simulations to ensure that the predicted retrieval latencies were on an appropriate time scale. Simulations systematically combined parameter values from across the range of values in Table 5.14. This method resulted in the construction of 3,000 different models, each with a unique parameter setting. As discussed by Dillon et al. (2013), this method of parameter exploration identifies key model predictions independent of idiosyncratic parameter combinations. 1,000 Monte Carlo simulations were run for each combination of parameters. Each trial included the full series of hypothesized retrievals, and each trial generated a predicted retrieval latency for the chunk with the highest probability of retrieval.

Table 5.14: Summary of ACT-R model parameters (Experiment 6).

Parameter	Value
Latency factor (*F*)	1.0
Total goal activation (*G*)	0.50 – 1.50
Noise (*ans*)	0.00 – 0.30
Maximum associative strength (fan)	1.00 – 2.00
Decay (*d*)	0 – 1.0
Maximum difference (*P*)	0 – -1.0

5.8.3 Materials and retrieval schedules

The empirical data of interest are the illusory licensing effects observed for NPI manipulation (Experiments 2-3) and the timing manipulation (Experiment 6). I simulated NPI licensing based on the constituent creation times and retrieval schedules described in Tables 5.15 – 5.22. Creation times were estimated from the empirical reading times from Experiments 3 and 5. Retrievals are linked to the introduction of a given constituent, e.g., a main clause VP, a relative clause VP, or a reflexive.

Table 5.15: Constituent creation times and feature composition for sentences involving the NPI *ever* from the ever/any comparison in Experiments 2-3 (Experiment 7).

	NP-Target	**NP-Distractor**	**VP1**	**VP2**	**NPI**
Time	1083	2657	3225	3831	4381
Category	NP	NP	VP	VP	ADV
Number	pl	pl	-	-	-
Negation	+/-	+/-	-	-	-
Command	+	-	-	+	+
Embedding	IP1	IP2	IP2	IP1	IP1

Table 5.16: Schedule of retrievals and cue sets for sentences involving the NPI *ever* from the ever/any comparison in Experiments 2-3 (Experiment 7). VP1-1 = retrieval of the relative clause subject for thematic binding inside the relative clause. VP1-2 = retrieval of the main clause subject to resolve the object gap inside the relative clause. VP2 = retrieval of the main clause subject for the main clause verb to resolve number agreement. NPI = retrieval for an NPI licensor.

	VP1-1	**VP1-2**	**VP2**	**NPI**
Time	3425	3425	4031	4581
Category	NP	NP	NP	NP
Number	-	-	pl	-
Negation	-	-	-	+
Command	-	+	+	+
Embedding	IP2	IP1	IP1	-

Table 5.17: Constituent creation times and feature composition for sentences involving the NPI *any* from the ever/any comparison in Experiments 2-3 (Experiment 7).

	NP-Target	NP-Distractor	VP1	VP2	VP3	NPI
Time	1083	2657	3225	3831	4399	4949
Category	NP	NP	VP	VP	VP	ADV
Number	pl	pl	-	-	-	-
Negation	+/-	+/-	-	-	-	-
Command	+	-	-	+	+	+
Embedding	IP1	IP2	IP2	IP1	IP1	IP1

Table 5.18: Schedule of retrievals and cue sets for sentences involving the NPI *any* from the ever/any comparison in Experiments 2-3 (Experiment 7). VP1-1 = retrieval of the relative clause subject for thematic binding inside the relative clause. VP1-2 = retrieval of the main clause subject to resolve the object gap inside the relative clause. VP2 = retrieval of the main clause subject for the main clause verb to resolve number agreement. VP3 = retrieval of the main clause subject for thematic binding at the main clause verb. NPI = retrieval for an NPI licensor.

	VP1-1	VP1-2	VP2	VP3	NPI
Time	3425	3425	4031	4599	5149
Category	NP	NP	NP	NP	NP
Number	-	-	pl	-	-
Negation	-	-	-	-	+
Command	-	+	+	+	+
Embedding	IP2	IP1	IP1	IP1	-

Table 5.19: Constituent creation times and feature composition for sentences involving a non-intervening parenthetical from the parenthetical phrase comparison in Experiment 6 (Experiment 7).

	NP1	VP1	NP-Target	NP-Distractor	VP2	NP4	VP3	NPI
Time	1585	3153	3236	4810	5378	6963	756 9	8119
Category	NP	VP	NP	NP	VP	NP	VP	ADV
Number	sg	-	pl	pl	-	sg	-	-
Negation	-	-	+/-	+/-	-	-	-	-
Command	-	-	+	-	-	-	+	+
Embedding	IP1	IP1	IP2	IP3	IP3	IP3	IP2	IP2

Table 5.20: Schedule of retrievals and cue sets for sentences involving a non-intervening parenthetical from the parenthetical phrase comparison in Experiment 6 (Experiment 7). VP1 = thematic subject binding inside the parenthetical phrase. VP2-1 = thematic subject binding inside the relative clause. VP2-2 = retrieval of the main clause subject for object gap resolution inside the relative clause. VP3 = retrieval of main clause subject for agreement licensing. NPI = critical retrieval for an NPI licensor.

	VP1	VP2-1	VP2-2	VP3	NPI
Time	2353	5578	5578	7769	8319
Category	NP	NP	NP	NP	NP
Number	-	-	-	pl	-
Negation	-	-	-	-	+
Command	-	-	+	+	+
Embedding	IP1	IP3	IP2	IP2	-

Table 5.21: Constituent creation times and feature composition for sentences involving an intervening parenthetical from the parenthetical phrase comparison in Experiment 6 (Experiment 7).

	NP-Target	NP-Distractor	VP1	NP3	VP2	NP4	VP3	NPI
Time	1083	2657	3225	4810	5416	7001	7569	8119
Category	NP	NP	VP	NP	VP	NP	VP	ADV
Number	pl	pl	-	sg	-	sg	-	-
Negation	+/-	+/-	-	-	-	-	-	-
Command	+	-	-	-	+	-	-	+
Embedding	IP1	IP2	IP2	IP2	IP1	IP3	IP3	IP2

Table 5.22: Schedule of retrievals and cue sets for sentences involving an intervening parenthetical from the parenthetical phrase comparison in Experiment 6 (Experiment 7). VP1-1 = thematic subject binding inside the relative clause. VP1-2 = retrieval of the main clause subject for object gap resolution inside the relative clause. VP2 = retrieval of main clause subject for agreement licensing. VP3 = thematic subject binding inside the parenthetical phrase. NPI = critical retrieval for an NPI licensor.

	VP1-1	VP1-2	VP2	VP3	NPI
Time	3425	3425	5616	7769	8319
Category	NP	NP	NP	NP	NP
Number	-	-	pl	-	-
Negation	-	-	-	-	+
Command	-	+	+	-	+
Embedding	IP2	IP1	IP1	IP3	-

5.8.4 Modeling results

Figure 5.13 shows the traces of the average ACT-R activation levels of the target, i.e. the head of the main subject NP, and intrusive licensor across a sentence involving the NPI *any* in a post-verbal position (Experiment 3). The traces show that the target is reactivated immediately prior to the critical NPI retrieval, giving it a relatively large activation advantage over the intrusive licensor at the point of NPI licensing.[24] This suggests an alternative explanation for the variable susceptibility to the illusion seen in the pre-/post-verbal comparisons: differences in passive memory dynamics due to a baseline activation bias for the target could selectively eliminate illusory licensing from the intrusive licensor for post-verbal NPIs. That is, reactivation of the target subject by the main clause verb might be the reason that the illusion is eliminated for NPIs in a post-verbal position. No such activation bias would be available for pre-verbal NPIs, since they appear before the verb, increasing susceptibility to illusory licensing.

[24] It is generally assumed in ACT-R models of sentence comprehension (e.g., Vasishth et al., 2008) that thematic binding involves verbs and heads of arguments, rather than verbs and entire arguments, e.g., a complex NP. This is neither an obvious nor an innocent assumption in the current case, as it impacts the relative activation of the correct and intrusive licensor positions.

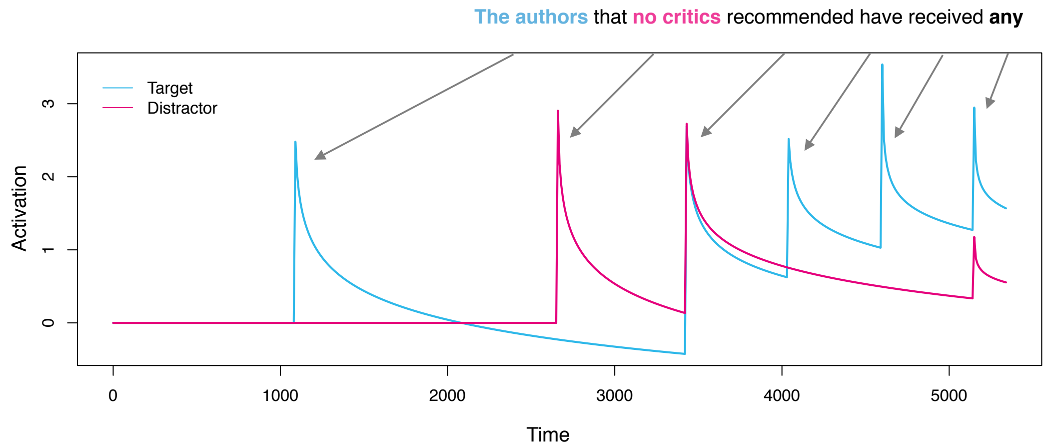


Figure 5.13: Average activation for target memory (blue) and intrusive licensor memory (red) for a sentence involving the post-verbal NPI *any*.

However, a comparison of the predicted and observed illusory licensing effects for the pre-verbal *ever* and post-verbal *any* comparison (Experiment 3) revealed that fluctuations of activations were not sufficient to capture the observed contrasts, as shown in Figure 5.14. ACT-R simulations predicted contrasting profiles, as reflected by a reduced illusory licensing effect for the post-verbal NPI *any*. However, this profile differs from the observed profile, which shows a disappearance of the illusory licensing effect for *any*. Figure 5.15 presents a graphical summary of the range of predictions under the different parameter settings.

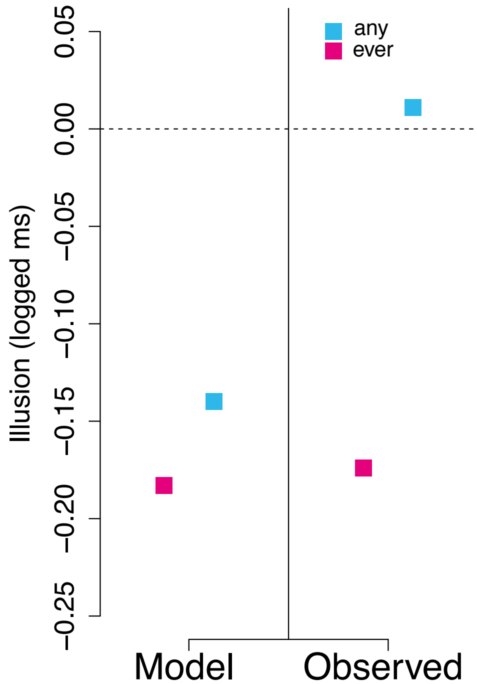


Figure 5.14: Comparison of ACT-R predicted illusory licensing effects and observed illusory licensing effects for pre-verbal *ever* and post-verbal *any* from Experiment 3. Observed illusory licensing effects calculated over the first spill-over region for both NPIs.

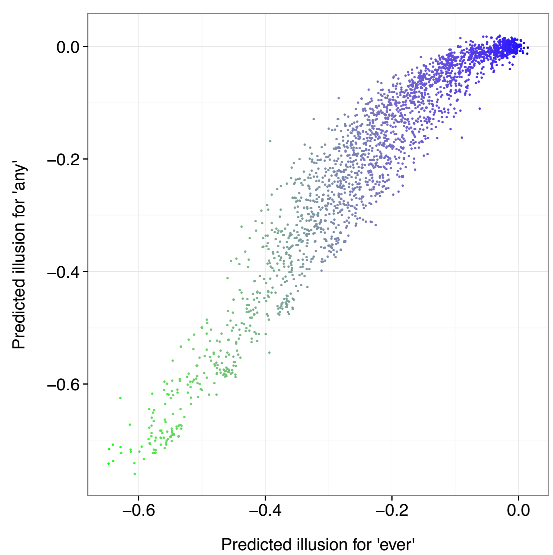


Figure 5.15: A graphical summary of the range of predictions under different assumptions for the NPI manipulation from Experiment 3. The illusory licensing effect is negative going, reflecting facilitated processing.

Simulations of the parenthetical phrase comparison (Experiment 6) provided an additional test of the impact of passive memory dynamics on the illusory licensing effect. An important feature of the parenthetical comparison was that the parenthetical phrase did not specifically highlight either of the potential licensor positions. Given that the ACT-R decay parameter is uniform across all elements of a representation, the relative activation levels of the potential licensor positions should be preserved, predicting similar profiles for sentences with an intervening parenthetical phrase and sentences with a non-intervening parenthetical phrase, as shown in Figures 5.16. However, as in the pre-/post-verbal NPI comparison, the predicted profile shown in Figure 5.17 differs from the observed profile, which shows a disappearance of the illusory licensing effect for sentences involving an intervening parenthetical. The differences between the predicted and observed profiles provides further evidence that the selective nature of the illusion cannot simply be due to differences in passive memory dynamics. Figure 5.18 presents a graphical summary of the range of predictions under the different parameter settings

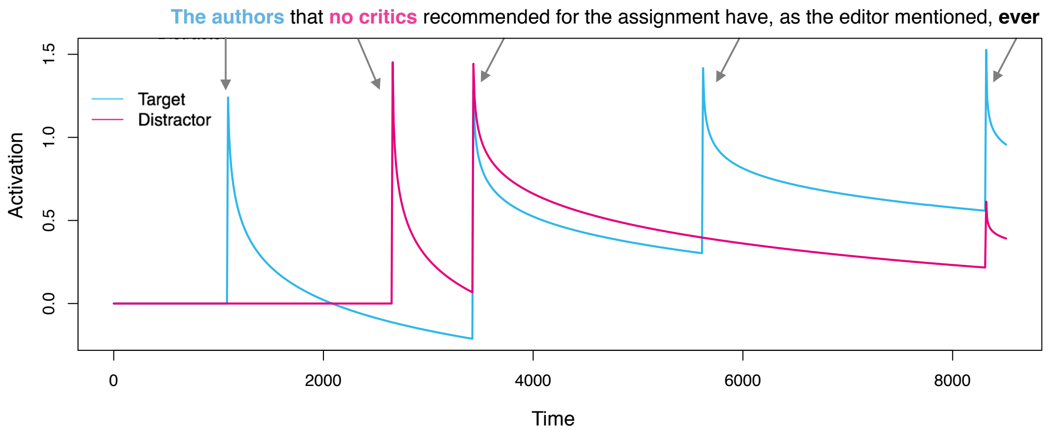


Figure 5.16: Average activation for target memory (blue) and intrusive licensor memory (red) for a sentence involving the an intervening parenthetical phrase, showing that the relative activation profiles for the target and intrusive licensor are preserved across the intervening material. Note that the relative activation levels of the target and distractor are preserved across the intervening parenthetical phrase.

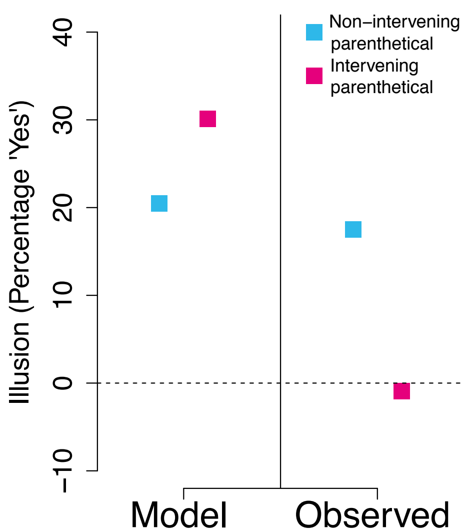


Figure 5.17: Comparison of ACT-R predicted illusory licensing effects and observed illusory licensing effects for the parenthetical phrase timing manipulation (Experiment 6).

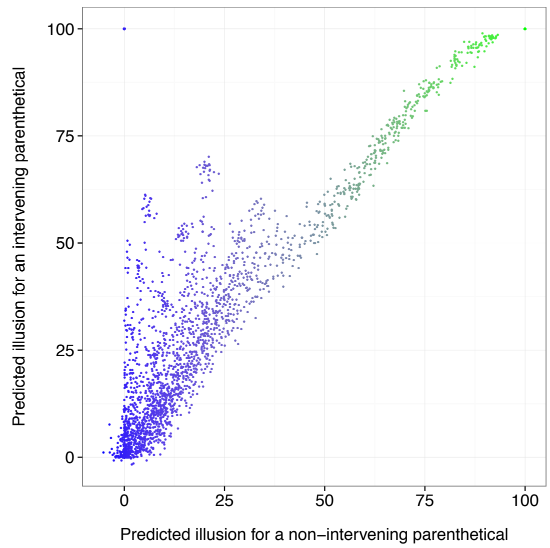


Figure 5.18: A graphical summary of the range of predictions under different assumptions for the timing manipulation from Experiment 6. The illusory licensing effect is positive-going, reflecting increased acceptability.

5.8.5 Discussion

ACT-R simulations showed that differences in passive memory dynamics can modulate the illusory licensing effect to some degree. However, the predictions across the range of parameter combinations revealed largely similar profiles for different positions, and comparisons to the empirical date revealed that the narrow range of differences predicted across the parameter space are not sufficient to capture the full profile of contrasting illusions. In particular, ACT-R simulations did not predict the disappearance of the illusion that was consistently observed across our comparisons.

The modeling results of Experiment 7, taken together with the behavioral results of Experiments 1-6, inform our understanding of the scope and source of illusory licensing effects in two ways. First, they show that the NPI illusion is highly selective, occurring reliably only in specific configurations. This selective profile is not predicted by existing accounts of illusory licensing effects, such as those proposed by Vasishth et al. (2008) and Xiang et al. (2009). These accounts attribute the illusion to general mechanisms that should apply whenever the parser attempts to license an NPI, and as such, they predict that the illusion should extend to a wide range of NPIs and contexts. However, the findings from the present study show that this is not the case. Rather, these findings suggest that it is the position of the NPI relative to the intrusive licensor that determines susceptibility to illusory licensing. Second, these results suggest that the selective nature of the NPI illusion must reflect changes across the representation beyond differences in passive memory dynamics, as these differences did not capture the observed contrasts.

As discussed earlier, existing studies on NPI illusions have probed for illusory licensing effects at different time points after the appearance of the NPI, and they have shown that sensitivity to the structural properties of a potential licensor grows as the lag from the NPI to the probe point increases. In contrast, across all of our comparisons, I held constant the point of probing relative to the NPI, and I found that sensitivity to the structural properties of a potential licensor also varied as a function of when the NPI appeared, such that sensitivity grew as the lag from the context containing the potential licensors increased.

In short, the key finding is that the material between the irrelevant licensor and the NPI determines susceptibility to the illusion, rather than the structural or linear position of the NPI in the sentence. However, it is an open question as to why this material might matter. It could reflect the impact of extended time on some component of the licensing mechanism, or the impact that specific linguistic items like intervening parenthetical phrases or main clause verbs have on the interpretation or encoding of the licensing context. The present results do not decide between these different possibilities.

5.8 Experiment 8

The goal of Experiment 8 was to further test the ability to systematically turn linguistic illusions on and off. Based on the findings from Experiment 7, it is possible that the passage of time/distance should lead to immunity for other types of linguistic illusions, such as those commonly observed for subject-verb agreement. Previous studies have shown that subject-verb agreement is highly susceptible to illusory licensing from structurally irrelevant material (e.g., Wagers et al., 2009; Dillon et al., 2013), showing an illusory licensing profile that is qualitatively similar to illusory NPI licensing (see the introduction to this chapter for discussion about the similarities between these illusions). If there is a homogenous cause for these illusions, as tacitly assumed in previous studies, then we might expect the contrasting illusory licensing profiles observed for NPIs to extend to subject-verb agreement illusions. I tested this prediction in Experiment 8 by extending the timing manipulation for NPIs to subject-verb agreement.

5.8.1 Participants

Participants were 18 native speakers of English who were recruited using Amazon's Mechanical Turk web-service. Participants were compensated $2.50. The experiment lasted approximately 15 minutes.

5.8.2 Materials

Forty-eight item sets of the form shown in Table 5.21 were constructed. The subject-verb agreement. Three experimental factors were manipulated: grammaticality, attraction, and the position of the parenthetical phrase. Following Wagers et al. (2009), grammaticality was manipulated by varying the number feature of the agreeing verb. The agreeing verb was always a present tense agreeing form of *be* (*was* in the grammatical conditions and *were* in the ungrammatical conditions). Attraction was manipulated by varying the number of the distractor noun. The head noun was always singular. Across all conditions, the subject head noun (i.e. the local subject) was held constant. The subject head noun was always modified by prepositional phrase that contained the distractor noun. A four to seven word parenthetical phrase appeared either at the beginning of the sentence (non-intervening) or between the prepositional phrase and the agreeing verb (intervening). In order to ensure that the parenthetical phrase did not specifically highlight either of the NP positions where the potential agreement licensors appeared, the parenthetical phrase never directly engaged or referred back to any component of the complex subject NP, and no component of the parenthetical phrase required access to this NP.

Table 5.23: Sample set of items for Experiment 8.

NON-INTERVENING	*PARENTHETICAL*	GRAMMATICAL WITH ATTRACTOR	As the janitor mentioned, the key to the cabinets probably was destroyed by the fire.
		GRAMMATICAL WITHOUT ATTRACTOR	As the janitor mentioned, the key to the cabinet probably was destroyed by the fire.
		UNGRAMMATICAL WITH ATTRACTOR	As the janitor mentioned, the key to the cabinets probably were destroyed by the fire.
		UNGRAMMATICAL WITHOUT ATTRACTOR	As the janitor mentioned, the key to the cabinet probably were destroyed by the fire.
INTERVENING	*PARENTHETICAL*	GRAMMATICAL WITH ATTRACTOR	The key to the cabinets, as the janitor mentioned, probably was destroyed by the fire.
		GRAMMATICAL WITHOUT ATTRACTOR	The key to the cabinet, as the janitor mentioned, probably was destroyed by the fire.
		UNGRAMMATICAL WITH ATTRACTOR	The key to the cabinets, as the janitor mentioned, probably were destroyed by the fire.
		UNGRAMMATICAL WITHOUT ATTRACTOR	The key to the cabinet, as the janitor mentioned, probably were destroyed by the fire.

Each participant read 144 sentences, consisting of 48 subject-verb agreement sentences and 96 filler sentences. The 48 sets of subject-verb agreement items were distributed across 8 lists in a Latin Square design. The filler sentences were of similar length and complexity to the subject-verb agreement sentences. Half of the filler sentences used determiners similar to those used in the subject-verb agreement sentences. Materials were balanced so that across the experiment half of the sentences were ungrammatical. The anomalies in the filler sentences comprised a variety of grammatical violations, including agreement errors, pronoun gender violations, and unlicensed verbal morphology.

5.8.3 Procedure

The same procedure was used as in Experiments 2, 4 and 6.

5.8.4 Data analysis

Data were analyzed using logistic mixed-effects models, since the dependent measure was categorical (i.e., ‘yes’ or ‘no’). I used helmert contrast for experimental fixed effects, and a fully-specified random effects structure, which included random intercepts and slopes for all fixed effects by participants and by items (Baayen et al., 2008; Barr et al., 2013). As in the NPI experiments, the effect of attraction in the ungrammatical conditions is of primary theoretical interest.

5.8.5 Results

Figure 5.19 shows the proportion of ‘yes’ responses for the six experimental conditions in Experiment 2. The results of the statistical analysis are presented in Table 5.22. Results showed reliable detection of an unlicensed agreeing verb, as grammatical sentences were more likely to be accepted than their ungrammatical counterparts. Judgments for the ungrammatical sentences across the parenthetical phrase manipulation showed qualitatively similar profiles, as both sets of sentences showed an illusory licensing effect (agreement attraction), as participants were more likely to accept ungrammatical sentences with an attractor than ungrammatical sentences without an attractor. These findings replicate previous findings for subject-verb agreement attraction (e.g., Wagers et al., 2009).

The individual illusory licensing profiles by participant are shown in Figure 5.20. Notice the overlap in distributions, which contrasts with separated distributions observed from the NPI experiments.

Since the contrast between NPI illusions and subject-verb agreement illusions is of primary theoretical interest for Experiment 8, I directly tested the illusory licensing effects between NPIs (the timing manipulation from Experiment 6) and subject-verb agreement. Contrasting illusory licensing profiles were observed for NPIs and subject-verb agreement, as reflected by interaction between illusory licensing and dependency within the ungrammatical conditions.

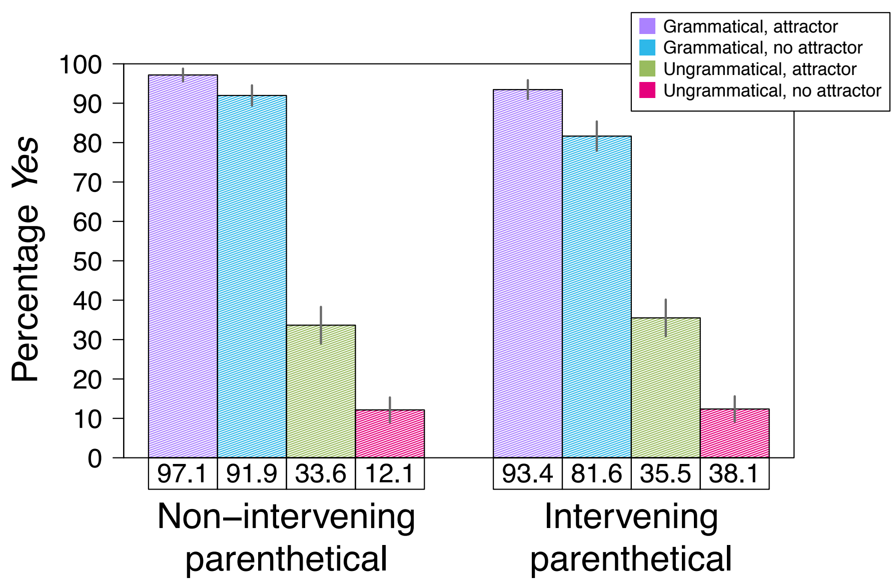
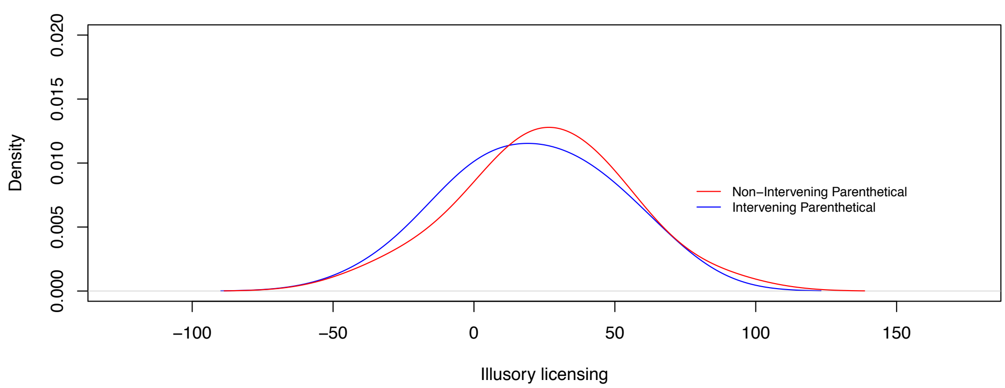


Figure 5.19: Speeded-acceptability judgments (Experiment 8). Mean percentage 'yes' responses for sentences with a non-intervening parenthetical phrase (left) and an intervening parenthetical phrase (right). Error bars indicate standard error of the mean.

Table 5.24: The four orthogonal contrasts for experimental factors (Experiment 2). Contrast 1 (C1) is the effect of grammaticality for sentences with a non-intervening parenthetical phrase. Contrast 2 (C2) is the effect of attraction for sentences with a non-intervening parenthetical phrase. Contrast 3 (C3) is the effect of grammaticality for sentences with an intervening parenthetical phrase. Contrast 4 (C4) is the effect of attraction for sentences with an intervening parenthetical phrase. All *t*-values with an absolute value greater than 2 are statistically significant, indicated in bold.

COMPARISON	$\hat{\beta}$	SE	*z*
C1	**4.99**	**0.60**	**8.30**
C2	**1.40**	**0.38**	**3.63**
C3	**3.71**	**0.46**	**7.98**
C4	**1.55**	**0.43**	**3.58**
Attraction × Dependency	**1.58**	**0.49**	**3.22**



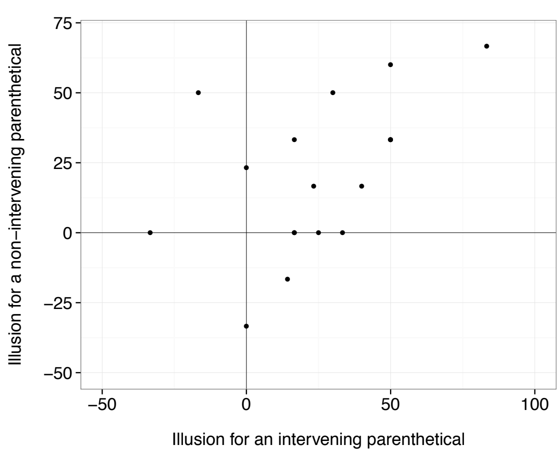


Figure 5.20: Distributions of attraction effects by participant for sentences with an intervening parenthetical phrase (red) and a non-intervening parenthetical phrase (Experiment 8). Bottom: Attraction effects by participant for sentences with an intervening parenthetical phrase and a non-intervening parenthetical phrase. Attraction is positive-going, reflecting increased acceptability.

5.8.6 Discussion

The purpose of Experiment 8 was to assess the generality of the ability to systematically turn linguistic illusions on and off. To accomplish this, I tested the prediction that the passage of time/distance should lead to immunity for other types of linguistic illusions by extending the parenthetical timing manipulation for NPI illusions to a qualitatively similar illusion involving subject-verb agreement. Results revealed a surprising contrast between these two illusions. Unlike for NPIs, the passage of time/distance does not lead to immunity from illusory licensing for subject-verb agreement. This contrast was supported by an interaction between illusory licensing effects in the ungrammatical sentences and dependency type.

Specifically I found that while the NPI illusion shows a fleeting time profile, such that it is present or absent based on position of a parenthetical phrase to vary the time between the licensors and the NPI (as shown in Experiments 2-6), the same time profile does not extend to subject-verb agreement, as the position of a parenthetical phrase did not modulate the agreement illusions, i.e., subject-verb agreement illusions do not disappear under the same conditions that prevent the NPI illusion.

The results from Experiment 7 align with the claims that agreement and NPIs (i) have distinct licensing mechanisms (discussed in Chapter 2), and (ii) their corresponding illusions arise from different underlying sources (e.g., Xiang et al., 2009; Xiang et al., 2013). Specifically, agreement involves a syntactic relation between lexical items or specific features, whereas NPI licensing involves syntactic constraints as well as semantic and pragmatic constraints (discussed in Chapter 2). Furthermore, agreement illusions are argued to reflect mis-retrieval from content-addressable memory, and while NPI illusions have been given a similar treatment, they have also been argued to reflect over-pragmatic accommodation. In short, the results from the current study favor recent accounts which claim that NPI and agreement illusions reflect a fundamental distinction in how we encode and navigate syntactic vs. semantic/pragmatic dependencies (e.g., Xiang et al., 2009; Xiang et al., 2013; but cf. Vasishth et al., 2008).

5.9 General discussion

5.9.1 The source and scope of illusory licensing effects

In the present study, I examined a linguistic illusion that involves illusory licensing of negative polarity items (NPIs), where comprehenders temporarily accept sentences with an illicit NPI in on-line measures, but later judge those same sentences as unacceptable after more reflection in off-line tasks. Experiments 1-3 provided the first direct comparison of the NPIs *ever* and *any*. Results revealed that while *ever* shows the illusion, *any* does not. I suggested that this contrast could reflect inherent properties of the NPIs or selective repair of *ever*, or that it could be a consequence of their differing sentential positions, e.g., pre-verbal *ever* vs. post-verbal *any*. Experiments 4 and 5 distinguished these alternatives by testing a single NPI *ever* in pre- and post-verbal positions. Results showed the same modulation of illusions seen for *ever* and *any*, favoring the positional account. Experiment 6 held constant the linear and structural position of the NPI *ever*, and manipulated the position of a parenthetical phrase to vary the time and distance between the NPI and the main subject. Results revealed that the illusion disappeared when a parenthetical phrase intervened between the potential NPI licensors and the NPI. Experiment 7 modeled these results, revealing that differences in passive memory dynamics were not sufficient capture the observed contrasts. Lastly, in Experiment 8, I tested whether the contrasts observed for NPIs extends to subject-verb agreement illusions using a timing manipulation similar to the one that I used for NPIs. Results from this experiment revealed that the fleeting time profile for NPI illusions does not extend to subject-verb agreement illusions.

The results of the present study inform our understanding of the scope and source of linguistic illusions in comprehension as follows. First, the results demonstrate that linguistic illusions are much more restricted than previously assumed, as evidenced by the fact that we can systematically turn NPI illusions on and off. These results, taken together with the findings reported in Chapters 3 and 4 for anaphora illusions, point to a new generalization which states that illusory licensing effects are highly selective, occurring only in specific configurations.

More specifically, the fleeting time profile for the NPI illusion implies that the illusion cannot simply be due to a noisy online implementation of the licensing mechanism, as suggested by existing accounts (e.g., Vasishth et al., 2008; Xiang et al., 2009). These accounts attribute the illusion to noisy on-line licensing mechanisms that should apply whenever a comprehender attempts to license an NPI. However, accounts that rely on a noisy licensing mechanism incorrectly predict that the NPI illusion should be rather general, which I found not to be the case. In particular, the results of the present study show that the presence or absence of the illusion is strongly impacted by the time or distance between the potential licensors and the NPI, rather than the structural or linear position of the NPI in the sentence. Across our comparisons, I consistently found that the NPI illusion disappeared when there was a greater lag in time/distance between the intrusive licensor. In short, additional factors that I describe in detail below must contribute to the categorical presence or absence of the illusion.

Lastly, despite the fact that psycholinguistic studies often treat NPIs and subject-verb agreement illusions similarly, the finding that the NPIs and subject-verb agreement show different time profiles with regard to illusory licensing effects suggests that there is not a homogenous cause for these two illusions. Based on their alignment with distinct formal licensing mechanisms, one attractive possibility that I discuss in detail below is that the variability is a consequence of general mechanisms (whether they be memory retrieval mechanisms or pragmatic inferencing mechanisms) that distinguish between the encoding of emerging/intermediate syntactic and semantic/pragmatic representations in memory.

5.9.2 Noisy semantic/pragmatic encodings and interpretation

Existing accounts of the NPI illusion, such as those proposed by Vasishth et al. (2008) and Xiang et al. (2009) have emphasized that NPI licensing is a function of (i) the licensing conditions on NPIs, (ii) the access mechanisms, and (iii) the encoding in memory of the prior sentence context. Under these accounts, the illusion is understood as a partial-match effect, and suggests that on-line processing mechanisms can access semantic licensing features such as negation independently regardless of the position of those features in the structured sentence representation. Importantly, these studies have tested for illusory licensing effects at different time points after the appearance of the NPI, and they have shown that the sensitivity of the access mechanisms to the structural properties of a potential licensor grows as the lag from the NPI to the point of sampling increases. That is, susceptibility to illusory licensing is impacted by the amount of time that a comprehender has had to process the NPI, with more time yielding greater grammatical accuracy. This contrast between immediate, online responses and later, offline responses can be captured by holding constant all components of the licensing function, i.e. (i)-(iii) above, while assuming that the access mechanisms are noisy, and hence yield more sensitive results over time with repeated access attempts using the same mechanism (see the computational modeling results in the appendix for corresponding evidence). These accounts predict that variation in the time or position of when the NPI is introduced should not strongly impact the illusion, given the assumption that the components of the licensing function (i.e., the licensing conditions, the access mechanisms, and the encoding of the context) are constant as a parse is extended.

In contrast to these previous studies, I focused on the period of time between the licensing context the introduction of the NPI. I consistently found that sensitivity to the structural properties of a potential licensor grew as the lag from the intrusive licensor increased, leading to a disappearance of the NPI illusion. These results suggest that some component of the licensing function is not constant across the representation. For example, one possibility is that the licensing conditions on NPIs could vary based on when the NPI is introduced. I believe this to be an unlikely possibility, especially given the results of parenthetical phrase comparison (Experiment 6). In this comparison, the NPI always appeared in exactly the same linear and structural position, yet I observed differential sensitivity to the structural properties of a potential licensor from the same position. This suggests that the structural or linear position of the NPI is not critical for eliminating the illusion, and hence it is unlikely that the variable susceptibility to the illusion is due to processes that are initiated when the NPI is introduced.

Another possibility is that the access mechanisms vary based on when the NPI is introduced. Recent research on the memory operations that are used to query linguistic structure suggests that comprehenders build complex hierarchical representations as they process a sentence, but that they have different ways of navigating those representations to build linguistic dependencies. For example, hierarchical representations in memory can be navigated either using a content-addressable retrieval, which relies on a combination of structural and non-structural information to query linguistic structure in memory, or using a structure-guided search operation (see Dillon, 2011; Dillon et al., 2013; Phillips et al., 2011, for discussion). I suggest that the use of multiple access mechanisms for NPI licensing is unlikely, again based on the findings from the parenthetical phrase comparison. If, for example, parenthetical phrases could trigger the use of a structure-guided access strategy, then we would expect to see grammatically accurate processing for sentences involving a non-intervening parenthetical phrase. However, this prediction was not borne out in the data, as sentences with a non-intervening parenthetical phrase showed a robust effect of illusory licensing.

A third possibility that could impact licensing is decay. Recent research on distance effects in on-line processing has shown that factors such decay and the number of intervening words are a key determinant of dependency resolution difficulty, with greater processing difficulty observed as the distance or number of words that intervene between dependent elements increases (Gibson, 1998; Gibson & Thomas, 1999; Gibson, 2000; Grodner & Gibson, 2005; Hawkins, 1994). This hypothesis predicts increased errors of illusory licensing across increased distances. However, the findings from the present study show the opposite effect, as errors of illusory licensing disappeared as the distance from the intrusive licensor to the NPI increased.

I argue that whereas the licensing conditions on NPIs and access mechanisms are constant, the encoding or interpretation of the prior context changes as a parse is extended. Specifically, I suggest two possibilities: (i) the initial interpretation of the licensing context is accurately and rapidly built and remains constant, but the format of its encoding changes over time, or (ii) the interpretation of the licensing context changes over time. According to either of these proposals, the fleeting NPI illusion reflects access to the internal stages of encoding and interpretative processes. Below, I step through the details of how capture the fleeting NPI illusion as an effect of changes over time in the encoding or interpretation of the licensing context.

One possibility is that phrasal encodings change over time from a format where individual semantic and structural licensing features can be readily accessed for partial-matching to one where the semantic licensing features are no longer accessible independently from the position of those features in the structured sentence representation. For example, as comprehenders process a sentence, they may periodically consolidate the independent features of an expanding hierarchical constituent structure into a single, unitized encoding in order to reduce processing load and conserve working memory resources. This feature-integration process might bind together the independent semantic and structural features to create a compressed encoding of the hierarchical constituent structure. As a result, the embedded irrelevant licensor would become opaque for causing illusory licensing, as the representations that encode this information would have to be recovered holistically since the individual constituent features are no longer transparently accessible. This shift in the format of the encoding would prevent further illusory licensing errors, and hence yields greater processing accuracy when the licensing context is probed later in the sentence. Under this view, the different licensing profiles observed at different points in time would reflect access to different stages of the encoding process. In Appendix B, I provide the details of an explicit computational model that relies on multiple-stage encodings to capture the fleeting NPI illusion. It is worth mentioning that preliminary results provide a close fit to the empirical data.

Alternatively, the fleeting NPI illusion could reflect changes over time in the initial parse of the licensing context, leading to slow or delayed detection of an illicit NPI in the illusory licensing conditions. For example, comprehenders could be uncertain about the scope of the irrelevant licensor until reliable cues about the structure of the licensing context are encountered. Such cues might include the main clause verb, for example. It is thus possible that encountering more reliable cues to the structure of the sentence forces the comprehender to revise their initial parse of the licensing context. Under this view, when the licensing context is sampled early, as in the pre-verbal conditions, it is possible that comprehenders had the incorrect parse or were uncertain about the scope of the irrelevant licensor to rule it out as a licensor for the main clause. Once more reliable cues have been encountered, the comprehender will be able to settle on the appropriate interpretation of the licensing context, aiding in the accurate processing of a subsequent illicit NPI. Under this view, no such difficulty would be expected in the grammatical and no licensor conditions since the scope of the main clause subject is clear, i.e., the presence or absence of negation on the head of the relative clause is clear regardless of whether it has a modifier.

A variant of this proposal would suggest that interpretation of the licensing context is delayed because the comprehender must construct the appropriate mental scenario that makes licensing NPIs by pragmatic inferencing felicitous. Evaluating whether or not the semantic context and accompanying pragmatic inferences can appropriately license an NPI likely involves a number of steps that take some amount of time to execute and evaluate. It is thus possible that when the licensing context is sampled early for NPI licensing, as in the pre-verbal NPI conditions, comprehenders have not yet had sufficient time to appropriately compute the necessary licensing inferences. In such a context, the comprehender may be forced to rely on incomplete or faulty inferences based on the presence of the embedded licensor, even though that licensor is in the incorrect position. No such effect would be expected in the no licensor condition due to the lack of a licensor.

The present results do not decide between the possibilities that the initial parse of the licensing context is rapidly and accurately constructed and remains constant, but the format of its encoding changes over time, or that the parse itself changes over time, leading to slow or delayed error detection. There are two ways to tease apart these possibilities:

One way to probe the parse of the licensing context and the status of the embedded negative NP would be to test whether the same type of embedded negative NP can spuriously license a bound-variable pronoun. For example, it may be informative to test whether embedded negative QPs like *no schoolgirl* in the sentence *The librarian that no schoolgirl liked had scolded her for an overdue book* are similarly disruptive for licensing bound-variable pronouns like *her*. Previous tests have shown that for sentences like the one above, feature-matching, but structurally irrelevant negative QPs are not considered for bound-variable licensing (Kush, 2013). However, these previous tests relied on contexts in which the pronoun was introduced after the parser has likely had sufficient time to construct the appropriate parse of the licensing context (note that a similar amount of time led to a disappearance of the NPI illusion for NPIs in similar position, e.g., post-verbal *any*). In order to probe the parse at different points in time, we would need to manipulate the timing of when the pronoun is introduced relative to the embedded negative QP. It may be possible to achieve this in a V2 language like German, where we can manipulate the time/distance between the pronoun and the licensing context by varying whether the main verb appears in V2 or verb-final position. If bound-variable anaphora continues to show a non-attraction profile after controlling for time/distance, then we can conclude that the fleeting NPI illusion is not simply a consequence of uncertainty about the scope of the embedded negative QP, since a non-attraction profile for bound-variable anaphora would provide reasonable evidence that comprehenders are certain about the scope of the embedded negation and that they can quickly act on that information.

Another way to test whether the parse of the licensing context remains constant or changes over time would be to probe for semantic equivalency judgments at various points following the licensing context. If the encoding of the licensing context changes but the parse remains constant, then responses to a probe sentence that is semantically equivalent to the base proposition should be accurate immediately following the licensing context and remain accurate across the sentence. If, however, the parse changes or is delayed, judgments might initially be inaccurate and improve as the sentence progresses.

There is also some independent evidence that might bear on whether or not the fleeting NPI illusion reflects changes over time in the format of the encodings or changes over time in the parse. However, this evidence is inconclusive at present. Below, I briefly describe the evidence for the two current proposals.

Evidence for changes in the parse and delayed interpretation

Perhaps the clearest case that involves parsing revisions of the sort that lead to delayed interpretation is garden-path sentences, e.g., *The horse raced past the barn fell* (Bever, 1970). Garden-path sentences are cases where comprehenders fail to build a certain structure in their initial parse due to a structural ambiguity. In these cases, encountering a reliable cue about the structure of the sentence may signal that the parser has made a wrong choice, triggering a subsequent revision. I do not think that the fleeting NPI illusion is a case of a garden-pathing. However, there is a general parallel that certain contexts can induce uncertainty for the parser. These contexts could involve a point of ambiguity or the structure of a complex, subject-modifying relative clause or downward entailing contexts, for example. In these cases, encountering a reliable cue to the structure of the sentence may trigger re-parsing of the prior context, leading to delayed or slow interpretation of subsequent material.

A similar phenomena more closely related to NPI illusions involves comparative illusions, e.g., *More people have been to Russia than I have* (Montalbetti, 1984; Townsend & Bever, 2001). Comparative illusions constitute a linguistic illusion because comprehenders often fail to detect the semantic incoherence. Wellwood and colleagues have recently argued that comparative illusions reflect mis-parsing. In their study, Wellwood et al. (submitted) found that the illusion is more robust with 'repeatable' predicates, e.g., *go to the gym*, than with ‘non-repeatable’ predicates, e.g., *won the lottery yesterday*. These findings suggest that the comparative illusion sentences are initially mis-parsed as event quantification rather than quantification over individuals. Importantly, these findings provide further evidence that mis-parsing can lead to delayed detection of a semantic anomaly that is parallel to spurious NPI licensing.

There is also evidence suggesting that the interpretation of object relative clauses similar to those used in the current study is slow or delayed. For example, Kowalski and Huang (2014) investigated cases where comprehenders appear to delay role assignment for object-relative clauses in sentences like *the bear that the horse pushed ate the sandwich*, as evidenced by increased reading times after the relative clause verb is encountered. Based on visual-world eye-tracking and working memory and cognitive inhibition measures, Kowlaski and Huang found that role assignment is delayed in object-relative clauses sentences because comprehenders are forced to revise their initial role assignments for the subject NPs after they encounter the verb. These findings are consistent with the hypothesis that comprehenders initially mis-parsed the embedded clause subject, i.e. the irrelevant licensor, or were uncertain about its status in the illusory licensing conditions, leading to delayed detection of the illicit NPI.

Lastly, it also has been recently suggested that the syntactic encoding of adjuncts in memory is less distinctive than arguments (e.g., Van Dyke & McElree, 2011). Based on the assumption that the relative clauses in our NPI materials have an adjunction structure, it is possible that the irrelevant negative licensor was less distinctly encoded in memory, which could increase the amount of time that is needed to evaluate whether or not it has main clause scope or to compute the appropriate semantic interpretation. No such difficulty is expected for the other conditions, as the status of the main clause subject as a core argument is clearly encoded.

Evidence for changing encodings of a single interpretation

Just as there is independent evidence for delayed semantic interpretation, there is independent evidence that suggests that the initial structure and interpretation of the sentence are accurately and rapidly computed. For example, the previous studies that have failed to find reflexive attraction effects (e.g., Sturt, 2003; Dillon et al., 2013) imply that the parser is capable of rapidly and accurately building a richly structured representation of the licensing context. Importantly, just as in the current study on NPI licensing, many of these previous studies on reflexives relied on contexts where the irrelevant licensor was contained inside a relative clause.

These results provide good evidence that comprehenders have assembled the correct syntactic structure of the licensing context, but they do not entail that they have computed the appropriate interpretation. Fortunately, there are several studies that provide clear evidence that the semantic interpretation is registered early, as evidence by rapid anomaly detection. For example, Boland and colleagues showed that comprehenders are able to rapidly detect violations of verb argument structure at the verb using a “stop making sense” task (Boland, Tanenhaus, Garnsey, & Carlson, 1995; see also McElree & Dosher, 1995). More fined-grained information about the speed of semantic interpretation comes from

ERP studies. For example, many studies have shown that rapid semantic interpretation can strongly constrain expectations about the upcoming input, as evidenced by an N400 effect for unpredicted, semantically incongruent words, e.g., *sock* in the sentence *I like my coffee with sugar and socks* (Kutas & Hillyard, 1980).

Lastly, although it does not provide direct evidence for or against my proposal for changing encodings, it is worth pointing out that similar proposals for multiple-stage encodings have been made in other cognitive domains, including vision and language. It is thus possible that multiple-stage encoding is a general cognitive capacity that could be exploited in the language domain. For example, in the psycholinguistic domain, there are a number of models that assume a modular architecture in which the parser structures incoming material in two stages (Frazier & Fodor, 1978; Kimball, 1973; Townsend & Bever, 2001; Abney, 1991; see also; Fodor, Bever, & Garrett, 1974; Whitney, 2004; but cf. Marslen-Wilson & Tyler, 1980). Although the models may differ with respect to the nature of the units that are shunted between the two stages, they all assume that the complete representation of a sentence is built in two stages, with one stage temporally prior to the other. For example, in the two-stage model proposed by Frazier and Fodor (1978), which is named the 'Sausage Machine', the parser first constructs a shallow representation of incoming material before shunting constituents off to a more compact store where they are combined into a complete structure for interpretation. My proposal for two-stage encoding shares two key insights with these models: (i) the division of parsing into multiple stages is a consequence of the functional time and memory pressures, and (ii) the division of parsing into two stages determines restrictions on the information that is available to each stage. One key difference, however, is that whereas Frazier and Fodor assume that the first-stage encoding is structurally similar to the second-stage encoding, I suggest that the encodings at different stages are structurally different.

Modular two-stage encoding schemes are also observed in other domains of cognition, such as visual processing, leading to changes in representational format that are parallel to what we have seen for NPI processing in the language domain. For example, in visual cognition, there is an initial stage of processing during which individual object features such as shape and color are encoded independently of each other. This first stage of encoding is followed by a feature-integration stage, where the separate, independently accessible features are consolidated into a single, unitized encoding of the object in visual memory. This two-stage encoding scheme in vision is explicitly characterized by the feature-integration theory of attention (Treisman & Gelade, 1980; Treisman & Schmidt, 1982; Treisman et al., 1977; see also Chun & Potter, 2000; Holcombe & Clifford, 2012; Treisman, 1996; Wolfe, 2007; Wolfe, 2012).

5.9.3 A pragmatic account of the fleeting NPI illusion

The semantic/pragmatic account proposed by Xiang and colleagues argues that the illusion reflects over-active pragmatic accommodation. According to this account, when parsing a sentence like *The authors that no critics recommended have ever P*, where *P* stands for the yet to be encountered predicate, the NPI *ever* may be spuriously licensed by the negative inference about the contrasting set of referents, namely that *the authors that some critics recommended have NOT P*. There are two ways to capture the fleeting NPI illusion under this pragmatic account. If the fleeting NPI illusion reflects changing encodings, is possible then that once *P* is known, the internal interpretative properties of the unwarranted inference, which might include the head of the relative clause, the members of contrast set, the embedded quantification, and the complement of the quantified set, become opaque, since they are no longer relevant or necessary for deriving the assertoric meaning of the base proposition that invites the unwarranted inference. Once opaque, these interpretative properties would lose their ability to spuriously license any NPIs that may be subsequently encountered in the input. If, on the other hand, the fleeting NPI illusion reflects changes in the initial interpretation, integrating *P* might aid in re-vising the interpretation of the licensing context, which would increase sensitivity to a subsequent illicit NPI.

However, it may be argued that the contrast between subject-verb agreement and NPI illusions is not at all a consequence of how we encode and access semantic vs. syntactic representations. Thus far, I have framed subject-verb agreement dependencies and NPI dependencies as model syntactic and semantic dependencies, respectively. However, based on these two data points, i.e. NPIs vs. agreement, it is not yet clear how representative these dependencies are of the syntax-semantics divide. For example, it could be that NPIs are idiosyncratic with respect to illusory licensing effects, and subject-verb agreement is representative of the broader class of linguistic dependencies. Alternatively, NPIs could be the representative dependency, and subject-verb agreement is an outlier, with respect to its sensitivity to illusions.

For example, if the fleeting NPI illusion reflects delayed interpretation, as suggested above, then the timing manipulation that we extended to subject-verb agreement may not have been an appropriate test to manipulate the agreement illusion, since the tail of the dependency, i.e., the agreeing auxiliary verb, was likely introduced before comprehenders had sufficient time to appropriately interpret the licensing context. As such, we may not have given subject-verb agreement a fair chance to exhibit a fleeting illusion, since the encoding of the prior context may not yet have been fully interpreted or transformed into a single, unitized encoding. Under this view, the contrast between subject-verb agreement and NPIs might not reflect differences in how the parser encodes and access syntactic vs. semantic representations, but rather differences due to the positions of the dependent elements relative to the trigger of representational change.

There is some evidence that is consistent with the possibility that the contrast between subject-verb agreement and NPI illusions is not a consequence of how the parser encodes and accesses syntactic and semantic information. For example, Shravan Vasishth (in his review of our paper) has pointed out that there cases where changes in the encoding appear to impact syntactic analysis. He pointed out that in contexts with a relative clause attachment ambiguity, with two candidate NPs (e.g., a complex "NP of NP" structure), it is possible to trigger a non-local attachment (the dispreferred attachment) by simply adding a comma before the relative clause, which could allow for opaqueness of the NP to set in.

By contrast, there are two pieces of evidence that favor the proposal that the contrasting illusions reflect differences in syntactic vs. semantic processing. First, as discussed above, previous research has shown that illusions remain even when structural cues are made clear (e.g., Xiang et al., 2006). Second, if the format of the syntactic encoding changes over time, then we would incorrectly predict similar profiles for reflexive anaphors in the 1- vs. 2-feature mismatch manipulation reported in Chapter 4, since antecedent retrieval takes place over the same structural representation of the licensing context in both cases.

An important task for future research is to test a wider range of syntactic and semantic dependencies to determine the appropriate generalization to capture the variability that we've uncovered in the present study. For example, a task for future work includes extending the timing manipulation that I used for NPIs to other syntactic illusions, like those involving reflexives (Chapter 4) or Case illusions (Bader, Meng, & Bayer, 2000; Slogget, 2013) and semantic illusions, like comparative illusions.

5.10 Conclusion

In this chapter, I investigated the source and scope of two linguistic illusions involving NPIs and subject-verb agreement. In previous research on these illusions, there had been a consensus that they both reflect a homogenous class of structural dependencies and that the time profile of the illusions varies only as a function of the amount of time after the introduction of the licensee (i.e., the NPI or the agreeing verb). The existing accounts (e.g., Vasishth et al., 2008; Xiang et al., 2009; Wagers et al. 2008) predict that the representation of existing structure is stable over time, and that variation in the time or position where the licensee is introduced in the sentence should not drastically impact the illusion.

I presented data from self-paced reading and speeded-acceptability judgments that show that illusory licensing effects for NPIs, which depend on semantic and pragmatic licensing mechanisms, show a fleeting time profile, such that it is present or absent depending on the timing of when the licensee is introduced. However, the same time profile does not extend to illusory licensing effects for subject-verb agreement, which depends on a structural, morpho-syntactic licensing mechanism.

These results inform our understanding of the source and scope of linguistic illusions in several ways. First, the fleeting time profile for NPIs provides further evidence that linguistic illusions are much more selective than previously assumed. Second, the fleeting time profile for NPIs provides further evidence against the generalization that attraction effects are directly linked to specific dependencies. Third, the finding that NPI and subject-verb agreement show different time profiles suggests that there is not a homogenous cause for linguistic illusions. Based on their alignment with distinct licensing mechanisms in the grammar, I argued that variability in the ability of the parser to accurately implement linguistic constraints in real-time processing is, in part, a consequence of processing mechanisms that are sensitive to changes in the encoding and interpretation of emerging syntactic and semantic representations.

Chapter 6 Conclusion

6.1 Summary

In this dissertation, I argued that our previous generalizations about the source and scope of linguistic illusions are incorrect. While previous research has focused on dependency-wise differences in illusory licensing effects, such as the contrasting illusions observed for subject-verb agreement and reflexive licensing, I argued in favor of a narrower focus on the micro-structure of the representations that are built during real-time processing and the mechanisms that are used to encode and access information in those representations.

To support this argument, I provided converging evidence from two case studies. In the first case study, I showed how to "turn on" a linguistic illusion for anaphor resolution, which is a phenomenon that has resisted illusions in the past. These effects are not predicted by existing accounts, but they were observed for two types of anaphoric dependencies involving null subject licensing and reflexive anaphors. Based on evidence from computational simulations, I showed that is possible to derive both the presence and absence of the illusion from within the same memory architecture using a cue-combinatorics scheme that prioritizes structural information in memory retrieval. I then proposed an account of why it is harder to obtain illusions in anaphor resolution than in subject-verb agreement, based on whether retrieval is triggered by error correction or normal resolution. The finding that anaphora and subject-verb agreement show qualitatively similar illusory licensing profiles supports recent proposals that sentence comprehension relies on a single direct-access memory mechanism that deploys all available information using a linear, weighted cue-combinatorics scheme (e.g., Van Dyke & McElree, 2011; Kush, 2013). More generally, these results suggest that variability in illusory licensing effects is, in part, a consequence of how different sources of linguistic information are combined by the memory architecture to guide memory access.

In the second case study, I showed how to "turn off" a robust illusion involving NPIs. Specifically, I showed that illusory licensing effects for NPIs, which depend on semantic and pragmatic licensing mechanisms, show a fleeting time profile, such that it is either present or absent depending on the timing of when the NPI is introduced in the sentence relative to the intrusive licensor. But I also showed that the same fleeting time profile does not extend to illusions involving subject-verb agreement, which depends on a morpho-syntactic licensing mechanism. This contrast supports the claim in the linguistics literature that these two dependencies are qualitatively different, as distinguished by their formal licensing mechanisms. However, the fleeting time profile for the NPI illusion is not predicted by existing accounts. Based on these findings, taken together with evidence from computational simulations, I argued that the selective nature of the NPI illusion is a consequence of noisy semantic/pragmatic encodings and interpretation. Furthermore, the finding that NPI and subject-verb agreement show different time profiles suggests that there is not a homogenous cause for these illusions. Based on their alignment with distinct formal licensing mechanisms, I argued that variability in linguistic illusions is, in part, a consequence of the parser's ability to distinguish changes in the encoding and interpretation of semantic vs. syntactic representations.

6.2 Implications for the parser-grammar relation

The findings that I have reported in this dissertation suggest an architecture of language that relies on noisy encoding and access mechanisms to navigate linguistic structure in real-time. In particular, I have uncovered evidence for a system that relies on (i) selective structural priority to access linguistic information stored in memory, and (ii) dynamic encodings and interpretations that are subject to fundamental changes over time.

Thus far, I have attempted to describe these mechanisms in a way that is consistent with the view that linguistic illusions are the product of a single-structure building system (the grammar) that is embedded in a general cognitive architecture (e.g., S. Lewis & Phillips, in press; Phillips & Lewis, 2013). Under this view, linguistic illusions are understood as a misalignment between the constraints of the grammar and the constraints that the general cognitive architecture places on how we encode and navigate structured representations. Specifically, failures in real-time processing, such as those involving linguistic illusions, are claimed to reflect the limitations of general-purpose memory and cognitive control mechanisms. That is, linguistic illusions can be explained by appealing to independently motivated properties of domain-general memory retrieval and cognitive control mechanisms, without recourse to an additional, dedicated parser that is separate from the grammar.

My proposals for selective structural priority in retrieval and noisy encoding and interpretative processes capture the selective illusions uncovered in this dissertation, but are they really consistent with the view that one can simply embed a grammar in a general cognitive architecture, without an additional parser? What status do these mechanisms hold in the overall architecture of language? Do they reflect ad hoc, task-specific encoding and access strategies, e.g., a separate "grammar of encoding and access", or do they reflect the properties of a task-independent grammar or general cognition? In this section, I assess whether the proposed accounts necessitate a task-specific grammar of encoding and access that cannot be attributed to domain-general cognitive properties or a task-independent grammar.

Many different task-specific parsing mechanisms have been proposed in the sentence processing literature, including proximity-based rules like Right Association (Kimball, 1973), 'rough-and-ready' or 'good enough' representations, pseudo-grammatical templates, or pre-compiled parsing-specific rules (Ferreira, 2003; Ferreira & Patson, 2007; Townsend & Bever, 2001), and violable structural constraints (Tabor, Galantucci, & Richardson, 2004). The evidence for these task-specific heuristics and strategies comes from a variety of sources, including garden-pathing and revision failures, sensitivity to probabilistic biases, linguistic illusions, delayed or slow grammatical analyses, and production-comprehension contrasts. All of these properties are controversial because they suggest the existence of a heuristic-based parser that is separate from the grammar and separate from domain-general memory and cognitive control systems. That is, the parser and grammar are distinct cognitive systems.

However, it has recently been argued that most, if not all, of these properties could be reduced to the interaction between a task-independent grammar and deeper principles of memory and cognitive control, instead of distinct cognitive systems. For example, Phillips and colleagues (e.g., Phillips et al., 2011; Phillips & Lewis, 2013; Lewis & Phillips, in press) have argued that what appear to be task-specific properties are actually a reflex of memory and cognitive control limitations. Under this hypothesis, many of the empirical findings that have been used to motivate a separate parser simply reflect either mis-parsing or noisy memory representations and cognitive control structures. Lewis and Phillips (in press) suggest four key properties that are expected with a single, task-independent grammar that is embedded in a general cognitive architecture:

i) Computations that are not yet complete ("internal stages of computation")

ii) Computations that fail to complete, due to resource limitations ("processing overload")

iii) Computations that complete, but inaccurately, due to a noisy architecture ("properties of memory access mechanisms")

iv) Computations that complete successfully, but that are later challenged by subsequent input ("garden path and revision failures")

Based on this debate, my proposals for selective structural priority and noisy encodings/interpretation face several challenges: If the task-specific mechanisms that have been previously proposed cannot be reduced to one of the properties outlined above in i)-iv), then my proposals for selective structural priority and changing encodings/interpretations are likely unproblematic, as there are more important challenges to be addressed, such as explaining how and why the representations and computations of real-time comprehension differ from those of the grammar. That is, there may be little benefit in accounting for the phenomena that I have uncovered in this dissertation if we cannot also capture all of the other effects that have been attributed to an additional, dedicated parser. If, however, the apparent task-specific principles can be reduced to the properties in i)-iv), then my proposals might be problematic for the single-system view if they really do necessitate task-specific strategies.

I argue that the proposed mechanisms can be explained under a single-system view without invoking ad hoc, task-specific parsing strategies. Below, I work through the details of how to capture each of the proposals under a single-systems account.

Changing encodings and interpretations (NPIs)

In Chapter 5, I argued that NPI illusions reflect changing encodings or interpretations, and that the fleeting time profile reflects access to the internal stages of the corresponding encoding and interpretation processes. Both possibilities (changing encodings/interpretation) reflect the properties of general-purpose memory and interpretative mechanisms, rather than the product of a task-specific strategy. For example, the proposal that fleeting NPI illusions reflect access to the internal stages of encoding processes could be reduced to property i) under a single-system view: Computations that are not yet complete ("Internal stages of computation"). Under this view, encoding processes take some amount of time and they may require multiple steps involving temporal feature integration. As such, it is possible that functional time and memory pressures might force the cognitive system to access the results of the intermediate steps of that computation, leading to the apparent mismatch between the representations revealed in real-time comprehension and those licensed by the grammar.

A compelling piece of evidence for the separation of the grammar and parser is the timing difference between rapid comprehension and slow grammatical judgments. This difference may be taken to suggest that the parser and grammar operate on independent time scales. Could the source of the fleeting time profile observed for the NPI illusion be a slow interpretative system that is separate from a rapid parser? I do not think so. The apparent mismatch between fast but inaccurate interpretation and slow but accurate interpretation could reflect a number of sources other than distinct cognitive systems. One possibility is that slow or delayed commitment to a semantic analysis in the contexts that elicit an NPI illusion could reflect repeated attempts at re-parsing the sentence. In chapter 5, I suggested that proper semantic interpretation is slow as a consequence of mis-parsing or uncertainty about the structural encoding of the subject NP and the modifier that contains the irrelevant negation, rather than independent cognitive systems that operate on different time scales, This proposal fits naturally with property iv) above: Computations that complete successfully, but that are later challenged by subsequent input ("garden path and revision failures"), without invoking a separate interpretation system.

Alternatively, semantic interpretation of the complex subject may be delayed because the comprehender must construct the appropriate mental scenario that makes licensing NPIs by pragmatic inferencing felicitous. For example, evaluating whether or not the semantic context and accompanying pragmatic inferences can correctly license the NPI likely involves a number of steps that take some amount of time, leading to delayed or slow detection of an illicit NPI. Crucially, this hypothesis relies on a single interpretative system that simply requires additional time to integrate different sources of information.

Although these accounts differ in their details (see Chapter 5), the key insight is that the fleeting time profile for NPI illusions can be accommodated by a single, task-independent interpretative system or general memory mechanisms without invoking additional, task-specific encoding and interpretative strategies.

Selective priority for structural information (anaphora)

In Chapter 4, I argued that the selective illusions observed for anaphora reflect a selective priority for structural information, which was achieved by preferentially weighting structural cues in retrieval. This proposal raises two separate, but related questions: (i) what is the status of structural priority in the overall architecture of language? and (ii) what is the origin of the cue weighting and is it a task-specific solution to the problem of how to capture structural priority?

Structural priority could have two sources: the grammar or a separate heuristic-based parser. For example, structural priority could be defined as an access strategy that is specific to comprehension or as a retrieval-specific instruction. This approach would imply the use of a separate "grammar of access". However, the notion of structural priority could follow directly from a task-independent grammar if the constraint on anaphora is defined in purely structural terms (e.g., like Principle A of the Binding Theory). If a constraint is already stated by the grammar in terms of relations defined over hierarchically structured representations, then invoking an independent parser that postulates exactly the same priority for structure does not extend empirical coverage, weakening the motivation for additional structure-building systems that are separate from the grammar.

If structural priority comes "free of charge" from the grammar, then is there a straightforward mapping from structural priority to cue-weighting, either in the memory retrieval architecture or as part of general cognition? I suggest that there is. First, cues are generally assumed to combine in a weighted fashion in many perceptual and cognitive domains (e.g., Trommershäuser et al., 2011). Furthermore, many leading models proposed in the memory literature assume a weighted cue-combinatorics scheme or some other type of weighting factor based on empirically motivated principles of working memory, e.g., individual differences (Anderson, 2007). That is, weighting is a stock component of the assumed memory architecture. For example, a weighted cue-combinatorics scheme is implemented in Lewis and Vasishth's (2005) the ACT-R model of sentence comprehension as well as Gillund and Shiffrin's (1984) Search of Associative Memory (SAM) model. Most importantly for my proposal of structural priority via cue weighting, recent research that has explicitly investigated the role of cue-weighting sentence comprehension suggests that structural cues are in fact given greater weighting than semantic cues during retrieval (Van Dyke & McElree, 2011; see also Kush, 2013).

So far I have shown that preferential cue weighting is a core property of domain-general memory mechanisms and that there is independent empirical evidence for the preferential weighting of structural information, without recourse to a parser that is separate from the grammar or domain-general memory mechanisms. One remaining issue to address now is how to achieve the correct weighting for the all-or-none illusory licensing effects observed for reflexive anaphors. This issue is perhaps the most challenging to address without invoking an ad hoc, task-specific solution.

I suggest that the specific weighting that is used to achieve the effect of all-or-none structural priority is a product of the domain-general utility mechanisms that govern reinforcement learning. Such mechanisms have a wide variety of applications in cognitive models and they play an important role in conflict resolution (see Anderson, 2007, for discussion). In the current context, the process of distinguishing relevant from irrelevant information in memory can be treated as case of conflict resolution. Under this view, each weighting scheme has a utility value and the general cognitive architecture has a reinforcement learning mechanism that is capable of learning the optimal weighting by reinforcing those schemes that are the most successful or have the highest utility. Such reinforcement learning mechanisms may be implemented using the utility function in a domain-general memory architecture like ACT-R, where the utility of a given weighting scheme is gradually adjusted to reflect how often it recovers a perfectly matched item in memory. Importantly, this approach is consistent with the recent argument that memory retrieval in sentence comprehension is a skill that may be developed and optimized as a function of skilled language use (e.g., Lewis & Vasishth, 2005).

There are two factors that could lead the utility mechanism to converge on a weighting that gives rise to the all-or-none behavior for reflexive dependencies. The first factor is inherent noise in the utility measures (see Anderson, 2007, for a description of the noise component), which can vary the particular weighting scheme that will be chosen for a single retrieval attempt. Second, and most important, the weighting scheme used in retrieval must balance the constraint hierarchy that governs a reflexive dependency. Specifically, the relationship between a reflexive and its antecedent is subject not only to structural constraints but also to a formal constraint on morphological feature concord. Both types of constraints may be actively used in retrieval to locate the proper antecedent, as evidenced by the presence of an illusion, but the constraint hierarchy as defined by the grammar states that morphological feature match does not matter unless the corresponding morphological features are in the correct structural position. It is possible then that these factors (noise + competition from multiple constraints) conspire, resulting in the selection of a weighting scheme that prioritizes structural information, but that shows some sensitivity to morphological feature match in restricted contexts. For example, an all-or-none sensitivity to morphological feature match could emerge when the architecture is strained or pressured, as in robustly ungrammatical contexts, like those involving multiple feature mismatches or low probe-to-target similarity, or when the system is placed under time or resource pressures. However, it is n important task for future research to determine whether this type of domain-general utility mechanism can actually lead all speakers to converge on the same outcome.

The proposed implementation of structural priority via cue weighting makes two key predictions: First, it predicts individual differences in susceptibility to illusions. For example, differences in working memory capacity or cognitive control resources could impact susceptibility to reflexive illusions, with low working capacity individuals expected to show increased illusory licensing effects, as their ability to converge on the optimal cue-combinatorics scheme may be weakened. Conversely, individuals with high working memory capacity could have developed advanced abilities to combine information optimally, predicting the ability to resist illusions even in contexts with low probe-to-target similarity.

Second, it predicts that there should be a stage of language acquisition where the learner has not yet had sufficient experience to determine the optimal cue-combinatorics scheme. This hypothesis would predict that children should show difficulty implementing certain structural dependencies, such as those involving reflexives, control, or pronouns. There is some evidence that suggests that children are misled into considering grammatically illicit referential relations that violate Principle B (Chien & Wexler, 1990).

There is also some evidence that suggests that there is a stage of language acquisition where children show difficulty implementing the structural constraints on adjunct control. My findings on adjunct control in adults, as reported in Chapter 3, suggest that adults have acquired the appropriate structural constraint on null subject licensing in adjunct control structures, and that they are able to appropriately use syntactic information to guide antecedent retrieval in well-formed contexts. However, in cases where structural information competes with non-structural information, particularly in contexts with low probe-to-target similarity, the priority for structural information is counter-balanced, increasing sensitivity to structurally irrelevant but feature matched items. Several studies on the acquisition of adjunct control, by contrast, have shown that children, even in well-formed contexts, appear to be misled into considering grammatically irrelevant licensors, suggesting that the priority for structural information in retrieval is not yet even in place (e.g., Hsu, Cairns, & Fiengo, 1985; McDaniel & Cairns, 1990). More specifically, these findings could be interpreted to suggest that children have the appropriate grammatical knowledge, but due to memory resource limitations, they have not yet acquired the optimal cue-combinatorics scheme or are unable to use it to acheve the behavior of an expert language user.

I have only provided a rough sketch of how one could capture the effects of noisy encodings/interpretation and selective structural priority via cue-weighting without invoking ad hoc, task-specific strategies or a separate "grammar of encoding and access". In both cases, invoking additional task-specific mechanisms do not extend empirical coverage when functionally equivalent mechanisms come "free of charge" from a task-independent grammar and domain-general memory and cognitive control structures. Furthermore, the ability to selectively "turn off" the linguistic illusions that has motivated these proposals suggests that linguistic illusions are not as diverse and arbitrary as expected under accounts that rely on a separate, heuristic- or strategy-based parser. Such multiple-systems accounts incorrectly predict illusions across-the-board. As such, a key empirical contribution of this dissertation was to show that the real-time comprehension does indeed display fine-grained sensitivity to the constraints of the grammar, which might not be expected if the architecture relies on shallow, rough-and-ready representations and parsing-specific strategies.

6.3 Conclusion

The nature of the memory architecture for the parser, including the representations that it encodes, the mechanisms that operate over those representations, and its relation to the grammar, is a central topic in psycholinguistic research. The work presented in this dissertation has attempted to contribute to our understanding of how the parser builds, maintains, and accesses information in recent memory for real-time language understanding. Specifically, I have argued that certain effects of real-time language processing, such as linguistic illusions, are a consequence of how the grammar interacts with domain-general memory and cognitive control structures, including combinatory cue-based retrieval and noisy encoding and interpretation processes. In sum, this dissertation contributes to a larger research program on memory and language processing by shedding new light on the mental procedures that allow us to encode and navigate linguistic structure.

Appendix A: Capturing the online/offline contrast

Just as visual illusions reflect conflicting visual percepts, linguistic illusions reflect conflicting judgments about a sentence at different points in time. Specifically, linguistic illusions reflect a mismatch between online/time-sensitive and offline responses. The mismatch between online and offline phenomena has been argued to reflect a distinction between the representations that are built in the service of rapid, incremental processing, and the representations that are licensed by slow grammatical processes, and it has provided motivation for a dual system architecture of language, in which the online comprehension and production mechanisms constructs representations that differ substantially from those licensed by the grammar (e.g., Ferreira, 2003; Ferreira & Patson, 2007; Townsend & Bever, 2001).

However, the reasons for the mismatch between online and offline responses remain poorly understood, and it could reflect a number of possibilities other than a dual system architecture. For example, a number of researchers (Dillon et al., 2013; S. Lewis & Phillips, in press; Phillips et al., 2011; Phillips & Lewis, 2013; Wagers et al., 2009; Wellwood, Pancheva, Hacquard, & Phillips, submitted; Xiang et al., 2009) have argued that apparent mismatches between online and offline responses reflect a single, procedural grammar that is implemented in a noisy cognitive architecture. Under this view, linguistic illusions are argued to arise from limitations of general-purpose memory retrieval and cognitive control mechanisms that create the opportunity for error.

Lewis and Phillips (in press) have recently suggested an attractive possible explanation for the mismatch between online and offline responses, in which the licensing mechanisms and the representation are held constant, without recourse to a two-systems view. They suggest that the mismatch between online and offline responses reflects improvement over time in the signal-to-noise ratio in the responses, rather than the use of distinct online and offline representations. They reasoned that if slow judgments involve repeated attempts at retrieval in a noisy memory architecture for the parser, then increased time for a judgment could lead to improved accuracy. For example, an outcome of an illusion that has a 255 probability of occurrence on a single retrieval trial will have a substantially reduced probability of being the dominant outcome over the course of multiple retrieval trials, leading to greater grammatical accuracy.

To demonstrate that iterative memory sampling leads to improved accuracy over time, I provide ACT-R computational simulations for a relatively uncontroversial linguistic illusion involving agreement attraction. Specifically, I implemented the ACT-R model as described in the dissertation for the agreement attraction sentence in (1), from Dillon et al. (2013).

(1) *The executive who oversaw the managers apparently were …

Constituent creation times and the retrieval schedule are described in Tables A.1 and A.2. The model parameters are described in Table A.3.

Table A.1: Constituent creation times and feature composition for the critical agreement attraction sentence.

	NP-Target	VP1	NP-Distractor
Time	1083	2142	3225
Category	NP	VP	NP
Number	sg	-	pl
Command	+	-	-
Embedding	IP1	IP2	IP2

Table A.2: Schedule of retrievals and cue sets for the critical agreement attraction sentence. VP1 = retrieval to resolve subject gap inside the relative clause. VP2 = critical agreement retrieval at the main clause verb.

	VP1	VP2
Time	3425	4531
Category	NP	NP
Number	-	Pl
Command	+	+
Embedding	IP1	IP1

Table A.3: Summary of ACT-R model parameters

Parameter	Value
Latency factor (*F*)	0.60
Total goal activation (*G*)	1.00
Noise (*ans*)	0.45
Maximum associative strength (fan)	1.50
Decay (*d*)	0.50
Maximum difference (*P*)	-0.50

100 Monte Carlo simulations were run to obtain a distribution of activation values for the target and a corresponding distribution of activation values for the distractor, with each Monte Carlo simulation representing a single memory retrieval trial. To simulate the effect of iterative memory sampling, the

model was run an additional 100 times, with each run repeating the same retrieval 20 times. For each run, the activation values for the target and distractor were averaged separately, obtaining two distributions. The results of the computational simulations are shown in Figure A.1.

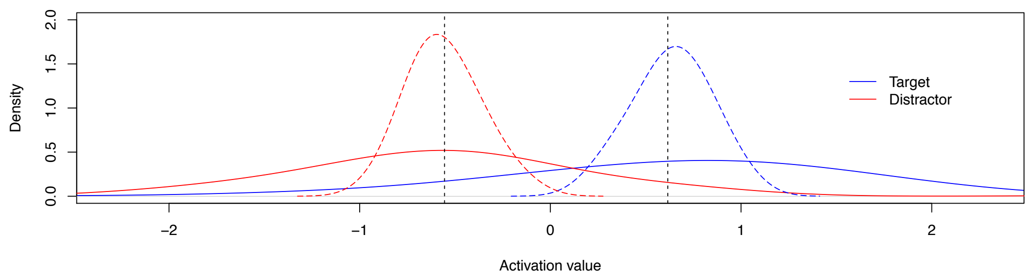


Figure A.1: Results of the computational model simulating the effect of iterative memory sampling. The distributions of activation values for the target are in blue, and the distributions of activation values are in red. Solid lines reflect distributions for a single memory retrieval trial, and dotted lines reflect distributions for iterative memory sampling.

The results of the computational simulations show a substantial overlap in the distributions of activation values for the target and the distractor for individual memory retrieval trials (solid lines). However, the area of the overlap decreases when the same retrieval is executed 20 times (dotted lines), with a regression toward the means (vertical black lines). These results show that with iterative memory sampling, the target gains a clear activation advantage over time. Since the probability of retrieving a given item is proportional to its activation value (in accordance with the equations in Chapter 5), the activation advantage that the target gains with iterative memory sampling will increase its likelihood of retrieval, thereby ensuring proper detect of the ungrammaticality over time.

In short, the computational simulations show that it is possible to capture the effect of improved accuracy over time by holding constant the representation, and assuming a noisy memory retrieval architecture that is capable of iterative sampling.

Appendix B: Modeling two-stage encoding

In Chapter 5, I suggested that the fleeting NPI illusion could be the product of a multiple-stage encoding scheme. There are several mathematical models of feature binding and compression that could be extended to explicitly characterize the computational character of multiple-stage encoding, including tensor-product variable binding (e.g., Smolensky, 1990), Spatter Code (e.g. Kanerva, 1994, 1996, 1997), vector-symbolic architectures (e.g., Gayler, 2003; Sommer & Kanerva, 2006), holographic reduced representation (Plate, 1991, 1994, 2003), and context-dependent thinning (e.g., Rachkovskij & Kussel, 2001). A feature shared by these models is that linguistic features and combinatorial structures are represented as high-dimensional vectors that are manipulated by operations that generate new high-dimensional vector representations (see Kanerva, 2009, for a review). For example, Plate (1991, 1994, 2003) proposed a model of "holographic reduced representations" (HRRs) in which the binding of two vector representations in hyper-dimensional space can be described as a compression of their tensor product to a vector representation that is of the same dimension as each of the sub-components, i.e. the size of the representation does not increase as more structure is added (see also Hinton, 1990). This property is potentially important for cognitive models of feature binding in language comprehension especially given the stringent limit on the amount of information that can concurrently occupy working memory.

Crucially, for the purpose of capturing the effects of multiple-stage encoding in sentence comprehension, an HRR is resistant to the kind of partial-match effects that can give rise to linguistic illusions: feature binding within the HRR framework creates a representation that is completely dissimilar to any of its bound features, and since the sub-component features are no longer transparently accessible, the representation must exhibit an all-or-none match to the retrieval probe in order to be retrieved from memory. However, there are numerous mathematical algorithms for generating reduced representations, such as convolution (Plate, 1991, 1994, 2003), element-wise multiplication (Gayler, 2003; Kanerva, 1994, 1996, 1997), and permutation-based thinning (Rachkovskij & Kussel, 2001). Below, I discuss the details I tensor products and Holographic reduced representations.

A tensor product scheme binds features together into a single, unitized encoding by taking the outer product of the vector representations for two features. For example, the feature vectors for thematic role and negation in (1a) may be combined as shown in (1b).

(1) a. Feature vectors

AGENT = [123]
NEGATION = [abc]

b. Tensor product binding

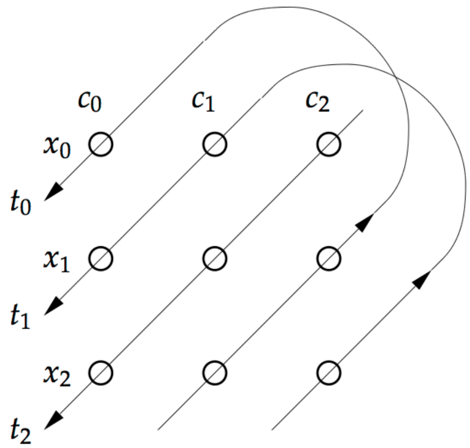
$$\begin{pmatrix}1\\2\\3\end{pmatrix} \otimes \begin{pmatrix}a\\b\\c\end{pmatrix} = \begin{pmatrix}1a & 1b & 1c\\2a & 2b & 2c\\3a & 3b & 3c\end{pmatrix}$$

However, there is one key problem with using a tensor product scheme to bind linguistic features together into a single encoding. In particular, as the number of features that are combined increases, the size of the tensor product code grows exponentially. The issue of the size of the code is particularly important for developing a cognitively plausible theory of memory encoding, especially given the stringent limitations on working memory capacity.

There are several solutions to the problem of code size, such as limiting the depth of the composition (Smolensky, 1991), discard elements (Metcalfe, 1982), or use infinite vector representations (Murdock, 1982, 1993). Plate (1991, 1994, 2003) proposes an alternative solution that uses Holographic Reduced Representations (HRRs), which rely on convolution-correlation based matrix memories. Rather than multiplication, HRRs use circular convolution to recursively bind information together, as defined by Equation (1).[25] This operation can be represented as in Figure B.1.

$$\mathrm{z} = \mathrm{x} * \mathrm{y} \ \text{ where } z_j = \sum_{k=0}^{n-1} x_k \, y_{j-k} \text{ for } j = 0 \text{ to } n-1 \tag{1}$$

[25] Convolution and correlation are used primarily in image and signal processing (Gabel & Roberts, 1973), and they are the core mathematical operations of holography, hence the term "holographic", and the term "reduced" refers to the compression of the representation, in the sense of Hinton (1990). Subscripts are interpreted as modulo *n*, giving the operation its circular nature.



$$\mathbf{t} = \mathbf{c} \circledast \mathbf{x}$$

$$\begin{aligned} t_0 &= c_0x_0 + c_2x_1 + c_1x_2 \\ t_1 &= c_1x_0 + c_0x_1 + c_2x_2 \\ t_2 &= c_2x_0 + c_1x_1 + c_0x_2 \end{aligned}$$

$$t_j = \sum_{k=0}^{n-1} c_k x_{j-k}$$

for $j = 0$ to $n - 1$
(Subscripts are modulo-n)

Figure B.1: Circular convolution represented as the compressed tensor outer product t of the argument vectors c and x for n=3. The values *j* and k represent the row and column indices. The convolution of the features x and y is calculated as the summation (represented by the lines) of the outer product elements along path of the wrapped diagonals. Figure from Plate (1994).

Importantly for reasons of memory capacity limitations, the size of the bound representation does not increase as more features are bound together, as the circular convolution of two *n*-dimensional vectors produces a vector with dimensionality *n* using modulo subscripts. Furthermore, circular convolution produces unique associations and the individual components of the resulting, compressed representation are not independently accessible unless the entire reduced description is decoded by correlation, as defined by Equation 2. These features are critical for capturing the effects of representational change that I proposed in Chapter 5, as the intact HRR will be resistant to the kind of partial-matching that can give rise to linguistic illusions: feature binding as the result of convolution creates a representation that is completely dissimilar to any of its bound features, and since the sub-component features are no longer transparently accessible, the representation must exhibit an all-or-none match to the retrieval probe in order to be successfully retrieved from memory.

$$\mathrm{w} = \mathrm{x} \# \mathrm{z} \text{ where } w_j = \sum_{k=0}^{n-1} x_{k-j}\, z_k \text{ for } j = 0 \text{ to } n-1 \qquad (2)$$

However, one relatively minor issue with convolution, as defined by Equation 1, is it can be computationally costly, since it involves taking the sum of products. Specifically, convolution with modulo subscripts takes $O(n^2)$ time to compute. One solution to this problem that Plate suggests is to perform convolution in the cyclic domain over Fast Fourier transform (FFT) of the component vectors, which involves element-wise multiplication of the two feature vectors. Equation 3 relates convolution and the Fourier transform, where $\odot$ is element-wise multiplication of two vectors. Convolution via FFT takes only $O(n \log n)$ time to compute.

$$\mathrm{x} \circledast \mathrm{y} = \mathrm{f}'(\mathrm{f}(\mathrm{x}) \odot \mathrm{f}(\mathrm{y})) \qquad (3)$$

In previous research, HRRs have been easily modeled in connectionist systems, as a consequence of their recursive and dimensionality properties. I take a new approach here, integrating HRRs into the ACT-R framework, as described by Lewis and Vasishth (2005), and as implemented in the dissertation. In particular, I used HRRs to simulate the effects of representational change, as described in Chapter 5. The model that I used is based on an integrated implementation in the R software environment of the core ACT-R equations, as described in Lewis and Vasishth (2005), as implemented in the dissertation, and the HRRs, as described in Plate (2003).[26] As an initial test of the model, I simulated presence and absence of the NPI illusion, specifically the fleeting time profile observed for a post-verbal NPI, e.g., *any* as in Experiments 2-3 of Chapter 5. Constituent creation times and the retrieval schedule are the same as used in Chapter 5. The representational change was implemented at the main clause verb (see the general discussion of Chapter 5 for discussion about possible triggers of representational change). For this initial test, I used the parameters described in Table B.1. 1000 Monte Carlo simulations were run.

Table B.1: Parameters for the ACT-R/HRR hybrid model of representational change.

Parameter	**Value**
Latency factor (*F*)	3.25
Total goal activation (*G*)	1.00
Noise (*ans*)	0.20
Maximum associative strength (fan)	1.50
Decay (*d*)	0.50
Maximum difference (*P*)	-0.60

The results of the computational simulations are shown in Figures B.2 and B.3.

[26] The library of HRR functions were written by Plate, and they are freely available online at: https://r-forge.r-project.org/projects/vsa/.

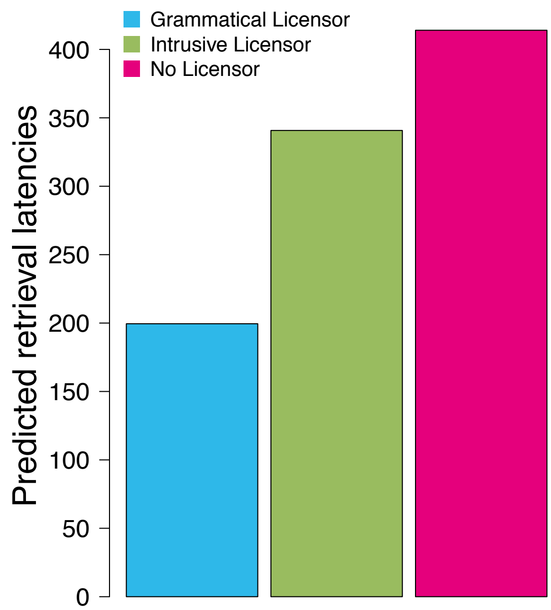


Figure B.2: Illusory licensing effect for the post-verbal NPI *any* before representational change has been triggered.

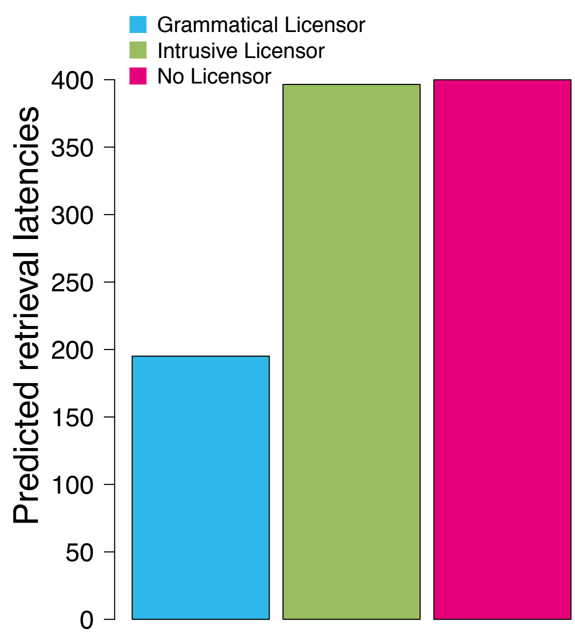


Figure B.3: Illusory licensing effect for the post-verbal NPI *any* after representational change has been triggered.

Results of the computational simulation experiment using the hybrid ACT-R/HRR model revealed contrasting illusory licensing profiles for the post-verbal *any*. The predicted profile is qualitatively similar to the observed profile, making this experiment a success. However, further investigation is needed to determine the overall success of this model. In particular, a task for future work is to test the model on additional structural configurations (e.g., the position and timing manipulations reported in Chapter 5). We must also determine the range of predictions of different assumptions about the trigger, which was the main clause verb in these simulations, and we must explore the range of predictions of different representational assumptions beyond HRRs, as discussed in Chapter 5.

Bibliography

Abney, S. (1991). Parsing by Chunks. In R. Berwick, S. Abney & C. Tenny (Eds.), *Principle-Based Parsing*. Dordrecht: Kluwer Academic Publishers.

Alcocer, P. (2011). *There is more than one way to search working memory.* (Unpublished Ph.D. qualification paper). University of Maryland, College Park, MD.

Alcocer, P., & Phillips, C. (2009). *A cross-language reversal in illusory agreement licensing*. University of California, Davis, CA: Poster at the 22nd CUNY conference on human sentence processing.

Alcocer, P., & Phillips, C. (2012 ms.). *Using relational syntactic constraints in content-addressable memory architectures for sentence processing.* Unpublished manuscript.

Alcocer, P., Phillips, C., França, A. I., & Maia, M. (2010). *Structure sensitive and insensitive retrieval of subjects in Brazilian Portuguese*. NYU, NYC: Talk at the 23rd CUNY conference on human sentence processing.

Altmann, E. M., & Gray, W. D. (2002). Forgetting to remember: The functional relationship of decay and interference. *Psychological Science, 13*, 27-33.

Anderson, J. R. (1974). Retrieval of propositional information from long-term memory. *Cognitive Psychology, 6*, 451-474.

Anderson, J. R. (2007). *How can the human mind occur in the physical universe?*. NYC: Oxford University Press.

Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., & Qin, Y. (2004). An integrated theory of the mind. *Psychological Review, 111*, 1036-1060.

Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Anderson, J. R., & Reder, L. M. (1999). The fan effect: New results and new theories. *Journal of Experimental Psychology: General, 128*, 186-197.

Anderson, J. R., & Schooler, L. J. (1991). Reflections of the environment in memory. *Psychological Science, 2*, 396-408.

Antón-Mendex, I., Nicol, J., & Garrett, M. (2002). The relation between gender and number agreement processing. *Syntax, 5*, 1-25.

Aoshima, S., Yoshida, M., & Phillips, C. (2009). Incremental processing of coreference and binding in Japanese. *Syntax, 12*, 93-134.

Baayen, R. H., Davidson, D., & Bates, D. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of memory and Language, 59*, 390-412.

Baayen, R. H. (2008). *Analyzing linguistic data: a practical introduction to statistics using R*. Cambridge, UK ; New York: Cambridge University Press.

Baddeley, A. D. (1986). *Working memory*. New York: Oxford University Press.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement, and working memory retrieval in sentence production: Evidence from gender and case in Slovak. *Journal of Memory and Language, 56*, 65-85.

Badecker, W., & 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*, 748-769.

Bader, M., Meng, M., & Bayer, J. (2000). Case and reanalysis. *Journal of Psycholinguistic Research, 29*, 37-52.

Baker, M. (2008). *The syntax of Agreement and Concord*. Cambridge, UK: Cambridge University Press.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for testing interactions in linear mixed-effects models. *Frontiers in Psychology, 4*, 328.

Bates, D., Maechler, M. & Bolker, B. (2011). lme4: Linear mixed-effects models using S4 classes.

Baumgartner, G. (1960). Indirekte Gröβenbestimmung der rezeptiven Felder der Retina beim Menschen mittles der Hermannschen Gittertäuschung [Abstract]. *Pflügers Archiv Für Die Gesamte Psychologie, 272* 21-22.

Berman, M. G., Jonides, J., & Lewis, R. L. (2009). In search of decay in verbal short-term memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 35*, 317-353.

Berwick, R., & Weinberg, A. (1984). *The grammatical basis of linguistic performance: Language use and acquisition*. Cambridge, MA: MIT Press.

Bever, T. (1970). The cognitive basis for linguistic structures. In R. Hayes (Ed.) *Cognition and language development.* New York: Wiley.

Bicknell, K., Tanenhaus, M. K., & Jaeger, F. (2014). *Listeners maintain and rationally update uncertainty about prior words in spoken comprehension*. The Ohio State University: Talk at the 27th CUNY conference on human sentence processing.

Bock, K., & Miller, C. A. (1991). Broken Agreement. *Cognitive Psychology, 23*, 45-93.

Boland, J., Tanenhaus, M. K., Garnsey, S. M., & Carlson, G. N. (1995). Verb argument structure in parsing and interpretation: evidence from *wh*-questions. *Journal of Memory and Language*, *34,* 774-806.

Box, G. E., & Cox, D. R. (1964). An analysis of transformations. *Journal of the Royal Statistical Society Series B (Methodological),* 211-252.

Broadbent, D. E. (1958). *Perception and communication*. London: Pergamon Press.

Brysbaert, M., & Mitchell, D. (2000). The failure to use gender information in parsing: A comment on van Berkum, Brown, & Hagoort (1999). *Journal of Psycholinguistic Research, 29*, 453-455.

Caplan, D. (1972). Clause boundaries and recognition latencies for words in sentences. *Perception and Psychophysics, 12*, 73-77.

Caplan, D., & Waters, G. (2013). Memory mechanisms supporting syntactic comprehension. *Psychonomic Bulletin Review,20*, 243-268.

Chien, Y.C., & Wexler, K. (1990). Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition, 1*, 225-295.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the "logicality" of language. *Linguistic Inquiry, 37*, 535-590.

Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge, MA: MIT Press.

Chomsky, N. (2000). Minimalist inquiries: The framework. In R. Martin, D. Michaels & J. Uriagereka (Eds.), *Step by Step: Essays in Minimalist syntax in Honor of Howard Lasnik*. Cambridge, MA: MIT Press.

Chomsky, N., & Lasnik, H. (1993). The theory of principles and parameters. In J. Jacobs, A. Von Stechow & W. Sterneveld (Eds.), *Syntax: An international handbook of Contemporary Research, Vol. 1*. Berlin: Mouton de Gruyter.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. Duncan Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology, Vol. 2*. New York: Wiley.

Chomsky, N. (1981). *Lectures on government and binding*. Berlin: Mounten de Gruyter.

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

Chow, W. Y., Lau, E. F., Wang, S., & Phillips. C. (submitted). Electrical brain potentials reveal temporal dynamics of word prediction during language comprehension.

Chun, M. M., & Potter, M. C. (2000). A two-stage model for multiple target detection in RSVP. *Journal of Experimental Psychology: Human Perception and Performance*, *21*, 109-127.

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

Clark, S. E., & Gronlund, S. D. (1996). Global matching models of recognition memory: How the models match the data. *Psychonomic Bulletin & Review, 3*, 37-60.

Clifton, C., Frazier, L., & Deevy, P. (1999). Feature manipulation in sentence comprehension. *Rivista di Linguistica*, *11*, 11-39.

Conroy, A., Takahashi, E., Lidz, J., & Phillips, C. (2009). Equal treatment for all antecedents: How children succeed with Principle B. *Linguistic Inquiry, 40*, 446-486.

Cunnings, I., & Felser, C. (2012). The role of working memory in the processing of reflexives. Language and Cognitive Processes, *28*, 188-219.

Cunnings, I., & Sturt, P. (2012). *The time-course of reference resolution in picture noun phrases: Evidence from eye-movements during reading*. CUNY, NYC: Poster at the 25th Annual CUNY Conference on Human Sentence Processing.

Dillon, B. (2011). *Structured access in sentence comprehension.* (Unpublished Ph.D. dissertation). University of Maryland, College Park.

Dillon, B., Chow, W. Y., Wagers, M., Guo, T., Liu, F., & Phillips, C. (submitted). The structure-sensitivity of search: Evidence from Mandarin Chinese.

Dillon, B., Levy, J., Staub, A., & Clifton, C. (2014). *Linear order effects in agreement: Evidence from English wh-questions*. The Ohio State University: Talk at the 27th CUNY conference on human sentence processing.

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

Drenhaus, H., Saddy, D., & Frisch, S. (2005). Processing negative polarity items: When negation comes through the backdoor. In S. Kepser, & M. Reis (Eds.), *Linguistic evidence: Empirical, theoretical, and computational perspectives* (pp. 145-165). Berlin: de Gruyter.

Eberhard, K., Cutting, J., & Bock, K. (2005). Making syntax of sense: number agreement in sentence production. *Psychological Review*, *112*, 531-559.

Ferreira, F. (2003). The misinterpretation of noncanonical sentences. *Cognitive Psychology*, *47*, 164-203.

Ferreira, F., & Patson, N. D. (2007). The 'Good Enough' Approach to Language Comprehension. *Language and Linguistics Compass*, *1*, 71-83.

Fodor, J. A., Bever, T. G., & Garrett, M. F. (1974). *The Psychology of Language*. New York: McGraw-Hill.

Franck, J., Vigliocco, G., & Nicol, J. (2002). Subject-verb agreement errors in French and English: The role of syntactic hierarchy. *Language and Cognitive Processes*, *17*, 371-404.

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

Gayler, R. W. (2003). Vector symbolic architectures answer Jackendoff's challenges for cognitive neuroscience. In P. P. Slezak (Ed.), *Proceedings of the joint international conference on cognitive science*. University of New South Wales.

Geier, J., Bernáth, L., Hudák, M., & Séra, L. (2008). Straightness as the main factor of the Hermann grid illusion. *Perception, 37*, 651-655.

Gelman, A., & Hill, J. (2007). *Data analysis using regression and multilevel/hierarchical models*. New York: Cambridge University Press.

Giannakidou, A. (1998). *Polarity sensitivity as (non)veridical dependency.* Amsterdam, Philadelphia: John Benjamins.

Giannakidou, A. (2001). The meaning of free choice. *Linguistics and Philosophy*, *24*, 659-735.

Giannakidou, A. (2011). Positive polarity items and negative polarity items: variation, licensing, and compositionality. In C. Maienborn, K. von Heusinger & P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*. Berlin: Mouton de Gruyter.

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

Gibson, E. (2000). Dependency locality theory: A distance-based theory of linguistic complexity. In A. Marantz, Y. Miyashita & W. O'Neil (Eds.), *Image, language, brain: Papers from the first mind articulation project symposium*. Cambridge, MA: MIT Press.

Gibson, E., Paintadosi, S. T., Brink, K., Bergen, L., Lim, E., & Saxe, R. (2013). A Noisy-Channel Account of Crosslinguistic Word-Order Variation. *Psychological Science, 24*, 1079-1088.

Gibson, E., & Thomas, J. (1999). Memory limitations and structural forgetting: The perception of complex ungrammatical sentences as grammatical. *Language and Cognitive Processes*, *14*, 225-248.

Gillespie, M., & Pearlmutter, N. (2011). Hierarchy and scope of planning subject-verb agreement production. *Cognition, 118*, 377-397.

Gillund, G., & Shiffrin, R. M. (1984). A retreival model for both recognition and recall. *Psychological Review, 91*, 1-67.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *26*, 1411-1423.

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97-104.

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

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

Grodner, D., & Gibson, E. (2005). Consequences of the serial nature of linguistic input. *Cognitive Science*, *29*, 261-290.

Hartsuiker, R., Schriefers, H., Bock, K., & Kikstra, G. (2003). Morphophonological influences on the construction of subject-verb agreement. *Memory & Cognition, 31*, 1316-1326.

Häussler, J., & Bader, M. (2012). *Locality and anti-locality effects in German: Insights from relative clauses*. CUNY, NYC: Poster at the 25th CUNY conference on human sentence processing.

Hawkins, J. A. (1994). *A performance theory of order and constituency*. New York: Cambridge University Press.

Herd, S. A., & O'Reilly, R. C. (2005). Serial visual search from a parallel model. *Vision Research, 45*, 2987-2992.

Hinton, G. E. (1990). Mapping part-whole hierarchies into connectionist networks. *Artificial Intelligence*, *46*, 47-75.

Hintzman, D. L. (1984). MINERVA2: A simulation model of human memory. *Behavior Research Methods, Instruments, and Computers, 16*, 96-101.

Hintzman, D. L. (1988). Judgments of frequency and recognition memory in a multiple-trace memory model. *Psychological Review, 95*, 528-551.

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes, 26*, 376-405.

Holcombe, A. O., & Clifford, C. W. G. (2012). Failures to bind spatially coincident features: comment on Di Lollo. *Trends in Cognitive Science*, *16*, 402.

Horn, L. R. (2010). Polarity after the Thirty Years War: What's the point? Seminar presentation at Rutgers University.

Hsu, J. R., Cairns, H. S., & Fiengo, R. (1985). The development of grammars underlying children's interpretation of complex sentences. *Cognition, 20*, 25-48.

Israel, M. (2004). The pragmatics of polarity. In L. Horn, & G. Ward (Eds.), *The handbook of pragmatics* (pp. 701-723). Oxford: Blackwell.

Jonides, J., Lewis, R. L., Nee, D. E., Lustig, C. A., Berman, M. G., & Moore, K. S. (2008). The mind and brain of short-term memory. *Annual Review of Psychology*, *59*, 193-224.

Kadmon, N., & Landman, F. (1993). Any. *Linguistics and Philosophy*, *16*, 353-422.

Kanerva, P. (1994). The spatter code for encoding concepts at many levels. In M. Marinaro, & P. G. Morasso (Eds.), *Proceedings of the International Conference on Artificial Neural Networks (ICANN '94)*. Springer-Verlag.

Kanerva, P. (1996). Binary spatter-coding of ordered K-tuples. In C. von der Malsburg, W. von Seelen, J. C. Vorbruggen & B. Sendhoff (Eds.), *Artificial Neural Networks - ICANN 96 proceedings (Lecture Notes in Computer science, vol. 1112)*. Berlin: Springer.

Kanerva, P. (1997). Fully distributed representation. *Proceedings of the 1997 Real World Computing Symposium*. Real World Computing Partnership.

Kanerva, P. (2009). Hyperdimensional Computing: An Introduction to Computing in Distributed Representation with High-Dimensional Random Vectors. *Cognitive Computation*, *1*, 139-159.

Kawamoto, A. (1988). Distributed representations of ambiguous words and their resolution in a connectionist network. In S. L. Small, G. W. Cottrell & M. K. Tanenhaus (Eds.), *Lexical ambiguity resolution: Perspectives from psycholinguistics, neuropsychology, and artificial intelligence*. San Mateo, CA: Morgan Kaufmann.

Kawasaki, N. (1993). *Control and arbitrary interpretation in English.* (Unpublished Ph.D. dissertation). University of Massachusetts, Amherst.

Kazanina, N., Lau, E., Lieberman, M., Yoshida, M., & Phillips, C. (2006). The effect of syntactic constraints on the processing of backwards anaphora. *Journal of Memory and Language, 56*, 384-409.

Kazanina, N., & Phillips, C. (2010). Multiple routes to pronoun resolution in Russian cataphora. *Quarterly Journal of Experimental Psychology, 63*, 371-400.

Kennison, S. M., & Trofe, J. L. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, *32*, 355-378.

Kimball, J. (1973). Seven principles of surface structure parsing in natural language. *Cognition*, *2*, 15-47.

King, J., Andrews, C., & Wagers, M. (2012). *Do reflexives always find a grammatical antecedent for themselves?* CUNY, NYC: Poster at the 25th CUNY conference on human sentence processing.

Knuth, D. (1965). *The art of computer program. Volume 1: Fundamental algorithms.* Reading, MA: Addison-Wesley.

Kohonen, T. (1980). *Content-addressable memories*. Berlin; New York: Springer-Verlag.

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

Kowalski, A., & Huang, Y-T. (2014). Working memory, inhibition, and grammatical role assignment. Presentation at the conference on Architectures and Mechanisms for Language Processing (AMLaP). Edinburgh.

Krifka, M. (1995). The semantics and pragmatics of weak and strong polarity items. *Linguistic Analysis*, *25*, 209-257.

Kush, D. (2013). *Respecting relations: Memory access and antecedent retrieval in incremental sentence processing.* (Unpublished Ph.D. Dissertation). University of Maryland, College Park.

Kush, D., Lidz, J., & Phillips, C. (submitted). Configuration-sensitive retrieval: Resisting interference in processing bound variable pronouns.

Kush, D., Lidz, J., & Phillips, C. (2012). *Interference-insensitive local anaphora resolution: Evidence from Hindi reciprocals.* CUNY, NYC: Poster at the 25th CUNY conference on human sentence processing.

Kutas, M., & Hillyard, S. A. (1980). Reading senseless sentences: Brain potentials reflect semantic incongruity. *Science*, *207*. 203-205

Ladusaw, W. A. (1979). *Negative polarity items as inherent scope relations.* (Unpublished Ph.D. Dissertation). University of Texas, Austin.

Lago, S., Alcocer, P., & Phillips, C. (2011). *Agreement attraction in Spanish: Immediate vs. delayed sensitivity*. Stanford University, CA: Poster at the 24th Annual CUNY conference on human sentence processing.

Lago, S., Shalom, D., Sigman, M., Lau, E., & Phillips, C. (2014). *Yo pienso, tu piensas: Cross-linguistic attraction effects in agreement comprehension*. The Ohio State University, OH: Poster at the 27th CUNY conference on human sentence processing.

Lago, S., Shalom, D., Sigman, M., Lau, E., & Phillips, C. (submitted). Agreement processes in Spanish comprehension.

Landau, I. (2001). *Elements of control: Structure and meaning in infinitival constructions*. Dordrecht, The Netherlands: Kluwer.

Levy, R., Bicknell, K., Slattery, T., & Rayner, K. (2009). Eye movement evidence that readers maintain and act on uncertainty about past linguistic input. *Proceedings of the National Academy of Sciences, 106*, 21086-21090.

Levy, R. (2008). A noisy-channel model of ration human sentence comprehension under uncertain input. *Proceedings of the 2008 Conference on Empirical Methods in Natural Language Processing* (EMNLP). Honolulu, HI. Association for Computational Linguistics.

Lewandowsky, S., & Murdock, B. B. J. (1989). Memory for serial order. *Psychological Review, 96*, 25-53.

Lewandowsky, S., & Oberauer, K. (2008). The word-length effect provides no evidence for decay in short-term memory. *Psychonomic Bulletin & Review, 15*, 875-888.

Lewis, R. L. (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research*, *25*, 93-115.

Lewis, R. L., & Badecker, W. (2010). *Sentence production and the declarative and procedural components of short term memory*. New York University, NYC: Talk at the 23rd conference on human sentence processing.

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

Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science*, *10*, 447-454.

Lewis, S., & Phillips, C. (in press). Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research,*

Linebarger, M. (1987). Negative polarity and grammatical representation. *Linguistics and Philosophy*, *10*, 325-387.

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

Lukyanenko, C., Conroy, A., & Lidz, J. (2014). Is she patting Katie? Constraints on promonimal reference. *Language, Learning, and Development, 6*, 258-282.

MacWhinney, B., & Bates, E., (1989). *The crosslinguistic study of sentence processing.* New York: Cambridge University Press.

MacWhinney, B., Bates, E., & Kliegl, R. (1984). Cue validity and sentence interpretation in English, German, and Italian. *Journal of Verbal Learning and Verbal Behavior*, *23,* 127-150.

Marslen-Wilson, W. & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition*, *8*, 1-71.

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

Martin, A. E., & McElree, B. (2009). Memory Operations That Support Language Comprehension: Evidence From Verb-Phrase Ellipsis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1231-1239.

Martin, A. E., & McElree, B. (2011). Direct-access retrieval during sentence comprehension: Evidence from Sluicing. *Journal of Memory and Language*, *64*, 327-343.

McDaniel, D. & Cairns, H. S. (1990). The processing and acquisition of control structures by young children. In L. Frazier, & J. de Villiers (Eds.), *Language Processing and Language Acquisition*. Dordrecht, Holland: Kluwer Academic Publishers.

McElree, B. (1998). Attended and non-attended states in working memory: Accessing categorized structures. *Journal of Memory and Language*, *38*, 225-252.

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 817-835.

McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation*. San Diego: Academic Press.

McElree, B., & Dosher, B. A. (1989). Serial position and set size in short-term memory: Time course of recognition. *Journal of Experimental Psychology: General*, *18*, 346-373.

McElree, B., & Dosher, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General, 122*, 291-315.

McElree, B., & Griffith, B. A. (1995). Syntactic and thematic processing in sentence comprehension: Evidence for a temporal dissociation. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *21*. 134-157.

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

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

McKone, E. (1995). Short-term implicit memory for words and non-words. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 21*, 1108-1126.

McKone, E. (1998). The decay of short-term implicit memory: Unpacking lag. *Memory & Cognition, 26*, 1173-1186.

McRoy, S. W., & Hirst, G. (1990). Race-based parsing and syntactic disambiguation. *Cognitive Psychology, 14*, 313-353.

Metcalfe, E. (1982). A composite holographic associative recall model. *Psychological Review, 89*, 627-661.

Miller, G. A., & Chomsky, N. (1963). Finitary models of language users. In R. Duncan Luce, R. R. Bush & E. Galanter (Eds.), *Handbook of mathematical psychology, Vol. 2*. New York: Wiley.

Murdock, B. B. J. (1982). A theory for the storage and retrieval of item and associative information. *Psychological Review, 89*, 609-626.

Murdock, B. B. J. (1993). TODAM2: A model for the storage and retrieval of item, associative, and serial-order information. *Psychological Review, 100*, 183-203.

Nairne, J. S., Van Arsdall, J. E., Pandeirada, J. N., Cogdill, M. & LeBreton, J. M. (2013). Adaptive memory: The mnemonic value of animacy. *Psychological Science, 24*, 2099-2105.

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

Nairne, J. S. (1990). A feature model of immediate memory. *Memory & Cognition, 18*, 251-269.

Nicol, J. (1988). *Coreference processing during sentence comprehension.* (Unpublished Ph.D. dissertation). MIT, Cambridge, MA.

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

Oberauer, K. (2002). Access to information in working memory: exploring the focus of attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 411-421.

Patil, U., Vasishth, S., & Lewis, R. (2011). *Early retrieval interference in syntax-guided antecedent search*. Stanford, CA: Talk at the 24th CUNY conference on human sentence processing.

Pearlmutter, N., Garnsey, S., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, *41*, 427-456.

Phillips, C. (2006). The real-time status of island phenomena. Language, 82, 795-823.

Phillips, C. (2013). *Encoding and navigating structured representations: Three recent surprises.* University of Connecticut: Talk at the 44th Meeting of the North East Linguistics Society (NELS).

Phillips, C., & Lewis, S. (2013). Derivational order in syntax: Evidence and architectural consequences. *Studies in Linguistics, 6*, 11-47.

Phillips, C., & Parker, D. (2014). The psycholinguistics of ellipsis. *Lingua.*

Phillips, C., Wagers, M., & Lau, E. F. (2011). Grammatical illusions and selective fallibility in real-time language comprehension. In J. Runner (Ed.), *Experiments at the Interfaces* (pp. 147-180). Bingley, UK: Emerald Publications.

Plate, T. (1991). Holographic Reduced Representations: convolution algebra for compositional distributed representations. In J. Mylopoulos, & R. Reiter (Eds.), *Proceedings of the 12th international joint conference on artificial intelligence (IJCAI).* San Mateo, CA.: Kaufmann.

Plate, T. (1994). *Distributed representations and nested compositional structure.* (Unpublished Ph.D. dissertation). University of Toronto.

Plate, T. (2003). *Holographic reduced representation: distributed representation of cognitive structure*. Stanford: CSLI.

Plaut, D. (1997). Structure and function in the lexical system: Insights from distributed models of word reading and lexical decision. *Language and Cognitive Processes, 12*, 765-805.

Pollard, C., & Sag, I. (1992). Anaphors in English and the scope of binding theory. *Linguistic Inquiry, 23*, 261-303.

Pollard, C., & Sag, I. (1994). *Head-driven phrase structure grammar.* Chicago: University of Chicago Press.

Potter, M. C. (1988). Rapid serial visual presentation (RSVP): A method for studying language processing. In D. E. Kieras, & M. A. Just (Eds.), *New methods in reading comprehension research*. Hillsdae, NJ: Erlbaum Press.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1985). *A comprehensive grammar of the English language*. New York: Longman.

R Development Core Team. (2014). R: A language and environment for statistical computing.

Rachkovskij, D. A., & Kussel, E. M. (2001). Binding and normalization of binary sparse distributed representations by context-dependent thinning. *Journal of Neural Computation*, *2*, 411-452.

Rayner, K., & Pollatsek, A. (1989). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry, 24*, 657-720.

Rohde, D. (MIT). *Linger* http://tedlab.mit.edu/~dr/Linger/.

Runner, J., Sussman, R. S., & Tanenhaus, M. K. (2006). Processing reflexives and pronouns in picture noun phrases. *Cognitive Science, 30*, 193-241.

Schrauf, M., Lingelbach, B., Lingelbach, E., & Wist, E. R. (1995). The Hermann grid and the scintillation effect. *Perception, 24*, suppl. 88-89.

Slogget, S. (2013). *Case licensing in processing: Evidence from German*. University of South Carolina: Poster at the 26th CUNY conference on human sentence processing.

Smolensky, P. (1990). Tensor product variable binding and the representation of symbolic structures in connectionist networks. *Artificial Intelligence*, *46*, 159-216.

Sommer, F. T., & Kanerva, P. (2006). Can neural models of cognition benefit from the advantages of connectionism? *Behavioral and Brain Sciences*, *29*, 50-51.

Staub, A. (2010). Response time distributional evidence for distinct varieties of number attraction. Cognition, 114, 447-454.

Steinman, S. B. (1987). Serial and parallel search in pattern vision? *Perception, 16*, 389-398.

Stowe, L. (1986). Evidence for online gap creation. *Language and Cognitive Processes, 1*, 227-245.

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

Tabor, W., Galantucci, B., & Richardson, D. (2004). Effects of merely local syntactic coherence on sentence processing. *Journal of Memory and Language, 50*, 355-370.

Townsend, D. J., & Bever, T. G. (2001). *Sentence comprehension : the integration of habits and rules*. Cambridge, Mass.: MIT Press.

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

Traxler, M. J., Pickering, M. J., & Clifton, C. (1998). Adjunct attachment is not a form of lexical ambiguity resolution. *Journal of Memory and Language, 39*, 558-592.

Treisman, A. (1996). The binding problem. *Current Opinion in Neurobiology*, *6*, 171-178.

Treisman, A., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive Psychology*, *12*, 97-136.

Treisman, A., & Schmidt, H. (1982). Illusory conjunctions in the perception of objects. *Cognitive Psychology*, *14*, 107-141.

Treisman, A., Sykes, M., & Gelade, G. (1977). Selective attention and stimulus integration. In S. Dornicå (Ed.), *Attention and performance VI: proceedings of the Sixth International Symposium on Attention and Performance, Stockholm, Sweden, July 28-August 1, 1975* (pp. 333-361). Hillsdale, NJ: Lawrence Erlbaum.

Trommershäuser, J., Körding, K., & Landy, M. S. (Eds.). (2011). *Sensory Cue Integration*. New York: Oxford University Press.

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

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

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

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

Van Gompel, R., Traxler, M. J., & Pickering, M. J. (2001). Reanalysis in sentence processing: Evidence against current constraint-based and two-stage models. *Journal of Memory and Language, 45*, 225-258.

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

Vasishth, S., & Drenhaus, H. (2011). Locality effects in German. *Dialogue and Discourse, 2*, 59-82.

Vasishth, S., Brüssow, S., Lewis, R. L., & Drenhaus, H. (2008). Processing polarity: how the ungrammatical intrudes on the grammatical. *Cognitive Science, 32*, 685-712.

Verhaeghen, P., Cerella, J., & Basak, C. (2004). A working memory workout: How to change the size of the focus of attention from one to four in ten hours or less. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30*, 1322-1337.

Vigliocco, G., & Franck, J. (2001). When sex affects syntax: Contextual influences in sentence production. *Journal of Memory and Language, 45*, 368-390.

Vigliocco, G., & Nicol, J. (1998). Separating hierarchical relations and word order in language production: Is proximity concord syntactic or linear? *Cognition, 68*, 13-29.

Von Fintel, K. (1999). NPI-licensing, Strawson-entailment, and context-dependenty. *Journal of Semantics*, *16*, 97-148.

Wagers, M., Lau, E., Stroud, C., McElree, B., & Phillips, C. (2009). *Encoding syntactic predictions: evidence from the dynamics of agreement*. University of California, Davis: The 22nd Annual CUNY Conference on human sentence processing.

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

Wagers, M., & Phillips, C. (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics*, *45*, 395-433.

Wellwood, A., Pancheva, R., Hacquard, V., & Phillips, C. (submitted). Deconstructing a comparative illusion.

Whitney, C. S. (2004). *Investigations into the neural basis of structured representations.* (Unpublished Ph.D. dissertation). University of Maryland, College Park.

Wolfe, J. (2007). Guided Search 4.0: Current progress with a model of visual search. In W. Gray (Ed.), *Integrated Models of Cognitive Systems*. New York: Oxford.

Wolfe, J. (2012). The binding problem lives on: comment on Di Lollo. *Trends in Cognitive Science*, *16*, 307-308.

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

Xiang, M., Dillon, B. W., & Phillips, C. (2006). *Testing the strength of the spurious licensing effect for negative polarity items*. New York: Talk at the 19th CUNY conference on human sentence processing.

Xiang, M., Grove, J., & Giannakidou, A. (2013). Dependency dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology,*

Yoshida, M. (2006). *Constraints and mechanisms in long-distance dependency formation.* (Unpublished Ph.D. dissertation). University of Maryland, College Park.

Zwarts, F. (1995). Nonveridical contexts. *Linguistic Analysis, 25*, 286-312.