THE ROLE OF SEMANTIC AND SYNTACTIC FEATURES IN

DEPENDENCY INTERFERENCE EFFECTS

by

Myung Hye Yoo

A dissertation submitted to the Faculty of the University of Delaware in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Linguistics

Spring 2022

THE ROLE OF SEMANTIC AND SYNTACTIC FEATURES IN

DEPENDENCY INTERFERENCE EFFECTS

by

Myung Hye Yoo

Approved: _____
Benjamin Bruening, Ph.D.
Chair of the Department of Linguistics and Cognitive Science

Approved: _____
John A. Pelesko, Ph.D.
Dean of the College of Arts and Sciences

Approved: _____
Louis F. Rossi, Ph.D.
Vice Provost for Graduate and Professional Education and
Dean of the Graduate College

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Rebecca Tollan, Ph.D.
Committee chair

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Satoshi Tomioka, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Zhenghan Qi, Ph.D.
Member of dissertation committee

I certify that I have read this dissertation and that in my opinion it meets the academic and professional standard required by the University as a dissertation for the degree of Doctor of Philosophy.

Signed: _____

Dave Kush, Ph.D.
Member of dissertation committee

ACKNOWLEDGMENTS

<h1>TABLE OF CONTENTS</h1>

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The processing of long-distance dependencies has been one of the primary means of investigating the linguistic components involved in working memory mechanisms behind the maintenance of *wh* fillers and retrieval of the stored filler at the gap site. This thesis investigates how semantic and syntactic features of the filler and an intervening element interact and have an impact on dependency formations. Specifically, this thesis explores the extent to which semantic and syntactic features can modulate the interference effects in dependency formations. First, I examine the effect of NP types concerning the filler and the intervening noun to probe how different semantic information of NP types interacts with contextual information in the processing of filler-gap dependencies. In three experiments, I found that semantic information on different NP types of an intervening noun plays a role in interferences during dependency formations and the semantic property of NP types is a crucial factor determining the susceptibility to contextual information. Secondly, I use the processing of subject-verb agreement dependencies involving the filler-gap dependency to probe how syntactic information, in terms of the argument status of a filler-gap dependency and an intervening element, modulates the difficulty of dependency formations. The processing of subject-verb agreement serves as a tool to explore the difficulty of dependency formations, which can elicit a temporary illusion of grammaticality/ungrammaticality. Two experimental studies reveal that the argument status of a filler-gap dependency can also modulate the processing of

subject-verb agreement. I propose two stages of subject-verb agreement formation, which consider the argument status of both the filler and intervening noun in the memory encoding and retrieval process.

Chapter 1

INTRODUCTION

1.1   General Introduction

Research on sentence comprehension has shown that information from many sources, including lexicon, sentence structures, discourse contexts, and extra-linguistic knowledge are integrated during sentence processing (Gibson & Pearlmutter, 1998; Tanenhaus & Trueswell, 1995). The purpose of research on sentence processing is to advance understanding of how parsers use and integrate sources of information in real-time sentence comprehension.

Sentence structures involving dependency formation are highly informative tools to address the question of how the sources influence the comprehension of complex sentences (Gibson, 1998, 2000; Grodner et al., 2000). Complex sentences associated with long-distance dependencies can cause processing difficulty due to memory constraints, which allows us to explore the process of structure building and memory mechanisms. This dissertation addresses the question of how semantic and syntactic information of noun phrases affect dependency formation ⌐ in sentences involving long-distance dependencies.

Long-distance dependencies involve constructions where an element occurs in a position and must be associated with a non-adjacent position in the sentences. Generative syntax considers the relevant element to undergo "movement" from its initial position to its surface position where it appears in the sentence. This process is sometimes referred to as displacement. Movement is assumed because the theta role of

the element is assigned from its initial position. The displaced element is linked to its original position via a dependency chain between the two positions.

Filler-gap dependencies, one type of long-distance dependencies, include grammatical functions like clefted sentences (as above), relative clauses, *wh*-questions, focusing, or topicalization. When parsers form a filler-gap dependency, it needs to interpret the filler (e.g., *the lawyer*) in the gap position and associate its thematic role with the verb (e.g., *interviewed*). For instance, an argument such as *the lawyer* in (1) – or, under several movement theories, a silent operator which is co-referential with the lawyer –can be moved to the front of the main clause to create a head noun in a cleft sentence. The fronted noun phrase (NP), which is the dependent element, is called *a filler* and the base position of the moved noun phrase, which is the controlling element, is called *a gap,* marked by underlining in (1).

(1) It was <u>the lawyer</u>$_i$ that *the client* interviewed ____$_i$ in a small office.

          FILLER   GAP

Other types of non-local dependencies involve morphological agreement such as subject-verb agreement, in which the verb and the (sometimes, non-adjacent) subject must agree in number, person, and gender features (collectively known as "phi-features"). For instance, the copular verb *was* in (2) requires a singular subject, which agrees with the singular copular verb *was* in number feature. The sentence (2) is syntactically well-formed because the number feature of the main subject *the new executive* matches the number feature of the main verb *was*.

(2) The new executive who oversaw the middle managers apparently was

dishonest about the company's profits.

<div align="right">(Dillon et al., 2013, p. 89)</div>

Similarly, the antecedent of reflexive and the corresponding reflexive must also agree in phi-features. Several studies used gender-biased nouns like *the soldier –* gender-biased to be male - as the antecedent of the reflexive, as in (3), to explore the processing of reflexives (Badecker & Straub, 2002; Sturt, 2003; Xiang et al., 2009)[1]. The number and gender feature of the reflexive must agree in singular and male features of its antecedent, *the tough soldier*. On the assumption that *the tough soldier* is stereotypically male, *herself* is unacceptable to be the corresponding reflexive of *the surgeon* in this sentence.

(3) <u>The tough soldier</u> that Fred treated in the military hospital introduced

*himself/*herself* to all the nurses.

<div align="right">(Xiang et al, 2009, p. 44)</div>

The processing of clefted sentences, an example of filler-gap dependencies, will be discussed in Chapters 2 and 3, and the processing of subject-verb agreement involving filler-gap dependencies, the other type of dependencies, will be discussed in Chapters 4 and 5.

---

[1] Xiang et al (2009) states that previous studies have shown that gender-stereotyped nouns were strong enough for parsers to detect mismatching reflexives upon reading it (Osterhout et al., 1997; Sturt, 2003).

1.2 Interference Effects in Memory Retrieval Mechanism

A number of psycholinguists have adopted a cue-based retrieval memory system (Lewis et al., 2006; Lewis & Vasishth, 2005; McElree, 2000; Ness & Meltzer-Asscher, 2017, 2019; J. A. van Dyke, 2007; J. A. van Dyke & Lewis, 2003; J. van Dyke & McElree, 2006, 2011) to account for the cognitive process of sentence processing.

1.2.1 Encoding of Representation

Items are argued to be stored in memory as chunks, a set of feature-value pairs (Anderson, 2004; Lewis et al., 2006; Miller, 1956; Nairne, 1990). The basic unit of a chunk is defined to be "the representational element that enters into novel relations with other elements (Lewis & Vasishth, 2005, p. 381)." Linguistic items are encoded in the working memory as a bundle of feature-value pairs, which include syntactic and semantic information (e.g., case, thematic role) as well as lexical and morphological information (e.g., category, number feature) when items are first encountered. That is, the semantic and syntactic feature-values are associated with a structural relation between two items. An example of encoding systems is shown in (4).

(4) Melissa knew that **the toy** from her uncle in Bogota **arrived** today

[encoding]

$_{NP}6$	$s7$
category: NP	category: S
head: *toy*	head: OPEN
case: nominative	specifier: $_{NP}6$
number: singular	complement: OPEN

(Lewis et al., 2006, p. 3)

Each category of chunks such as NPs or verbs is represented during incremental processing. For instance, the encoding of *the toy* in (4), is associated with a representation of the NP itself and a representation of the expectation of the predicate. Its lexical and morphological information such as word class as NP, definiteness, singular, inanimate features are stored as $_{NP}6$ in (4) and its syntactic node in the syntactic structure in which the head is empty was stored as $s7$ in (4).

1.2.2   Retrieval in Working Memory

When we consider real-time sentence comprehension, the task of the parser is to first store linguistic information in working memory at the initial encoding stage. Working memory in sentence processing, however, may not be able to hold all the multiple chunks over time in a limited focus of attention. That is, chunks encoded as a set of feature-value pairs in working memory are assumed to be stored in long-term memory, while the upcoming chunks are encoded in working memory. Thus, a dependent element from long-term memory is retrieved at the time when it needs to form the dependency.

The retrieval process, for instance, occurs at the verb *arrived* in (5) ((4) was repeated).

      (5) Melissa knew that **the toy** from her uncle in Bogota **arrived** today.

<div align="right">(Lewis et al., 2006, p. 3)</div>

When the parsers encounter the verb *arrived,* they must associate it with its subject NP to comprehend the sentence. Thus, parsers trace back the subject from long-term memory using retrieval cues in response to the verb, *arrived*. The retrieval cues (e.g., NP, subject) activate features of *the toy* at the verb, *arrived*. This retrieval mechanism is necessary for parsers to integrate the verb with the appropriate subject in order to comprehend the meaning of a sentence and to form filler-gap dependencies or agreement dependencies between a subject and a verb. During the retrieval process of dependency formations, elements intervening between the dependent elements may contribute to the difficulty of retrieval. Now, I discuss how interference of an intervening element arises during dependency formations and present two different kinds of interferences observed in empirical studies.

1.2.3   Interference Effects

According to the cue-based retrieval mechanism, multiple representations that either wholly or partially match the retrieval cues may contribute to interferences on retrieval (Lewis et al., 2006; Lewis & Vasishth, 2005; McElree, 2000; Ness & Meltzer-Asscher, 2017, 2019; J. A. van Dyke, 2007; J. A. van Dyke & Lewis, 2003; J. van Dyke & McElree, 2006, 2011). Interference effects arise when features of another

item match the feature values of a target item, which are activated in parallel to be retrieved[2].

> (6) It was the lawyer$_i$ that *the client* interviewed ___$_i$ in a small
>
>   FILLER   INTERVENOR   GAP
>
> office.

In (6), the features of the filler, *the lawyer*, such as its definite, animate, singular features, are encoded as a chunk in memory once the parser first encounters the filler. Then, the stored linguistic information of the filler must be integrated at the gap site following the verb *interviewed*. The parsers have to seek a noun phrase that has an animate feature from their memory upon encountering the main verb *interviewed* as the object of the verb. However, if the parser needs to process another NP intervening between the dependent elements, then this may incur an additional processing cost. For instance, an NP such as *the client* in (6) intervenes between the filler, *the lawyer*, and the gap position, after the main verb. I will refer to this intervening NP as an "intervenor". This intervenor, *the client,* shares similar features with the filler, as another NP which bears an animate feature. Therefore, the retrieval cues of the filler activate the features of both the filler and the intervenor, which affects the retrieval process by making the retrieval of the filler harder or increasing

---

[2] Interference effects can appear either at the stage of encoding or retrieval, or both. At this point, I mainly introduce interferences during retrieval process and discuss encoding interferences in the next Chapter.

the probability that an intervenor is erroneously retrieved. The interference effect can be of two kinds: *inhibitory* or *facilitatory* interference.

Inhibitory interference occurs when fully or partially matching features of an intervenor disrupt the retrieval of an intended target. A number of studies have provided evidence for similarity-based interference effects during the computation of dependencies (Cunnings & Felser, 2013; P. C. Gordon et al., 2001, 2004, 2006; L. A. Jäger et al., 2015; Villata & Franck, 2020).

(7) a. The worker was surprised that the resident who was living near the
        dangerous <u>warehouse</u> *was complaining* about the investigation.
      b. The worker was surprised that the resident who said that the
        <u>warehouse</u> was dangerous *was complaining* about the investigation.

J. A. van Dyke & Lewis (2003), for example, presented interference effects in terms of syntactically matching features. The NP *the warehouse* in (7a) is not a subject, whereas it is a subject in (7b). Under the assumption that syntactic features play a role in retrieval, *the warehouse* in (7b), which is in a subject position, incurs more processing load on retrieval of the target item, *the resident*. They found a significantly higher reading time at the verb in (7b) than in (7a). Reading times at the verb region were compared because the integration of subject-verb agreement between the target subject (i.e., *the resident*) and the verb (i.e., *was complaining*) must be formed upon encountering the verb for integration. Similarly, semantically matching features of an intervenor with a target item can also lead to processing overload (P. C. Gordon et al., 2001, 2004; J. A. van Dyke, 2007; J. van Dyke & McElree, 2011),

which will be discussed in the next Chapter in details. In brief, an intervenor can play a role as a feature-matching distractor due to its matching cues to the target item. The similarity-based interference effects will be discussed in Chapters 2 and 3.

The intervenor, on the other hand, can rather facilitate dependency processing when a target item mismatches its dependent element in retrieval cues but a feature-matching intervenor is present between the dependency, as in (8a). In grammatically ill-formed sentences, for instance, Wagers et al (2009) showed that reading times at spillover regions after the critical verb were faster when an intervenor agrees with the verb in number feature than when it does not. The main verb *were* agrees with an intervenor *the cells* in number feature in (8a), rather than the target main subject *the key* in (8b). The feature matching intervenor, *the cells* in (8a), led to a temporary illusion of grammaticality and facilitated the processing of the subject-verb dependency.

(8) a. *The key to *the cells* unsurprisingly *were* rusty from many years of disuse.

b. *The key to *the cell* unsurprisingly *were* rusty from many years of disuse.

A series of other studies have found this facilitatory interference effects in sentence productions across languages (Dillon et al., 2013; Sturt & Kwon, 2017; Tucker et al., 2015; Wagers et al., 2009). Facilitatory interference effects on subject-verb agreement dependencies (also referred to as attraction effects) will be discussed in Chapters 4 and 5.

The current research investigates how the linguistic features of an intervenor, situated between the dependency formation chains, exert influence on the resolution of dependencies. The studies on memory retrieval during sentence comprehension have discussed the interference effects of an intervenor to better understand factors that cause memory loads associated with the retrieval mechanism (Badecker & Lewis, 2008; Lewis et al., 2006; Lewis & Vasishth, 2005; Patil et al., 2016; J. A. van Dyke & Lewis, 2003). Retrieval cues under this retrieval mechanism are associated with the semantic and syntactic features of items to be retrieved. It is, however, still unclear the extent to which semantic and syntactic features of the displaced noun (i.e., the filler), and the intervenor noun interact during dependency formations. In terms of interference effects, the extent to which the gradient status of fine-grained semantic and syntactic features can modulate the interference effects remains unsettled.

1.3   Outline of the Dissertation

The goal of this thesis is to explore how linguistic information interacts in working memory during the processing of dependency formations. Theoretical studies have discussed discourse hierarchies and argument hierarchies typologically, but psycholinguistic studies have not yet fully resolved how the semantic and syntactic cues in relation to the hierarchies can play a role in retrieval mechanisms. On this view, I propose to investigate how semantic and syntactic cues contribute to the interference of intervenors in the processing of long-distance dependencies. The thesis comprises five further chapters including a concluding chapter.

First, I examine the role of discourse hierarchies concerning the filler and the intervenor in the processing of filler-gap dependencies. **Chapter 2** presents an

overview of processing theories that propose the effects of NP types on sentence processing. I discuss memory-based models, such as similarity-based interference effects, the binary and continuous metrics under the framework of the DLT (Gibson 1998, 2000), and expectation-based models, such as surprisal theory.

Chapter 3 introduces previous experimental findings on the effect of the intervenor NP types on dependency formations and investigates how semantic factors of different NP types and their interaction with contextual support influence dependency formation in terms of NP types, by looking at the processing of English clefted sentences, a type of filler-gap dependency, in three reading time experiments. Results reveal that the difference in the accommodability of a referent of full NPs like definite and indefinite intervenors in contexts plays a role in the interference during dependency formation, meaning that the effect of intervenor NP type is not simply due to morphological cues of NPs. Pronouns, on the other hand, had a strong advantage over indefinites despite the accommodation costs to a referent. The implications of these results are discussed.

Secondly, I examine the role of argument hierarchies of the main subject and an intervenor in the processing of subject-verb agreement dependencies. These experiments can offer a better understanding of the extent to which the syntactic features can modulate the interference effects in dependency formations. **Chapter 4** presents theoretical backgrounds on memory mechanism and attraction effects that have been widely discussed in subject-verb agreement dependencies. I focus on "agreement attraction" errors in sentence processing, so-called the illusion of grammaticality and ungrammaticality. In addition, previous experimental findings on subject-verb dependencies are explored.

Chapter 5 focuses on investigating the processing of subject-verb number agreement dependency in grammatical and ungrammatical sentences in two separate experiments. The overall findings of this study show that the argument status of both the filler and intervenor modulates the attraction effects in subject-verb agreement dependency formation. This finding necessitates a more nuanced account of agreement attraction effects, which considers the relative distinctiveness of both filler and attractor in memory encoding and agreement dependency formation, rather than the absolute distinctiveness of the attractor alone.

Chapter 6 discusses the implications of overall findings of how semantic and syntactic features of an intervenor mediates the interference during dependency formation and concludes the thesis.

1.3.1   Semantic Features of An Intervenor (Chapters 2 & 3)

In terms of NP types, all types of NP such as pronouns, definites, and indefinites can be an argument of the verb as long as their animacy features can be associated with it because the verb does not set any restriction on NP types. In the following sentence, either pronoun *her*, the definite NP *the girl*, or the indefinite NP *a girl* can be the object argument of the verb met because all these NP types satisfy the animate feature to be an argument of the verb *met*.

(9)   I met {her/a girl/ the girl} after class.

Gordon et al (2001, 2004) and Gordon et al (2006) showed an interference effect of the same lexical NP types in dependency formation and interpreted it as evidence of cue-based retrieval mechanism. For example, it took the longest reading

time at the critical regions for the definite filler to form a dependency when the NP type of an intervenor was definite, the same NP type as the filler, compared to a pronoun or indefinite intervenor.

(10) The barber that {*the lawyer/you/Joe*} admired climbed the mountain.

However, each NP type not only differs in terms of morphological forms but also has different accessibility in the discourse, including different interaction with contexts such as accommodability of presuppositions. In complexity rating and self-paced reading studies, Warren & Gibson (2002, 2005) found the effect of semantic properties of NP types in discourse, following the Givenness hierarchy (GH) of NP types in discourse (Gundel, 2010; Gundel et al., 1993, 2012) within the framework of Gibson (1998, 2000)'s distance-based dependency locality theory (DLT). Gibson's DLT proposes that an intervenor that is newly introduced in discourse induces processing costs in forming dependencies. Warren & Gibson (2002) found the effect of more fine-grained semantic properties of an intervenor in the discourse during dependency formations, following the GH. They applied the GH, a more gradient metric, to the spirit of the DLT, instead of the binary metric proposed by Gibson's DLT. Their experiments showed that the least processing cost incurred when an intervenor is highly accessible in discourse (e.g., *pronouns*), more processing cost when it is a name, the most processing cost when it is a definite description. This thesis, therefore, aims to explore whether the interference effect of intervenors as observed by Gordon et al (2001) is truly a similarity effect or is in fact a more fine-grained discourse-level of the semantic hierarchy and involved with pragmatic factors.

1.3.2   Syntactic Features of An Intervenor (Chapters 4 & 5)

In terms of the syntactic argument status of NPs, several studies on subject-verb dependencies have provided evidence for the view of interference effects of a structurally illicit intervenor on the formation of subject-verb agreement dependencies (Dillon et al., 2013; Franck et al., 2002; Parker, 2017; Parker & An, 2018; N. Pearlmutter et al., 1999; J. van Dyke & McElree, 2011; Wagers et al., 2009). However, mixed findings of interference effects have been observed across the different argument statuses of the intervenor in a wide range of configurations. Recent empirical research on the subject-verb agreement has observed that effects of an intervenor can be modulated by the syntactic prominence of its argument status (i.e., core arguments such as *subject* and *object positions*) (L. Jäger et al., 2015; Parker, 2017; Parker & An, 2018; J. van Dyke & McElree, 2011), but the extent to which the syntactic prominence plays a role in filler-gap dependency formations remains unresolved in empirical studies. The recent research findings suggest that the subject-verb dependencies are susceptible to the status of an intervenor such as its argument hierarchy or thematic-semantic properties associated with a verb. All of these studies explored different configurations with different experimental designs and methodologies. Importantly, however, they focused solely on the status of an intervenor, but not the argument status of both a filler and an intervenor. The main question that this current dissertation addresses is how the effect of intervenor NP types is mediated by the argument status of both (i) a filler and (ii) an intervenor.

Chapter 2

THE SEMANTIC/DISCOURSE FEATURES OF AN INTERVENOR

This chapter focuses on the effect of NP types of an intervenor in dependency formations, beginning with an introduction of existing processing theories of processing mechanisms.

2.1 Processing theories on the effect of NP types in sentence processing

This section introduces two main classes of sentence processing theories that can account for the effect of an intervenor during filler-gap dependency formation: memory-based theories and expectation-based theories. First, the core idea of memory-based theories is that parsers have to store linguistic information during incremental processing and retrieve the necessary information from working memory in order to form the dependencies. Similarity-based interference effect (P. C. Gordon et al., 2001, 2004, 2006) and the dependency locality theory are presented (Gibson, 1998, 2000). Secondly, expectation-based theories are based on the notion that ease of processing is derived from the high frequency and probability of linguistic elements in syntactic structures. The surprisal model (Levy, 2008) is presented as one of the expectation-based theories.

2.1.1 Memory-based Theories

Memory-based theories account for processing difficulty with respect to the limitation of resources in working memory. Parsers store inputs in memory and retrieve an element for dependency formation during incremental sentence processing.

When inputs have substantial information to hold in memory or other inputs interfere with the retrieval, parsers may need processing efforts to retrieve the dependent element due to limited storage of information.

2.1.1.1 Similarity-based Interference Effect

This similarity-based interference effect, or *inhibitory interference* (L. Jäger et al., 2017; Kwon & Sturt, 2016b)*,* supports a cue-based retrieval mechanism (Lewis et al., 2006; Lewis & Vasishth, 2005). Under the cue-based hypothesis, the similarity-based interference effect occurs if increasing the similarity of an intervenor to the retrieval cues for seeking a target item increases processing load at the point of retrieval to form dependencies.

This similarity-based interference, however, can also arise at the stage of encoding a representation in memory when a target item shares features with other elements, so-called *encoding interference* (J. Barker et al., 2001; Hofmeister & Vasishth, 2014; Kush et al., 2015; Oberauer & Lange, 2008; Villata et al., 2018). Some memory-load paradigm studies found interferences upon reading an overlapping NP, which, thus, provided evidence for encoding interference at the overlapping NP region (Kush et al., 2015). In terms of NP types, the processing at an intervenor should be more difficult when an intervenor is similar to the filler than when it is dissimilar to the filler. However, previous studies on the effect of NP types during filler-gap dependency formations found the encoding interferences at the retrieval of the filler, or the verb region (P. C. Gordon et al., 2001, 2004; Warren & Gibson, 2002, 2005). Therefore, it has been difficult to disentangle retrieval and encoding interference in several empirical studies on dependency formations such as gender/number agreement between subject and verb or filler-gap dependencies because interferences during

dependency formations have been observed at the verb region, which is associated with the integration (P. C. Gordon et al., 2002; Hofmeister & Vasishth, 2014)[3]. This finding can be accounted for by the fact that encoding interference can affect retrieval of the target by decreasing the distinctiveness of a target item in memory representation.

In terms of NP types, Gordon and colleagues found interference of NP types in dependency formations in a series of studies (P. Gordon et al., 1993; P. C. Gordon et al., 2001, 2004, 2006). Gordon et al (2001), for instance, contrasted the degree of processing difficulty between subject and object relative clauses as in (11) by manipulating the NP types of an intervenor—definites, pronouns, and names. They found more processing difficulty between subject and object relative clauses during the integration of the filler and its gap when the NP type of an intervenor (e.g., *the lawyer*) was more similar to the filler (e.g., *the barber*) than when the NP type of an intervenor was less similar (e.g., *you, Joe*). They attributed the processing difficulty of a definite intervenor to the similarity of its features to the filler.

(11) a. The barber that the lawyer/you/Joe admired climbed the mountain.

[ORC]

b. The barber that admired the lawyer/you/Joe climbed the mountain.

[SRC]

---

[3] Interference effects at an intervenor can be also associated with expectation-based accounts, which I discuss in this section later. I, therefore, view that encoding interferences of difference NP types of an intervenor contribute to retrieval process at the critical verb region.

The effect of different NP types between the filler and the intervenor observed by Gordon et al (2001) cannot be accounted for by retrieval cues of the verb because the object of the verb does not necessarily have to be a certain NP type. At the point of the verb, *admired*, the appropriate target NP, which is an object of the verb, needs to be retrieved by using retrieval cues. The retrieval cues activate features such as a category NP (or DP) or animate features, but do not activate the features of NP types such as definiteness because definiteness is not required to be an object of the verb, *admired*. Therefore, the effect of NP types on interferences is attributed to encoding interferences, which eventually affects retrieval of the filler at its gap site.(L. Jäger et al., 2017; Villata & Franck, 2020).

As encoding interferences, the definite intervenor (e.g., *the lawyer*) that matches the features of a target item (e.g., *the barber*), which is also definite, degrades a representation of the target item due to feature competition between two similar elements. This degradation arises immediately when the similar intervenor is encountered before the retrieval site. The encoding interference of an intervenor, which happens at the intervenor region, on the encoding of the filler could have contributed to processing loads of retrieving the filler. That is, the effects of similarity between a target and an intervenor in terms of NP types serve as encoding interferences, which weakens the representation of the filler in memory and eventually resulted in interference at retrieval of the filler.

Several researchers have also revealed similarity-based interference effects in the real-time processing of other types of dependencies such as subject-verb dependencies (Franck et al., 2015; J. A. van Dyke, 2007; J. A. van Dyke & Lewis,

2003) and reflexive-antecedent dependencies (Badecker & Straub, 2002; Felser et al.,

2009)[4].

2.1.1.2   Dependency Locality Theory

2.1.1.2.1   Binary Metric

Gibson (2000)'s distance-based dependency locality theory (DLT) accounts for

the processing of complex sentences in terms of the use of computational resources.

Gibson (2000) discusses resources from two aspects of language comprehension: (i)

storage of structure, and (ii) integration cost. The storage of structure is associated

---

[4] For instance, Franck et al's (2015) study on subject-verb agreement dependencies showed more processing difficulty when an intervenor agrees with the verb than when it disagrees with the verb in number feature in French relative clauses as in (1), even the intervenor did not intervene between subject and the verb in syntactic structures.

(1) a. Matching intervenor

Jérôme/parle/à la prisonnière/que/le gardien/sort/parfois/dans la cour.
Jérôme/speaks/to the prisoner-**SG**/that/the guard-SG/takes-SG
out/sometimes/in the yard.

b. Mismatching intervenor
Jérôme/parle/aux prisonnière**s**/que/le gardien/sort/parfois/dans la cour.
Jérôme/speaks/to the prisoners-**PL**/that/the guard-SG/takes-SG
out/sometimes/in the yard.

Similarly, Badecker & Straub's (2002) study on reflexive-antecedent processing showed more processing cost when an intervenor matches with the verb than when it mismatches with the verb in gender feature as in (2).

a. Matching intervenor
John thought that *Bill* owed him another chance to solve the problem.

b. Mismatching intervenor
John thought that *Beth* owed him another chance to solve the problem.

with holding structural information of incomplete dependencies. These storage costs depend on the distance or locality between the head of a dependency and its foot, or a filler and gap. The integration cost components, as the second computational resource, involve both structural and discourse processing costs, which connect an input word into the structure built thus far.

The process of structural integration, or dependency formation, has been accounted for by the distance between the heads of the two maximal projection XPs in the structure, as shown in Figure 1.
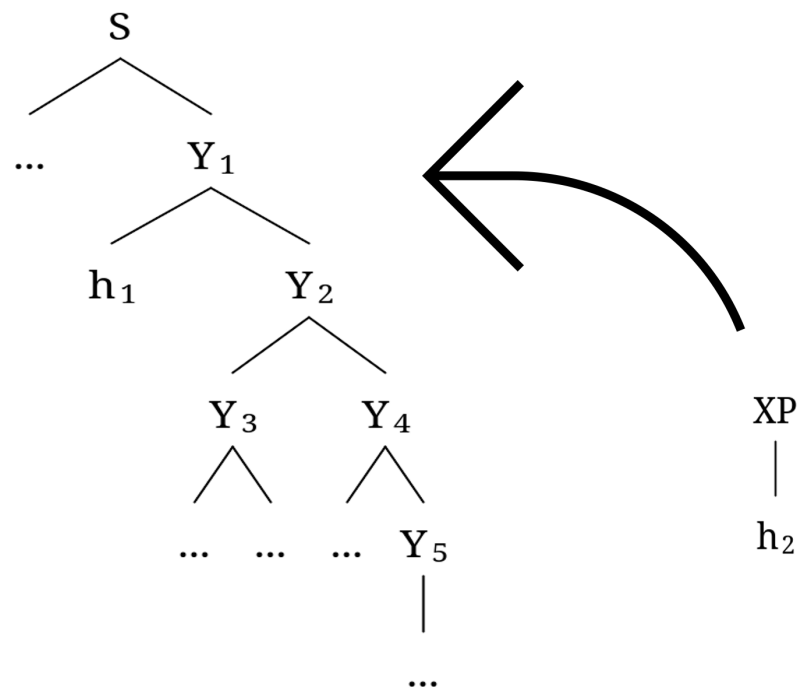
Figure 1:   Structural integration of a maximal projection XP of a newly input head h_2 to an attachment site headed by a head h_1 in the structure for the input so far. (Gibson, 2000, p. 104)

For instance, a previous syntactic head h_1 is retrieved from memory when parsers integrate it with a new maximal projection, XP, headed by h_2. The activation of h_1 is decayed as intervening words are processed and integrated into the structure. This difficulty of the structural integration depends on discourse processing costs as well as structural processing costs[5]. Therefore, the accessibility of the referent of an intervenor in the discourse affects the difficulty of dependency integration. Following Gibson (1998), DLT proposes a binary metric of discourse processing costs such that building new discourse referents, occurring between the elements being integrated, consumes more resources than accessing previously introduced or already existing referents in discourse, as shown in (12) and (13).

(12) DLT Simplified discourse processing cost

Discourse processing cost (the cost associated with accessing or constructing the discourse structure for the maximal projection of the input word head h_2):

1 energy unit (EU) is consumed if h_2 is the head of a new discourse referent; 0 EUs otherwise.

(Gibson, 2000, p. 104)

(13) DLT Structural integration cost

The structural integration cost associated with connecting the syntactic structure for a newly input head h_2 to a projection of a head h_1 that is part of the current structure for the input is dependent on the complexity of the computations that took place between h_1 and h_2. For

---

[5] I will mainly discuss discourse processing costs for structural integration, or dependency formation, which are related with the main topic for this Chapter.

simplicity, it is assumed that 1 EU is consumed for each new discourse referent in the intervening region.

(Gibson, 2000, p. 105)

In other words, this theory predicts that an intervenor associated with a new discourse referent between the formation of the filler-gap dependency requires more processing load than an intervenor associated with a pre-existing discourse referent. Gibson's binary metric in discourse integrations proposes that integrating across first- and second-person pronouns, which are given in the discourse by default consumes fewer resources and requires less processing than integrating other new discourse referents.

2.1.1.2.2 Continuous Metric

The accessibility of referents in discourse (Ariel, 1988, 1990; Arnon, 2010; von Heusinger, 2007) has been discussed in more gradient status in discourse. A referent for an NP which is high on the hierarchy is activated in short-term memory, and thus easily accessible, whereas a referent for an NP low on the hierarchy is less accessible because it is found in long-term memory or introduced as a new referent. In this chapter, I introduce two discourse hierarchies, namely The Givenness Hierarchy (Gundel, 2010; Gundel et al., 1993, 2012) and the Accessibility Hierarchy (Ariel 1988, 1990; Givon 1983). While the two theories are similar in spirit to the gradient status of a referent in discourse, they crucially differ in the underlying concept of accessibility of referents.

The Givenness Hierarchy: Gundel (2010), Gundel et al (1993, 2012).

Gundel (2010) and Gundel et al (1993, 2012) propose a discourse hierarchy, namely the Givenness Hierarchy (GH). The GH highlights that the GH itself is based on the *manner* of accessibility: how and where the addressee mentally accesses an appropriate representation, not the degree of referential accessibility proposed by Ariel (1988, 1990) and (Givon, 1983), which will be discussed in the next section. This GH theory takes account of multiple other factors such as cognitive and pragmatic elements that contribute to the cognitive status of a referent. Referring expressions signal the discourse status of the referent, which is associated with the addressee's expected background knowledge in memory and cognitive status to encode them. The GH theory makes indirect predictions that highly accessible forms for cognitive status restrict possible interpretations, compared to less accessible forms, thereby leading to ease of processing. Gundel et al (1993) and Gundel (2010) illustrate how different forms encode the cognitive status of the addressee in English as in (14) and (15):

(14) Givenness Hierarchy

in focus	> activated	> familiar	> uniquely identifiable	> referential	> type identifiable
{it}	that this this N	{that N}	{the N}	{indefinite this N}	{a N}
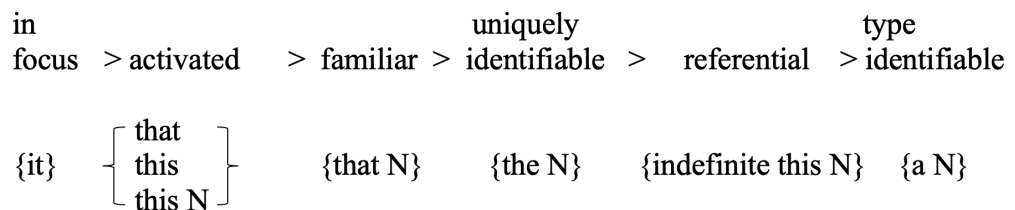
(Gundel, 1993, p. 275)

(15) Examples of coding guidelines for cognitive status

A referent can be assumed to be _in focus_ if

a. the addressee is intently looking at it.

b. it was introduced in a syntactically prominent position in the immediately preceding sentence.

A referent can be assumed to be at least *activated* if

a. it is present in the immediate extralinguistic context.

b. it is mentioned in the immediately preceding sentence.

A referent can be assumed to be *at least familiar* if

a. it is part of general cultural knowledge shared by speaker and addressee.

b. it is mentioned in the previous discourse.

A referent can be assumed to be at least *uniquely identifiable* if

a. a unique representation can be constructed via a "bridging inference" to a recently activated referent.

b. a unique representation can be constructed solely based on information conceptually encoded in the phrase.

An interpretation can be assumed to be *referential* if

a. it is evident from the context that the speaker intends to refer to a specific entity.

b. it is mentioned subsequently in the discourse.

An interpretation can be assumed to be *type identifiable* if

the addressee understands the sense of the phrase (the descriptive/conceptual content it encodes) and can therefore be assumed to be able to associate an appropriate type representation.

(Gundel, 2010, p. 154)

For example, (16) shows that different referring expressions can be used to refer to the same entity but differ how a referent is to be accessed, having different restrictions of possible interpretations.

(16)

a. A restudy of pareiasaurs reveals that *these primitive reptiles* are the nearest relatives of turtles. [M.S.Y. Lee, The origin of the Turtle Body Plan. Science, 1993, p. 1649]

b. A restudy of pareiasaurs reveals that *the primitive reptiles* are the nearest relatives of turtles. [M.S.Y. Lee, The origin of the Turtle Body Plan. Science, 1993, p. 1649]

c. A restudy of pareiasaurs reveals that *they* are the nearest relatives of turtles. [M.S.Y. Lee, The origin of the Turtle Body Plan. Science, 1993, p. 1649]

(Gundel, 2010, p. 151)

In (16a), the proximal demonstrative determiner NP, *these primitive reptiles,* encodes the cognitive status of *reptiles* "activated" in memory. Since this NP is activated in the current consciousness, it is easily interpreted as referring to *pareiasaurs*. The definite description in (16b), *the primitive reptiles*, however, is not as easily accessible as (16a). The definite article *the* in (16b) requires more processing effort to build more conceptual content, encoding "referential". Therefore, it can refer to *pareiasaurs* in the context but allows an alternative reading. It is accessible to refer to the class of primitive reptiles: generic interpretation. The proximal demonstrative

determiner encodes a higher and more restrictive cognitive status than the definite article because it is activated in working memory. The NP in (16a), thus, is more accessible and restricts other possible interpretations than the NP in (16b), which eventually leads to less processing cost of the NP in (16a) than the NP in (16b).

The pronoun *they* in (16c), on the other hand, is most accessible, even more than the proximal demonstrative determiner NP in (16a) since it is already in the addressee's focus of attention with minimal processing effort of conceptual content. Therefore, the pronoun in (16c) is almost automatically interpreted as *pareiasaurs*, and it is in fact the only accessible referent in this context.

In sum, the cognitive status of the pronoun in (16c) is most accessible and requires the least processing effort. The other two forms, the proximal demonstrative determiner NP in (16a) and the definite description in (16b) are less accessible than the pronoun because they are not in the focus of attention. They both require more processing of conceptual content than the pronoun. The NP headed by the proximal demonstrative determiner, however, is more accessible than the NP headed by the definite article because the former form is activated in short-term memory and does not allow more than one possible interpretation, unlike the latter form (the definite description).

It should be noted, however, that the GH does not guarantee that the referring forms that are high on the hierarchy are more accessible to those low on the hierarchy. Forms that are lower on the hierarchy can be more accessible than those that are higher on the hierarchy when the encoded cognitive status of the forms that are lower on the hierarchy requires less processing effort in a given context, as shown in (17).

(17) Ms. Eisen exhausted her unemployment benefits before her
check was restored by *a federal extension*. [The New Poor,
Peter S. Goodman, The New York Times, 2/21/2010]

(Gundel et al, 2012, p. 256)

Gundel et al (2012) presented examples of the underspecification of the cognitive status of referring forms from newspapers, as in (17). It was noted that the forms that are low on the hierarchy can be underspecified for the forms that encode higher cognitive status, which entails all lower cognitive statuses. For instance, anything "referential", the higher cognitive status on the hierarchy, is also "type-identifiable", the lower cognitive status on the hierarchy. Therefore, "type-identifiable" can be underspecified for "referential".

In (17), the phrase *a federal extension* is viewed as referential, assuming that readers would retrieve an existing representation of the referent. The cognitive status "type identifiable" encoded in the indefinite NP, which is associated with a representation of the type of object/person, is underspecified for the higher cognitive status of "referential", which is associated with a representation of a particular object. The indefinite NP does not exclude the possibility of being "referential" because any object referential is also "type identifiable". Therefore, the lower cognitive status of "type-identifiable" can be underspecified for the higher cognitive status of the intended referent "referential" as long as the interpretation of the lower cognitive status is restricted. That is, the indefinite NP (type-identifiable) in (17) is felicitous although the cognitive status of the intended referent is "referential" (higher cognitive status) because simply identifying the type is sufficient to build the representation of the intended referent and understand the meaning of the sentence. The retrieval of an existing referent from long-term memory can rather cause more processing effort than

simply identifying the type in this case. Therefore, GH predicts that the type-identifiable interpretation can be more accessible and favored although the cognitive status of "type-identifiable" is lower on the GH than "referential", depending on contexts.

Warren & Gibson (2002) incorporate the GH into DLT such that the gradient status of an intervenor influences the processing of dependencies, instead of Gibson's binary distance metric. The lower the intervenor is on the hierarchy, the more difficult it is to form dependencies. This will be discussed in more detail in Chapter 3, section 3.1.1.

Accessibility Hierarchy: Ariel (1988, 1990), Givon (1983).

The Accessibility Hierarchy (Ariel, 1988, 1990; Givon, 1983) has been viewed as the same as the GH (Kaiser & Trueswell, 2008; Warren & Gibson, 2002) due to the identical order of accessibility. However, the Accessibility Hierarchy (AH) is distinct from the GH with respect to the assumption it makes for the role of accessibility. The GH is associated with the manner of accessibility, which is about *how and where* the parser encodes information on the referent.

Unlike the GH, the AH is based on the *degree* of accessibility, which directly maps referring expressions onto the accessibility hierarchy; different forms straightforwardly correspond to different degrees of accessibility. Instead of accounting for reference by the notion of context, Ariel (1998) proposes that "natural languages primarily provide speakers with means to code the ACCESSIBILITY of the referent to the addressee. Accessibility, in its turn, is tied to context types in a definitely NON-arbitrary way (p.68)." Therefore, the inherent definition of each

referring form itself is the determinant for its processing such that the highly accessible referents require less processing cost to be retrieved than those of low accessibility.

The underlying assumption of this hierarchy is that "the less accessible the representation of a given entity, the lower the Accessibility marker is used. (Ariel 1988: 76)." Accordingly, Ariel predicts that referring expressions that are lower on the hierarchy (e.g., proper names, definite descriptions) mark relatively lower degrees of accessibility than those signaling higher on the hierarchy (e.g., pronoun). The following scale by a degree of Accessibility is systematized without contextual effect.

> (18) Joan Smith, the president > Joan Smith > The president >
> Smith > Joan > That/this hat we bought last year > That hat >
> This hat > That > This > SHE > she > herself > ∅

> (Ariel, 1988, p. 84)

The high Accessibility markers like pronouns are associated with minimal formal complexity and vague contents, whereas the low Accessibility markers like full NP have rich and less vague semantic information. When a referent is highly accessible in contexts, the minimally complex forms are enough to access the intended interpretation, whereas the semantically richer information of more complex forms is not necessary.

Givon (1983) proposes a hierarchical order of referring expressions, involving syntactic structures, in line with the idea of Ariel's (1988) AH.

(19)

Most continuous/accessible topic

▲ zero anaphora
| unstressed/bound pronouns or grammatical agreement
| stressed/independent pronouns
| R(ight)-dislocated DEF-NPs
| neutral-ordered DEF-NPs
| L(eft)-dislocated DEF-NPs ('contrastive topicalization')
| cleft-focus constructions
▼ referential indefinite NPs

Most discontinuous/inaccessible topic
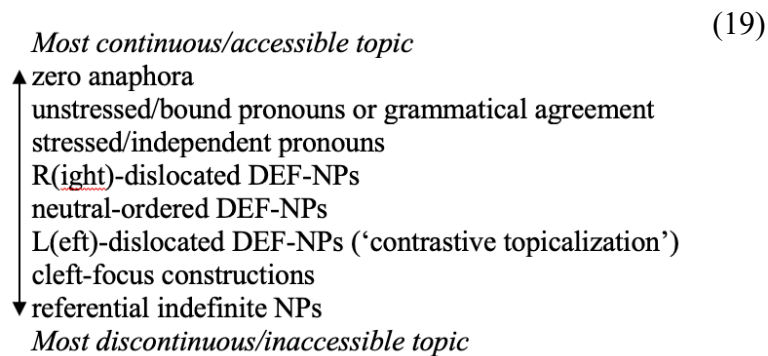
(Ariel, 1988, p. 84)

Ariel and Givon's AH appears to be identical with Gundel's GH in that the AH presents the same hierarchical order of referring forms with the GH. The AH and GH, however, make different predictions in terms of contextual factors. The AH predicts a robust advantage of forms that are higher on the hierarchy than those that are lower on the hierarchy, regardless of contextual factors. This prediction is based on the assumption that the forms themselves inherently mark their own degree of accessibility and speakers/writers select a particular form of a certain level of accessibility to convey their intended interpretations. The GH, on the other hand, does not predict that forms that are higher on the hierarchy are always more accessible than those that are lower on the hierarchy. The forms that are lower on the hierarchy may be more accessible than those that are high on the hierarchy if the more restrictive interpretation of higher cognitive status is not useful or relevant in contexts of use. The multiple factors interact and determine the cognitive status of a referent and thus contribute to the degree of accessibility of referents in different contexts.

I introduce the AH, in addition to the GH, because of their different approaches to contextual effects, which are associated with Experiments 2 and 3. I incorporate the

AH into DLT as Warren & Gibson (2002) incorporated the GH into DLT such that the AH of an intervenor can have an effect on dependency formations.

2.1.2 Expectation-based Theories: Surprisal Theory

As an alternative explanation of the main source of processing difficulty in sentence comprehension, several studies have focused on the role of expectations. Some experimental studies have supported the claim that probabilistic information plays the main role in language acquisition and comprehension (Corley & Crocker, 2000; Jacoby & Brooks, 1984; M. C. MacDonald et al., 1994b, 1994a; M. MacDonald & Christiansen, 2002; Spivey-Knowlton & Sedivy, 1995; Trueswell, 1996).

According to a surprisal model, parsers use probabilistic knowledge from previous experience in incremental sentence processing to predict a lexicon following the lexicon they currently encountered, and eventually establish possible interpretation. This expectation is a crucial factor that determines processing difficulty. Several theories have implemented computational models to make a general prediction of the online processing difficulty for each word in a sentence (Futrell et al., 2020; Hale, 2001; Levy, 2008; Mitchell et al., 1995; Roark et al., 2009).

Surprisal, introduced by Attneave (1959), is the logarithm of the reciprocal of a probability. Hale (2001) proposes the concept of surprisal as a measure of processing complexity. Surprisal is associated with a probabilistic notion of linguistic structures under information theory. Along with Hale (2001), in Levy (2008)'s surprisal model, the surprisal of a word is a crucial factor of processing difficulty. It serves as a bottleneck between representational contents built during sentence comprehension and processing difficulty. The idea is that if a predictable word (in its context and syntactic

configuration) is encountered during sentence processing, then the processing cost is reduced, compared to when an unpredictable word is encountered (Hale, 2001; Jurafsky, 2003). Several reading-time studies have found results that are consistent with the predictions of this model (Boston et al., 2008; Demberg & Keller, 2008; Roark et al., 2009; Smith & Levy, 2013).

Levy (2008) assumes the effects of surprisal during the retrieval and integration of long-distance dependencies as well as the online processing. It should be noted, however, that the prediction of processing difficulty under Levy's surprisal model is not associated with the time point at the integration of dependencies, in contrast to the memory-based theories' prediction of difficulty at the integration (usually at the verb). According to the surprisal model, the cost of low expectation for an input word arises upon encountering the word in the course of incremental processing. Some empirical studies found the difficulty of uncommon words occurred at the point where a rare construction or lexicon is first encountered (Demberg & Keller, 2008; Gennari & MacDonald, 2008).

In terms of the processing of NP types in object-extracted relative clauses, for instance, the processing difficulty aligns with lexical probabilities of an intervenor. A less frequent word in syntactic configuration and its context is more surprising than a more frequent and common word. This effect of surprisal is predicted to occur upon encountering the intervenor. Levy (2008) discusses the surprisal of NP types of embedded subject NPs in object-extracted RCs, as in (20), which was examined by Gordon et al (2004).

(20) a. The salesman that **{the/an} accountant** contacted spoke
very quickly.
(Definite/Indefinite)

b. The salesman that **(the) accountants** contacted spoke very
quickly.   (Definite/Bare
Plural)

c. The salesman that **{the accountant/everyone}** contacted
spoke very quickly.
(Definite/Quantifier)

(Levy, 2008, p. 38)

In a corpus study, Gordon et al (2004) found definite NPs to be more frequent than indefinite and bare NPs in the embedded subjects. This pattern is consistent with the idea that definite NPs are relatively more prototypical than indefinite NPs as the subject of a sentence or clause (Abeillé et al., 2020; Givon, 1984). Subjects serve as background constituents, which typically provide given information or topic of a sentence rather than a new discourse referent (Erteschik-Shir, 1973; Goldberg, 2006; Takami, 1992; van Valin, 1995). Therefore, indefinite NPs are less common as subjects of a sentence or clause in that they are peripheral in discourse and used to introduce a new referent, compared to definite NPs (Warren & Gibson, 2002; Heim 1982, 1983).

This data provides a piece of evidence to predict the surprisal of NP types in the embedded subjects. Given a higher frequency of definite NPs than indefinite and bare NPs in the embedded subjects, the surprisal model predicts that definite NPs should be easier to process than indefinite and bare NPs, at the embedded subject in object-extracted relative clauses. In (20c) contrast, the frequency can be measured in terms of the contrast between open-class and closed-class. The closed-class, *everyone*,

is likely to have a higher frequency than the relatively open-class, definite descriptions. Therefore, *everyone* is predicted to show less difficulty than *the accountant* in (20c). In brief, the processing difficulty for object-extracted relative clauses is predicted to be at the embedded subject when it is indefinite or bare NPs than when it is definites in (20) because the high expectation of definite NPs is not met.

  Having detailed the relevant background theories, I turn now to Chapter 3. In Chapter 3, I present previous empirical studies of the effect of NP types on dependency formation and the results of three experiments that I conducted.

<div align="center">**Chapter 3**</div>

<div align="center">**THE PROCESSING OF FILLER-GAP DEPENDENCIES**</div>

In this chapter, I introduce previous empirical studies on the effect of NP types of an intervenor on dependency formation and motivations of experimental studies that I conducted. I present three separate experiments, using reading-time measurement. For each experiment, I report the method and results of the experimental study and discuss the implication of the results of the experiment.

3.1   Experiment 1: The Effect of NP types in Null Contexts

3.1.1   Previous studies on the effect of NP types during dependency formation

As discussed in the introduction of similarity-based interference effects, Gordon et al. (2001, 2004, 2006) revealed that different NP types affect the difficulty of processing filler-gap dependency. Their findings were attributed to a similar NP type of an intervenor with the filler in English relative clauses, which may give rise to a similarity-based interference effect: a processing slowdown when the NP type of an intervenor matches the NP type of the dependent element or the filler. Gordon et al (2001) investigated the difference in the processing difficulty between object and subject relative clauses in English as in (21):

> (21) a. Object relative clause: The banker [that *the barber* praised __]
>   climbed the mountain.
>   b. Subject relative clause: The banker [that __ praised *the barber*]

climbed the mountain.

The processing difficulty of object relative clauses (ORCs) than subject relative clauses (SRCs) has been well observed using various experimental methods: self-paced reading times (King & Just, 1991), ERP (King & Kutas, 1995), eye-tracking (Traxler et al., 2002), and fMRI (Caplan et al., 2002; Just et al., 1996). This processing asymmetry has been attested across languages: Dutch (Frazier, 1987), German (Schriefers et al., 1995), Chinese (L. Jäger et al., 2015; Vasishth et al., 2013), Japanese (Ueno & Garnsey, 2008), and Korean (Kwon, 2008; Kwon et al., 2006), as well as English (Caplan et al., 2000).

Memory-based theories account for this asymmetrical difficulty of processing in terms of a relatively higher memory-load of ORCs than SRCs [6]. For instance, in (21), the unintegrated filler (i.e., *the banker*) must form its dependency at gap site after the embedded verb (i.e., *praised*) in ORCs, crossing an intervenor (i.e., *the barber*), while the dependency between the filler and its gap is formed before the verb (i.e., *praised*) in the SRCs. Therefore, SRCs form dependencies in the absence of an intervenor.

Gordon et al (2001) used this processing asymmetry between ORCs and SRCs as a probe to investigate how NP types of an intervenor modulate dependency formation. They predicted that the NP type of an intervenor would modulate the processing difficulty of dependency formation in ORCs. Under this prediction, they

---

[6] Various aspects have approached to account for the advantage of the processing SRCs: syntactic storage (Chomsky & Miller, 1963), perspective shift (MacWhinney & Pleh, 1988), Distance locality theory (Gibson, 2000), and frequency (M. MacDonald & Christiansen, 2002).

measured two critical words, which are in bold type in (22), following King and Just's (1991) design[7]. The two critical regions are involved with dependency formation and assignment of the thematic role of the filler at its gap site.

(22) a. The barber that <u>the lawyer/you/Joe</u> **admired climbed** the mountain. [ORC]

b. The barber that admired **<u>the lawyer/you/Joe</u> climbed** the mountain. [SRC]

Gordon et al's (2001) studies showed reduced ORCs-SRCs differences when the NP type of the intervenor was a name or pronoun, as compared to the definite description at the two critical regions, respectively. Their findings support the similarity-based interference in that different NP types of an intervenor were read faster than the definite description, which is the same NP type as the filler.

However, the discourse hierarchy (Ariel, 1988, 1990; Arnon, 2010; Gundel, 2010; Gundel et al., 1993, 2012), coupled with Gibson (2000)'s DLT, might provide an alternative account for Gordon et al's (2001) findings. On the view of the GH, the greater ease in processing name or pronoun intervenors than definite intervenors can be attributed to the higher accessibility of name and pronoun than definites (repeated below).

(23) Givenness Hierarchy

---

[7] These critical regions presented the crucial difference in ease of processing between SRC and ORCs (King & Just, 1991). The first critical region of ORCs is the verb in the embedded clause, while it is the embedded subject for SRCs. The second critical region is the verb of the main clause in both ORCs and SRCs.

```
     in                          uniquely              type
  focus  > activated  > familiar > identifiable  >  referential  > identifiable

          ⎡ that  ⎤
  {it}    ⎢ this   ⎥        {that N}      {the N}      {indefinite this N}   {a N}
          ⎣ this N ⎦
```
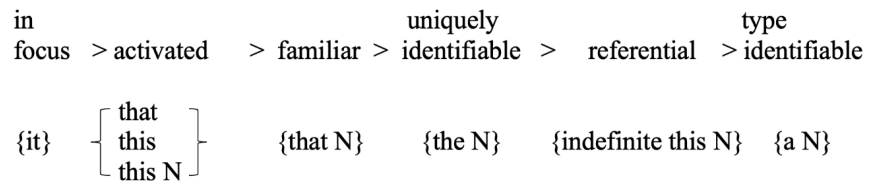
<div style="text-align:right">(Gundel, 1993, p. 275)</div>

Proper names (e.g., *Bob*) and indexical pronouns (e.g., *you*) are more central in discourse as compared with definite descriptions (e.g., *the salesman*), which are more peripheral. The former referential forms (i.e., proper names, pronouns) reduced the processing asymmetry between ORCs and SRCs, while the latter categories did not do so. Under this account, the ease of processing filler-gap dependencies with intervening proper names or pronouns relative to those with intervening definite NPs could be attributed to the profile of the intervenor in discourse.

Warren & Gibson's (2002, 2005) experiment provided evidence supporting the underlying assumptions of Gibson's (1998, 2000) DLT and further observed the sensitivity to the gradient status of an intervenor in discourse across six levels as in (24), consistent with the GH.

(24) $1^{st}/2^{nd}$ pronouns > 3rd-person pronouns > first names > full
     names > definite descriptions > indefinite descriptions

<div style="text-align:right">(>: easer to process)</div>

<div style="text-align:right">(Warren & Gibson, 2002, p. 87)</div>

Warren & Gibson (2002) test whether the GH theory plays a role in the comprehension of complex sentences containing filler-gap dependencies in both complexity rating and self-paced reading tasks as shown in (25).

(25) a. Complexity rating study

The old lady who the government assistance program which {you, the reporter, a reporter, Bill Clinton, Brad} praised had saved did not have enough money to heat her house.

 b. Self-paced reading study

The consultant who {we/Donald Trump/ the chairman/a chairman} called advised wealthy companies…

They found a significant correlation between the GH and the complexity rating/reading time at critical regions, the main verb (i.e., *called*), and the following spillover region (i.e., *advised*): a monotonically increasing trend. The complexity rating becomes higher and the reading time becomes longer as the intervenor becomes more peripheral in discourse. Under the spirit of the DLT, Warren & Gibson (2002) proposed that discourse processing cost of dependency formation is sensitive to the gradient status of an intervenor in the discourse, such that the greater degree of the peripheral status of a referent in the intervening region leads to consuming more processing costs for dependency formations.

3.1.2   Motivation for Experiment 1

Gordon et al (2001) accounted for their findings in terms of the similarity-based interference effect, whereas Warren & Gibson (2002) suggested that a gradient

status of referents in discourse modulates the interference effect, underlying the discourse hierarchy of an intervenor. I now present Experiment 1 which examines how the property of NP types modulates dependency formations, aiming to tease these competing accounts apart.

In view of findings in Warren & Gibson (2002), the findings of Gordon et al (2001, 2004) can have an alternative explanation for the easier processing of name or pronoun intervenors than definite intervenors. It can be attributed to different discourse-levels of NP types such that both pronoun and name NP types are higher and more central than definite descriptions in the semantic hierarchy. Furthermore, Gordon et al (2001) conducted separate experiments for each level of NP on the GH and thus, could not directly compare and analyze across conditions. Therefore, their experiments did not capture the fact that the reading times on the critical region were faster in the pronoun condition than the name condition, which is in fact the pattern predicted by the discourse hierarchy.

Thus, in Experiment 1, I explore which property of NP types contributes to determining the difficulty of dependency formation—(i) semantic similarity between an intervenor and a filler, (ii) a more fine-grained discourse-level of the semantic hierarchy of NPs, and/or (iii) frequency effects of NP types in the structure, using a self-paced reading experiment.

The current study aims to further explore the interaction between the NP types of the filler and an intervenor. For this purpose, clefts were used, following Warren & Gibson (2005), because clefts impose fewer restrictions than relative clauses in that the filler can be a linguistic expression like a pronoun or name, as opposed to being

restricted to a definite NP (Gordon et al, 2001)[8]. Previous research, including Gordon et al (2001) and Warren & Gibson (2002), has mainly focused on the NP type of the intervenor, and did not manipulate the NP type of the filler (i.e., the head noun of a relative clause): they used a definite filler across all conditions. In the relative clause constructions that they used, in fact, a pronoun or name is not grammatical as the head noun of relative clauses[9].

Among processing models introduced in Chapter 2, only similarity-based interference account predicts the interaction between the NP type of the filler and an intervenor because it is associated with matching/mismatching NP type of an intervenor to the filler. Other processing models mainly discuss processing difficulty based on NP types of an intervenor only. Therefore, no interaction between the NP types of the filler and an intervenor is not predicted.

Furthermore, in order to properly evaluate the contribution of the discourse hierarchy in intervention effects, I test indefinite descriptions instead of proper names. This is because the pronoun is higher than the definite description on the GH, whereas the indefinite description is lower than the definite description, recalling the givenness hierarchy (Warren & Gibson, 2002, approx.), as in (26).

---

[8] Gordon et al (2001) tested clefts in Experiment 4 and manipulated the NP types of both clefted and the embedded NP position into definite descriptions and names but excluded another condition, which is critical to tease apart the two possible accounts.
[9] This is because the purpose of a relative clause is to modify the head noun so as to make it uniquely identifiable. Since a pronoun or name is already uniquely identifiable, it cannot appear in this environment, as the relative clause would contain redundant information.

(26) **pronouns** > first names > full names > **definite descriptions** >

indefinite descriptions

This experiment had a 2 x 3 design, which crossed two types of NP in the clefted (NP1) position and three types of NP in the embedded NP (NP2) position: [definite descriptions, indefinite descriptions] x [definite descriptions, indefinite descriptions, pronouns], as shown in (27).

(27) It was {the actor/an actor} who {we/the director/a director} graciously

NP1   NP2

thanked before the show.

In this experiment, object-extracted cleft sentences were mainly tested to examine the interference effect of NP2, as an intervenor, in the resolution of dependency[10]. The pronoun types used in this experiment are 1st person, 2nd person singular pronouns (e.g., *I, you*), and the 1st person plural pronoun (e.g., *we*), all of which are central and assumed to be obligatory participants in the discourse.

3.1.3 Materials

Experimental materials consisted of 24 sets of items. 18 sentence sets were created based upon materials that both Gordon et al (2001) and Warren & Gibson

---

[10] In subject-extracted cleft sentences, NP2 is encountered after the resolution of dependency. In this configuration, NP2 is no longer between the dependency formation, and thus no longer an intervenor.

(2002) used, and further six items were added to them. The 24 sets of six conditions were distributed across six lists in a Latin Square design (each participant saw 4 items in each of the 6 conditions) and combined with 26 filler sentences, for a total of 50 sentences. 18 filler sentences consisted of target items for Experiment 4 in Chapter 5, while the other 8 filler sentences were constructed to be similar to experimental items in length.

Comprehension questions were presented immediately after each target sentence. All the correct answers to items targeted NP1. For example, each target sentence, as in (20), was followed by the following question: "What was the profession of the person who was thanked?". The six experimental conditions were counterbalanced to prevent participants from adopting the patterns of target sentences and comprehension tasks.

3.1.4 Participants

Fifty native speakers of English were recruited using Amazon's Mechanical Turk web service. All participants provided informed consent and were compensated $5.00. The experiment lasted approximately 25 min. Three of the participants were excluded from the analysis because their performance across all comprehension questions was lower than 80%.

3.1.5 Procedure

The task was a self-paced reading using a moving window paradigm. The experiment used the online software platform Ibex (Drummond, 2018). To ensure participants did not learn to expect a one-word display, a word or phrase was presented at a time. Each word or phrase was displayed in the middle of the screen.

The crucial region and spillover regions were always presented separately for analysis. The display pattern was constant across conditions of each item in the Latin Square design. Participants pressed the spacebar to proceed to the next word or phrase. After the last word or phrase of each sentence, the full sentence of a comprehension question appeared. Participants were instructed to press the spacebar to see the multiple-choice answer set. Prior to beginning the experiment, three practice trials were given to familiarize participants with the task.

3.1.6 Predictions

In this section, I present predictions for Experiment 1 by processing models introduced in Chapter 2. Recall the example sentence (repeated below). At the verb region (e.g., *thanked*), the filler (e.g., *the actor/an actor*) is integrated to associate its thematic role with the verb (Warren & Gibson, 2002, 2005) and the dependency between the filler and its gap is formed. Therefore, the verb region is the critical region to explore the difficulty of dependency formation.

> (28) It was {the actor/an actor} who {we/the director/a director} graciously
>
> NP1   NP2
>
> thanked before the show.

Memory-based theories

3.1.6.1 Similarity-based Interference Effect

The similarity-based interference hypothesis attributes the processing difficulty of dependency formation to the similarity of linguistic information when parsers

retrieve the dependent element from working memory. Gordon et al (2001) accounted for their findings in terms of similarity-based interferences such that "sentence processing was easier when the two NPs were of different types than when they were of the same type (p. 1419)". According to this similarity-based interference account, matching NPs between NP1 and NP2 should lead to longer reading times on the verb (e.g., *thanked*) than other conditions. That is, the processing should be more difficult when both NP1 and NP2 are the same NP types, either definite or indefinite descriptions, than when NP1 and NP2 are different NP types, as illustrated in (29):

(29) Prediction of the similarity-based interference effect[11]

Definite-Indefinite, Definite-Pronoun   <   Definite-Definite,

Indefinite-Definite, Indefinite-Pronoun <   Indefinite-Indefinite

(<: harder to process)

3.1.6.2   Dependency Locality Theory

In this model, processing difficulty is defined in terms of the use of computational resources (Gibson 2000). The more storage and integration cost the element needs, the more difficult it is to process. The integration cost arises from structural and discourse processing costs in working memory, which is computed based on an energy unit (EU). The structural processing cost is controlled in

---

[11] Definite-Indefinite, for example, indicates the condition where the NP type of NP1 is a definite description and NP2 is an indefinite description. I assume that encoding interference contributes to interference on retrieval, at the verb region (L. Jäger et al., 2017; Villata & Franck, 2020). If encoding interference of an intervenor is not predicted, a main effect of NP1 may not be predicted in that definiteness feature is not a retrieval cue to the verb.

Experiment 1, and, thus, the discourse processing cost is mainly discussed for predictions for the Experiment.

Binary distance metric

Gibson (2000) proposes that the accessibility of the referent in the discourse affects the difficulty of dependency integration. New discourse referents, occurring between the elements being integrated, require more processing costs than previously accessed discourse referents (Haviland & Clark, 1974; Murphy, 1984). In other words, this theory predicts that the intervenor associated with a new discourse referent between the formation of the filler-gap dependency requires more processing load than the intervenor associated with a pre-existing discourse referent. Gibson's DLT proposes a binary distance metric such that first-and second-person pronouns require less processing cost in that they are already given in the discourse, compared to other full NPs, which need to build new discourse referents.

Under the binary distance metric, the two main levels of processing difficulty, based on Gibson (1998), EU is assumed to be consumed for an intervening new discourse referent during integration. In Experiment 1, I used 1st and 2nd person pronouns (e.g., *I, you, we*) for the pronoun condition. They become a default referent even in a null context because a speaker/writer and a listener/reader exist in every discourse (Chafe, 1987; Enc, 1983). Thus, integrating the clefted element to the object position of the verb should be easier when crossing a pronoun, or an existing referent, than when crossing a new discourse referent such as definite and indefinite descriptions.

(30) Prediction of DLT based on the binary distance metric

Pronoun < Definite, Indefinite   (<: harder to process)

Continuous metric: The Givenness Hierarchy

Gundel (2010) and Gundel et al. (1993, 2012)'s GH proposes more of a gradient hierarchy in the discourse rather than a binary concept. They propose a discourse hierarchy based on the information about how and where the addressee mentally accesses an appropriate representation. The GH is associated with "the addressee's assumed memory and attention state in relation to the intended referent" (Gundel, 2010, p. 151). The GH makes indirect predictions that highly accessible forms lead to ease of processing in that the cognitive status of those forms more restricts other possible interpretations than the cognitive status of low accessible forms.

Warren & Gibson (2002) provided evidence for the susceptibility to the gradient status of intervenors in sentence comprehension, supporting the effect of the GH under the DLT. Recall the discourse hierarchy they tested:

(31) $1^{st}/2^{nd}$ pronouns > 3rd-person pronouns > first names > full names > definite descriptions > indefinite descriptions

The referential forms listed in order on the left are central and constrain possible interpretations, while those that appear on the right are peripheral in discourse and allow more possible interpretations. Building a more central referent is predicted to be easier to process than building a more peripheral referent. According to the

discourse hierarchy within the frame of Gibson's DLT, more central and accessible intervenors in discourse should be less complex and cause less processing load than peripheral ones during the integration of filler-gap dependencies. The GH, therefore, predicts a main effect of NP2 type: the fastest reading times when NP2 is a pronoun (e.g. *we*), longer when it is a definite description (e.g. *the director*), and longest when it is an indefinite description (e.g. *a director*).

      (32) Prediction of DLT based on the GH

          Pronoun < Definite < Indefinite   (<: harder to process)

Continuous metric: Accessibility Hierarchy

      The AH is a form-specific constraint that directly associates different forms of referring expressions with different degrees of activated information in memory, or different accessibility of mental representations of a referent. The referents that are high on the hierarchy are held in short-term memory, while those that are low on the hierarchy tend to be stored in long-term memory. Therefore, highly accessible referring expressions require less processing cost to be implemented than less accessible forms. Under this mechanism, the AH predicts that pronouns are more accessible than full NPs, definite and indefinite descriptions. Although Ariel's AH does not include indefinites on its hierarchy, I assume higher accessibility of definites than indefinites, following Givon (1983)'s hierarchy that shares the spirit of the AH. This ranking is based on the general usage of definites and indefinites. Definite descriptions tend to be topics that are uniquely identifiable and familiar with, while indefinites are topics introduced for the first time. The prediction of the AH for

processing different NP types of an intervenor within the framework of the DLT is as follows:

(33) Prediction of the Accessibility Hierarchy

Pronoun < Definite < Indefinite   (<: harder to process)

3.1.6.3 Expectation-based Theories

Expectation-based theories are based on frequency or probability, which depends on statistical information and the parser's experience in the language (Arnold, 1998, 2001). Surprisal theory assumes that processing difficulty is associated with the frequency information of a word in syntactic configuration. Therefore, this model predicts that more frequent NP types in the embedded subject position are less surprising than less frequent NP types. The more common and frequent NP types should be easier to process than less common NP types upon parsing the NP types. which suggests interference effects at the point of parsing an intervenor.

Previous literature has shown that pronouns are unmarked or typical NP types in subject positions in main and embedded clauses, being a given referent, whereas lexical NPs are marked to be a subject. Lexical NPs are rather unmarked objects because they tend to represent new information. (Arnold, 2010; Arnold et al., 2003; Arnon, 2010). A large corpus study of spoken English (Reali & Christiansen, 2007; Roland et al., 2007; Roland & Jurafsky, 2002) revealed that the majority of the embedded subject in object relative clauses was a first or second person pronoun. In a corpus study of spoken English, German, and Hebrew by children (Arnon, 2010; du Bois, 2003; Francis et al., 1998; Kidd et al., 2007), it was found that children showed a

similar pattern such that a pronominal embedded subject was most frequently produced in object relatives.

In terms of a comparison between definite and indefinite descriptions, definite descriptions are found to be more frequent than indefinite descriptions in sentential or clausal subjects (Givon, 1984; Leonetti, 2004). In a corpus study by Gordon et al (2004), they compared the frequency of NP types between a definite and an indefinite article in the embedded subject of object relative clauses. They observed that the embedded subject was much more likely to be definite than indefinite articles. This pattern was highly significant and consistent across three kinds of corpora (*Brown*, *Childes*, and *Switchboard*).

Given the findings from previous corpus studies on the frequency of NP types, experience-based theories would predict a much greater incidence of pronouns than lexical NPs, either definite or indefinite in the embedded subject. Among the lexical NPs, definite descriptions are relatively more prototypical than indefinite descriptions as the embedded subject.

(34) Prediction of experience-based theories

Pronoun <<< Definite < Indefinite   (<: harder to process)

This processing difficulty is predicted to occur at the time of encountering an intervenor, which is an embedded subject. The predictions of processing theories concerning the processing difficulty of the dependency formation are summarized in Table 1.

Table 1:   Prediction by each model for the effect of NP types in sentence
             processing

Models	Prediction
Memory-based account	
Similarity-based interference	Definite-Indefinite, Definite-Pronoun Indefinite-Definite, Indefinite-Pronoun < Definite-Definite & Indefinite-Indefinite (at the verb)
Binary metric	Pronoun < Definite = Indefinite (at the verb)
Continuous metric: Givenness Hierarchy	Pronoun < Definite < Indefinite (at the verb)
Continuous metric: Accessibility Hierarchy	Pronoun < Definite < Indefinite (at the verb)
Expectation-based account (Surprisal model)	Pronoun <<< Definite < Indefinite (at the embedded subject)

3.1.7   Results

A total of 47 participants were included in the analysis. The mean accuracy
rate on the comprehension questions was 96.2%. Reading times above 3SD for each
condition in each region were replaced with the mean for that condition for the
relevant region, affecting less than 2% of the data[12]. 1128 critical trials were
generated, and trials that resulted in incorrect responses were excluded from data
analysis (n=40). The main regions of interest are the critical verb region, the two
words following the verb (spillover regions 1 and 2, respectively) to explore the

---

[12] Outliers were replaced with the mean to maintain the date set of regions, items, and
participants for analysis.

processing load at the integration of the filler to its gap site. "NP2+adverb" region, which includes an intervenor and adverb following the intervenor, was also observed to explore processing loads once a different NP type of an intervenor was encountered. "NP2 and adverb" regions were one of the regions that were not presented separately to prevent participants from learning a one-word display. Although they were presented at once, I presume that results are driven by the NP2 region because identical adverbs were shown across conditions. The response time and accuracy, as offline processing, were also observed.

Data were analyzed with linear mixed-effects models with crossed random effects for participants and items (Baayen et al., 2008) using the *lme4* package (Bates, 2010) in the R software, version 3.4.4. (Bates, Maechler, et al., 2015). This was conducted using parsimonious models (Bates, Kliegl, et al., 2015), beginning with the maximal random effects structure that would converge, and simplifying it by removing random slopes that did not significantly improve the model. The predictors NP1 type (2 levels) and NP2 type (3 levels) were contrast-coded using centered Helmert contrasts. For NP1 type, definite was coded as -1/2 and indefinite was coded as +1/2[13]. For NP2 type, the first coefficient, PRONOUN vs. NP, contrasted pronouns (coefficient: +2/3) with NPs consisting of determiner + N (coefficient: -1/3 for definite NP; -1/3 for indefinite NP). This contrast asks about the effect of having a pronoun in the NP2 position versus having an NP with a determiner in this position. The second coefficient, DEFINITENESS, contrasted definite NPs with indefinite NPs (DEFINITENESS coefficients: 0 for Pronoun, -1/2 for definite NP, and +1/2 for

---

[13] Helmert contrasts was used only for 3 levels.

52

indefinite NP): this contrast asks the effect of definiteness. P values were computed via a Satterthwaite approximation, using the *LmerTest* (Kuznetsova et al., 2017) function in R.

Figure 2 reports region-by-region reading times for the indefinite NP1 condition and Figure 3 reports the same for the definite NP1 condition. Table 2 presents each reading time of regions of interest. At the verb region, (the critical region), there was no main effect of NP1 type (β=1.15, *SE*=10.77, *t*=0.11, *p*=.92). In terms of the NP2 type, the pronoun condition was approximately 65ms faster than the indefinite condition and 90ms faster than the definite condition. A test of the main effect of NP2 showed a significant effect in that the pronoun condition took significantly shorter reading time than the averaged definite and indefinite conditions (β=-72.66, *SE*= 16.92, *t*=-4.30, *p* < .001)[14]. No interaction was significant (all *ps* > .49).

At the Spillover region 1, a marginal effect of NP2 was observed such that pronoun conditions were read faster than other conditions (β=-17.61, *SE*=9.92, t=-1.78, *p* =.07). There was no significant interaction (all *ps* > .36). At the Spillover region 2, a marginal interaction between NP1 type and Pronoun vs full NPs (β=-31.22, *SE*=20.08, *t*=-1.55, *p* =.1). Planned comparisons suggest that this trend was driven by the comparison when NP1 is indefinite (β=-30.08, *SE*=14.24, *t*=-2.11, *p* =.03), but not when it is a definite (*p*=.93). That is, the pronoun condition was read marginally faster

---

[14] Surprisingly, the overall reading times of definite conditions in NP2 type averaged 20ms faster than the overall reading times of indefinite conditions. The effect of definiteness between the definite and indefinite conditions was not significant (β=-13.43, *SE*=24.98, *t*=-0.54, *p*=.59). This subtle effect of definiteness was observed across NP1 types, however.

than full NPs pooled. This longer-lasting effect for indefinite NP1 than definite NP1 suggests that the NP1 effect appears after the critical region. The significant difference in reading times was observed only when NP1 was indefinite at Spillover regions.

Table 2:   Mean reading time (*ms*) of critical and spillover regions

Average response time (*ms*)	Intervenor NP types			
	Pronoun (NP2)	Definite (NP2)	Indefinite (NP2)	Grand Total
Verb Region				
Definite (NP1)	414	497	471	460
Indefinite (NP1)	405	497	481	461
Grand Total	410	497	476	461
Spillover 1 Region				
Definite (NP1)	399	413	422	411
Indefinite (NP1)	401	433	429	421
Grand Total	400	422	426	416
Spillover 2 Region				
Definite (NP1)	465	455	474	464
Indefinite (NP1)	452	471	489	471
Grand Total	458	463	482	468

In the "NP2 and adverb" region, a main effect of NP2 type (intervenor) was also observed. Table 3 shows the mean reading times at the "NP2 and adverb" region. First, pronoun intervenors led to a significantly faster reading time than full NPs pooled (β=-168.04, *SE*=32.09, *t*=-5.24, *p* < .001). Interestingly, the contrast between

definite and indefinite intervenors also revealed that indefinite intervenors led to significantly slower reading time than definite intervenors (β=-78.15, *SE*=32.64, *t*=-2.40, *p* = .02). No main effect of NP1 type or interaction was found (all *ps*>.49)[15]. This effect of NP2 type at the NP2 region presents a gradient difficulty of NP2 type—*pronoun* < *definite* < *indefinite* (<: harder to process), unlike the integration region at the verb. The earlier effects of NP2 types once encountering the NP2 region could have been associated with the expected probability of the NP type under expectation-based accounts.
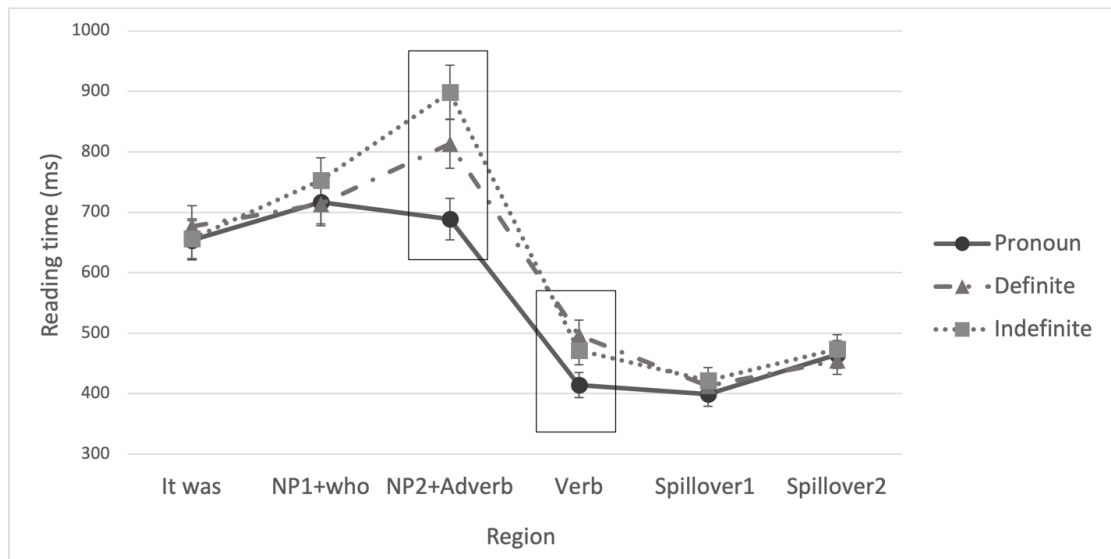
Figure 2:   Mean reading time (*ms*) by region for definite NP1 conditions. Boxes indicate the regions where a main effect was observed.

---

[15] There is a trend towards an interaction between NP1 types and NP2 types (definite and indefinite) at spillover regions, but it may not be significant due to lack of power.
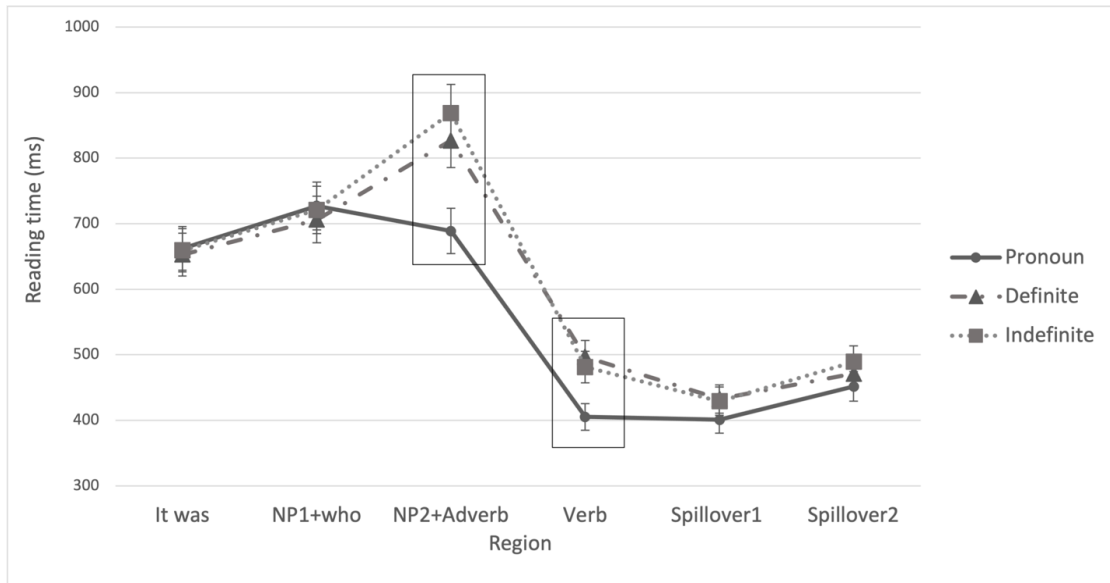
Figure 3: Mean reading time (*ms*) by region for indefinite NP1 conditions. Boxes indicate the regions where a main effect was observed.

Table 3: Mean reading time (*ms*) of the NP2+Adverb region

Average response time (*ms*)	Intervenor NP types			
	Pronoun (NP2)	Definite (NP2)	Indefinite (NP2)	Grand Total
NP2 + Adverb Region				
Definite (NP1)	689	813	898	799
Indefinite (NP1)	689	827	869	794
Grand Total	689	820	884	797

Figure 4 reports average response times of the comprehension questions across the six conditions. The overall reading time of indefiniteness was longer than definiteness (30ms) in NP1 (1549ms vs 1521ms), but the difference between the types

of NP1 was not statistically significant. However, there was a reliable effect of NP2 type. The pronoun conditions averaged more than 300ms faster than the other two conditions, which was significant (β=-375.38, *SE*=70.25, *t*=-5.34, *p*<.001). The comparison between definite and indefinite conditions revealed no significant difference, and there were no interactions (all *ps* >.41). This pattern is consistent with the result of the mean response time of comprehension questions. In brief, there was a main effect of the type of NP2 but not NP1. Participants had less difficulty when the NP2, an intervenor, was a pronoun than when NP2 was a definite or indefinite description in the dependency resolution.
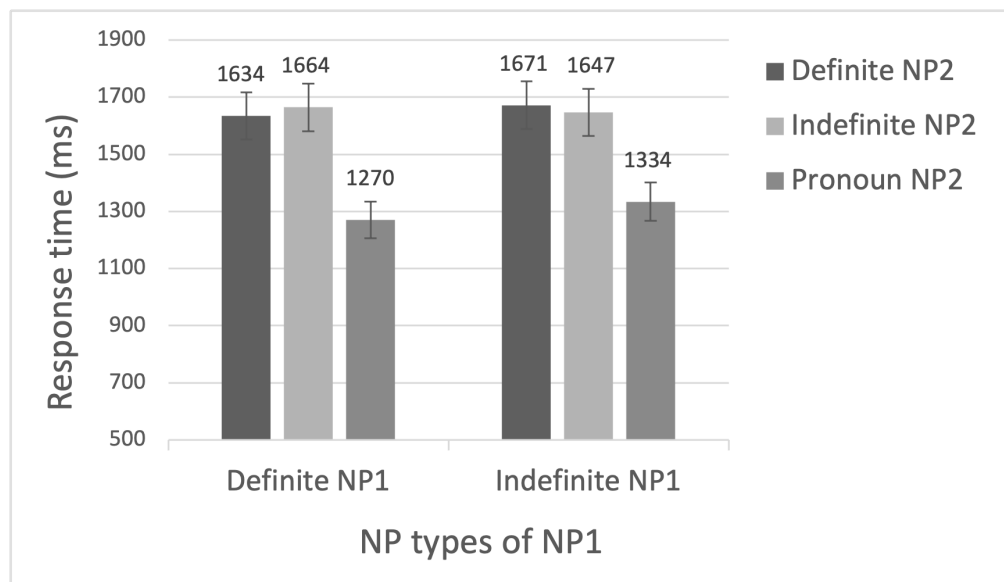
Figure 4:   The mean response times to the comprehension questions (*ms*) by condition. Error bars indicate 95% Confidence intervals.

The mean accuracy rate for comprehension questions by conditions did not reveal significant effects, which are shown in Table 4.

Table 4:   The mean accuracy rate for comprehension questions (%) by condition

	Intervenor NP type (NP2)			
	Pronoun	Definite	Indefinite	Total
Clefted indefinite (NP1)	97	97	95	96
Clefted definite (NP1)	98	97	95	97
Total	98	97	95	96

3.1.8 Discussion

Differences between the NP types of filler

I set out to investigate the effect of NP types in the resolution of dependency in clefts. This experiment found no main effect of NP1, the filler. The discourse status of the filler did not modulate the processing of the filler-gap dependency, as opposed to previous studies on other syntactic constraints. Several studies on filler-gap dependencies have provided evidence that the semantic or/and syntactic status of the filler plays a role in the processing of long-distance dependencies such as *wh*-questions (Donkers et al., 2013; Goodall, 2015), relative clauses, as well as syntactic constraints on syntactic well-formedness, such as extraction from "island" configurations in the sense of Ross (1967). For instance, semantically richer fillers (i.e., *the ruthless military dictator*; *which soldier*) were retrieved faster at the gap site than less rich fillers (i.e., *the dictator*; *which person*) (Hofmeister, 2007, 2011; Hofmeister & Sag, 2010). (i) Semantically and syntactically more complex D-linking

which NP than *who* (Donkers et al., 2013), and (ii) syntactically more prominent argument fillers than adjuncts (Hofmeister & Sag, 2010) reduced the processing time of dependency resolution. Unlike these factors of the fillers, the formation of the filler-gap dependency appears to be not susceptible to the NP types of the filler. In addition, the interference effect was also not mediated by the NP types of the filler.

Differences between the NP types of an intervenor

Evaluation of theories

The similarity-based interference effect predicts higher reading times in the definite-definite and indefinite-indefinite conditions at the critical region, which was not observed in this experiment[16]. No interaction between the NP types of the filler and an intervenor was observed. Overall, the reading time was the slowest when the intervenor was definite: definite-definite and indefinite-definite conditions. Therefore, the overall findings do not support the similarity-based interference effects.

The overall findings instead may provide support for Gibson's binary discourse-based integration cost metric in that the pronoun intervenor condition took significantly less processing load than the indefinite or definite intervenor conditions. The 1st or 2nd pronoun is a default referent in the discourse even in null contexts, whereas both definite and indefinite descriptions introduce new referents. This result supports the findings of Gordon et al (2001) and Warren & Gibson (2002) and confirmed the ease of processing pronouns, relative to descriptions.

---

[16] If we assume the encoding interference at an intervenor region, the processing difficulty should be observed when the NP types of the filler and the intervenor is the same at "NP2+adverb" region. This pattern, however, was not observed at the intervenor region.

These results are also closely aligned with the GH and AH in that the pronoun, which is the highest on the hierarchy, led to a faster reading time than full NPs, which are lower on the hierarchy. However, no effect of gradient status of an intervenor was observed; there was no significant difference between definite and indefinite intervenors in the reading time at the critical verb region. Furthermore, it rather presented a descriptively longer reading time for definite conditions than indefinite conditions, as opposed to the predictions of discourse hierarchies. According to the GH and AH, definite descriptions are on the higher hierarchy than indefinite descriptions. Therefore, the GH and AH under the DLT predict more processing load when intervenors are indefinites than when they are definites during dependency formation, which is inconsistent with the findings.

The main effect of intervenor types at the NP2 and adverb region can be accounted for by the expectation-based accounts. In previous corpus studies, the frequency of pronouns was much greater than full NPs, such as definites and indefinites, in a subject position of both main and embedded clauses. The corpus studies (see discussion in section 2.1.2) regarding the frequency between definites and indefinites reported a higher frequency of definite NPs than indefinite NPs in the embedded subject position (Givon, 1984; Leonetti, 2004). This pattern was found at the NP2 or the intervenor region in the current experiment: the pronoun intervenors were read significantly faster than the full NPs, and the definite intervenors were read significantly faster than indefinite intervenors[17]. The result supports the surprisal

---

[17] The faster reading time of a pronoun intervenor may be attributed to its shorter length than full NPs. However, I assume that the result is not simply due to its length, given the definite and indefinite contrast despite of the similar lengths.

model in that the violation of expected NP types contributes to processing difficulty, immediately upon reading the word, which was at the time parsers encounter the intervenor in this experiment.

The overall findings of processing difficulty at two different regions are consistent with the result of recent eye-movement experiments on the processing difficulty of ORCs than SRCs (Staub, 2010). He found more regressive saccades at the embedded NP of ORCs than SRCs and longer reading time at the verb of ORCs than the verb of SRCs.

In sum, a main effect of NP2 types was found in the critical verb and spillover regions. The overall findings were not fully accounted for either by the discourse hierarchy or the similarity-based interference effect. It provides support for Gibson's (2000) binary discourse-based integration cost metric in that the pronoun intervenor condition was read significantly faster than the indefinite or definite intervenor conditions. A main effect of NP2 types was also observed at the earlier time-course of processing before the integration: NP2 and adverb region, which is accounted for by expectation-based accounts. Therefore, I conclude that both experience-based expectations and processes of NP type in working memory contribute to the process of dependency formations - but critically, at different time-courses - during sentence processing.

Remaining issue with the binary metric

The results of the current experiment are not fully accounted for by either similarity or the gradient status of an intervenor in discourse. The overall theories also predict faster processing of pronouns than full NPs, which was indeed the finding

61

here. The binary metric is most compatible with the experimental results in that it only predicts the difference between pronouns, as an old referent, and full NPs, as a newly introduced referent. One caveat, however, is that this experiment and other previous experiments have investigated the effect of intervenor NP types on dependency formation without prior contexts. This arbitrary setting does not capture the lexical properties of referential expressions in contexts, which is the case where the forms are used in natural language processing.

Definite descriptions, for instance, are generally considered to refer to a recently activated referent or a particular entity in the discourse, while indefinite descriptions typically do not seek a referent but introduce a new discourse referent (I. R. Heim, 1982). In the current study, one problem with directly comparing definites and indefinites is the absence of supporting contexts (i.e., sentences were provided to participants out of the blue). A setting without contexts may give rise to a higher processing load of definites, but not indefinites because definites (but not indefinites) usually refer to old or established referents in the discourse. Thus, parsers are likely to automatically look for a referent when they encounter a definite. Since no contexts were given, they might fail to find the referent, which could be the byproduct of processing difficulty. Indefinites, on the other hand, introduce a new referent and thus do not trigger a search for the referent (Heim, 1982). Parsers do not have to trace back and no additional processing load is required for indefinite. In brief, the high accommodation cost with a particular referent may rather cause additional processing load for a definite, but not an indefinite.

A further investigation of definiteness in contexts is also necessary to evaluate the system of the binary metric, which assumes processing costs of a "new" discourse

referent. The binary metric, which is the most compatible theory with the experimental results, is based on the computation of discourse processing costs under the DLT. The costs are computed in terms of the number of new discourse referents, following the findings of Gibson & Warren (1998) and Warren & Gibson (1999). Their findings suggest that integrating across a first- or second-person pronoun (an indexical pronoun) is easier than integrating across new discourse referents such as a proper name, 3rd person pronoun lacking a referent, and full NPs in null contexts. The indexical pronouns are viewed as old referents in that their referents are present in the current discourse, whereas other complex NPs, including definites and indefinites, are viewed as new discourse referents because they are newly introduced in the discourse. In the presence of an appropriate referent in contexts, the definite description refers to a referent already given in the context. Therefore, I can explore whether the definite description would have less processing cost when it refers to an old referent, no longer introducing a new referent.

The investigation of definiteness in truly given contexts providing an appropriate referent allows us to better comprehend how the property of definiteness mediates the interference effect during the dependency formation—whether the difficulty of definite intervenors in typologically arises from the cost of accommodation.

3.2   Experiment 2: Definites in Contexts

3.2.1   Motivation of Experiment 2

The results of Experiment 1 raised the questions of (i) the lexical property of definites, and (ii) the system of binary metrics. To explore these remaining issues,

Experiment 2 focuses on examining definite intervenors in contexts. The presence of contexts advances our understanding of how the nature of NP types interacts with discourse information and mediates the interference during dependency formation. Specifically, it allows us to test the possibility that the definite description, in particular, was difficult to process in null contexts due to accommodation costs of presupposition failure driven by the absence of contextual support. The definites in the absence of a prior context that would provide a (unique) referent may have taken high accommodation costs (Kirsten et al., 2014; Löbner, 1985; Tiemann et al., 2011)[18].

This experiment on the definite descriptions also enables us to investigate the binary metrics between old and new referents. Experiment 2, therefore, tests whether the presence of an appropriate referent in a prior context attenuates the difficulty of definite descriptions, which tend to refer to an old referent. The definite descriptions do not introduce a new referent and, thus, no need to build a new referent, unlike it did in Experiment 1. The definite description becomes an old referent given in the discourse. It will serve as a diagnostic for the binary metric whether the two levels of difficulty are truly old and new referents in the discourse or inherent differences in lexical properties between pronouns and full NPs.

---

[18] It is controversial to account for semantic features of definite descriptions. The necessary property of definiteness has been argued to be two main factors—(i) uniqueness (C. Barker, 2004; Löbner, 1985; Russell, 1905; Strawson, 1950), and (ii) familiarity (I. R. Heim, 1982; Kamp, 1981), or both (Beaver & Coppock, 2015; Farkas, 1978; Rawlins, 2005; Roberts, 2003; Schwarz, 2009) —in order to account for the full distribution of definite descriptions. Whichever condition turns out to be necessary or sufficient, both factors implies that the definite descriptions are intended to identify a particular entity in contexts, which is not satisfied in null contexts.

I therefore further explore the mechanism of processing definite descriptions in the discourse, providing contexts—either satisfying or not satisfying the accommodation of an appropriate referent. Their processing cost is compared to the processing cost of indefinite descriptions. The comparison of these two NP types allows us to test the possibility of accommodation failure of presupposition for definite descriptions because indefinite descriptions do not carry presupposition of uniqueness or familiarity, unlike definite descriptions. If the absence of a referent in null contexts caused accommodation failure for the definite descriptions in Experiment 1, then the definite description should lead to less processing cost than the indefinite description when accommodation of presupposition is satisfied with the right contextual support.

3.2.2　Methods

3.2.2.1　Materials

The structures of clefted sentences were identical to those used in Experiment 1, but some of the target items consist of different professions of NP1 or NP2 from those in Experiment 1 for the purpose of natural context settings. The presupposition accommodation costs in contexts were manipulated in two ways, either felicitous or presupposition failure, in reference to the definite description. The felicitous condition provides contexts that include a unique referent, which makes it easy for the definite description to presuppose a referent. The presupposition failure condition provides contexts that have two possible referents to be referred to. This condition is infelicitous for the definite because it fails to identify its particular referent. The

65

indefinite description, on the other hand, is felicitous because it does not presuppose a unique referent.

This experiment has a 2x2 design, crossing NP types of an intervenor and the cost of presupposition accommodation for definites. The definite intervenor is compared to the indefinite intervenor, which does not need to accommodate to presuppose a unique referent. The NP type of the clefted element (NP1 in Experiment 1) was held constant, as a definite description, across conditions, given that no main effect of NP1 type was observed at the critical region in Experiment 1.

If accommodation cost contributed to the processing difficulty of the definite condition in null contexts, the processing cost of definites is predicted to differ across context types. In felicitous contexts that satisfy the presupposition accommodation, the definite condition should be easier to process than the indefinite condition. In presupposition failure conditions, on the other hand, the definite intervenor is predicted to require more processing load than the indefinite intervenor due to accommodation failure. A sample set of items is shown in Table 5.

Table 5:   Sample item set for Experiment 2

	Definite NP2	Indefinite NP2
The felicitous condition for **DEF** NP2: John, an actor, and Matt, **a director**, were at the coffee shop. They met Andy, **a writer**, who graciously thanked John for his incredible performance in the latest movie.	Target TRUE/FALSE sentence: It is the actor [who **the** writer graciously thanked for the incredible performance].	Target TRUE/FALSE sentence: It is the actor [who **a** writer graciously thanked for the incredible performance].
The presupposition failure condition for **DEF** NP2: John, an actor, and Matt, **a writer**, were at the coffee shop. They met Andy, **another writer**, who graciously thanked John for his incredible performance in the latest movie.		

In felicitous conditions, two people who have a different profession, like a director and a writer in the example, are given in context. This context makes an intervenor uniquely identifiable, which is felicitous with a definite referent to refer to 'a director' because it is an already existing unique representation in memory. That is, the definite intervenor has no processing cost to accommodate the presupposition of uniqueness.

In presupposition failure conditions, on the other hand, two people have the same profession, like a director. The addressee can associate an appropriate type representation, which is sufficient for use of the indefinite *a director,* but not the definite *the director* A definite intervenor is infelicitous due to the accommodation

failure of uniqueness presupposition. The processing difficulty of definite distractors, which need to accommodate the presupposition, may arise from the absence of a unique referent.

Experimental materials consisted of 16 sets of 4 items in each 4 condition, presented in 4 lists according to Latin square. I included 16 filler items in a similar format to the target items. For instance, three referents with their own professions were presented, followed by a word display of the last statement to ask the judgment of its truth value, as in (35).

> (35) Brian, a violinist, was playing in a concert with Kate, a cellist, and Joe, a pianist. Everyone agreed to have Brian the lead performer for the concert.
>
> Target sentence: It was the cellist who became the lead performer.
> Question: Is the last statement true or false?

3.2.2.2 Participants

Participants were 40 native speakers of English who were recruited from Amazon's Mechanical Turk web service. Participants were compensated $3.75. None had participated in Experiment 1. All participants indicated that they were native speakers of English (as with Experiment 1, payment was not contingent upon the response to this question, so there was no incentive to answer dishonestly).

3.2.2.3 Procedure

In the experiment, a full context was given all at once, followed by a by-region display of the target sentence. A true-false question such as "Is the last statement true or false?" appeared once the last word or phrase of each sentence is presented. All other procedures were identical to Experiment 1.

3.2.3 Predictions

In this section, I present predictions for Experiment 2 by processing models introduced in Chapter 2, namely, the similarity-based interference account, the binary and continuous metrics under the DLT, and expectation-based accounts. I mainly discuss predictions for felicitous conditions for definite intervenors because most theories do not make prediction regarding accommodation costs. I will mainly discuss presupposition failure conditions in discussion section[19].

3.2.3.1 Similarity-based Interference Effects

The similarity-based interference hypothesis supports the idea that the most critical factor that determines the processing difficulty is the similar features of an intervening element to the target item, causing interference. Accordingly, this model accounts for the difficulty of an intervenor in terms of its matching NP type to a target item. In Experiment 2, the NP1, or the target item, is a definite description across conditions. Therefore, it predicts the processing difficulty of a definite intervenor due to the matching NP type to the target filler, compared to an indefinite intervenor.

---

[19] The surprisal model does not directly make a prediction on the results of presupposition failure conditions, but we can anticipate its prediction on the conditions in terms of the probability of a word in contexts.

Longer reading times on the verb (that is, at the dependency formation site itself, as with Experiment 1) should be observed when an intervenor is a definite than when it is indefinite.

(36) Prediction of the similarity-based interference effect
Indefinite < Definite   (<: harder to process)

3.2.3.2   Dependency Locality Theory

Integration-based DLT is based on the notion that structural and discourse integration costs consume resources in working memory. In Experiment 2, the syntactic configurations are held constant across conditions. Thus, structural processing costs are controlled. The crucial factor that plays a role in processing difficulty in this experiment is the discourse integration cost of an intervenor during the processing of the filler-gap dependency.

Binary distance metrics

Gibson (1998, 2000) proposes that discourse processing costs of dependency formations are associated with the accessibility of an intervenor. They simplified the accessibility of an intervenor into two levels of processing difficulty based on the findings of Gibson & Warren (1998): introducing new referents imposes a processing cost, whereas old referents cause less processing difficulty. Therefore, this model hypothesizes that the integrations crossing indexical pronouns (e.g., *I, you, we*), which are always present in every context, require less processing effort than integrations crossing new referents during dependency formations. Thus, this binary metric posits no difference in processing difficulty of dependency formations between definite and indefinite intervenors. In Experiment 2, both definite and indefinite intervenors

identify a referent introduced in a prior context. They all refer to an old referent in the discourse across conditions. Therefore, no difference in reading times at the critical region is predicted between definite and indefinite intervenors.

(37) Prediction of DLT based on the binary distance metric
       Definite = Indefinite

Continuous metric: The Givenness Hierarchy

Warren & Gibson (2002) build upon the binary metrics of Gibson's (2000) work by adopting the GH proposed by Gundel (2010) and Gundel et al (1993, 2012). Their empirical studies showed the effect of a gradient hierarchy in the discourse, rather than a binary concept, on the discourse processing costs in integrations. The GH basically assumes the cognitive status of referring expressions that are high on the hierarchy tends to be more accessible than the cognitive status of those that are low on the hierarchy. The forms that are low on the hierarchy are less accessible in that they are more likely (i) to be stored in long-term memory, which causes more processing efforts, and (ii) to allow more possible interpretations.

The GH, however, points out that the cognitive status of referring expressions that is lower on the hierarchy, however, can be more accessible than the cognitive status of those higher on the hierarchy, depending on contexts. This can happen when the cognitive status of a lower hierarchy form is sufficient to construct the representation of an intended referent in contexts. Recall the sample context where a unique referent is introduced in Experiment 2, repeated here in (38).

(38) Felicitous conditions:

      John, an actor, and Matt, a director, were at the coffee shop. They met

Andy, **a writer**, who graciously thanked John for his incredible

performance in the latest movie.

Target TRUE/FALSE sentence:

    (i) It is the actor [who **the** writer graciously thanked for the

        incredible performance].

    (ii) It is the actor [who **a** writer graciously thanked for the

        incredible performance].

The context introduces one particular writer, *Andy*, who becomes an intervenor in the target sentence. The cognitive status encoded to the definite description, *the writer*, in the target sentence is "uniquely identifiable", and, therefore, it is felicitous to use the definite description. The cognitive status of the indefinite description, *a writer*, is "type identifiable" at most. It encodes a type representation with the property of being a writer. The type-identifiable interpretation allows referring any referent whose profession is a writer. This possible interpretation, however, is restricted, and instead, parsers can associate a unique representation, *Andy*, because there is only one writer given in a prior context. The type-identifiable interpretation is enough to refer to the intended referent, *Andy*, presented in the context. Therefore, simply identifying the type in this context is not predicted to be more difficult than identifying a unique referent. The "uniquely identifiable" interpretation may involve more processing effort than necessary to retrieve an existing representation from memory. The GH predicts that the type-identifiable interpretation (indefinites) is felicitous and sufficient to build

the representation of the intended referent, which would incur less processing effort than the uniquely identifiable interpretation (definites).

(39) Prediction of DLT based on the GH

Indefinite < Definite (<: harder to process)

Continuous metric: The Accessibility Hierarchy

The AH assumes that referring expressions directly represent the degree of accessibility. This account assumes that pragmatic factors such as general knowledge, physical surroundings, and previous linguistic materials are all incorporated into the forms within the framework of the accessibility hierarchy (Ariel, 1988, 1994). Ariel (1990) claims that different forms mark the different amount of processing cost in the retrieval of the intended entity.

(40) An addressee is instructed to retrieve a mental representation which may be characterized by reference to the individual features associated with it ('wise', 'short'), but always also with a feature establishing its current Accessibility to him.

(Ariel, 1990, p. 16)

Therefore, this theory generalizes the order of accessibility such that highly accessible referents incur less processing cost than less accessible referents, regardless of other factors. The AH proposed by Ariel does not clearly compare the accessibility of definite and indefinite descriptions. Following Givon (1983)'s hierarchy, in line with the spirit of the AH, the AH would predict that a definite intervenor is more

accessible than an indefinite intervenor. Thus, definite intervenors are predicted to be easier to process than indefinite intervenors.

(41) Prediction of DLT based on the AH

      Definite < Indefinite   (<: harder to process)

3.2.3.3 Expectation-based Models

This model predicts that more common and frequent words are easier to process than less common words. The processing difficulty at a given word is determined by the probability of the word within the syntactic configuration and the meaning of a sentence. Given that definites are used to identify a unique referent, whereas indefinites identify a type of entity, it is predicted that definite conditions are easier to process than indefinite conditions in felicitous conditions, and the reversed pattern is predicted in presupposition failure conditions. This main effect should be observed at an earlier time point upon encountering the NPs, the intervenor region.

(42) Prediction of DLT based on the GH (at the NP2 region)

    Felicitous condition: Definite < Indefinite

    Presupposition failure condition: Indefinite < Definite

                              (<: harder to process)

The predictions of processing theories with regard to the processing difficulty of the dependency formation are summarized in Table 6.

Table 6:   Prediction by each model for the effect of NP types in sentence processing

Models	Prediction
Memory-based account (at the main verb region)	
Similarity-based interference	Indefinite < Definite
Binary metric	Definite = Indefinite
Continuous metric: Givenness Hierarchy	Indefinite < Definite
Continuous metric: Accessibility Hierarchy	Definite < Indefinite
Expectation-based account: Surprisal model (at the NP2 region)	Felicitous conditions: Definite < Indefinite (<: harder to process) Presupposition failure conditions: Indefinite < Definite (<: harder to process)

3.2.4 Results

7 out of 40 participants were excluded due to lower performance accuracy than 80% across all comprehension questions, including fillers. The data of 33 participants, thus, were analyzed. Of the 528 critical trials generated, the trials which generated incorrect responses (n=83) were excluded from further analysis. The main regions of interest were identical to the three regions as with Experiment 1: the critical verb and the following two regions. The same methods were used for outlier removal and modeling procedure as reported in Experiment 1. NP2 region was also analyzed to investigate the effect of NPs once encountering an intervenor. The offline

processing—response time to comprehension tasks and the accuracy—was also observed.

Data were analyzed with linear mixed-effects regression models performed with the *lme4* package in the R software and conducted using parsimonious models as the analysis conducted in Experiment 1. The predictors, context type (2 levels) and intervenor NP type (2 levels), were contrast-coded. For accommodation costs, the felicitous condition was coded as +1/2 and the presupposition failure condition was coded as -1/2. For intervenor NP type, definite NPs were contrasted with indefinite NPs. The definite NPs were coded as +1/2 and the indefinite NPs were coded as -1/2.

The reading time on the main verb revealed no main effects either of accommodation costs or intervenor NP type (*ps* > .20), but the interaction was significant (β=-99.55, *SE*=39.88, *t*=-2.50, *p*=.01). Figure 5 reports region-by-region reading times by condition. Table 7 illustrates the mean reading times of critical and spillover regions. That is, the effect of definiteness of intervenors differed according to accommodation costs. Planned comparisons showed no effect of intervenor type in the felicitous condition (*p* = .4) but in the presupposition failure condition, the indefinite intervenor condition was read significantly faster than the definite intervenor condition (431ms vs 523ms, β=-77.11, *SE*= 28.20, *t*=-2.73, *p*=.006). No significant main effects or interaction was observed in spillover 1 and spillover 2 regions (*ps* > .17).

At the NP2 region, there was no significant effect or interactions (all *ps* > .28). Instead, the significant results were revealed in the spillover region (adverb). At the adverb region, I observed a main effect of context type (506ms vs 571ms, β=69.17, *SE*=32.99, *t*=-2.10, *p*=.04); that is, the adverb was read significantly faster in the presupposition failure condition than in the felicitous condition. The planned

comparisons also revealed a marginal effect of intervenor NP types in the presupposition failure condition such that the indefinite condition was read marginally faster than the definite condition (470ms vs 539ms; β=-50.99, SE=35.51, t=-1.44, p=.15). I assume that the faster reading time of the indefinite condition in the presupposition failure condition contributed to the effect of accommodation costs.

I now turn to data regarding the answer to the comprehension question. Figure 6 reports response time to the comprehension question. Here, unlike with reading times, no main effects or interaction was observed (ps > .43). The mean accuracy rate for comprehension questions was 84% and the mean accuracy by conditions is shown in Table 8. A logistic mixed-effects regression model showed no main effects (all ps >.33) but a significant interaction of accommodation costs and intervenor NP types (β=-0.95, SE=0.50, t=-1.92, p=.05). Planned comparisons revealed that this interaction was driven by a main effect of intervenor NP types in the presupposition failure condition (β=-0.72, SE=0.37, t=-1.98, p=.04). When the intervenor is definite, a Speed-Accuracy Tradeoff was found only in presupposition failure conditions. The definite intervenor led to a significantly longer reading time at the verb but higher accuracy than the indefinite intervenor in presupposition failure condition. This inverse result between speed and accuracy was found in other dependencies (Foraker & McElree, 2007; McElree, 2000; McElree et al., 2003).
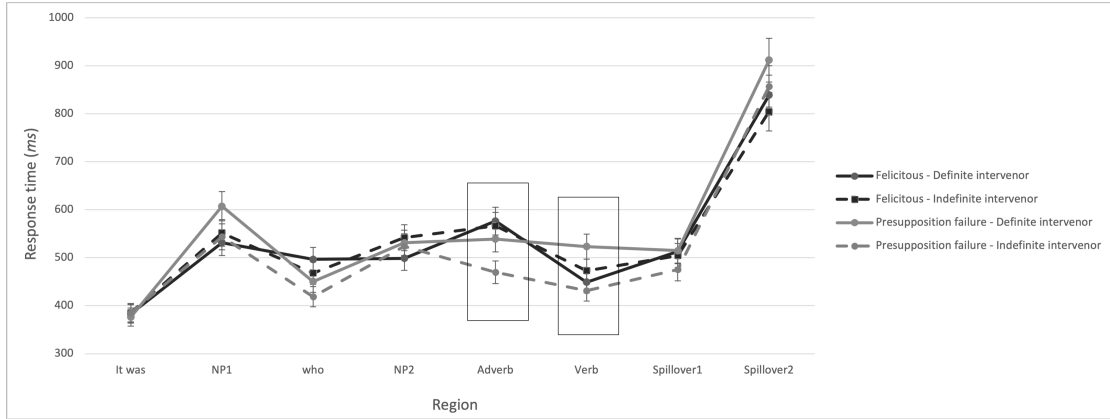
Figure 5:   Mean reading time (*ms*) by region by condition. Boxes indicate the regions where a main effect was observed.

Table 7:   Mean reading time (*ms*) of critical and spillover regions

Average response time (*ms*)	Intervenor NP types		
	Definite	Indefinite	Grand Total
Verb Region			
Felicitous condition	449	473	461
Presupposition failure condition	523	431	480
Grand Total	488	453	471
Spillover 1 Region			
Felicitous condition	513	504	509
Presupposition failure condition	515	476	496
Grand Total	514	490	502
Spillover 2 Region			
Felicitous condition	839	805	822
Presupposition failure condition	912	857	886
Grand Total	877	830	854
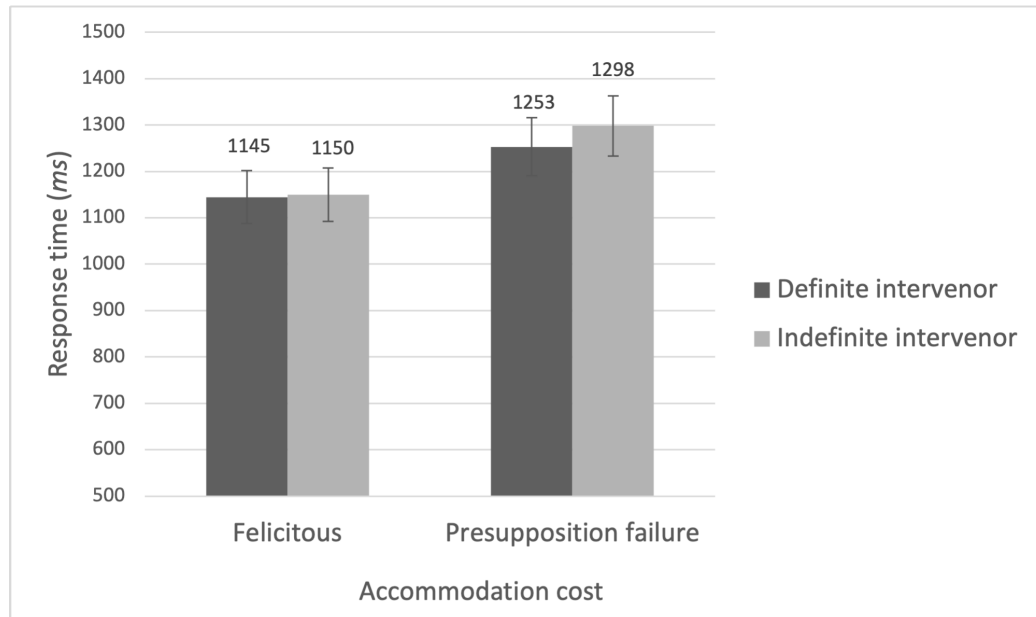
Figure 6:   The mean response times to the comprehension questions (*ms*) by
             condition. Error bars indicate 95% Confidence intervals.

Table 8:   The mean accuracy rate for comprehension questions (%) by condition

	Intervenor type		
	Definite	Indefinite	Total
Felicitous condition	82	85	83
Presupposition failure condition	89	81	84
Total	86	83	84

3.2.5   Discussion

Evaluation of theories

I evaluate each model based on the reading time results presented in this

chapter. According to similarity-based interference effects, the main factor that

mediates the interference effects on dependency formations is similar NP types of an intervenor to the target NP. That is, the matching NP type of an intervenor to the target element is more difficult to process than a mismatching intervenor because the matching cues interfere with the retrieval of the target element in working memory. In terms of the experimental result, the similarity-based interference account predicts the definiteness effect in the definite-indefinite contrast due to its matching NP type to NP1, the clefted element (also, definites). The result, however, showed no significant difference in the reading time at the critical region between definite and indefinite intervenors when a unique referent was introduced in a prior context. The result rather revealed descriptively faster reading time when an intervenor is a definite than when it is indefinite.

A broader perspective of similarity-based interference effects of NP types can be compatible with the findings in felicitous conditions (Gordon et al., 2004) They proposed that similarity interferences of NP types of an intervenor are based on the property of common nouns. Gordon et al (2004) conducted a follow-up study based on Gordon et al (2001), collecting more data on other NP types of an intervenor such as indefinites (e.g., *an accountant*), generic expressions (e.g., *the accountants*), quantified expressions (e.g., *everyone*), and semantically lean nouns (e.g., *the person*). Putting the result of Gordon et al (2001) together, they showed that proper names, pronouns, and quantified expressions reduced the processing difficulty between SRC and ORC, compared to definite types than definite descriptions. The other NP types such as definites, indefinites, generics, and semantically lean nouns did not make a difference between SRC and ORC. The sample item is presented in (43) and the summary of overall findings is shown in Table 9.

(43) SRC: The salesman that the accountant contacted spoke very quickly.

ORC: The salesman that contacted [the accountant/an accountant/the accountants/the person/everyone] spoke very quickly.

Table 9:   Summary of results from Gordon et al (2001, 2004)

Baseline (the head noun)	Reduced ORCs-SRCs difference	No reduced ORCs-SRCs difference
Definite Description "the salesman"	Proper name ("Bob") Indexical pronoun ("you") Quantified expression ("everyone")	Indefinite description ("a salesman") Generic ("salesmen") Semantically lean noun ("the person")

They grouped NP types into the property of common nouns. The NPs which have the property of common nouns (e.g., *definites, indefinites, generics*) and those that do not (e.g., *proper names, pronouns, everyone*) differ in representational similarity. The processing difficulty of the former NPs was attributed to their similarity to the filler NP type, sharing common noun features in the memory representation. That is, similarity-based interference effects occur when an intervenor has a property of common nouns, the same as the filler. Accordingly, the result of Experiment 2 can be accounted for by the property of a common noun. Indefinite and definite intervenors shared the property of common nouns with the indefinite filler,

unlike pronoun intervenors. Indefinites and definites, as common nouns, required more processing load than pronouns that do not share the features of a common noun.

This broad perspective of similarity-based interference account predicts that both definite and indefinite descriptions belong to the same category, having the property of common nouns. This account is compatible with the result shown in the context providing a unique referent in that no significant difference between definite and indefinite intervenors was observed. In presupposition failure conditions, however, we observed another source of difficulty in dependency formation: the difference in the accommodability of presupposition of an intervenor in contexts.

In the binary metric, the two main levels of discourse processing costs are predicted such that each new discourse referent consumes 1 energy unit, whereas an old discourse referent does not. Only an indexical pronoun, as an old referent, is easy to process in a null context because its referent is always included in the current discourse. Other NP types such as a proper name and full NPs consume energy resources, constructing a new discourse referent. According to the binary metric, both definite and indefinite NPs force the parser to construct a mental representation of a new referent, unlike indexical pronouns. Given that an old referent was given in a prior context in Experiment 2, this model predicts the processing costs of both definite and indefinite intervenors. As the similarity-based account, it does not concern with the distinct property of a definite article *the* from an indefinite article *a* in terms of accommodation cost of presupposition and how the processing costs for accommodations of presupposition modulate dependency formations.

The GH theory does not predict that different forms directly reflect the accessibility of cognitive statuses and assumes the interaction with discourse and

pragmatic factors. For instance, in the contexts where other possible interpretation of a lower form is restricted, the cognitive status of lower forms on the hierarchy can be underspecified because the cognitive status of high forms on the hierarchy entails that of lower forms. The experimental condition which introduced a unique referent in a prior context can be an example of this case. Since this context excludes the possibility of referring to other referents, the type-identifiable interpretation (indefinites) is enough to refer to the unique referent described in the context. Therefore, simply identifying the type can be more accessible than identifying a unique referent. The results do not provide direct support for the GH under the DLT because there was no indication of a difference in processing difficulty between definite and indefinite conditions when a unique referent was provided in contexts.

The AH proposes that the retrieval of an intended referent is dependent on the degree of accessibility such that a higher degree of accessibility indicates relatively easy processing of retrieval. This degree of accessibility is a crucial factor that determines the referring forms. In the AH (Ariel, 1988, 1990), definite descriptions belong to low accessibility referring expressions and mainly discussed definite descriptions and proper names within the category of low accessibility referring expressions. Ariel (1988) does not directly compare the accessibility of indefinite descriptions to that of definite descriptions but discusses Givon (1983)'s accessibility hierarchy, which shares a similar idea with Ariel's AH. In Givon's (1983) hierarchy, indefinite NPs are assumed to be less accessible than definite NPs. I, therefore, assume that the AH would predict more processing cost of indefinite intervenors than definite intervenors, based on the notion that the referring expressions represent their own degrees of accessibility. The experimental results, however, did not support the

predictions of the AH: the statistical analysis revealed no significant difference in reading time at the critical region when contexts were felicitous, providing a unique referent.

Expectation-based theories are based on the probability of words, which can be accounted for by frequency in corpus analysis and parsers' experience in language. Accordingly, this model is on the assumption that common and frequent words are easier to process than less common words. Given that definite intervenors are infelicitous in conditions where two potential referents exist, definite intervenors are not likely to be used in this case. A comparison between the frequency of definites and indefinites in this context, therefore, would show that indefinites are more frequent and common to refer to a non-unique referent among two referents. This effect was shown at the adverb region, right after the intervenor region, such that the indefinite condition is read marginally faster to process than the definite condition when contexts do not provide a unique referent. The marginal difference observed in the spillover region, following the intervenor region, may be due to the accommodation process to a referent in a prior context upon reading the intervenor. In addition, the main effect of accommodation costs at the adverb region was also driven by a fast reading-time of the indefinite condition in the contexts without a unique referent, compared to the other three conditions. In terms of frequency, it is predicted that the parsers found easy processing of the indefinite condition due to its frequent usage in presupposition failure conditions. The use of definites is infelicitous when there are two possible referents, and thus definites are hardly expected, compared to indeifnites.

In brief, the similarity effect account (Warren & Gibson, 2004) and binary metric (Gibson, 2000) can account for the similar processing load of definite and

indefinite conditions in felicitous conditions providing a unique referent because they belong to the same category in both accounts. The processing difficulty of definites in presupposition failure conditions providing two potential referents, compared to indefinites, on the other hand, clearly showed that differences in accommodability in discourse context mediates the difficulty of dependency formations, in addition to the factors discussed in processing theories. Given the effect of accommodation costs observed in the presupposition failure condition, experimental results in felicitous conditions can be also accounted for by the different properties of definites and indefintes in terms of accommodation for presupposition. I discuss this issue in the next subsection.

Implications for presupposition accommodations

Definite descriptions carry presuppositions of unique or familiar referents, while indefinite descriptions do not carry presuppositions of uniqueness or familiarity because they are often used to introduce new referents (Haviland & Clark, 1974; I. R. Heim, 1982). Previous literature has examined the definite and indefinite determiners regarding the uses of presuppositions in different contexts and experimental settings. Most of the literature used different tasks, involving decision-making processes, such as picture selections (Bade & Schwarz, 2019b, 2019a, 2021; Masharov, 2008; Schwarz, 2007), production tasks (Schneider et al., 2019, 2021), and different methods such as visual world eye-tracking (Bade & Schwarz, 2019b, 2019a; Schwarz & Tiemann, 2012) or ERP (Kirsten et al., 2014; Tiemann, 2014; Tiemann et al., 2011), and mouse-tracking (Schneider et al., 2019), including reading time studies (Clifton 2013). These studies have paid attention to examining the processing differences

between definites and indefinites in simple sentences in order to focus on their underlying mechanism regarding the accommodation of presuppositions. There are no further experimental investigations, to my knowledge, of how the accommodation of presuppositions of definite/indefinite intervenors mediates dependency formations. The current study provides insights into the strong role of accommodation of an intervenor's presuppositions during dependency formations, even though the accommodation of presuppositions is not associated with the target item, or the filler.

In presupposition failure conditions providing two potential referents, the use of a definite intervenor was infelicitous because it failed to seek its unique referent in the discourse, leading to a significantly longer reading time than an indefinite intervenor. This result confirms the common assumption that definites trigger a search for a particular referent (Clifton, 2013; I. Heim, 1991; I. R. Heim, 1982). Definite descriptions take up processing effort to build the representation of the referent when no referent is given in the context. The absence of a referent in null contexts, however, is presumed to be not as difficult as the presence of two potential antecedents, which causes the failure of the accommodation. In null contexts, parsers can accommodate the presupposition by building the representation of a unique referent in their minds, whereas in the presence of two referents, parsers have no other way to posit a unique referent and fail to accommodate the presupposition due to two potential referents. Therefore, it becomes infelicitous to use definites, which leads to the robust processing difficulty of definites, compared to indefinites. The use of an indefinite is not infelicitous when two potential referents exist in contexts because it does not take up accommodation costs to presuppose a unique referent, unlike a definite. For the use of indefinites, it is sufficient to identify a type of person that has the profession it

describes. In brief, the failure of accommodation for definite conditions in contexts with no unique referent gave rise to significantly more difficult processing than indefinite conditions.

In felicitous contexts, on the other hand, the definite condition showed a descriptively faster reading time than the indefinite condition but it was not significant. No significance between definite and indefinite conditions can be also accounted for by their different properties of presuppositions. That is, the uniqueness information did not contribute to a higher processing cost for indefinites than definites. This could be because indefinites do not carry presuppositions of uniqueness or familiarity. Indefinites, thus, were compatible with referring to a unique referent that has been introduced in the context. They do not necessarily have to be associated with a new discourse referent or a non-unique referent. This finding provides support for the weaker epistemic status of indefinites' so-called "anti-uniqueness" or "non-uniqueness" inferences than the strong uniqueness presupposition of definites (Alonso-Ovalle et al., 2011; Chemla, 2008; I. Heim, 1991; Percus, 2006; Sauerland, 2008).

Indefinites are assumed to bear the anti-uniqueness inference such that there is not exactly one unique referent. For example, in the sample target sentence of the current experiment ("*It is the actor who a writer graciously thanked for the incredible performance.*"), the indefinite description "a writer" yields the inference that there is more than one writer. This anti-uniqueness presupposition of indefinites was not as strongly used as the uniqueness presupposition of definites is used for picture selection tasks in several studies (Bade & Schwarz, 2019a, 2021). Following this theory, the experimental result of no significant processing cost between definite and indefinites

in the contexts providing uniqueness information, thus, can be attributed to the weak anti-uniqueness presupposition of indefinites. That is, for instance, the presupposition that there is not exactly one unique writer in the sample sentence is weak enough to be compatible in contexts where there is one unique writer. The usage of indefinites is not a strong disadvantage even when the anti-uniqueness presupposition is not met, having a unique referent.

Previous works of literature have examined the usage of definite and indefinite descriptions in the main subject position, as a topic, which describes a referent in certain situations or contexts. They have paid attention to exploring the processing of felicitous and infelicitous uses of definites in different contexts. However, experimental investigations have not delved into a deeper question of how the presupposition accommodation of a definite mediates interference effects on the resolution of dependencies, to the best of my knowledge. Experiment 2 further shows the influence of presupposition accommodation properties of full NPs on online processing as well as sentence comprehension; the presupposition accommodation properties of definites and indefinites in the intervening position between the dependent elements modulate the resolution of dependencies.

3.3   Experiment 3: Pronouns in Contexts

3.3.1   The Motivation of Experiment 3

The first experiment explored the effect of intervenor NP types with no reference to contexts. The experimental setting in a null context raised a possibility of the effect of pragmatic factors such as the presupposition accommodation on processing a definite intervenor. The second experiment tested this possibility by

manipulating the difficulty of the accommodability in a given context and compared it to an indefinite. Experiment 2 showed a strong effect of the accommodability conditions. Experiment 3 aims to further explore how the accommodation property of a pronoun intervenor modulates dependency formation. The question is whether the processing cost of accommodation has an impact on the dependency formation for pronouns as it does for definites and indefinites despite the advantage of pronouns, being most accessible and the highest on the discourse hierarchy. The investigation of the 3rd person pronoun (e.g., *he, she*) serves as a diagnostic for accommodation costs of a pronoun because it must accommodate the presupposition of its referring entity.

In terms of the status of referents in the discourse hierarchy, definite and indefinite descriptions observed in Experiment 2 have been assumed to lead to more processing costs than pronouns. Previous literature has provided support for less processing cost of indexical pronouns than other complex NPs. For instance, Gibson & Warren (1998) and Warren & Gibson (2002) included a 3rd person pronoun in their experiments and provided support for easier processing of indexical pronouns than other NPs, including a 3rd person pronoun, in a null context. This difference has been attributed to the fact that the 3rd person pronoun is new to the discourse and carries presuppositions of a referent in the discourse. Therefore, I investigate whether the presence of its referent in a given context promotes the ease of processing the 3rd person pronoun like indexical pronouns.

3.3.2   Materials

The identical target sentences and similar contexts to Experiment 2 were used, but some of the professions or names and sentences in contexts were revised for the

purpose of natural context settings. The context was manipulated in two ways, either low or high level of accommodation costs to presuppose a referent. In the context with a low accommodation cost, a unique referent is provided to be an intended referent, whereas, in the other context with a high accommodation cost, there are two potential referents to be identified in target sentences. This experiment has a 2x2 design, crossing NP types of an intervenor and the level of accommodation costs. The NP type of the clefted element (NP1 in Experiment 1) was held constant, as a definite description, across conditions, as in Experiment 2.

If the processing of the pronoun condition is sensitive to accommodability, the opposite pattern between pronoun and indefinite conditions should be observed across context types. As illustrated in Experiment 2, indefinites do not need to accommodate uniqueness or familiarity presuppositions. Therefore, indefinite conditions should be easier to process than pronoun conditions when contexts require a high processing load to accommodate a referent. In the other condition with low accommodation cost, on the other hand, the pronoun condition may be easier to process than the indefinite condition as it did in Experiment 1.

Table 10: Sample item set for Experiment 3

	Pronoun NP2	Indefinite NP2
Context with low accommodation cost: Jane, a writer, and Alice, another writer, were at the coffee shop. **Jane** graciously thanked Matt, an actor, for his incredible performance in the latest movie.	Target TRUE/FALSE sentence: It is the actor [who **she** graciously thanked for the incredible performance].	Target TRUE/FALSE sentence: It is the actor [who **a writer** graciously thanked for the incredible performance.]
Context with high accommodation cost: Jane, a writer, and Alice another writer, were at the coffee shop. **One of them** graciously thanked Matt, an actor, for his incredible performance in the latest movie.		

In both contexts, two people who share the same gender feature and profession are given in context. For instance, Jane and Alice are typically females and they have the same profession, a writer. The following sentence refers to one of the two people introduced in the first sentence. It is the referring expression that differs across context types.

The context type which takes up low accommodation cost provides a unique referent by specifying either of their names, while the other context type takes more processing cost to accommodate the referent because the following sentence does not specify a unique referent and use a phrase, "one of them".

Experimental materials consisted of 16 sets, presented in 4 lists according to Latin square. I included the same 16 filler items with items used in Experiment 2. The gender of referents and the order of the intended referents were balanced across items.

3.3.3   Participants

Participants were 40 native speakers of English who were recruited from Amazon's Mechanical Turk web service. Participants were compensated $3.75. None had participated in Experiment 1 and 2. All participants indicated that they were native speakers of English (as with Experiments 1 and 2).

3.3.4   Procedure

All other procedures were identical to Experiment 2. A full context was shown at a time followed by a word or phrase display of a target sentence. The same type of true-false question appeared after the last word or phrase of a target sentence. The same 16 filler items used in Experiment 2 were used for this Experiment.

3.3.5   Predictions

The predictions of processing models for the effect of NP types are presented in this section.

Similarity-based interference effects

According to the broader perspective of similarity-based interference accounts by Gordon et al (2001, 2004), the processing difficulty of indefinites compared to

pronouns is predicted due to the similar features of the definite filler and indefinite intervenor, sharing the property of common nouns. Recall that Gordon et al (2004) suggested the similarity-based interference effects of NP types in terms of the category of common nouns such that sharing the same feature of being common nouns to the filler causes processing difficulty. Indefinite intervenors are common nouns as the definite filler, whereas 3rd person pronoun intervenors are not common nouns, unlike the definite filler. Therefore, the indefinite intervenors are expected to present more processing load than 3rd person pronoun intervenors in the current experiment.

The binary metric

Under the DLT, the processing difficulty of discourse integrations depends on the accessibility of the referent in the discourse. The more integration costs are consumed when the referent is less accessible in the discourse (Garrod & Sanford, 1994; Haviland & Clark, 1974), (Following Gibson (1998), Gibson (2000) assumes building a new representation of a referent in the discourse consumes resources during integration. In the current experiment, both 3rd person pronoun and indefinite intervenors refer to the person introduced in a prior context. Therefore, the parsers do not have to construct a new referent when the intervenor is encountered during sentence comprehension. On this view, both 3rd person pronoun and indefinite intervenors are not predicted to be a new discourse referent and do not give rise to integration costs.

The continuous metric: The Givenness Hierarchy

The discourse integration costs based on the GH also depend on the accessibility of the referent in the discourse but are expected to be more gradient levels of processing difficulty. The GH is associated with how accessible the cognitive status of the referent is; the addressers have knowledge on the cognitive status encoded by referring forms, interacting their knowledge with the addressee's knowledge and attention state of the forms in the discourse. According to the GH, a 3rd person pronoun is activated in short-term memory and focus of attention in the discourse by referring to the topic of the preceding context. The pronoun is most accessible, signaling that the intended referent is in focus. An indefinite description, on the other hand, signals that a representation of the type of an entity is identifiable. This indefinite description needs to construct more conceptual content than a pronoun, which is much less accessible than the pronoun. Therefore, the GH predicts the easier processing of the pronoun than the indefinite description even though both forms are felicitous to use in contexts.

The continuous metric: The Accessibility Hierarchy

The AH (Ariel, 1988, 1990) is based on the degree of accessibility under the assumption that the referring forms indicate their own degree of accessibility with no reference to contexts. This model predicts that the degree of accessibility of pronouns is highly accessible, compared to full NPs such as definites and full names. Indefinites are assumed to be less accessible than definites under the hierarchy of Givon (1983), which shared the spirit of the AH proposed by Ariel (1988, 1990). I, thus, assume that, considering the degree of accessibility, the 3rd person pronoun causes more processing load than the indefinite.

3.3.6 Results

4 out of 40 participants were excluded due to lower performance accuracy than 80% across all comprehension questions, including fillers. The data of 36 participants, thus, were analyzed. Of the 576 critical trials generated, the trials which generated incorrect responses (n=82) were excluded from further analysis. The critical verb and the following two regions were analyzed as the main regions of interest as with Experiment 1 and 2[20]. All other methods of data analysis were identical to previous experiments.

In linear mixed-effects regression models, the predictors, context type (2 levels) and intervenor NP type (2 levels), were contrast-coded. For context type, the context with low accommodation cost was coded as +1/2 and the context with high accommodation cost was coded as -1/2. For the intervenor NP type, pronoun NPs were contrasted with indefinite NPs. The pronoun NPs were coded as +1/2 and the indefinite NPs were coded as -1/2.

At the main verb, a main effect of intervenor NP types (β=-60.70, *SE*=21.64, *t*=-2.81, *p*=.008) and no main effect of the level of accommodation costs and interaction was found (*ps* >.30). For intervenor NP types, the pronoun conditions were

---

[20] A main effect of intervenor NP types was observed in the NP2 and adverb regions, such that pronoun intervenors were processed faster than indefinite intervenors when the intervenors were encountered and the following region. This effect can be due to shorter length of pronouns than indefinites. In order to properly measure the reading time at these regions, log-transformation is necessary for data analysis (Box & Cox, 1964; Vasishth et al., 2013). I will focus on the analysis of critical regions in the thesis.

read faster than indefinite conditions across different accommodation costs. The planned comparisons confirmed the main effect of intervenor NP types (for contexts with low accommodation cost: β=-57.60, SE=28.58, t=-2.25, p=.04; for contexts with high accommodation cost: β=-65.04, SE=28.91, t=-2.25, p=.02). Figures 7 reports region-by-region reading times by context type and Table 11 presents the reading times of the critical and spillover regions.

Concerning the Spillover regions, there was a main effect of accommodation costs (β=59.31, SE=27.80, t=2.13, p=.04) at the Spillover 1 region, meaning that the overall reading time in low accommodation cost conditions providing a unique referent was longer than in high accommodation cost conditions without a unique referent. This processing load of high accommodation cost conditions might be due to the process of consolidating the unique referent in the sentence meaning.

I now turn to data regarding the answer to the comprehension question. For response time (see Figure 8), a significant effect of intervenors (β=333.05, SE=153.76, t=2.15, p=.04) and a marginal effect of accommodation costs (β=-217.40, SE=123.54, t=-1.76, p=.09) were found. The main effect of intervenor types indicates that response times were longer when the intervenor was a pronoun than when it was an indefinite (1593ms vs 1281ms). In terms of the level of accommodation costs, the overall reading time was longer in contexts with high accommodation costs than in contexts with a low accommodation cost. No interaction between intervenor NP type and context types was found (ps >.99)

The mean accuracy rate for comprehension questions was 86% and the mean accuracy by conditions is shown in Table 12. I observed a marginal effect of the intervenor NP types (88% vs 84%), in which sentences with a pronoun intervenor

generated higher accuracy rates than those with an indefinite intervenor (β=-0.34, SE=0.26, t=-1.32, p=.18). Planned comparisons confirmed that the marginal effect of intervenor NP types was present only in contexts with high accommodation cost (β=0.62, SE=0.37, t=1.67, p=.09); in contexts with high accommodation cost, the accuracy rate for the pronoun condition was higher than the indefinite condition.
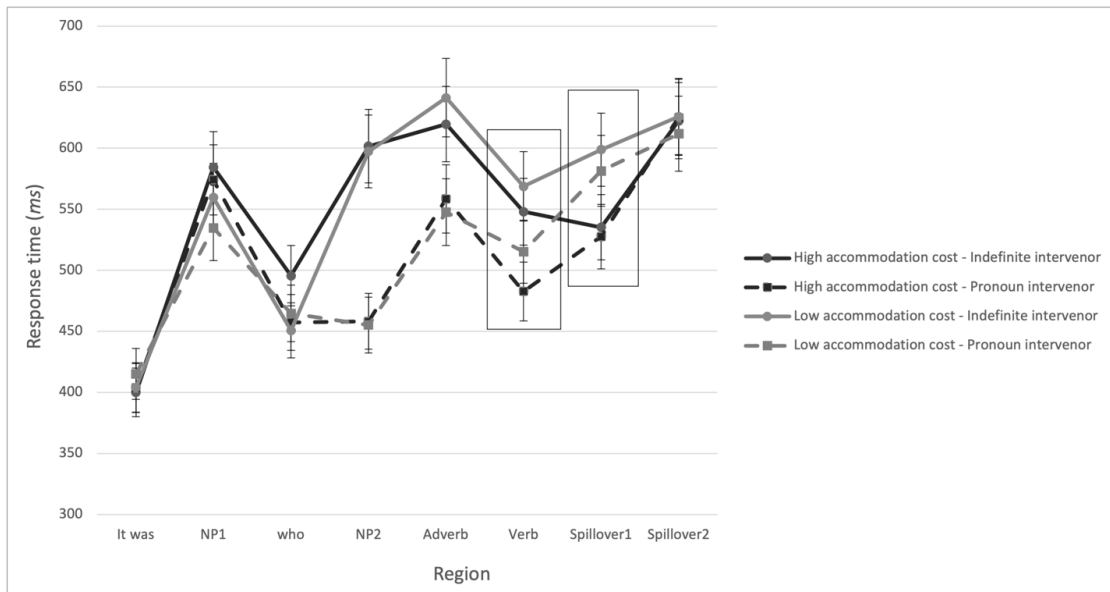
Figure 7:   Mean reading time (*ms*) by region by condition. Boxes indicate the regions where a main effect was observed.

97
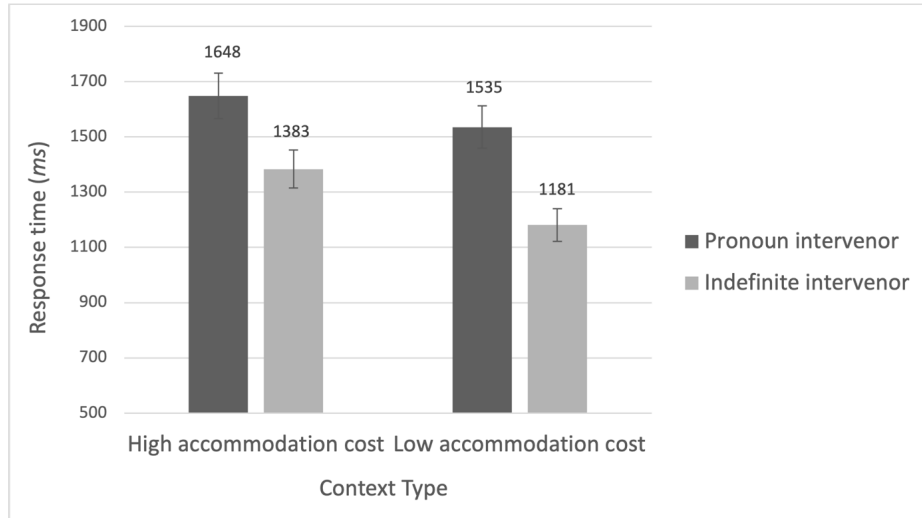
Figure 8:   The mean response times to the comprehension questions (*ms*) by
condition. Error bars indicate 95% Confidence intervals.

Table 11:   Mean reading time (*ms*) of critical and spillover regions

Average response time (*ms*)	Intervenor NP types		
	Pronoun	Indefinite	Grand Total
Verb Region			
High accommodation cost	483	548	514
Low accommodation cost	515	569	542
Grand Total	499	558	528
Spillover 1 Region			
High accommodation cost	528	535	531
Low accommodation cost	581	599	590
Grand Total	554	567	560
Spillover 2 Region			
High accommodation cost	625	622	624
Low accommodation cost	612	629	619
Grand Total	619	624	621

Table 12:   The mean accuracy rate for comprehension questions (%) by condition

	Intervenor type		
	Pronoun	Indefinite	Total
High accommodation cost	90	83	86
Low accommodation cost	85	85	85
Total	88	84	86

In brief, the results of offline processing suggest that (i) indefinite intervenors are easier to retrieve discourse contexts than pronoun intervenors because of more lexical information on indefinite intervenors, and (ii) in contexts with low accommodation cost, a faster reading time and accuracy different was observed, compared to contexts with a high accommodation cost. In the latter contexts, a Speed-Accuracy Tradeoff was found, similar to Experiment 2.

3.3.7   Discussion

The goal of Experiment 3 was to investigate how the accommodability of the 3rd person pronoun intervenor, which is known to be high on the discourse hierarchy, modulates the filler-gap dependency formation in contexts.

Evaluation of theories

In the similarity-based account proposed by Gordon et al (2004), the processing difficulty arises from the similarity of the filler and intervenor in terms of the features of common nouns. The 3rd person pronoun NP type differs from the definite NP, or the filler, in that the pronoun is not a common noun. The indefinite NP, on the other hand, has the features of a common noun as the filler, denoting the set of

entities having the same profession. Accordingly, the similarity-based interference account made the prediction that indefinite intervenors sharing similar features with definite fillers contribute to the processing load, compared to pronoun intervenors, which is consistent with the experimental results.

Both the GH and AH within the framework of the DLT can account for this result, as can the similarity-based account. Under the GH, the pronoun is the most accessible referring expression because it tends to refer to the previously mentioned or existing referent in the current focus of attention. Unlike pronouns, the full NPs such as indefinite descriptions must construct conceptual contents, which causes more processing efforts than pronouns. Therefore, pronouns are predicted to be easier to process than indefinites when both NP types are felicitous and contextually supported. It is hardly possible for the cognitive status of full NPs to be more accessible than that of pronouns because the encoded cognitive status of the pronouns is in the current focus with a minimum cost of building conceptual content in any given context.

The AH (Ariel, 1988, 1990) predicts that the degree of accessibility of pronouns is highly accessible, compared to full NPs such as definites and names. Indefinites are assumed to be less accessible than definites under the hierarchy proposed by Givon (1983), which shared the spirit of the AH (Ariel,1988, 1990). I, thus, assume that the 3rd person pronoun causes more processing load than the indefinite considering the degree of accessibility. This AH is compatible with the overall experimental results.

The results, however, did not support the binary metric in any obvious way. Given that both pronoun and indefinite intervenors are felicitous and refer to an old referent in all context types, this model predicts similar processing costs between

pronoun and indefinite conditions. This prediction cannot account for the difference between pronoun and indefinite conditions such that pronoun intervenors contributed to easier processing than indefinite conditions, regardless of the accommodability costs.

To summarize, the results of Experiment 3 showed that the pronoun intervenors led to easier processing than indefinite intervenors across contextual types. These results support memory-based models: *similarity-based interference effects*, the continuous metrics, the *GH* and the *AH* under the DLT, excluding the binary metric under DLT.

Accommodation of presupposition

Most theories discussed in this thesis make a general prediction that indexical pronouns elicit less processing cost than other complex/full NPs. Other theories involving contextual effects, such as Centering Theory (Brennan, 1995; B. Grosz et al., 1995; J. Grosz et al., 1983; Walker et al., 1998), Discourse-prominence theory (P. Gordon et al., 1993; P. Gordon & Hendrick, 1998), and Information Load Hypothesis (Almor, 1999, 2000, 2004; Almor & Nair, 2007), support the easy processing of indexical pronouns compared to other full NPs.

The advantage of indexical pronouns in processing has been robustly attested in previous literature (P. Gordon & Scearce, 1995; Greene et al., 1992), including studies that performed production tasks (Nice & Dietrich, 2003) and children's comprehension tasks (Haendler et al., 2015b, 2015a). Along with the previous observations, Experiment 1 also found significantly easier processing for indexical pronouns than full NPs, definite and indefinite conditions. An indexical pronoun

generates a default representation of its referent, even without reference to contexts, because it refers to a default participant in the discourse, either a speaker/writer or a lister/reader. The first- and second pronouns are inherently associated with default participants in any discourse, even without contexts. Alternatively, the easy processing of an indexical pronoun, however, can be attributed to a pragmatical factor because no/little processing effort is necessary to pragmatically accommodate a referent due to the default presence of a referent, compared to other complex NP types.

The investigation of the 3rd person pronoun, which requires presupposition accommodation but belongs to pronoun NP types, allows us to tease apart this confounding factor. Experiment 3 investigated how the processing cost of a 3rd person pronoun intervenor is modulated by presupposition accommodation in contexts: whether the 3rd person pronoun intervenor in the discourse is easy to process as indexical pronouns if an intended referent is provided or incurs processing cost of accommodation as other full NPs despite the existence of a referent. The results showed a faster reading time at the critical regions when the intervenor is a pronoun than when it is indefinite, regardless of the level accommodation cost. The results provided support for the advantages of pronouns in processing even when the pronoun requires more cost of accommodating its referent than the indefinite.

The faster online processing observed for all types of pronouns, as compared with full NPs in the presence of a referent suggests that the processing advantage is the overall properties of pronouns, not just indexical pronouns. The accommodation costs of pronouns in the intervening position between two dependent elements were not powerful enough to modulate the difficulty of processing cost, unlike full NPs. Pronoun conditions were easier to process than indefinite conditions as long as there is

an existing referent in discourse. I assume that pronouns are prone to be less susceptible to accommodation costs because they are most accessible on the discourse hierarchy. The strong advantage of pronouns over full NPs in processing suggests that pronouns involve a default interpretation of tracing back the previous context to find a potential referent (Clark & Sengul, 1979) and undergo an automatic process of presupposition accommodation to the current discourse topic, like a pronoun-as-cue strategy (Gerrig, 1986; Greene et al., 1992).

In contrast to the definite intervenors in Experiment 2, pronoun intervenors led to easier processing than indefinite intervenors across contexts, and the cost of accommodation did not play a role in the difficulty of dependency formations. This finding contrasts with the findings from previous studies on production, which found that speakers tend to use full NPs more than pronouns when two targets were given, compared to when the unique target was given (Arnold & Griffin, 2007; Fukumura et al., 2011; Koolen et al., 2011), and discourse effect on pronoun resolutions (P. Gordon & Scearce, 1995; Greene et al., 1992). The little cost of accommodation for pronouns in Experiment 3 can be attributed to different tasks and syntactic positions of the referring forms. The tasks in previous studies are production tasks (e.g., creating a sentence) and the target referring forms are the subjects in the main clauses, which do not include filler-gap dependencies. Experiment 3 suggests that the cost of accommodation does not modulate interferences during dependency formations when an intervening NP is a pronoun, which is most accessible and the highest on the discourse hierarchy.

Pronouns, which are highly accessible, are more likely to be less influenced by pragmatic factors such as accommodations in a given context than full NPs, which are

less accessible and low on the hierarchy. The notion of accessibility can be associated with the Information Load Hypothesis (ILH) in that both models assume the easier processing of pronouns than full NPs when the referent is in the current focus of attention and activated in working memory. According to the ILH, referring expressions that carry more information than necessary in the discourse pose a processing burden (Almor, 1999, 2000, 2004; Almor & Nair, 2007). For instance, the low degree of information of pronouns is enough when the referent is in focus, whereas the extra information of full NPs such as definite descriptions is necessary and useful when the referent is not focused and must be reactivated.

It can be noted that this hierarchical scale, which captures the effects of different types of intervenors on the processing difficulty of filler gap dependencies, exhibits a similar hierarchical scale posited to capture syntactic typology in terms of morphological marking. Research in syntactic typology has long observed that different kinds of NPs can exhibit different kinds of morphological marking, even when the grammatical function is held constant. This phenomenon is known as "Differential Argument Marking", and it is most commonly discussed insofar as it pertains to grammatical objects, so-called *Differential Object Marking (DOM)* (Aissen, 2003; Bossong, 1991; Comrie, 1979, 1981; Croft, 1988, 1990; Dalrymple & Nikolaeva, 2011; Danon, 2006; de Hoop, 1996; de Swart, 2007; Enc, 1991; Kalin, 2018; Lazard, 1998; Moravcsik, 1978; Richards, 2010; Rodriguez-Mondonedo, 2007; Silverstein, 1976; Torrego, 1998). The relevant scales for DOM are mainly discussed as animacy and definiteness.

(44) Animacy/person scale:

First or second person > third-person pronoun > name > human > animate > inanimate

(45) Specificity/definiteness scale:

Pronoun > name > definite > specific > nonspecific

In terms of the discourse hierarchy, referents on the highly accessible end of the scale (i.e., in-focus referring expressions such as pronouns or names) invoke *lower* processing demand as intervenors, whereas referents on the less accessible end of the scale (i.e., referential-only referring expressions such as definite and indefinite NPs) invoke *higher* processing demand as intervenors. Likewise, prominence scales such as topicality, animacy/person, and specificity/definiteness scales have been posited to account for the distribution of overt marking of objects. The referents on the unmarked end of the scale (for an object) are less likely to be overtly marked typologically, whereas those on the marked end of the scale (for an object) are more likely to be overly marked typologically[21]. This DOM indicates referents that are marked objects are likely to be morphologically more complex than those which are unmarked objects. Kalin (2018) proposes that referents on the marked end of the scale have a greater need to be licensed in the syntax, having more structures, than those on the unmarked end of the scale, based on (Pesetsky & Torrego, 2007)'s feature-sharing model. This morphological and structural system of DOM is in line with semantic

---

[21] That is, what is unmarked for subject is marked for objects, and vice versa—so called *markedness reversal* (Aissen, 2003; Battistella, 1990; Croft, 1990).

features of referents in the discourse hierarchy in that referents which are low on the discourse hierarchy need to build more conceptual content and contain more lexical information than those that are highly accessible on the hierarchy.

The pronoun advantage of processing over other full NP types in the current Experiment can be also connected to the cut-off point on the scale for DOM. For instance, some languages such as Romanian (Farkas, 1978) and Catalan (Comrie, 1979) distinguish pronouns from non-pronouns such that pronoun objects are obligatorily case marked. The exact cut-off point on the discourse hierarchy in terms of processing difficulty remains to be further investigated, but the absence of susceptibility to accommodation costs and the strong advantage of a pronoun, compared to full NPs, implies a categorical cut-off point may exist (between pronouns and full NPs), in line with the DOM system.

3.4   Interim Summary and Future Direction

The first aim of the thesis is to explore how semantic information and local discourse contexts mediate the resolution of filler-gap dependencies. Specifically, Chapter 2 and 3 aim to comprehend how the difficulty of dependency formation depends on the types of NPs in real-time processing (T. Bever, 1970; T. G. Bever, 1974; Gibson, 1998; P. C. Gordon et al., 2001, 2004; Kac, 1981; Kluender, 1998).

In Experiment 1-3, I investigated the processing of English cleft sentences across different NP types of the clefted element and intervenor, using reading time measures. In Experiment 1, I focused on investigating the effects of NP types of an intervenor, namely pronoun, definite and indefinite NPs, during the processing of clefted sentences in null contexts. The results revealed a significant difference

between pronoun intervenors and the other full NPs, definite and indefinite

intervenors. Experiment 2 focused on definite-indefinite contrasts in given contexts,

suggesting the difficulty of dependency formations arises from accommodation costs.

The results showed that the lexical property of an intervenor to accommodate a

particular referent played a role in dependency formation. In the next experiment, I

mainly investigated a 3rd person pronoun intervenor, which is highly accessible and

central in discourse but must accommodate the presupposition of a referent in the

discourse. The result of the experiment shows that a 3rd person pronoun intervenor is

not highly influenced by accommodation costs from a prior context, unlike a definite

intervenor. The overall results of experiments showed how the accessibility status of

NPs interacts with contextual support of accommodations.

Memory-based accounts such as similarity-based interference effects, binary

and continuous metrics, and expectation-based accounts have been discussed to

account for the experimental results. None of the processing theories fully captured the

overall findings of experiments, however. The compatibility of theories with each

experimental result is summarized in Table 13.

Table 13:   The summary of the compatibility of theories with overall findings

	Experiment 1 Pronoun – Definite – Indefinite	Experiment 2 Definite vs Indefinite	Experiment 3 3rd person Pronoun vs Indefinite
Similarity-based interference	$\sqrt{}/\Delta^{22}$	$\sqrt{}$	$\sqrt{}$
Binary metric	$\sqrt{}$	$\sqrt{}$	$\times$
Givenness Hierarchy	$\times$	$\times$	$\sqrt{}$
Accessibility Hierarchy	$\times$	$\times$	$\sqrt{}$
Expectation-based account	$\sqrt{}$	$\sqrt{}$	--

In brief, Experiments 1 and 2 did not support the effect of gradient status of NP types, such as the GH or AH. Experiment 3 showed that the binary metric between old and new discourse referent did not account for the effect of intervenor NP types on interference during dependency formations, either. A broader category of similarity in terms of common nouns (Gordon et al., 2004) appears to be most compatible with overall results (excluding presupposition failure conditions in Experiment 2), but Experiment 2 showed other sources of processing difficulty, which are pragmatic factors such as accommodation costs of an intervenor, also play a role in the difficulty of dependency formation. The similarity-based account attributes the processing

[22] A narrow perspective of the similarity-based interference effect of NP types (Gordon et al, 2001) is not compatible with the result because there was no effect between indefinite and definite conditions. However, a broader NP category under the similarity-based interference effect by Gordon et al (2004) provide support for the result in that definites and indefinites belong to the same category of common nouns.

difficulty of intervenors that contain common nouns to the similar NP types to the filler, which is also a common noun. It, however, remains to be unclear because the other NP type of the filler (e.g., *pronoun, proper name*), which is not a common noun, has not been explored. Given the strong advantage of pronoun in processing, the investigation of sentences where both the filler and an intervenor are pronouns would tease apart whether it is truly due to the similarity or the difficulty of intervenors containing common nouns. I leave this issue open for future research.

The distinction between common nouns and other NP types (Gordon et al, 2004) in fact exhibits a similar notion of accessibility to the GH. Gordon et al (2004) highlight that common nouns semantically denote sets of entities sharing the same property (e.g., *definites, indefinites, generics*), and thus refer indirectly, as opposed to pronouns and proper names, which refer more directly. Common nouns, which correspond to forms that are low on the GH, allow more possible interpretation than the other NP types, which correspond to forms that are high on the GH because common nouns refer indirectly and allow possible interpretations that satisfy their own properties. This distinct property of forms is consistent with the GH, which captures more restricted interpretations of forms that are high on the hierarchy (e.g., *pronoun, determiner*) than those that are low on the hierarchy. The restrictive interpretation of forms that are high on the GH leads to little processing cost without a working load to consider other possible interpretations.

The overall findings of experiments suggest the binary metric, or categorical processing difficulty between high and low forms on the hierarchical ranking, namely, *pronoun* vs *full NPs*. The hierarchical ranking of an intervenor plays a role in interference effect on the resolution of dependencies. The experiments showed

consistent results such that highly accessible NP types, such as pronouns, induce faster dependency formation than less accessible NPs, such as definites and indefinites across context types. This hierarchical ranking also affects how susceptible the interference effect is to pragmatic factors such as accommodability in contexts: the highly accessible NPs, *pronouns*, are likely to be less susceptible to accommodation costs, whereas less accessible NPs, *full NPs*, are more likely to be influenced by accommodation costs in contexts. This different pattern implies the processing advantage of pronouns despite accommodation costs in contexts. To better understand the precise cut-off point on the hierarchy, other NP types, such as *proper names*, which are assumed to be between the pronoun and full NPs on the hierarchy, need to be further investigated.

As an extension of the current study, other comparisons of two NP types of an intervenor can be observed. First, 3rd person pronouns and definites can be compared to explore how accommodation costs modulate dependency formations. Both NP types must accommodate to refer a particular referent. If pronoun conditions are easier to process than definites regardless of accommodation costs in contexts, it confirms that highly accessible NPs are less susceptible to accommodation costs, compared to less accessible NPs. Secondly, the comparison of indexical pronouns and 3rd person pronouns in the same context type allows us to verify that all pronoun types have the same property of ease in processing if no interaction between them is revealed.

Experimental results also presented the effect of intervenor NP types at another region before the resolution of dependency. Expectation-based models can account for the effect of different intervenor NP types at the different time course of processing— the intervenor region, not the verb region which involves the integration of a long-

distance dependent. In Experiment 1, the definiteness of the filler was manipulated, but no significant difference was observed in this "NP1+who" region (β=-10.82, SE=13.88, t=-.78, p=.44), which appears right before the intervenor region. This result implies that parsers make use of their probabilistic knowledge on the upcoming intervenor NP types, which appear in the embedded subject position, although they do not predict the NP types in the clefted filler position[23].

Given the findings at both an intervenor and the verb regions, this observation can be accounted for by both experience-based expectations and retrieval processes in working memory. This is, however, still unclear how much it can be interpreted as the retrieval processing of filler-gap dependencies. This is because the surprisal of an unpredicted NP type at the intervenor region could have been continuous in incremental sentence processing, which could have impacted the later processing of the verb. In Experiment 1, for instance, the verb follows right after the "intervenor+ Adverb" region and in Experiment 2, an effect was found in the adverb region, which follows the intervenor region, for presupposition felicitous conditions. Further research is necessary to disentangle the two factors by examining, for sentence structures that do not involve dependency formations as in (46a) whether the same pattern of reading times is observed at the verb region with that of the verb involving dependency formation as in (46b).

(46) a. It was true that {we/the director/a director} graciously thanked…

---

[23] In corpus studies, the embedded subject position showed the same pattern with main subject position. Therefore, it can't be attributed to the special encoding of the subject in the relative clause.

b. It was Tom that {we/the director/a director} graciously thanked…

I will leave this question for future work.

In terms of the response time on comprehension questions, a different pattern was observed from the online reading time. Online working memory load for dependency formation is what Caplan, Alpert, & G. (1998) call *interpretive processing* and the response time of the comprehension question in the present study can be considered as *post-interpretive processing*. In interpretive processing, parsers extract meanings from linguistic information, such as semantic/syntactic features at the sentence. In this study, for example, pronoun intervenors led to a robust tendency to easier processing than intervenors containing full NPs. In post-interpretive processing, parsers plan actions and reasoning, recalling the stored information about the sentence. Unlike interpretive processing, sentences containing a pronoun intervenor generated a longer response time than those with full NPs when parsers retrieved the full interpretation of the sentence to perform truth-value judgment tasks. It can be attributed to the type of task, which is involved with the discourse-level of comprehension as well as the proposition retrieval. When parsers were asked to judge the truth-value of the target sentence, the full NPs which contain more conceptual information could have facilitated their decisions by providing the explicit information. This result is in line with previous studies on syntax processing (Caplan & Waters, 1999; Carpenter & Just, 2013; Waters et al., 1995) and abstract/concrete words (Holmes & Langford, 1976; Schwanenflugel et al., 1988; Schwanenflugel & Shoben, 1983) for recall or decision tasks: The complex syntactic structure or abstract

sentence/words required additional processing load when recall the sentence/word or perform a decision task.

I admit that overall experiments have potentially low power due to a small number of participants. Higher powered replications would be necessary as future work to confirm an effect observed in Experiments in this thesis.

In the following two chapters, I turn to explore how syntactic features of an intervenor modulate interferences in the processing of subject-verb dependencies.

Chapter 4

THE SYNTACTIC FEATURES OF AN INTERVENOR

4.1 Attraction Effects

A large body of research has revealed facilitatory interference effects, usually called "attraction" effects, which are considered "facilitatory" because they give rise to a momentary illusion that an ungrammatical sentence is, in fact, syntactically well-formed (Ness & Meltzer-Asscher, 2017, 2019; J. A. van Dyke, 2007; J. A. van Dyke & Lewis, 2003; J. van Dyke & McElree, 2006, 2011); this facilitatory processing in ungrammatical sentences is also called an *illusion of grammaticality*. Such illusion effects occur when a target item mismatches its dependent in number features, but a feature-matching attractor is present between the dependency, as in (8), repeated below, a typical example of attraction effects (J. L. Nicol et al., 1997; N. J. Pearlmutter, 2000)[24]. In (47), the number feature of the main subject (i.e., *the key*) does not match the main verb (i.e., *are*). Instead, the attractor (i.e., *the cabinets*) matches the plural number feature of the main verb.

(47) *The key to the cabinets **are** rusty.

---

[24] The facilitatory interference effects in the history of literature on subject-verb agreement dependency formation has been called "attraction effects" and thereby an intervening NP between the dependency has been called as an attractor. In order to keep in line with the literature, I refer to an intervenor as *an attractor* in Chapters 4 and 5.

This illusion of grammaticality has been widely observed across different syntactic structures and different types of agreements: subject-verb agreement (Dillon et al., 2013; Franck et al., 2002, 2006, 2010, 2015; Hammerly et al., 2019; J. van Dyke & McElree, 2006; Wagers et al., 2009), gender marking (Badecker & Straub, 2002; Cunnings & Felser, 2013; J. Nicol & Swinney, 1989; Sturt, 2003; Xiang et al., 2009), negative polarity item (NPI) licensing (Xiang et al., 2009, 2013), and honorific marking (Kwon & Sturt, 2016a). A large body of studies on subject-verb dependency formations has shown that structurally illicit attractors (such as "the cabinets" in (48), which is not a syntactically licit target for agreement) give rise to facilitatory attraction effects (Dillon et al., 2013; Franck et al., 2002; Parker, 2017; Parker & An, 2018; N. Pearlmutter et al., 1999; Sturt & Kwon, 2017; Tucker et al., 2015; J. van Dyke & McElree, 2011; Wagers et al., 2009)

Recent studies have focused largely on investigating the function of syntactic cues as retrieval cues in the processing of subject-verb agreement. In terms of argument or thematic hierarchy, subjects (typically agents or experiencers as thematic roles) are known to be most prominent and highest on the hierarchy, compared to other types of NPs (Comrie, 1981; Levin & Hovav, 2007). This prominence of subjects on the hierarchy can influence the encoding of representations such that noun phrases in subject positions are distinctively encoded compared to noun phrases in other syntactic positions. The extent to which how much subjects are more distinctively encoded than other NPs is still unclear in empirical studies. I begin with the question in Chapters 4 and 5 whether the distinctive encoding of a subject is invariant or is relative to the argument status of other NPs across different syntactic configurations. For instance, a subject of a transitive verb appears with an object,

another core argument of the verb, as *the athletes* in (48a). A subject of an intransitive verb, on the other hand, may appear with an adjunct, non-core argument, as *the athletes* after the preposition *with* in (48b).

> (48) a. The coach trained *the athletes*.
>
>   b. The coach trained with *the athletes*.

The comparison between (48a) and (48b) enables us to explore how the different statuses of other NPs affect the encoding of a subject. For this purpose, I used relative clauses where the second NP, like *the athletes*, is displaced and leaves a gap as in (49).

> (49) a. The athletes who the coach trained…was/were…
>
>   b. The athletes who the coach trained with…was/were…

In this comparison, the subject *the coach* remains constant in a subject position but the initial argument status of another NP differs, either core argument or non-core argument. The attraction effect in the processing of subject-verb agreement dependency was used as a probe to examine how the encoding of the subject would be impacted by the different status of another NP in sentence structures.

Research on the syntactic profiles of the relevant NPs during subject-verb agreement formations have observed that the effects of an attractor can be modulated by the syntactic prominence of its argument status; that is, core arguments, standardly taken to be *subject* or *object*, induce attraction effects, whereas other NPs (e.g.,

oblique adjuncts) do not (Cunnings & Sturt, 2018; Engelmann et al., 2015; Parker, 2017; Parker & An, 2018; J. van Dyke & McElree, 2011), as showed using a variety of experimental methodologies. However, a commonality of these studies is that they focus exclusively on the status of the *attractor*, whereas the role of the status of the filler (e.g., the wh pronoun which is coreferential with the target itself) at its gap site (e.g., object, oblique adjunct), remains uninvestigated. This thesis therefore asks how the argument status of a filler mediates the online processing of subject-verb agreement, focusing on the contrast between direct object fillers as in (51) and oblique (i.e., prepositional) adjunct fillers, at gap sites, as in (50b).

(50) a. Direct object filler

The athlete [who[OBJECT] the coach(es)[SUBJECT] trained __] definitely was/were prepared for the championship game.

b. Oblique adjunct filler

The athlete [who[OBLIQUE] the coach(es)[SUBJECT] trained with __] definitely was/were prepared for the championship game.

This study contrasts with earlier studies in that the status of the *attractor* itself is held constant (i.e., it is always a subject).

4.2 Background: Previous Studies on the Argument Status of an Attractor

4.2.1 Subject-verb Thematic Binding: Manipulating Animacy

In research looking at retrieval in subject-verb dependencies, (J. van Dyke & McElree, 2011) investigated how the syntactic status of the attractor affects retrieval, and observed that prominence in terms of grammatical role plays a role in the interference effect[25]. As shown in the following examples, the attractor's syntactic position was either a subject, as in (51a), or a direct object, as in (51b). The animacy of the attractor was also manipulated, such that animate attractors (e.g., *the witness*) matched the required animacy feature of subjects thematically associated with the verb (e.g., *compromised*) and were therefore plausible as targets. Inanimate attractors, on the other hand (e.g., *the motion*), were implausible as targets because they cannot be interpreted as thematic agents of the clause-final verb.

(51) a. [Subject attractor]

The attorney who the judge realized had declared that ***the motion/the witness*** was inappropriate <u>compromised</u>.

b. [Object attractor]

The attorney who the judge realized had rejected ***the witness/ the motion*** in the case <u>compromised</u>.

(Van Dyke & McElree, 2011)

---

[25] Interference effects in broad perspectives can be either facilitatory or inhibitory interference effects. Van Dyke & McElree (2011) presents inhibitory interference effects such that a processing cost increases when an attractor matches the target in features. I focus on the susceptibility of argument status of attractors on interference effects, regardless of interference patterns. I discuss this issue on discussion.

The authors used eye-tracking to examine interference effects of subject and object attractors as in (51) and found that animate (but not inanimate) subject attractors gave rise to an interference effect at the verb *compromised*, a finding attributed to animate NPs being considered as its potential thematic agent. On the other hand, object attractors did not trigger interference effects (regardless of animacy). The authors evaluated this result against a previous finding by Van Dyke (2007), which revealed an interference effect of oblique PP attractors, as in (52): Van Dyke found a faster reading time at the critical main verb (i.e., *moaned* in 52) when the oblique PP attractor was inanimate due to its distinctive animate feature from the target item (i.e., *the lady* in 52).

(52) [PP attractor]

The pilot remembered that the lady who was sitting near ***the smelly man/seat*** <u>moaned</u> about a friend.

(Van Dyke, 2007)[26]

---

[26] Van Dyke (2007) did not manipulate the plausibility of a verb as Van Dyke & McElree (2011) did by substituting the main verb. Instead, the corresponding subject attractor was compared:

*e.g. The pilot remembered that the lady who said that **the man/seat** was smelly moaned about a friend.*

I discuss only a PP attractor in the current paper because the result of the subject attractor is akin to Van Dyke & McElree (2011)'s finding. Van Dyke & McElree (2011) compared subject and object attractors, respectively with constructions where attractors precede the filler, but I focus on constructions where attractors intervene between the dependency formation.

Van Dyke & McElree (2011) attributed this difference between different NPs of differing grammatical functions to relative syntactic prominence of these NPs within an argument hierarchy of grammatical function (e.g., E. L. Keenan & Comrie, 1977), in which subjects (e.g., *the motion/witness* in 51a) and objects (e.g., *the motion/witness* in 51b) - collectively referred to as 'core arguments') - are more distinctive than oblique PPs (e.g., *the smelly man/seat* in 52), because, as arguments of a verb, they are more prominent when computing the propositional content of a sentence. Modifying oblique adjuncts, including PPs like "near the smelly man/seat", are less prominent than core arguments because the theta roles of the nominals in these constructions (e.g., *the smelly man/seat*) are not directly assigned by the main verb but by the preposition itself (e.g., *near*), and thus, they argued to be less important in building a mental representation (Bresnan, 2001; Frazier & Clifton, 1996). In terms of this differing prominence of grammatical functions, Van Dyke & McElree (2011) proposed that syntactic encoding of an attractor is susceptible to the distinctiveness of grammatical function of that attractor in the memory, which leads to different aspects of interference effects according to how saliently the attractor is encoded.

Concerning the contrast between object attractors (which do not trigger an interference effect) and oblique PP attractors (which do), the authors reasoned that core arguments provide more salient cues than oblique adjuncts PPs, such that core arguments are prominent enough to prevent interference effects because they produce a more salient mismatch with the syntactic retrieval cues. Oblique arguments, on the other hand, are less salient in the memory encoding system and are thus more likely to trigger an interference effect because they produce a less salient mismatch with the

syntactic retrieval cues. Under this assumption, the absence of the interference effect of the direct object condition was accounted for by its prominent argument status shown in (51), whereas the interference effect of the PP condition was accounted for by its less prominent argument status as in (52).

In order to account for the presence of an interference effect with subject attractors, Van Dyke & McElree further proposed that not *all* core arguments are prominent enough to be resistant to attraction effects. The grammatical function of subject attractors, unlike object attractors, matches with the *main* subject, with which the agreement dependency with the verb must be formed. This matching syntactic cue of subject attractors strengthens their prominence even further, compared to object attractors due to their same syntactic position with the target item (in the main subject position). Therefore, the subject attractors were more prominent than object ones and were, in fact, *too* prominent to be precluded from retrieval, resulting in interference effects, which the authors called 'syntactic gating'. Van Dyke & McElree's proposal is therefore in line with theoretical claims that subjects are more prominent than direct objects in discourse and grammatical hierarchy (Chafe, 1976; Crawley et al., 1990; Keenan & Comrie, 1977) as well as empirical studies on the resolution of reflexive anaphora (Engelmann et al., 2015; Patil et al., 2016) and the resolution of ambiguous pronouns (Gernsbacher & Hargreaves, 1988; P. C. Gordon et al., 1993).

4.2.2 Subject-verb Agreement Dependency: Manipulating Number

Building on Van Dyke & McElree's (2011) manipulation of the semantic feature of an attractor as a diagnostic of interference effects, Parker and An (2018) examined morphological subject-verb agreement, as in (53). Their study manipulated (i) the number feature of the attractor (singular versus plural) and (ii) the number

feature of the verb (also singular versus plural). The authors conducted two reading time experiments, using direct object attractors as a baseline (e.g., 50b). In the first experiment, PP attractors were compared to direct object attractors (i.e., 50c vs. 50b), whereas subject attractors were compared to this same baseline in the second experiment (i.e., 53a vs 53b).

(53) a. [Subject attractor]

The celebrity who *the journalist(s)* insulted certainly was/*were upset about the claims.

b. [Object attractor]

The waitress who sat *the girl(s)* unsurprisingly was/*were unhappy about all the noise.

c. [PP attractor]

The waitress who sat near *the girl(s)* unsurprisingly was/*were unhappy about all the noise.

Parker & An (2018) observed a main effect of attractor number when attractors were in PP position at the spillover region (after the main verb), such that the reading time for plural attractors was shorter than for singular attractors, in ungrammatical sentences. No main effect was observed when attractors were in either subject or object positions. The finding regarding PP and object attractors is consistent with Van Dyke & McElree (2011); that is, object attractors do not cause attraction whereas PP

attractors do. However, Parker & An (2018) observed no attraction effect of subject attractors, unlike Van Dyke & McElree (2011).

In brief, they found the attraction effect of an oblique PP attractor, but not attractors in *any* core argument position. Their results also differed from Van Dyke & McElree (2011) in terms of the time course of attraction effects: all main effects appeared at spillover regions, not at the critical main verb region (i.e., *was, were*). A plural attractor in PP position generated significantly faster reading times at the spillover regions in ungrammatical sentences, reflecting the facilitatory agreement attraction effect observed in previous studies (Dillon et al., 2013; Franck et al., 2015; Lago et al., 2015; Wagers et al., 2009b); a longer reading time of ungrammatical sentences, compared with grammatical sentences, was also revealed at spillover regions in all conditions.

4.2.3 Summary

The findings of both Van Dyke (2007) and Parker & An (2018) are consistent with the generalization that PP attractors give rise to interference effects. Their results, however, do not align as concerns subject attractors. Van Dyke (2007) and Van Dyke & McElree (2011) found interference effects of subject attractors, whereas Parker & An (2018) did not (just as was the case for object attractors). Van Dyke (2011) attributed the interference effect of subject attractors to semantic and syntactic properties of the attractors, both of which match those properties of the target item. The retrieval cue of subject attractors matches the target item in that both are in the subject position and plausible with the main verb. The absence of interference effect of subject attractors in Parker & An (2018) was accounted for by salience, just as object attractor, such that both subject and object attractors, as core arguments, are

salient enough to be ruled out and are resistant to attraction effects. Subject attractors, like object attractors, are prominent enough to be distinct from the target subject even though their syntactic cues match with it. Therefore, no attraction effect of subject attractors was revealed as was with object attractors.

Table 14 provides a summary of the findings of these studies. They found consistent findings that PP attractors caused attraction effects, whereas direct object attractors did not. Subject attractors, however, revealed inconsistent results across studies.

Table 14:   Summary of results from Van Dyke (2007), Van Dyke & McElree (2011), and Parker & An (2018)

Interference effects	Van Dyke (2007), Van Dyke & McElree (2011)	Parker & An (2018)
Subject attractor	✓	✗
Direct Object attractor	✗	✗
PP (oblique) attractor	✓	✓

This chapter has introduced the aspects of attraction effects on subject-verb number agreement dependencies and presented the main background literature on them. In the following chapter, I present two experiments, looking at how the processing of subject-verb dependencies is influenced by the argument status of *both* the filler and an intervenor.

Chapter 5

THE PROCESSING OF SUBJECT-VERB AGREEMENT DEPENDENCIES

5.1 Motivation of Experiments 4 and 5

Parker & An (2018) argue that different retrieval cues and dependency-types as possible factors that gave rise to their differing results from those of Van Dyke & McElree (2011). Van Dyke & McElree (2011) tested the role of semantic cues, like animacy, which is associated with a verb during filler-gap dependency formation. Parker & An (2018), on the other hand, examined subject-verb agreements, which are more associated with morpho-syntactic number cues (i.e., singular/plural); different types of cues might be associated with different weights or levels on processing (Kempe & MacWhinney, 1999; Lewis et al., 2006; Lewis & Vasishth, 2005; MacWhinney et al., 1984; Stoops et al., 2014; J. A. van Dyke & Lewis, 2003).

I note that the constructions used in both Van Dyke & McElree (2011) and Parker & An (2018) have confounding structural factors, which might have impacted the saliency of retrieval cues. First, Van Dyke & McElree (2011) included an additional embedded clause between the subject-verb agreement dependency formation in the subject condition only, which could have led to different processing mechanisms and more processing load than simple relative clause constructions in the object condition: the elements within the same clause are known to be in competition for the same mechanism in processing, whereas those in a distinct clause are under a different mechanism (Bock & Cutting, 1992; Franck et al., 2002; Frazier & Clifton, 1989; Kluender & Kutas, 1993). In addition, syntactically more complex structures

125

and longer distances between dependency formations cause processing load (E. Gibson, 1998, 2000; Edward Gibson & Warren, 2004; Grodner, Watson, & Gibson, 2000). Second, Parker & An (2018)'s manipulations contrasted not only the syntactic status of the *attractor*, but also that of the *wh filler*: unlike in Van Dyke & McElree (2011)'s materials, in which the filler was consistently interpreted as the subject of an embedded verb, the filler in Parker & An's materials was interpreted as *an object* in the subject attractor condition, but as *a subject* in the object and oblique attractor conditions. Thus, their absence of any contrast between the subject and object attractor conditions could be interpreted either as a reflection of the differing statuses of the attractor (subject versus object), or the differing statuses of the filler (object versus subject) at its gap site, or a combination of these two factors.

The role of the argument status of a *wh filler* phrase has indeed received considerable attention in research on the processing of long-distance dependencies such as in wh-questions (Donkers et al., 2013; Goodall, 2015), relative clauses, as well as constraints on syntactic well-formedness, such as attempted dependency formation in so-called "island" configurations (that is, stretches of structure in which dependency formation is ungrammatical; Ross, 1967). In terms of the processing of subject-verb agreement dependencies, however, this topic has thus far been uninvestigated. The current study, therefore, examines whether the prominence of argument status of the *fillers* themselves plays a role in the resistance to attraction effects of attractors, in addition to the argument status of the attractors as observed by Van Dyke & McElree (2011) and Parker & An (2018). This examination will enable us to better comprehend how the argument status of the dependency and the attractor, as well as their interaction, modulates the retrieval mechanism. I administered two self-paced reading

tasks: Experiment 4 employed *ungrammatical* sentences, in which verb agreement was plural whereas the main subject was singular (and hence, there was a number mismatch). Experiment 5 used *grammatical* sentences with singular verb agreement, thus matching with the main subject, which was also singular.

5.2 Experiment 4: Ungrammatical plural agreement

The current experiments build upon Parker & An (2018)'s study on attraction effects in the comprehension of subject-verb agreements, aiming to examine whether and how the argument status of a *wh* filler can modulate the attraction effect and how this interacts with the status of an attractor. I examine sentences such as (54), in which the singular main subject ("*The athlete*") is mismatched in number with the plural main verb ("*were*"). Following Parker & An's (2018) design, the subject is modified by a relative clause which requires the formation of a filler-gap dependency between the relative pronoun ("*who*") and the gap (after "*trained*" in 54).

(54) *The athlete [who the coach trained __] definitely were prepared for the championship game.

Keeping the argument status of the *attractor* ("*the coach*" in 54) constant (i.e., it is always the *subject* of the relative clause), I manipulated the argument status of the filler at its gap site within the relative clause (it was either an object argument or an oblique PP adjunct of the relative clause).

Experimental items (adapted from Parker & An, 2018) consist of relative clauses containing an attractor, which is an (embedded) subject in the relative clause. A full set of experimental conditions is presented in Table 15. The experiment crossed

(i) the number feature of the subject attractor (singular vs plural; two levels) with (ii) the argument status of the filler at a gap site (object vs oblique, to which I added a third level, called "long oblique" in order to control for a confound in dependency length). This latter manipulation, therefore, had three levels. In the first level, the filler was the direct object, as a core argument, of the embedded verb at its gap site within a relative clause, which I call *object condition*. In the second level, the filler was an oblique adjunct of the embedded verb within a relative clause, being an object of a preposition. I call this condition the *oblique condition*. Comparing these two conditions is our critical comparison because it asks about the susceptibility of attraction effects to the argument status of the filler at its gap site.

However, as well as differing in terms of argument status of the filler/gap, these two conditions also differ in terms of the dependency length between the main subject and the verb with which the agreement dependency is formed: there is one additional word (i.e., the preposition "with" in Table 15) in the oblique conditions, which is not present in the object conditions. This means that any difference between the object and oblique conditions could be attributed either to their differences in grammatical functions or simply to agreement dependency *length*. In order to accommodate this issue, I included a third level, in which adverbial phrases (e.g., "on Sunday") were included after gaps in an oblique adjunct position. In this condition, which I call "long-oblique", the dependency length between subject and verb was extended even further by the addition of extra (adverbial) material; however, the argument status of the filler at the gap site is identical to the oblique condition. Therefore, if the oblique condition patterns with the long-oblique condition, then any difference between the object condition and the oblique condition can be safely

attributed to grammatical function, not to the length of the subject-verb agreement

dependency.

Table 15:   A sample set of items for Experiment 4

Object condition

SG attractor
The athlete [who the coach trained __] definitely **were** prepared for the
championship game.

PL attractor
The athlete [who the coach<u>es </u>trained __] definitely **were** prepared for the
championship game.

Oblique condition
SG attractor
The athlete [who the coach trained with __] definitely **were** prepared for the
championship game.

PL attractor
The athlete [who the coach<u>es</u> trained with __] definitely **were** prepared for the
championship game.

Long-oblique condition

SG attractor
The athlete [who the coach trained with __ on Sunday] definitely **were** prepared for
the championship game.

PL attractor
The athlete [who the coach<u>es</u> trained with __ on Sunday] definitely **were** prepared
for the championship game.

5.2.1   Materials

18 sets of 3 items in each of the 6 conditions were distributed across six lists in

a Latin Square design. These were coupled with 32 filler sentences with similar

lengths to the items. 24 of them were cleft sentences (used as stimuli for a different experiment) as in (55a) and 8 of them consist of relative clauses as target items as in (55b). Twelve of them were made ungrammatical via mismatched tense agreement, as in (55a), and the other twelve were maintained as grammatical (e.g., 55b).

(55) a. *It was the tailor who the customer gratefully thanks at the banquet.

b. The car that had been stolen by the criminal fortunately was found in an empty parking lot.

Comprehension questions appeared immediately after each item. For example, the question that followed the items shown in Table 15 was "What was the profession of the {(person/people)} who trained the athlete?". Participants selected a correct answer from a multiple-choice set of three options. Items were presented in a pseudo-randomized order, with no more than two critical items adjacent (and any two adjacent critical items were always of different conditions).

5.2.2 Participants

Participants were 48 native speakers of English who were recruited from Amazon's Mechanical Turk web service. All participants provided informed consent and were compensated $5.00. All participants were included in the analysis because their performance accuracy across all comprehension questions (including fillers) was higher than 80%, and all participants indicated that they were native speakers of English (payment was not contingent upon the response to this question, so there was no incentive to answer dishonestly).

130

5.2.3 Procedure

The task was a self-paced reading using a non-cumulative moving window display. The experiment used the online software platform Ibex Farm (Drummond, 2018). Each word or phrase was presented in the middle of the screen. The crucial verb region and spillover regions were always displayed independently, and the display pattern was constant across conditions of each item in the Latin Square design. Participants were asked to press the spacebar to see the next word or phrase. Once the last word or phrase of each sentence was presented, a comprehension question appeared separately as a full sentence. The multiple-choice answer set was given after the comprehension questions; no feedback was provided. Three practice trials were included at the beginning. Participants were informed that some of the sentences were not quite well-formed (Hammerly et al., 2019) but that all were understandable. The experiment lasted approximately 25 minutes.

5.2.4 Predictions

Under the previous accounts of interference effect in terms of the status of attractors themselves, all conditions are expected to present consistent results, either the presence or absence of interference effects, because all attractors remain in the same position (i.e., subject position). As a core argument, subject attractors are prominent enough to prevent interference effects, and thus, no effects of attractors across conditions would be observed. A second prediction, in terms of "syntactic gating" proposed by Van Dyke & McElree (2011), would be the presence of interference effect of subject attractors because all fillers across conditions are in the main subject position, which is identical to the position of the attractors in the surface

structure (recall that "syntactic gating" predicts that it is the matching syntactic position between the filler and an attractor that causes interference effects).

On the other hand, if the status of the filler at its gap site modulates the processing of the subject-verb agreement dependency, the object conditions will show a different result from the oblique and long-oblique conditions, due to the differing status of the filler at its gap site (the filler remains in the subject position across conditions). Applying previous processing models (e.g., Van Dyke & McElree, 2011; Parker & An, 2018) for the syntactic cues of attractors to the status of the filler at its gap site, I predict that interference effects should obtain in both of the oblique conditions (and critically, to the same extent, if agreement dependency length is not the decisive factor) and not in the object condition, because obliques are less saliently encoded, and are therefore more susceptible to interference.

5.2.5   Results

Two participants were excluded from analysis as their mean comprehension accuracy was below 80%, leaving a total of 46 participants for data analysis. Thus, 828 critical trials were generated. Trials that generated incorrect responses (n = 90) were excluded from further analysis. Outliers – identified as reading times 3SD above the mean for each condition in each region - were replaced with the mean for that condition for the relevant region. This process affected less than 2% of the data.

I analyzed three regions of interest: the critical verb (critical region), and the two words following the verb (spillover regions 1 and 2, respectively). For each region, I analyzed the reading times by fitting a mixed-effects linear regression with crossed random effects for participants and items (Baayen et al., 2008), using the *lme4*

package (Bates, Maechler, Bolker, & Walker, 2015). Following Parker & An (2018), I analyzed the raw, untransformed reading times since research on interference effects of these types suggests that data transformations (e.g., log-transformation) can obscure these effects (Lago et al., 2015; Staub, 2010b; Tucker & Almeida, 2017; Villata et al., 2018). For both experiments reported in this paper, I used the maximal random effect structure justified by the experiment design that would allow for model convergence (Barr et al., 2013). P values were calculated via a Satterthwaite approximation, as computed using the *LmerTest* function.

The argument status of the *wh* filler at the gap site (3 levels) was contrast-coded using centered Helmert contrasts. The first coefficient, GRAMMATICAL FUNCTION, contrasted object (coefficient: +2/3) with oblique and long oblique conditions, pooled (coefficient: -1/3 for oblique conditions; -1/3 for long oblique conditions). This contrast asks about the effect of the *wh* filler at the gap site: a core argument versus a non-core argument. The second coefficient, LENGTH, contrasted oblique conditions (coefficient: -1/2) with long oblique conditions (coefficient: +1/2): this contrast asks the length effect in the two oblique conditions (the object condition does not participate in this comparison; its coefficient is 0). Our second manipulation, the number feature of the attractor ("NUMBER", 2 levels), was coded as +1/2 when the attractor was singular and -1/2 when it was plural.

Figure 9 shows the mean word-by-word reading times for the object conditions, and figures 10 and 11 show the same for oblique conditions and long-oblique conditions, respectively. I begin by examining the critical region (i.e., the copula verb "were"). Here, the main effect of argument status in terms of the comparison between object and oblique *wh* filler dependencies was marginal

("GRAMMATICAL FUNCTION"; β=19.44, *SE*=10.83, *t*=1.79, *p*=.084), but the second contrast comparison between oblique and long oblique ("LENGTH") was significant (437 ms vs 401 ms; β=-35.52, *SE*=11.38, *t*=-3.12, *p*=.0043), indicating that verbs with the mismatched agreement are read faster when there are more intervening words between the mismatching verb and the target subject. There was no main effect of NUMBER (*t* = .88 *p* = .38); that is, the copula verb was not read significantly faster when the attractor was singular ("*the coach*") compared with when it was plural ("*the coaches*"), overall. The interaction of NUMBER and GRAMMATICAL FUNCTION, however, was significant (β=56.67, *SE*=25.25, *t*=2.25, *p*=.033), whereas the interaction of NUMBER and LENGTH was not (*t*=.03, *p*=.98).

To further examine the former interaction, I conducted nested planned comparisons asking, for each level of argument status (object, oblique, long oblique), whether there was a difference in reading times when the attractor was singular (and thus, *mismatched* in number with the copula verb) compared with when it was plural (and thus, *matched* in number with the verb). Importantly, only in the *object* condition- in which the copula verb was read *faster* when the attractor was plural than when it was singular - was this comparison significant (411 ms vs 462 ms; β=47.29, *SE*=21.25, *t*=2.23, *p*=.032); this corresponds to the boxed region in Figure 9). This same difference showed no signs of significance in the oblique condition (444 ms vs 430 ms; *t* = -.41; *p* = .69; see box in see Figure 10) or in the long oblique condition (404 ms vs 397 ms; *t* = -.61; *p* = .55; see box in Figure 11). This indicates that the attraction effect of a *subject* attractor which is matched with the verb (here: a plural attractor matched with plural agreement) obtains when the relevant filler-gap dependency path (along which the attractor is situated) pertains to an object gap (i.e.,

core argument) and not when it pertains to an oblique gap (i.e., non-core argument), regardless of length. This effect instantiates the classic "illusion of grammaticality" found in prior literature on agreement attraction (recall that the sentences in this experiment were in fact, ungrammatical).

Concerning the Spillover regions, I found no significant or marginal effects in Spillover 1 (all $ps > .28$); but at Spillover 2, I found an interaction of LENGTH and NUMBER ($\beta=44.58$, $SE=20.57$, $t=2.16$, $p=.036$). This indicates evidence that agreement dependency length does affect number agreement processing, but only after the main verb has been encountered. There were no other significant effects at Spillover 2 (all other $ps > .32$).
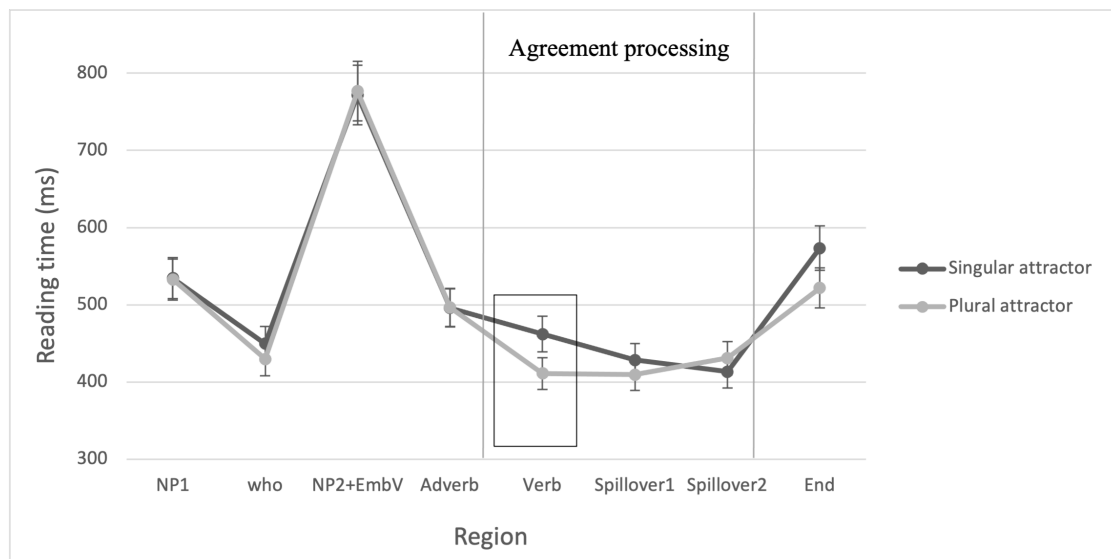
Figure 9: Region-by-region reading times for the object conditions (singular attractor as a dashed line; plural attractor as a solid line). The critical region is boxed. Error bars indicate 95% Confidence intervals.
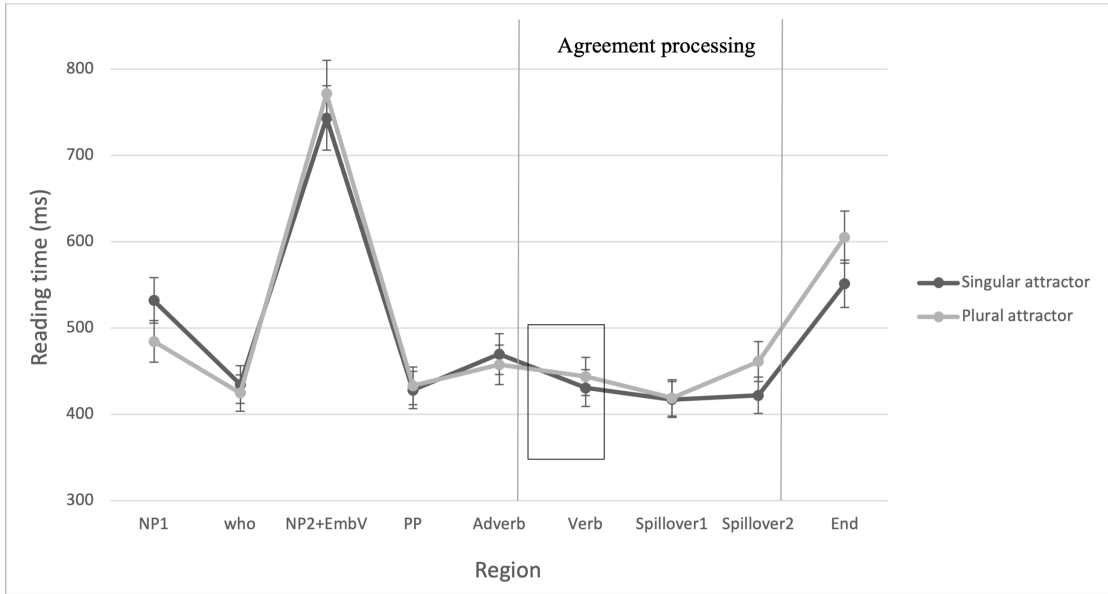
Figure 10:   Region-by-region reading times for the oblique conditions (singular
attractor as a dashed line; plural attractor as a solid line). The critical
region is boxed. Error bars indicate 95% Confidence intervals.
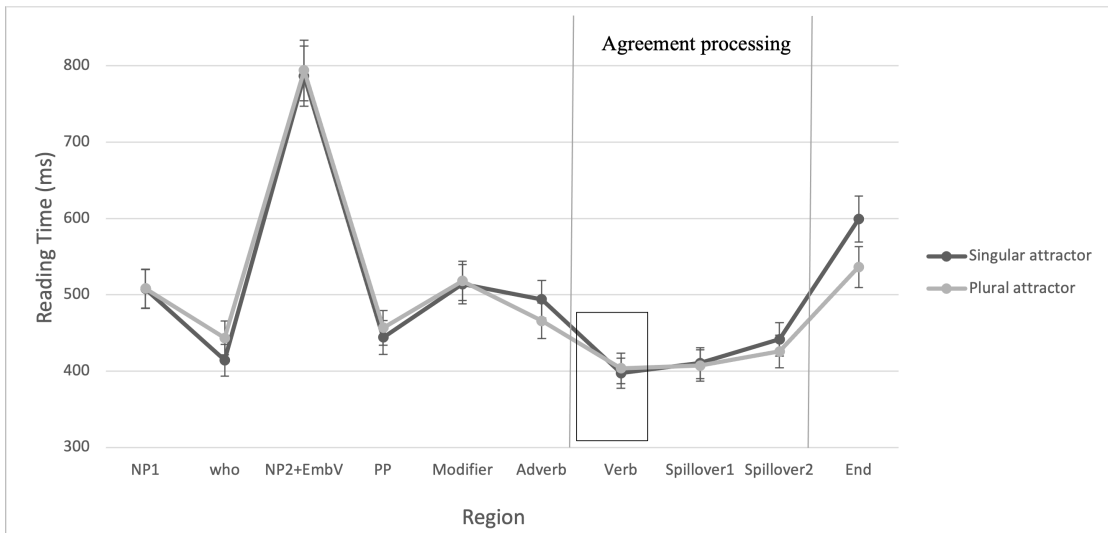
Figure 11:   Region-by-region reading times for the long oblique conditions (singular
attractor as a dashed line; plural attractor as a solid line). The critical
region is boxed. Error bars indicate 95% Confidence intervals.

For response times (see Figure 12), I observed a significant interaction of LENGTH and NUMBER (β=-326.61, SE=143.18, t=-2.28, p=.023), wherein response times in the long oblique conditions, which were longer for sentences in which the attractor was plural than when it was singular (1893 ms vs 1741 ms), as compared with the inverse trend in oblique conditions (1736 vs 1936); however, nested planned comparisons did not reveal significant or marginal effects of NUMBER for either level (both ps >.1). Beyond this, no other effects were significant (all other ps >.53).
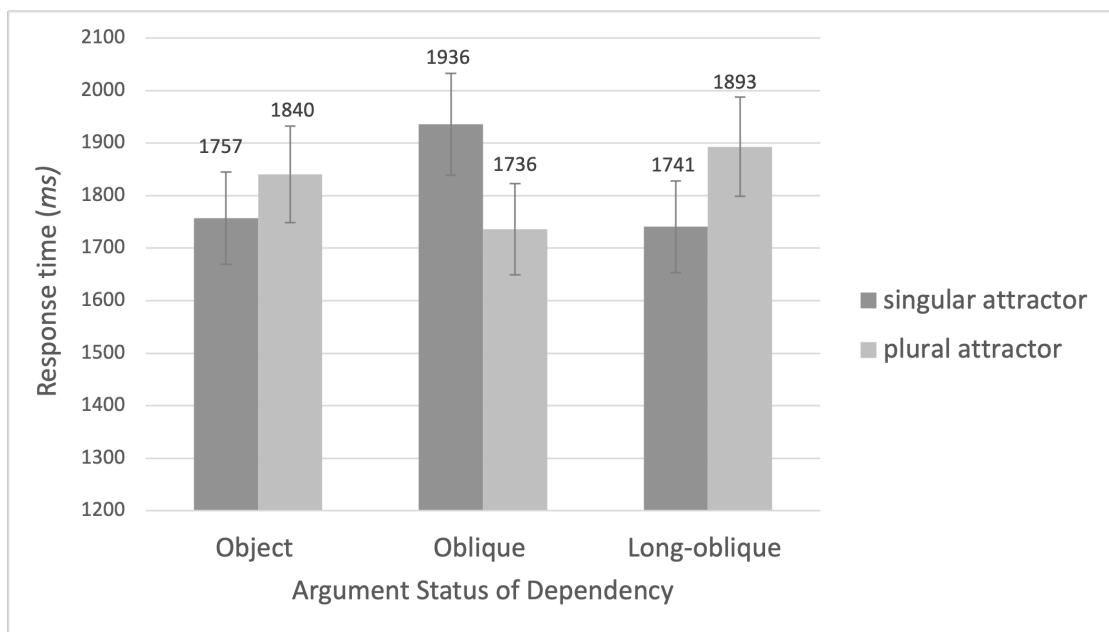
Figure 12:   Average response times to the comprehension questions (*ms*) by condition. Error bars indicate 95% Confidence intervals.

The mean accuracy rate for comprehension questions by conditions is shown in Table 16. A logistic mixed-effects regression model showed a main effect of

LENGTH, in which sentences in the long oblique conditions were answered less accurately than those in the oblique conditions (87% vs 92%; β=-.63, SE=.32, z=-1.98, p=.047), and a main effect of NUMBER, in which sentences with a plural attractor generated lower accuracy rates than those with a singular attractor (88% vs 91%; β=.5, SE=.25, z=1.96, p=.05). I also found a significant interaction of LENGTH and NUMBER (β=-1.42, SE=.64, z=-2.25, p=.025). Planned comparisons revealed that this interaction was driven by the oblique condition, in which sentences with a singular attractor were significantly more accurate than those with a plural attractor (95% vs 88%; β=1.22, SE=.5, t=2.47, p=.014). There were no other significant effects (all other ps >.76).

Table 16: The mean accuracy rate for comprehension questions (%) by condition

	Attractor number	
	Plural attractor	Singular attractor
Object condition	87	91
Oblique condition	88	95
Long-oblique condition	88	86

To summarize the offline effects, I find – contrary to our observations regarding online reading time comprehension – that dependency length affects post-reading interpretation. Long oblique sentences (i.e., those with oblique gaps which were followed by additional adverbial material) generated longer response times when

the attractor was plural (and thus, matched in number features with the verb) compared with when it was singular (and thus mismatched). Long oblique sentences also generated lower accuracy in terms of response to comprehension questions than (short) obliques, and in this latter condition, sentences with singular attractors generated higher accuracy than those with plural attractors.

Importantly, I also found that, overall, sentences with singular attractors generated higher accuracy than those with plural attractors. These effects indicate that plural attractors may cause greater processing difficulty than singular ones, regardless of the number feature of the main verb, at the post-reading interpretative stage. I reason that this is likely due to the additional morphological and complexity of plural attractors (Eberhard, 1997; Wagers et al., 2009).

5.2.6   Discussion

The results of Experiment 4 provide evidence that the attraction effect is susceptible to the argument status of a *wh* filler gap dependency that contains it. Specifically, an attraction effect occurs only when the filler is the direct object – that is, a core argument – of the embedded verb. In this object condition, the attraction effect of attractors is observed at the main verb region, which is read faster when the attractor is plural (potentially leading to a grammaticality illusion). Recall that the argument status of the attractor was held consistent across conditions: it was always a subject as the filler in the subject position. The previous theories of interference

effects such as syntactic gating (Van Dyke & McElree, 2011) alone, therefore, predicts that such attraction effects should obtain consistently.

In terms of the offline results, first, I note that the pattern of the accuracy rates for the comprehension questions in Experiment 4 indicates that plural attractors led to greater processing difficulty than singular attractors, at some post-interpretative level. I reason that this additional processing load arises due to greater memory storage costs associated with a plural noun phrase compared with a singular noun phrase because plurals are more morphologically and semantically complex (Eberhard, 1997; N. Pearlmutter et al., 1999; Wagers et al., 2009), and therefore, building a mental representation of a plural NP requires more resources. Second, I found longer response times to comprehension questions when the subject-verb agreement dependency itself was of greater length (i.e., in the long oblique condition). This suggests that agreement dependency length is a relevant factor in post-interpretative processing when the subject and verb mismatch in number features (thus, yielding an ungrammatical sentence): this could suggest that the greater the distance between a subject and a mismatched verb, the more difficult the subsequent representation of the sentence is to build.

There is one aspect of our results, however, that conflicts with previous empirical literature: the *object* (filler) conditions of the current experiment in fact correspond to the "subject attractor" conditions in Parker & An (2018). However, unlike our current study, Parker & An (2018) found no significant attraction effect in ungrammatical (or indeed in grammatical) sentences of these types. There is

140

nonetheless an important difference between the current experiment and Parker &

An's (2018) study: whereas Parker & An's 2x2x2 factorial design included sentence

grammaticality (grammatical agreement vs. ungrammatical agreement) as a

manipulation, our current experiment employed exclusively *un*grammatical sentences

(thus, our 3x2 design involved no manipulation of sentence grammaticality). Thus,

whereas participants in Parker & An's study were exposed to both balanced

proportions of grammatical and ungrammatical sentences, participants in our

experiment read for the most part (e.g., 75% of the time), *ungrammatical* sentences

(the well-formed fillers comprised 25% of the total tokens).

I consider that, unlike in Parker & An's study, participants in the present study

were deploying a more shallow, or "good-enough", processing mode (Ferreira et al.,

2002; Ferreira & Lowder, 2016; Frank & Bod, 2011) wherein the parser comprehend

information structures which seem "just good enough", rather than fully analyze

sentence structures. I reason that this shallow processing mode can be activated when

participants are faced with stimuli comprising chiefly ungrammatical sentences

because this steady ill-formedness does not make them actively engage in the process

of construing grammaticality judgments. Instead, a more comprehensive analysis of

sentence structure would occur when deep(er) processing mechanisms are activated, as

in Parker & An (2018), and I reason that the additional attentiveness associated with

such a "deep" level does *not* lead the parser to erroneously encode subject attractors as

potential targets for verb agreement.

Fortunately, our hypotheses pertaining to both our main finding from Experiments and the conflict with Parker & An's (2018) results as arising from the deployment of a "good-enough" processing mode in our Experiment 4, converge to make a clear prediction: the effect in which a subject attractor impacts agreement dependency formation only when the gap in relative clauses containing the attractor is a core argument (i.e., object), should be reversed in grammatical sentences, yielding "slowdown effect of an attactor" (i.e., the inverse of an "agreement attraction" effect). That is, when the verb bares singular agreement (and is therefore *matched* with the target subject) I expect to find that this mismatching verb is read *slower* in sentences with a plural subject attractor, but only when the gap in relative clauses containing the attractor is in *object* position, like in (56a; cf. 50a). If the grammatical status of the gap is indeed at stake, then this inhibitory illusion effect should be absent when the gap is in oblique position (56b, 56c; cf. 50b).

(56) a. Direct object filler
The athlete [who[OBJECT] the coach(es)[SUBJECT] trained __] definitely **was** prepared for the championship game.

b. Oblique adjunct filler
The athlete [who[OBLIQUE] the coach(es) [SUBJECT] trained with __] definitely **was** prepared for the championship game.

c. Oblique adjunct filler ("long oblique")
The athlete [who[OBLIQUE] the coach(es) [SUBJECT] trained with __ on Sunday] definitely **was** prepared for the championship game.

Importantly, large sample studies have recently proposed that such illusions of ungrammaticality arise *only* in good-enough processing mode, and not in deep

processing mode (Laurinavichyute & Malsburg, 2019, 2021). Our Experiment 5, therefore, tests exclusively grammatical sentences such as in (56), with the prediction that slowdown effect of an attractor will obtain (i) only in the object filler conditions, and (ii) only if good-enough processing mode is deployed.

5.3   Experiment 5: Grammatical singular agreement

5.3.1   Participants

Thirty-six native speakers of English were recruited using Amazon's Mechanical Turk web service. None had participated in Experiment 4. All indicated that they were native speakers of English and were compensated with $5.00 (as with Experiment 4, payment was not contingent upon the response to this question, so there was no incentive to answer dishonestly).

5.3.2   Materials

All experimental design and item orders were identical to those used in Experiment 4, but this time, they were all grammatical - with singular verb agreement ("*was*") – as shown in Table 17. This, therefore, was matched with the number feature of the target subject (which was also always singular). Importantly, I used only grammatical filler items. Thus, the fillers were identical to those used in Experiment 4, except that all of them were revised to be grammatical (for example, the erroneous present tense *thanks* in the prior example sentence in (27) was modified to past tense form *thanked*; some fillers from Experiment 4 were removed, resulting in a study of slightly shorter duration). Our motivation for including exclusively grammatical fillers

was based on an observation that previous studies which have reported illusions of ungrammaticality tested grammatical sentences only (Acuña-Fariña et al., 2014; Adani et al., 2010; Franck et al., 2008; Villata et al., 2018; Villata & Franck, 2020; see Hammerly et al., 2019b for a comprehensive study of response bias in agreement attraction).

Table 17:   A sample set of items for Experiment 5

Object condition
SG attractor
The athlete [who the coach trained __] definitely **was** prepared for the championship game.
PL attractor
The athlete [who the coach**es** trained __] definitely **was** prepared for the championship game.
Oblique condition
SG attractor
The athlete [who the coach trained with __] definitely **was** prepared for the championship game.
PL attractor
The athlete [who the coach**es** trained with __] definitely **was** prepared for the championship game.
Long-oblique condition
SG attractor
The athlete [who the coach trained with __ on Sunday] definitely **was** prepared for the championship game.
PL attractor
The athlete [who the coach**es** trained with __ on Sunday] definitely **was** prepared for the championship game.

5.3.3 Procedure

This experiment used a self-paced reading task, following the same procedure and analysis used in Experiment 4. The experiment lasted approximately 20 minutes.

5.3.4 Results

Of the 648 critical trials generated, the trials which generated incorrect responses ($n = 73$)[27] were excluded from further analysis. I analyzed the same four regions of interest as with Experiment 4, following the same outlier removal, contrast coding, and modeling procedure as reported in Section 5.2.5.

Figure 13 shows the mean word-by-word reading times for the object conditions, and figures 14 and 15 show the same for oblique conditions and long-oblique conditions, respectively. At the critical region, I observed a marginal main effect of argument status in terms of the comparison between object and oblique *wh* filler dependencies ("GRAMMATICAL FUNCTION"; β=26.29, SE=13.62, t=1.93, p=.064), hinting at a trend in which the agreeing copula verb "*was*" was read slower when the target subject is modified by a relative clause containing a subject attractor and an *object* (i.e., core argument) filler-gap dependency, as compared with when the filler-gap dependency is oblique. The second contrast comparison between the two oblique conditions ("LENGTH") was not significant ($p = .53$); thus, like Experiment 4, this provides no evidence that linear distance between the target subject and the agreeing verb affected resolution of the agreement dependency.

---

[27] A logistic-mixed effects regression revealed no significant effects of accuracy across conditions, and no interactions (all ps > .12).

As with Experiment 4, there was no main effect of NUMBER (p = .21); that is, the copula verb was not read significantly faster when the attractor was singular ("*the coach*") compared with when it was plural ("*the coaches*"), overall. However, I observed a significant interaction of NUMBER and GRAMMATICAL FUNCTION (β=-52.15, SE=17.16, t=-3.04, p=.0025), indicating that the object and oblique dependencies responded differently to each other according to whether the attractor was singular or plural (the interaction of NUMBER and LENGTH was not significant; p = .24). As with Experiment 4, I examined this interaction by conducting nested planned comparisons, asking, for each level of argument status (object, oblique, long oblique), whether there was a difference in reading times when the attractor was singular (and thus, *matched* in number with the agreeing copula verb) compared with when it was plural (and thus, *mismatched* in number with the verb). Here, our predictions from Experiment 4 were borne out in that our effects were now mirrored: the *object* condition - in which the copula verb was read slower when the attractor was *plural* than when it was *singular* - was this comparison significant (462 ms vs 417 ms; β=-45.99, SE=14.68, t=-3.13, p=.002; see the box in Figure 13). This same difference showed no signs of significance in either the oblique condition (420 ms vs 417ms; t = .47; p = .72; see box in see Figure 14) or in the long oblique condition (401 ms vs 424 ms; t = -.98; p = .23; see box in Figure 15). This indicates that the attraction effect of a *subject* attractor which is mismatched with the agreeing verb (here: a plural attractor mismatched with singular agreement) obtains when the relevant filler-gap dependency path (along which the attractor is situated) pertains to an object gap (i.e., core argument) and not when it pertains to an oblique gap (i.e., non-core argument), regardless of length.

I now turn to the spillover regions. At Spillover 1 (i.e., the predicative adjective following the copula verb), I observed the same significant interaction of LENGTH and NUMBER (β=43.51, *SE*=21.11, *t*=2.06, *p*=.048) that I saw at Spillover 2 in Experiment 4. This further indicates that linear distance between the main subject and the agreeing verb indeed has an effect, but one which emerges after the copula verb has been processed. There were no other significant or marginal effects (all *p*s >.15). This time, Spillover 2 also showed no further significant or marginal effects (all *p*s >.29).
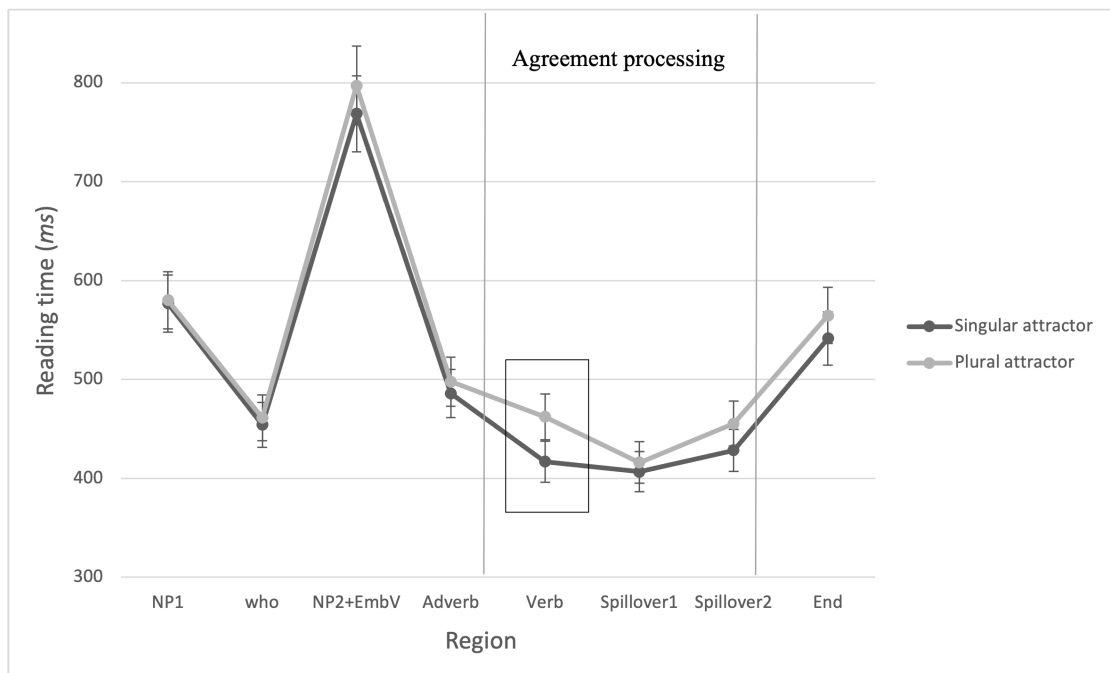
Figure 13: Region-by-region reading times for the object conditions (singular attractor as a dashed line; plural attractor as a solid line). The critical region is boxed. Error bars indicate 95% Confidence intervals.
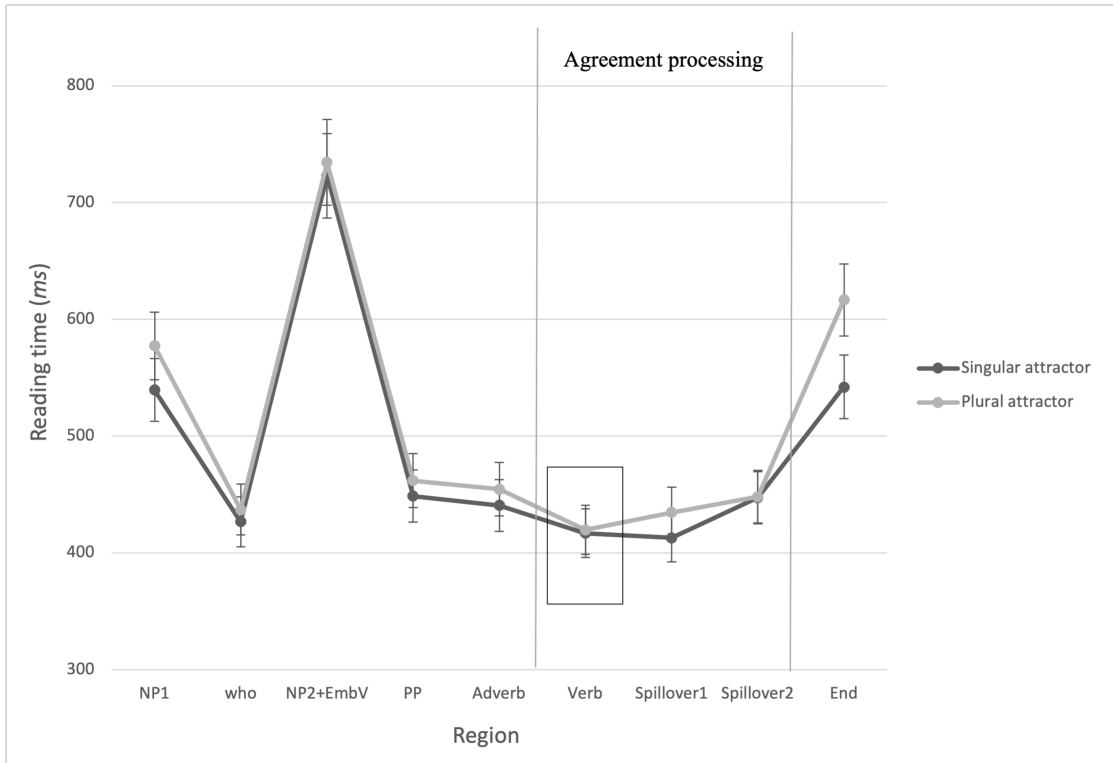
Figure 14: Region-by-region reading times for the oblique conditions (singular attractor as a dashed line; plural attractor as a solid line). The critical region is boxed. Error bars indicate 95% Confidence intervals.
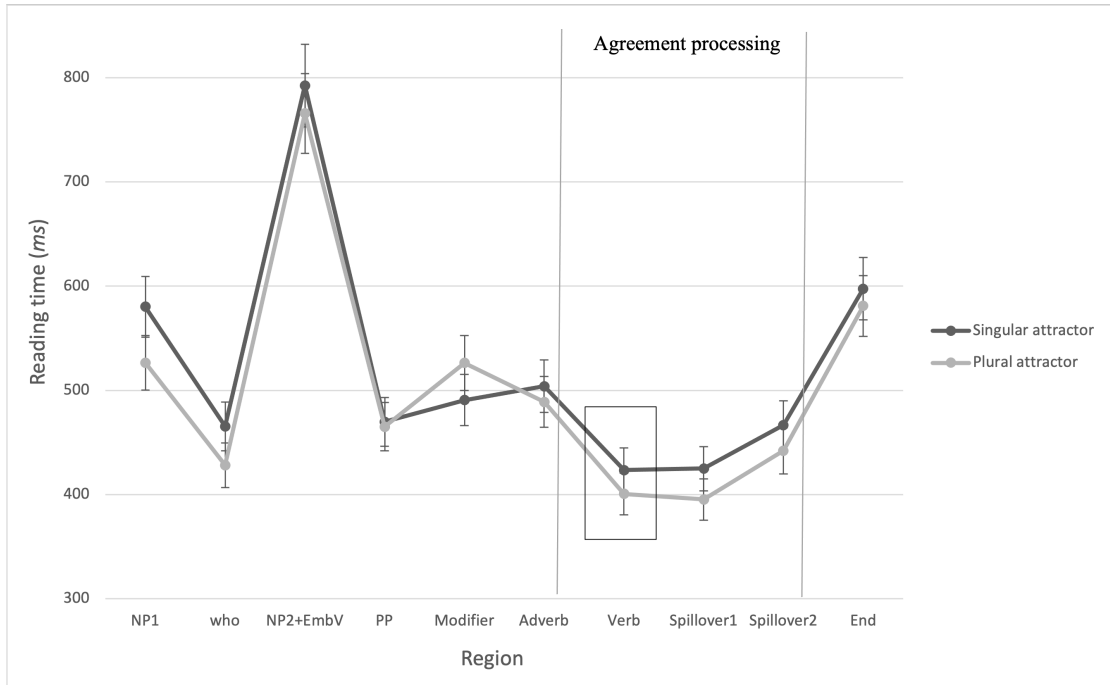
Figure 15: Region-by-region reading times for the long oblique conditions (singular attractor as a dashed line; plural attractor aa s solid line). The critical region is boxed. Error bars indicate 95% Confidence intervals.

I turn lastly to time taken to (correctly) answer the comprehension question ("response time"). Figure 16 reports response times of six conditions. Here, unlike with reading times, there was no main effect of argument status (all both $ps > .85$), but there was a marginal main effect of number, with response times to sentences containing a plural attractor longer than those for sentences with a singular attractor (1883 ms vs. 1724 ms; β=-168.09, SE=86.25, t=-1.95, $p = .063$). This trend suggests that plural attractors cause interference at the post-interpretative processing level

(Caplan et al., 1998), when the target subject matches the verb in number feature, and regardless of the grammatical function of the filler-gap dependency.
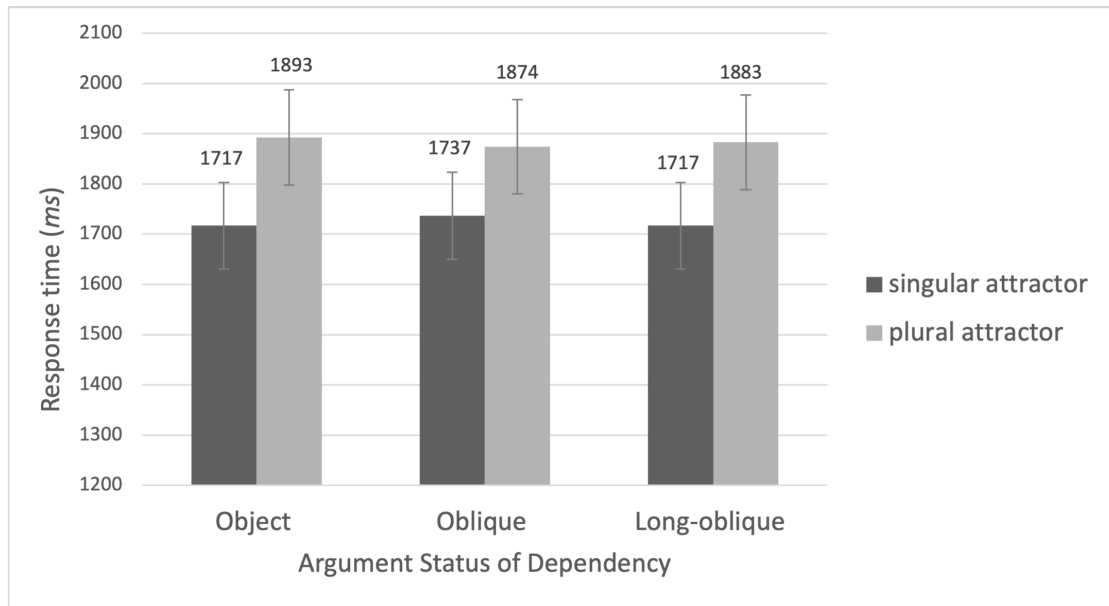
Figure 16:   Average response times to the comprehension questions (*ms*) by condition. Error bars indicate 95% Confidence intervals.

5.3.5   Discussion

The results of Experiment 5 further support our conclusion from Experiment 4; that is, that subject-verb agreement interference effects are susceptible to the argument status of the filler at its gap site: specifically, these effects are invoked by a subject attractor occurs only when the filler is the *direct object* (i.e., a core argument) of the embedded verb at its gap site within the relative clause. Just as I observed an "attraction" effect with ungrammatical sentences in Experiment 4, I see the inverse "slowdown effect" in Experiment 5: the reading time at a main agreeing verb is longer

when an attractor *mis*matches the main verb in the number feature as compared with when it matches. Critically, these two manifestations of an interference effect emerge only (i) when the dependency pertains to object position, and (ii) when the "good-enough" processing mode is deployed.

Response times to the comprehension questions, on the other hand, show consistent results across conditions: a longer response time for plural attractors than singular ones. This dovetails with the accuracy rates from Experiment 4, in which accuracy was lower when the attractor was plural than when it was singular; I reason that this is likely due to the processing cost of plurals because plurals are morphologically and semantically more complex than singulars: plurals are marked compared to singulars, requiring the presence of plural features (Eberhard, 1997; Wagers et al., 2009) and the concept of multiple entities.

5.4   General Discussion

Previous studies have provided evidence that the syntactic prominence of an *attractor* modulates dependency formation. The goal of the current study was to explore whether the argument status of a filler-gap dependency within which the attractor is situated, can modulate subject-verb dependency formation. In terms of online sentence comprehension, I find that when, and only when, the filler is a core argument (i.e., direct object), do attraction effects of a subject attractor occur. These manifest as a facilitatory "agreement attraction" effect when the sentence is grammatically ill-formed (faster reading times for plural attractors; Experiment 4), and an "slowdown effect" when the sentence is grammatically well-formed (slower reading times for plural attractors; Experiment 5). Prior literature (Van Dyke &

McElree, 2011; Parker & An, 2018) has claimed that if an attractor is a core argument, that is, a subject or a direct object, then it is salient enough to prevent attraction effects. If an attractor is in PP position, on the other hand, it is more likely to induce interference effects due to its less prominent thematic role in the sentence. Van Dyke & McElree (2011) further proposed the concept of syntactic gating in such a way that the matching syntactic cue from the verb, or the subject position, may trigger an interference effect. The current study raised the question of how the syntactic prominence of a filler at its gap site impacts the computation of subject-verb agreement dependencies.

The pattern of the reading times at the main verb was not predicted by theories of interference effects which rely solely on the argument status of an attractor (e.g., Van Dyke & McElree, 2011). First, in terms of the argument status of the attractor, all conditions were predicted to reveal attraction effects since the argument status of the attractor and, the *wh* filler in the matrix clause, remained constant, as a subject, across conditions, and thus, "syntactic gating" should manifest consistently. Second, in terms of the argument status of the filler at its gap site, strong attraction effects should be observed in *oblique* conditions and not in object conditions, because, if we follow the previous account of interference, then the encoding of an oblique is predicted to be less distinctive, and thus interference effect should arise precisely here. Therefore, our results point towards a memory mechanism for the processing of subject-verb agreements that is more grammatically sophisticated than previously assumed. It necessitates a more fine-grained account, involving the syntactic status of *both* the attractor *and* the *wh* filler gap-dependency which contains the attractor. In the current study, a subject-verb agreement dependency was more likely to be interfered with by

an attractor when the filler is a core verbal argument (i.e., a direct object) than when the filler was a more peripheral argument (i.e., an oblique or prepositional object) at its gap site.

However, I also suggested that this asymmetry in terms of filler-gap status (object vs. oblique) is operative only in "good-enough" processing mode (Ferreira et al., 2002; Ferreira & Lowder, 2016): whether the same asymmetry holds when "deep" processing mode is deployed remains an open question. While I leave a full exploration of this issue to future research, I note that Parker & An's (2018) study critically did *not* find attraction effects in sentences with direct object filler-gaps, and reason that this may be because good-enough processing was not operative in their study, due to the coupling of grammatical with ungrammatical sentences. Our account that follows, therefore, pertains minimally to "good enough" processing modes; I leave open the question of whether and how it relates to deep processing.

A two-stage model of subject-verb agreement dependency formation

I begin by situating our current findings in the context of previous studies (Table 18), and I notice that a general pattern emerges. First, when the attractor is an *oblique* (i.e., a non-core argument), an attraction effect arises, as replicated across multiple studies. Second, when the attractor is an *object*, then no attraction effect arises, as also replicated substantially. When the attractor is a *subject*, however, then the emergence of attraction effect depends upon the grammatical role of the filler-gap dependency in the relative clause containing the attractor: if the filler is oblique, then attraction effects are prevented. However, if the filler is a core argument (subject or object) then attraction effects arise.

Table 18:   Summary of interference effects grouped by roles of filler and attractor, in previous studies and the current study.

Role of Filler-gap in RC containing attractor	Role of clause-mate attractor	Attraction effect	References
Subject	Oblique	✓	Van Dyke (2007); Parker & An (2018)
Oblique	Subject	✗	**Current study;**
Subject	Subject	✓	Van Dyke & McElree (2011)
Object	Subject	✓	**Current study;** (Parker, Lago, & Phillips, 2015)[28]
Subject	Object	✗	Van Dyke & McElree (2011); Parker & An (2018)

I interpret this body of findings in terms of a framework which makes two separate stages of distinction. The first stage is a distinction between *core arguments* (i.e., subjects and objects), and *non-core arguments* (i.e., oblique adjuncts) in the memory encoding system. In syntactic terms, this amounts to a distinction between NPs whose thematic role is dependent on, or "assigned by", the verb proper. In (57a), for example, the verb "train" assigns two thematic roles – one to the subject "the coach", which is interpreted as an agent, and one to its object complement "the

---

[28] This thesis manipulated the number feature of an attractor and found facilitatory attraction effects of a subject attractor when the filler was an object position at its gap site in two separate experiments. This construction is analogous to the object conditions in the current paper.

> e.g., *The doctor that the researcher(s) described meticulously was/were certified after debunking the urban myth in the new scientific journal.*

athlete", which is interpreted as a patient. The subject and object are therefore *core arguments*. Conversely, in (57b), "train" assigns only one thematic role: agent, to the NP in subject position (and thus, this subject is a core argument). The second NP in (57b), "the athlete", is a complement not of the "train" but of the preposition "with", upon which it is both structurally and interpretatively dependent. Therefore, "the athlete" in (57b) is a *non-core argument*.

(57)   a. The coach [SUBJECT] **trained** the athlete [OBJECT]

AGENT   PATIENT

b. The coach [SUBJECT] **trained** (with the athlete [OBLIQUE])

AGENT
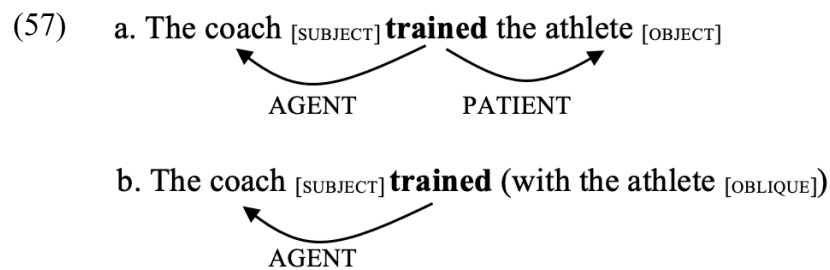
This first stage is responsible for distinctiveness in terms of the initial encoding of NPs in memory. Subsequently, the second stage of distinction is associated with the verb agreement with the main subject in retrieval mechanism. This second stage of processing takes place when the parser encounters the verb and attempts to form agreement; in this stage, the parser retrieves information on each NP, and actively traces syntactically matching cues from the verb (e.g., plural). It is per this second stage that verb agreement is attempted with – optimally – some NP in a *subject* position, and such an attempt leads to interference effects in instances when the attractor is a subject, but in which its clause-mate filler is also a core argument. To illustrate, we lay out this model in (58).

(58)

Two-stages of the subject-verb agreement formation

Stage 1 (argument encoding stage):   Core argument   > non-core argument.

Stage 2 (agreement processing stage):   Subject > object.

I begin with *Stage 1* in the memory encoding system. In line with Van Dyke & McElree (2011), I conjecture that the degree of prominence of an NP in terms of its argument status corresponds to the degree of *distinctiveness* in memory: a more prominent NP is more distinct, and thus, better encoded in memory, than a less prominent one. Also following Van Dyke & McElree (2011), I maintain a view in which attraction effects arise when an attractor is *non*-distinctively encoded. Building on this previous literature, however, I propose that distinctiveness is computed in terms of the prominence of both the attractor *and* its local clause-mate filler, alongside which it is evaluated. As shown in Table 18, attraction effects arise when the attractor is *less* distinctively encoded (i.e., as a non-core argument) than its clause-mate filler, (i.e., when the attractor is an oblique, non-core argument, and its clause-mate filler is a subject or object core argument), as in both Van Dyke (2007) and Parker & An (2018). If, conversely, the attractor is *more* distinctively (i.e., as a core argument) encoded than the filler (i.e., when it is a non-core argument), then attraction effects do not arise, as shown in the "oblique" conditions in the current study. However, when an attractor is of *equal* prominence to the filler (that is, both are core arguments), then both an attractor and the filler become potential candidates to agree with the main verb in retrieval at *Stage 2*: the agreement processing stage.

156

If the filler and its clause-mate attractor are both core arguments, then the sentence parser is able to consider forming verb agreement with, in principle, either one of them. However, the one which is the more optimal candidate for the agreement dependency, which is the subject, is more strongly attempted to agree with the verb due to its syntactically matching cue to the verb. If, under *Stage 2*, both the filler and attractor are equal (e.g., both are subjects), then the agreement is attempted with both. This attempt at double-agreement causes additional processing load (i.e., the verb takes longer to compute because there are two potential candidates that it could agree with), which corresponds to Van Dyke & McElree's (2011) "syntactic gating". Otherwise, the agreement is attempted with the more prominent *subject*. When the *filler* is the subject alone, which has a syntactically matching cue to the verb, the attractor (as it is thus the object), does not trigger an attraction effect (as was the finding in both Van Dyke & McElree, 2011, and Parker & An, 2018). However, when it is the attractor that is the subject, the agreement is temporally attempted with the attractor due to its syntactically matching cue, as in the "object" condition in the current study; this gives rise to either the attraction effect (as in our Experiment 4) or the slowdown effect (as in our Experiment 5).[29]

Online and offline processing

---

[29] In terms of surface position, all fillers considered in the current paper - and previous studies - are in the main subject position. I maintain that, when the syntactic role of the filler at its gap site and its surface structure matches (i.e., both subjects) then the filler either (i) becomes more salient in the encoding stage (as per stage 1) and/or (ii) provides an *even* stronger syntactic matching cue to the verb in retrieval (as per stage 2). If, however, the syntactic role at its gap site mismatches its role at the surface structure, such that the syntactic role of its gap site is not a subject (e.g., object or oblique), as in the present study, then its surface position being a subject no longer strengthen the salience of the filler.

Unlike the findings of the reading time at the critical region in real-time processing, the response time of comprehension questions in the current study revealed a significant effect of the plural attractors: response times to the comprehension questions were longer when the attractor was plural compared with when it was singular. The processing cost of plural attractors occurred even when the number feature of the main verb was plural (Experiment 4). This finding supports an explanation for this attraction effect, which Wagers et al. (2009) pointed out as the processing cost of plurals relative to singulars. Plurals are morphologically more complex than singulars, requiring the semantic composition of plurals with predicates. This process may incur an additional processing load for plural attractors. Several studies have shown that number features affect the lexical decision in processing (Baayena et al., 1997; Gimenes & Brysbaert, 2016; Lau et al., 2007; New et al., 2004). In some studies on sentence comprehensions N. Pearlmutter et al. (1999) and Julie Franck et al. (2015) have also observed a statistically significant attraction effect of a plural attractor, which the result of the current project aligns with. Particularly, Franck et al. (2015) found different results in different tasks, like between a self-paced reading and grammaticality judgment. Given these findings in previous studies, different tasks can be involved in different processes. Lexical decision tasks and acceptability judgment tasks, for instance, are associated with decision-making and reasoning, all of which are consistent with the comprehension questions presented in the current study.

The different pattern of the response time on the comprehension question and the reading time on the critical region suggests that effects occurred at different time points. In interpretive processing where parsers extract meanings from

semantic/syntactic features at the sentence or discourse level, the susceptibility of the argument status of the filler at its gap site was observed only in online processing. In post-interpretive processing where parsers recall the whole meaning of a sentence, the effect of the argument status of the filler at its gap site was absent once these integrations were constructed. The effect of number, on the other hand, mainly affects the processing of a conceptual representation while recalling the information about the sentence and performing the truth-value judgment. That is, the subject-verb agreement dependency was more sensitive to the number feature of a local attractor in post-interpretive processing than in sentence-level, interpretive processing.

Language comprehension and production

The role of the argument status of a referent for its mental representation has been widely discussed in language production, such that speakers or listeners pay more attention to referents in a prominent position. Many researchers have proposed that syntactic prominence is associated with language production (Bock, 1987; Bock & Warren, 1985; McDonald et al., 1993; Prat-Sala & Branigan, 2000). The prominence of syntactic category and the argument status of the verb (i.e., subcategorization frame) have been shown to be reflected in the production. Previous literature assumed that more readily retrieved information is generated prior to the less accessible information (Melinger & Dobel, 2005; Pickering & Branigan, 1998). For example, a readily retrieved referent is likely to be initially generated as a subject, which is the most prominent argument status.

The interference effects observed for the object conditions in the present study during real-time online processing is consistent with a few previous research on

159

interferences found in language production: The processing of number agreements (Bock & Cutting, 1992; Bock & Eberhard, 1993; Fayol et al., 1999; Hartsuiker & Barkhuysen, 2006) and gender agreements between the subject and the verb (Badecker & Kuminiak, 2007) was influenced by the features of an intervenor in sentence production. These empirical studies showed an interference effect of a plural intervenor, indicating that the parser was attempting to match the verb to the number feature of an intervenor. The role of this effect was introduced as the principle of proximity (Quirk et al., 1972) or linear distance hypothesis (Hsiao & Gibson, 2003; Tarollo & Myhill, 1983). This perspective of interference effects emphasizes the locality of an intervenor to the verb in sentence production. When an intervenor disagrees in number with the subject, parsers tend to falsely produce the verb agreement with the intervenor. In a similar fashion, the object condition of the current sentence-comprehension study showed the effect of an intervenor number. The reading time at the verb depended on the number feature of intervenors in the object conditions: a shorter reading time when the intervenor matches the verb in number and a longer reading time when the intervenor mismatches the verb in number, regardless of the sentence grammaticality.

This similarly fashioned phenomenon addresses an issue on whether the cognitive mechanisms between comprehension and production are shared or they are two distinct mechanisms. The current study is in line with the recent claim that posits a single mechanism for building the structural representation of sentences in comprehension and production (Kempen, 2000; Momma et al., 2017; Momma & Phillips, 2018) and suggests further research on how speakers generate sentences which require dependency formation, depending on the argument status of the filler

and an intervenor. The similarities in the attraction effects of subject-verb agreements, however, do not confirm that the cues used in comprehension and production are identical (Badecker & Lewis, 2008). In comprehension, for instance, morphological cues (e.g., number) can be used for retrieving the subject, whereas in production, speakers first select a verb form and do not use a process to retrieve morphological cues. Therefore, a more fine-grained account for cognitive mechanisms between production and comprehension is needed for future work.

More on the argument processing model

The current study overall is in line with the proposal that the interference is modulated by the argument status of NPs and motivates a broader study of argument processing in language comprehension. The findings of this study are closely aligned with a neurocognitive model of argument processing, the so-called Extended Argument Dependency Model (Bornkessel & Schlesewsky, 2006). The architecture of this Extended Argument Dependency Model (eADM) includes hierarchically organized three processing phases. It mainly discusses the processing of core arguments associated with the verb, stressing the role of the prominence of arguments.

Phase 1 proceeds basic constituent structure building without argument interpretation in relation to the verb. Phase 2 involves argument interpretations associated with processing a thematic role and a position in an argument hierarchy. This stage posits the computation of prominence based on a restricted set of hierarchical features, such as animacy, definiteness, and person. Given a separate stage of prominence assignments, this neurocognitive model proposes that the prominence

161

of an argument plays a crucial role in language processing. This argument interpretation process is completed in Phase 3, encompassing information from further domains such as discourse and pragmatic contexts.

This model is in line with the proposed model in this thesis in that it posits a separate stage of the computation of argument based on its prominence. In terms of the processing of subject-verb agreements in the presence of an attractor, Phase 2 under this eADM model can encompass argument interpretations of both the filler and the attractor associated with the main verb. In addition, the computation of the attractor's prominence compared to the filler will be associated with Phase 2 process.

5.5   Summary

In two self-paced reading experiments, I examined whether and how interference effects in the processing of subject-verb agreement are modulated by the grammatical function of the filler-gap dependency which contains the attractor, while holding constant the grammatical role of the attractor itself. Following previous studies, I find that an attractor in the subject position of a relative clause modifying the target subject indeed modulates processing of number agreement, but only when the relevant gap – associated with the filler – is also a core argument, and not when it is an (oblique) adjunct. I propose that this necessitates a nuanced model of subject-verb agreement processing, in which memory encoding of attractor NPs in terms of their distinctiveness (which is sensitive to a broad distinction between core argument vs non-core argument status) happens independently of the parser's attempt(s) at forming an agreement between an NP and the relevant verb (which is sensitive to the distinction between subject and object status). Most importantly, our findings indicate that encoding of attractor NP is sensitive not only to its own grammatical role, but to

other grammatical roles borne within its immediate clausal environment. I also draw further attention to distinctions between good-enough and deep processing levels, and maintain that our findings pertain minimally to a good-enough level; the question of whether and how this differs in deep processing remains open.

Chapter 6

CONCLUDING REMARKS

This dissertation aims to fill gaps in both syntactic/semantic and psycholinguistic theory concerning the processing of dependency formations. On this view, this thesis has explored how semantic and syntactic information contributes to interference effects of intervenor NP in the processing of long-distance dependencies.

• *Interference effects*. The semantic information pertaining to an intervenor and its accommodation costs in the discourse, and depending on the intervenor NP types, can mediate the processing of long-distance dependencies. The argument status of an intervenor mediates the dependency formation by interacting with the argument status of the filler.

• *NP types of an intervenor*. The referential hierarchy of NP givenness plays an important role in the difficulty of dependency formations. The binary metric between pronoun (high on the hierarchy) and full NPs (low on the hierarchy) was shown to be operative in these studies presented here, such that pronoun intervenors lead to easier processing than intervenors like definite and indefinite NPs. In addition, the robust ease in processing pronoun was revealed regardless of accommodation costs. The lexical properties of accommodation costs for full NPs, on the other hand, are crucial factors in determining the difficulty of dependency formations.

• *The argument status of an intervenor*. The relative distinctiveness of the argument status of an intervenor, compared to the filler, crucially modulates the interference effect on the resolution of the subject-verb number agreement

dependency. I propose two stages of subject-verb agreement formation. A more prominent argument, like a core argument, is better encoded in memory; when both the filler and intervenor are core arguments, the NP in the subject position is strongly attempted to agree with the verb in retrieval. When both are subjects, an additional processing load is predicted due to two syntactically matching cues.

REFERENCES

Abeillé, A., Hemforth, B., Winckel, E., & Gibson, E. (2020). Extraction from subjects: Differences in acceptability depend on the discourse function of the construction. *Cognition*, *204*.

Acuña-Fariña, J. C., Meseguer, E., & Carreiras, M. (2014). Gender and number agreement in comprehension in Spanish. *Lingua*, *143*, 108–128.

Adani, F., van der Lely, H., Forgiarini, M., & Guasti, M. (2010). Grammatical feature dissimilarities make relative clauses easier: a comprehension study with Italian children. *Lingua*, *120*, 2148–2166.

Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory*, *21*, 435–483.

Almor, A. (1999). Noun-phrase anaphora and focus: the informational load hypothesis. *Psychological Review*, *106*, 748–765.

Almor, A. (2000). Constraints and mechanisms in theories of anaphor processing. In M. Crocker, M. Pickering, & C. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 341–354). Cambridge University Press.

Almor, A. (2004). A computational investigation of reference in production and comprehension. In J. Trueswell & M. Tanenhaus (Eds.), *Approaches to studying world-situated language use: Bridging the language-as-product and language-as-action traditions* (pp. 285–301). MIT Press.

Almor, A., & Nair, V. (2007). The form of referential expressions in discourse. *Language and Linguistics Compass*, *10*, 84–99.

Alonso-Ovalle, L., Menndez-Benito, P., & Schwarz, F. (2011). Maximize presupposition and two types of definite competitors. In S. Lima, K. Mullin, & B. Smith (Eds.), *Proceedings of NELS 39* (pp. 29–40).

Anderson, J. (2004). An integrated theory of mind. *Psychological Review*, *111*, 1036–1060.

Ariel, M. (1988). Referring and accessibility. *Journal of Linguistics*, *24*, 65–87.

Ariel, M. (1990). *Accessing noun-phrase antecedents*. Routledge.

Ariel, M. (1994). Interpreting anaphoric expressions: A cognitive versus a pragmatic approach. *Journal of Linguistics*, *30*(1), 3–42.

Arnold, J. (1998). *Reference form and discourse patterns*.

Arnold, J. (2001). The effect of thematic roles on pronoun use and frequency of reference continuation. *Discourse Processes*, *31*(2), 137–162.

Arnold, J. (2010). How speakers refer: The role of accessibility. *Language and Linguistics Compass*, *4/4*, 187–203.

Arnold, J., Fagnano, M., & Tanenhaus, M. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research*, *32*(1), 25–36.

Arnold, J., & Griffin, Z. (2007). The effect of additional characters on choice of referring expression: Everyone counts. *Journal of Memory and Language*, *56*, 521–536.

Arnon, I. (2010). Rethinking child difficulty : The effect of NP type on children's processing of relative clauses in Hebrew. *Journal of Child Language*, *37*, 27–57. https://doi.org/10.1017/S030500090900943X

Attneave, F. (1959). *Applications of information theory to psychology: A summary of basic concepts, methods and results*. Holt-Dryden.

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

Baayena, R. H., Dijkstrab, T., & Schreuder, R. (1997). Singulars and plurals in Dutch: Evidence for a parallel dual-route model. *Journal of Memory and Language*, *37*(1), 94–117.

Bade, N., & Schwarz, F. (2019a). An experimental investigation of antipresuppositions. In A. Creemers & C. Richter (Eds.), *University of Pennsylvania Working Papers in Linguistics* (pp. 31–40).

Bade, N., & Schwarz, F. (2019b). (In-)definites, (anti-)uniqueness, and uniqueness expectations. In A. Goel, C. Seifert, & C. Freksa (Eds.), *Proceedings of the 41st Annual Conference of the Cognitive Science Society,* (pp. 119–125). Cognitive Science Society.

Bade, N., & Schwarz, F. (2021). New data on the competition between definites and indefinites. *Proceedings of ELM*, 15–26.

Badecker, W., & Kuminiak, F. (2007). Morphology, agreement and working memory retrieval in. *Journal of Memory and Language*, *56*, 65–85.

Badecker, W., & Lewis, RL. (2008). *Agreement and sentence formulation: the role of working memory (retrievals) in language production*.

Badecker, W., & Straub, K. (2002). The Processing Role of Structural Constraints on the Interpretation of Pronouns and Anaphors. *Journal of Experimental Psychology*, *28*(4), 748–769. https://doi.org/10.1037//0278-7393.28.4.748

Barker, C. (2004). Possessive weak definites. *University of Massachusetts Occasional Papers*.

Barker, J., Nicol, J., & Garrett, M. (2001). Semantic factors in the production of number agreement. *Journal of Psycholinguistic Research*, *30*, 91–114.

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, *68*(3), 255–278.

Bates, D. (2010). *Lme4: Mixed-Effects Modeling with R*.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *ArXiv Preprint*.

Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *ArXiv Preprint*.

Battistella, E. (1990). *Markedness: The evaluative superstructure of language*. State University of New York Press.

Beaver, D., & Coppock, E. (2015). Novelty and familiarity for free. In T. Brochhagen, F. Roelofsen, & N. Theiler (Eds.), *Proceedings of 20th Amsterdam Colloquium* (pp. 50–59).

Bever, T. (1970). The cognitive basis for linguistic structures. In *Cognition and the development of language* (pp. 279–362). Wiley.

Bever, T. G. (1974). The ascent of the specious, or there's a lot we don't know about mirrors. In D. Cohen (Ed.), *Explaining linguistic phenomena* (pp. 173–200). Hemisphere.

Bock, J. K. (1987). An effect of the accessibility of word forms on sen- tence structures. *Journal of Memory and Language*, *26*, 119–137.

Bock, J. K., & Cutting, J. C. (1992). Regulating mental energy: Performance units in language production. *Journal of Memory and Language*, *31*, 99–127.

Bock, J. K., & Eberhard, K. M. (1993). Meaning, sound and syntax in English number agreement. *Language and Cognitive Processes*, *8*, 57–99.

Bock, J. K., & Warren, R. K. (1985). Conceptual accessibility and syn- tactic structure in sentence formulation. *Cognition*, *21*, 47–67.

Bornkessel, I., & Schlesewsky, M. (2006). The extended argument dependency model: A neurocognitive approach to sentence comprehension across languages. *Psychological Review*, *113*(4), 787–821.

Bossong, G. (1991). Differential object marking in Romance and beyond. In D. Kibbee & D. Wanner (Eds.), *New analyses in Romance linguistics* (pp. 143–170).

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

Box, G., & Cox, D. (1964). An analysis of transformations. *Journal of the Royal Statistical Society: Series B (Methodological)*, *26*(2), 211–243.

Brennan, S. (1995). Centering attention in discourse. *Language and Cognitive Processes*, *102*, 137–167.

Bresnan, J. (2001). *Lexical Functional Syntax*. Blackwell.

Caplan, D., Alpert, N., & G., Waters. (1998). Effects of syntactic structure and propositional number on patterns of regional cerebral blood flow. *Journal of Cognitive Neuroscience*, *10*, 541–552.

Caplan, D., Alpert, N., Waters, G., & Olivieri, A. (2000). Activation of Broca's area by syntactic processing under conditions of concurrent articulation. *Human Brain Mapping*, *9*(2), 65–71.

Caplan, D., Vijayan, S., Kuperberg, G., West, C., Waters, G., Greve, D., & Al, E. (2002). Vascular responses to syntactic processing: Event-related fMRI study of relative clauses. *Human Brain Mapping*, *15*(1), 26–38.

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

Carpenter, P. A., & Just, M. A. (2013). The role of working memory in language comprehension. In *Complex information processing* (pp. 51–88). Psychology Press.

Chafe, W. L. (1976). Giveness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and Topic*. Academic Press.

Chafe, W. L. (1987). Cognitive constraints on information flow. In R. Tomlin (Ed.), *Coherence and grounding in discourse* (pp. 21–51). John Benjamins.

Chemla, E. (2008). An epistemic step for antipresuppositions. *Journal of Semantics*, *25*(2), 141–173.

Chomsky, N., & Miller, G. A. (1963). Introduction to the formal analysis of natural languages. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (2nd ed., pp. 269–321). John Wiley & Sons.

Clark, H., & Sengul, C. (1979). In search of referents for noun phrases and pronouns. *Memory & Cognition*, *7*, 35–41.

Clifton, C. (2013). Situational context affects definiteness preferences: Accommodation of presuppositions. *Journal of Experimental Psychology*, *39*(2), 487–501.

Comrie, B. (1979). Definite and animate direct objects: A natural class. *Linguistica Silesiana*, *3*, 13–21.

Comrie, B. (1981). *Language universals and linguistic typology: Syntax and morphology*. University of Chicago Press.

Corley, S., & Crocker, M. W. (2000). The modular statistical hypothesis: Exploring lexical category ambiguity. In M. W. Crocker, M. Pickering, & C. J. Clifton (Eds.), *Architectures and mechanisms for language processing* (pp. 135–160). Cambridge University Press.

Crawley, R. A., Stevenson, R. J., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research*, *19*, 245–264.

Croft, W. (1988). Agreement vs. case marking and direct objects. In N. Barlow & C. Ferguson (Eds.), *Agreement in natural language: Approaches, theories, descriptions* (pp. 159–179). Center for the Study of Language and Information.

Croft, W. (1990). *Typology and universals*. Cambridge University Press.

Cunnings, I., & Felser, C. (2013). The role of working memory in the processing of reflexives. *Language and Cognitive Processes*, *28*(1–2), 188–219. https://doi.org/10.1080/01690965.2010.548391

Cunnings, I., & Sturt, P. (2018). Retrieval interference and semantic interpretation. *Journal of Memory and Language*, *102*, 16–27.

Dalrymple, M., & Nikolaeva, I. (2011). *Objects and information structure*. Cambridge University Press.

Danon, G. (2006). Caseless nominals and the projection of DP. *Natural Language & Linguistic Theory*, *24*, 977–1008.

de Hoop, H. (1996). *Case configuration and noun phrase interpretation*. Garland.

de Swart, P. (2007). *Cross-linguistic variation in object marking*.

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

Dillon, B., Mishler, A., Sloggett, S., & Phillips, C. (2013). Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language*, *69*(2), 85–103. https://doi.org/10.1016/j.jml.2013.04.003

Donkers, J., Hoeks, J. C. J., & Stowe, L. A. (2013). D-linking or set-restriction? Processing Which- questions in Dutch. *Language and Cognitive Processes*, *28*, 9–28.

Drummond, A. (2018). *Ibex Farm*. http://spellout.net/ibexfarm

du Bois, J. (2003). Argument structure : Grammar in use. In J. du Bois, L. Kumpf, & W. Ashby (Eds.), *Preferred argument: Grammar as architecture for function* (pp. 11–60). Benjamins.

Eberhard, K. M. (1997). The marked effect of number on subject–verb agreement. *Journal of Memory and Language*, *36*, 147–164.

Enc, M. (1983). Anchored expressions. In M. Barlow, D. Flickinger, & M. Westcoat (Eds.), *Proceedings of the west coast conference of formal linguistics* (pp. 79–88). Stanford Linguistics Association.

Enc, M. (1991). The semantics of specificity. *Linguistic Inquiry*, *22*(1), 1–26.

Engelmann, F., Jäger, L., & Vasishth, S. (2015). Cue confusion and distractor prominence can explain inconsistent interference effects. *Paper Presented at the 28th CUNY Conference on Human Sentence Processing*.

Erteschik-Shir, N. (1973). *On the nature of island constraints*.

Farkas, D. (1978). Direct and Indirect Object Reduplication in Romanian. *The 14th Regional Meeting of the Chicago Linguistic Society*, 88–97.

Fayol, M., Hupet, M., & Largy, P. (1999). The acquisition of subject–verb agreement in written French: From novices to experts' errors. *Reading and Writing: An Interdisciplinary Journal*, *11*, 153–174.

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

Ferreira, F., Bailey, K. G., & Ferraro, V. (2002). Good-enough representations in language comprehension. *Current Directions in Psychological Science*, *11*(1), 11–15.

Ferreira, F., & Lowder, M. W. (2016). Prediction, information structure, and good-enough language processing. In *Psychology of Learning and Motivation* (Vol. 65, pp. 217–247).

Foraker, S., & McElree, B. (2007). The role of prominence in pronoun resolution: Active versus passive representations. *Journal of Memory and Language*, *56*(3), 357–383.

Francis, H. S., Gregory, M. L., & Michaelis, L. A. (1998). Are lexical subjects deviant? *Proceedings of the 35th Chicago Linguistic Society*, 85–98.

Franck, J., Colonna, S., & Rizzi, L. (2015). Task-dependency and structure-dependency in number interference effects in sentence comprehension. *Frontiers in Psychology*, 1–15.

Franck, J., Lassi, G., Frauenfelder, U. H., & Rizzi, L. (2006). Agreement and movement: A syntactic analysis of attraction. *Cognition*, *101*, 173–216.

Franck, J., Soare, G., Frauenfelder, U. H., & Rizzi, L. (2010). Object interference in subject–verb agreement: The role of intermediate traces of movement. *Journal of Memory and Language*, *62*, 166–182.

Franck, J., Vigliocco, G., Antón-Méndez, I., Collina, S., & Frauenfelder, U. H. (2008). The interplay of syntax and form in sentence production: A cross-linguistic study of form effects on agreement. *Language and Cognitive Processes*, *23*(3), 329–374.

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*(4), 371–404.

Frank, S., & Bod, R. (2011). Insensitivity of the human sentence-processing system to hierarchical structure. *Psychological Science*, *22*(6), 829–834.

Frazier, L. (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory*, *5*(4), 519–559.

Frazier, L., & Clifton, C. (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes*, *4*, 93–126.

Frazier, L., & Clifton, C. (1996). *Construal*. MIT Press.

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

Futrell, R., Gibson, E., & Levy, R. (2020). Lossy-context surprisal: An information-theoretic model of memory effects in sentence processing. *Cognitive Science*, *44*, 1–54.

Garrod, S., & Sanford, A. (1994). Resolving sentences in a discourse context: How discourse representation affects language understanding. In M. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 675–698). Academic Press.

Gennari, S. P., & MacDonald, M. C. (2008). Semantic indeterminacy in object relative clauses. *Journal of Memory and Language*, *58*(2), 161–187.

Gernsbacher, M. A., & Hargreaves, D. I. (1988). Accessing sentence participants: The advantage of first mention. *Journal of Memory and Language*, *27*, 699–711.

Gerrig, R. (1986). Process and products of lexical access. *Language and Cognitive Processes*, 187–195.

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

Gibson, E. (2000). The dependency locality theory: a distance-based theory of linguistic complexity. In Y. Miyashita, A. P. Marantz, & W. O'Neil (Eds.), *Image, language, brain* (pp. 95–126). MIT Press.

Gibson, E., & Pearlmutter, N. (1998). Constraints on sentence comprehension. *Trends in Cognitive Sciences*, *2*(7), 262–268.

Gibson, E., & Warren, T. (1998). *Discourse reference and syntactic complexity*. MIT.

Gibson, E., & Warren, T. (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax*, *7*(1), 55–78.

Gimenes, M., & Brysbaert, M. (2016). The Processing of Singular and Plural Nouns in English, French, and Dutch: New Insights From Megastudies. *Canadian Journal of Experimental Psychology*, *70*(4), 316–324.

Givon, T. (1983). Topic continuity in discourse: an introduction. In T. Givon (Ed.), *Topic Continuity in Discourse: A Quantitative Cross-Language Study* (pp. 1–41). John Benjamins.

Givon, T. (1984). *Syntax: A functional–typological introduction* (Vol. 1). John Benjamins Publishing.

Goldberg, A. (2006). *Constructions at Work*. Oxford University Press.

Goodall, G. (2015). The D-linking effect on extraction from islands and non-islands. *Frontiers in Psychology*, *5*(Article 1493).

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.

Gordon, P. C., Hendrick, R., & Johnson, M. (2001). Memory interference during language processing. *Journal of Experimental Psychology*, *27*(6), 1411–1423. https://doi.org/10.1037//0278-7393.27.6.1411

Gordon, P. C., Hendrick, R., & Johnson, M. (2004). Effects of noun phrase type on sentence complexity. *Journal of Memory and Language*, *51*, 97–114. https://doi.org/10.1016/j.jml.2004.02.003

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*, *32*(6), 1304–1321. https://doi.org/10.1037/0278-7393.32.6.1304

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

Gordon, P., Grosz, B., & Gilliom, L. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science*, *17*, 311–347.

Gordon, P., & Hendrick, R. (1998). The representation and processing of coreference in discourse. *Cognitive Science*, *22*(4), 389–424.

Gordon, P., & Scearce, K. (1995). Pronominalization and discourse coherence, discourse structure and pronoun interpretation. *Memory & Cognition*, *23*(3), 313–323.

Greene, S., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology*, *18*(2), 266–283.

Grodner, D., Watson, D., & Gibson, E. (2000). Locality effects on sentence processing. *Paper Presented at the 13th CUNY Sentence Processing Conference*.

Grosz, B., Joshi, A., & Weinstein, S. (1995). Towards a computational theory of discourse interpretation. *Computational Linguistics*, *21*, 203–225.

Grosz, J., Joshi, A., & Weinstein, S. (1983). Providing a unified account of definite noun phrases in discourse. In Association for Computational Linguistics (Ed.), *21st Annual Meeting of the Association for Computational Linguistics* (pp. 44–50). Association for Computational Linguistics.

Gundel, J. K. (2010). Reference and accessibility from a Givenness Hierarchy perspective. *International Review of Pragmatics*, *2*, 148–168.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, *69*(2), 274–307.

Gundel, J. K., Hedberg, N., & Zacharski, R. (2012). Underspecification of Cognitive Status in Reference Production : Some Empirical Predictions. *Topics in Cognitive Science*, *4*, 249–268. https://doi.org/10.1111/j.1756-8765.2012.01184.x

Haendler, Y., Kliegl, R., & Adani, F. (2015a). Discourse accessibility constraints in children's processing of object relative clauses. *Frontiers in Psychology*, *6*, 1–17. https://doi.org/10.3389/fpsyg.2015.00860

Haendler, Y., Kliegl, R., & Adani, F. (2015b). The impact of pronoun type and grammatical skills on children's processing of object relative clauses. *BUCLD 39 Proceedings*.

Hale, J. (2001). A probabilistic Earley parser as a psycholinguistic model. *North American Chapter of the Association for Computational Linguistics (NAACL)*, 159–166.

Hammerly, C., Staub, A., & Dillon, B. (2019). The grammaticality asymmetry in agreement attraction reflects response bias: Experimental and modeling evidence. *Cognitive Psychology*, *110*, 70–104.

Hartsuiker, R. J., & Barkhuysen, P. N. (2006). Language production and working memory: The case of subject-verb agreement. *Language and Cognitive Processes*, *21*(181–204).

Haviland, S., & Clark, H. (1974). What's new? Acquiring new information as a process in comprehension. *Journal of Verbal Learning and Verbal Behavior*, *13*, 512–521.

Heim, I. (1991). Artikel und definitheit. In A. Stechow & D. Wunderlich (Eds.), *Semantics: An international handbook of contemporary research* (pp. 487–535). De Gruyter.

Heim, I. R. (1982). *The semantics of definite and indefinite NPs*.

Hofmeister, P. (2007). Retrievability and gradience in filler-gap dependencies. In M. Elliott, J. Kirby, O. Sawada, E. Staraki, & S. Yoon (Eds.), *Proceedings of 43rd Regional Meeting of the Chicago Linguistics Society* (pp. 109–123).

Hofmeister, P. (2011). Representational complexity and memory retrieval in language comprehension. *Language and Cognitive Processes*, *26*(3), 109–123.

Hofmeister, P., & Sag, I. A. (2010). Cognitive constraints and island effects. *Language*, *86*, 366–415.

Hofmeister, P., & Vasishth, S. (2014). Distinctiveness and encoding effects in online sentence comprehension. *Frontiers in Psychology*, *5*.

Holmes, V., & Langford, J. (1976). Comprehension and recall of abstract and concrete sentences. *Journal of Verbal Learning and Verbal Behavior*, *15*, 559–566.

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

Jacoby, L. L., & Brooks, L. R. (1984). Nonanalytic cognition: Memory, perception and concept learning. In G. Bower (Ed.), *Psychology of learning and motivation* (Vol. 18, pp. 1–47). Academic Press.

Jäger, L. A., Engelmann, F., & Vasishth, S. (2015). Retrieval interference in reflexive processing: experimental evidence from Mandarin, and computational modeling. *Frontiers in Psychology*, *6*(May), 1–24. https://doi.org/10.3389/fpsyg.2015.00617

Jäger, L., Chen, Z., Li, Q., Lin, C.-J. C., & Vasishth, S. (2015). The subject-relative advantage in Chinese: Evidence for expectation-based processing. *Journal of Memory and Language*, 97–120.

Jäger, L., Engelmann, F., & Vasishth, S. (2017). Similarity-based interference in sentence comprehension: Literature review and Bayesian meta-analysis. *Journal of Memory and Language*, *94*, 316–339.

Jurafsky, D. (2003). Probabilistic modeling in psycholinguistics: Linguistic comprehension and production. In R. Bod, J. Hay, & S. Jannedy (Eds.), *Probabilistic linguistics* (pp. 39–95). MIT Press.

Just, M. A., Carpenter, P. A., Keller, T. A., Eddy, W. F., & Thulborn, K. R. (1996). Brain activation modulated by sentence comprehension. *Science*, *274*(5284), 114–116.

Kac, M. B. (1981). Center-embedding revisited. *In Proceedings of the Third Annual Conference of the Cognitive Science Society*, 123–124.

Kaiser, E., & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes*, *23*(5), 709–748. https://doi.org/10.1080/01690960701771220

Kalin, L. (2018). Licensing and Differential Object Marking : The View from Neo-Aramaic. *Syntax*, *21*(2), 112–159. https://doi.org/10.1111/synt.12153

Kamp, H. (1981). A theory of truth and semantic representation. In J. Groenendijk, T. Janssen, & U. Stokhof (Eds.), *Formal Methods in the Study of Language* (pp. 277–322). Mathematical Centre.

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

Kempe, V., & MacWhinney, B. (1999). Processing of morphological and semantic cues in Russian and German. *Language and Cognitive Processes*, *14*(2), 129–171.

Kempen, G. (2000). Could grammatical encoding and grammatical decoding be subserved by the same processing module? *Behavioral and Brain Sciences*, *23*, 38–39.

Kidd, E., Brandt, S., Lieven, E., & Tomasello, M. (2007). Object relatives made easy: A cross-linguistic comparison of the constraints influencing young children's processing of relative clauses. *Language and Cognitive Processes*, *22*(6), 860–897.

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

Kirsten, M., Tiemann, S., Seibold, V., Hertrichm I., Beck, S., & Rolke, B. (2014). When the polar bear encounters many polar bears: event-related potential context effects evoked by uniqueness failure. *Language, Cognition and Neuroscience*, *29*(9), 1147–1162.

Kluender, R. (1998). On the distinction between strong and weak islands: a processing perspective. In P. Culicover & L. McNally (Eds.), *Syntax and semantics: The limits of syntax* (Vol. 29, pp. 241–279). Academic Press.

Kluender, R., & Kutas, M. (1993). Subjacency as a processing phenomenon. *Language and Cognitive Processes*, *8*, 573–633.

Koolen, R., Gatt, A., Goudbeek, M., & Krahmer, E. (2011). Factors causing overspecification in definite descriptions. *Journal of Pragmatics*, *43*, 3231–3250.

Kush, D., Johns, C. L., & van Dyke, J. (2015). Identifying the role of phonology in sentence-level reading. *Journal of Memory and Language*, *79*, 18–29.

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in linear mixed effects models. *Journal of Statistical Software*, *82*(13), 1–26.

Kwon, N. (2008). *Processing of Syntactic and Anaphoric Gap-Filler Dependencies in Korean: Evidence from Self-Paced Reading Time, ERP and Eye-Tracking Experiments*.

Kwon, N., Polinsky, M., & Kluender, R. (2006). Subject preference in Korean. *Proceedings of the 25th West Coast Conference on Formal Linguistics (WCCFL 25)*, 1–14.

Kwon, N., & Sturt, P. (2016a). Attraction effects in honorific agreement in Korean. *Frontiers in Psychology*, *7*, 1–13.

Kwon, N., & Sturt, P. (2016b). Facilitatory intrusion effects in subject-verb honorific agreement in Korean. *A Poster Presented at CUNY Sentence Processing*.

Lago, S., Shalom, D., Sigman, M., Lau, E., & Phillips, C. (2015). Agreement attraction in Spanish comprehension. *Journal of Memory and Language*, *82*, 133–149.

Lau, E., Rozanova, K., & Phillips, C. (2007). Syntactic prediction and lexical surface frequency effects in sentence processing. *University of Maryland Working Papers in Linguistics*, *16*, 163–200.

Laurinavichyute, A., & Malsburg, T. von der. (2019). Agreement attraction effects in the comprehension of grammatical sentences. *Presented at 32nd Annual CUNY Conference on Human Sentence Processing*.

Laurinavichyute, A., & Malsburg, T. von der. (2021). Agreement attraction in grammatical sentences does arise, but only in the good-enough processing mode. *Presented at the Workshop on Linguistic Illusion in Sentence Processing*.

Lazard, G. (1998). *Actancy*. Mouton de Gruyter.

Leonetti, M. (2004). Specificity and differential object marking in Spanish. *Catalan Journal of Linguistics*, *3*, 75–114.

Levin, B., & Hovav, M. R. (2007). Reflections on the Complementarity of Manner and Result. *ACADEMIA: Accelerat Ing t He World's Research.*

Levy, R. (2008). Expectation-based syntactic comprehension. *Cognition*, *106*(3), 599–604.

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

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

Löbner, S. (1985). Definites. *Journal of Semantics*, *4*, 279–326.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994a). Lexical nature of syntactic ambiguity resolution. *Psychological Review*, *101*(4), 676–703.

MacDonald, M. C., Pearlmutter, N. J., & Seidenberg, M. S. (1994b). Syntactic ambiguity resolution as lexical ambiguity resolution. In C. Clifton, L. Frazier, & K. Rayner (Eds.), *Perspectives on sentence processing* (pp. 123–154). Hillsdale.

MacDonald, M., & Christiansen, M. (2002). Reassessing working memory: Comment on Just and Carpenter (1992) and Waters and Caplan (1996). *Psychological Review*, *109*(1), 35–54.

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.

MacWhinney, B., & Pleh, C. (1988). The processing of restrictive relative clauses in Hungarian. *Cognition*, *29*, 95–141.

Masharov, M. (2008). *Reference resolution and discourse salience*.

McDonald, J. L., Bock, J. K., & Kelly, M. H. (1993). Word and world order: Semantic, phonological, and metrical determinants of serial position. *Cognitive Psychology*, *25*, 188–230.

McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, *29*(2), 111–123. https://doi.org/10.1023/A:1005184709695

McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, *48*(1), 67–91.

Melinger, A., & Dobel, C. (2005). Lexically-driven syntactic priming. *Cognition*, *98*(B), 11–20.

Miller, G. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, *63*, 81–97.

Mitchell, Don. C., Cuetos, F., Corley, M. M. B., & Brysbaert, M. (1995). Exposure-based models of human parsing: Evidence for the use of coarse-grained (nonlexical) statistical records. *Journal of Psycholinguistic Research*, *24*(6), 469–488.

Momma, S., & Phillips, C. (2018). The relationship between parsing and generation. *Annual Review of Linguistics*, *4*, 233–254.

Momma, S., R, K., LR, S., & Phillips, C. (2017). Timing of syntactic and lexical priming reveals structure-building mechanisms in production. *Presented at 30th Annual CUNY Conference on Human Sentence Processing*.

Moravcsik, E. (1978). Agreement. In *Universals of human language, vol. 4, Syntax* (pp. 331–374). Stanford University Press.

Murphy, G. (1984). Establishing and accessing referents in discourse. *Memory and Cognition*, *12*(5), 489–497.

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

Ness, T., & Meltzer-Asscher, A. (2017). Working memory in the processing of long-distance dependencies: Interference and filler maintenance. *Journal of Psycholinguistic Research*, *46*, 1353–1365.

Ness, T., & Meltzer-Asscher, A. (2019). When is the verb a potential gap site? The influence of filler maintenance on the active search for a gap. *Language, Cognition and Neuroscience*, *34*(7), 936–948.

New, B., Brysbaert, M., Segui, J., Ferrand, L., & Rastle, K. (2004). The processing of singular and plural nouns in French and English. *Journal of Memory and Language*, *51*(568–585).

Nice, K., & Dietrich, R. (2003). *Task-sensitivity of animacy effects: Evidence from German picture descriptions*.

Nicol, J. L., Forster, K. I., & Veres, C. (1997). Subject–verb agreement processes in comprehension. *Journal of Memory and Language*, *36*(4), 569–587.

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

Oberauer, K., & Lange, E. B. (2008). Interference in verbal working memory: distinguishing similarity-based confusion, feature overwriting, and feature migration. *Journal of Memory and Language*, *58*, 730–745.

Osterhout, L., Bersick, M., & McLaughlin, J. (1997). Brain potentials reflect violations of gender stereotypes. *Memory and Cognition*, *25*, 273–285.

Parker, D. (2017). Selective agreement attraction effects: Not all phrases are equally attractive. *Paper Presented at the 30th CUNY Conference on Human Sentence Processing*.

Parker, D., & An, A. (2018). Not all phrases are equally attractive: Experimental evidence for selective agreement attraction effects. *Frontiers in Psychology*, *9*, 1–16.

Parker, D., Lago, S., & Phillips, C. (2015). Interference in the processing of adjunct control. *Frontiers in Psychology*, *6*, 1–13.

Patil, U., Vasishth, S., & Lewis, R. (2016). Retrieval interference in syntactic processing: the case of reflexive binding in English. *Frontiers in Psychology*, *7*, 329.

Pearlmutter, N., Garnsey, S. M., & Bock, K. (1999). Agreement processes in sentence comprehension. *Jornal of Memory and Language*, *41*(3), 427–456.

Pearlmutter, N. J. (2000). Linear versus hierarchical agreement feature processing in comprehension. *Journal of Psycholinguistic Research*, *29*(1), 89–98.

Percus, O. (2006). Antipresuppositions. In A. Ueyama (Ed.), *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science, Japan Society for the Promotion of Science* (pp. 52–73).

Pesetsky, D., & Torrego, E. (2007). The syntax of valuation and the interpretability of features. In S. Karimi, V. Samiian, & W. Wilkins. (Eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation* (pp. 262–294). John Benjamins.

Pickering, M., & Branigan, H. (1998). The representation of verbs: evidence from syntactic priming in language production. *Journal of Memory and Language*, *39*, 633–651.

Prat-Sala, M., & Branigan, H. P. (2000). Discourse constraints on syntactic processing in language production: A cross-linguistic study in English and Spanish. *Journal of Memory and Language*, *42*, 168–182.

Quirk, R., Greenbaum, S., Leech, G., & Svartvik, J. (1972). *A Grammar of Contemporary English*. Longman.

Rawlins, K. (2005). Possessive definites and the definite article. *UCSC Qualifying Paper*.

Reali, F., & Christiansen, M. H. (2007). Processing of relative clauses is made easier by frequency of occurrence. *Journal of Memory and Language*, *57*, 1–23. https://doi.org/10.1016/j.jml.2006.08.014

Richards, N. (2010). *Uttering trees*. MIT Press.

Roark, B., Bachrach, A., Cardenas, C., & Pallier, C. (2009). Deriving lexical and syntactic expectation-based measures for psycholinguistic modeling via incremental top-down parsing. *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, 324–333.

Roberts, C. (2003). Uniqueness in definite noun phrases. *Linguistics and Philosophy*, *26*(3), 287–350.

Rodriguez-Mondonedo, M. (2007). *The syntax of objects: Agree and Differential Object Marking*.

Roland, D., Dick, F., & Elman, J. (2007). Frequency of basic English grammatical structures: A corpus analysis. *Journal of Memory and Language*, *57*, 348–379.

Roland, D., & Jurafsky, D. (2002). Verb sense and verb subcategorization probabilities. In S. Stevenson & P. Merlo (Eds.), *The lexical basis of sentence*

processing: Formal, computational and experimental issues (pp. 325–346). John Benjamins.

Ross, JR. (1967). *Constraints on variables in syntax.*

Russell, B. (1905). On denoting. *Mind*, *14*, 479–493.

Sauerland, U. (2008). Implicated presuppositions. In A. Steube (Ed.), *Sentence and Context* (Vol. 15). Mouton de Gruyter.

Schneider, C., Bade, N., Franke, M., & Janczyk, M. (2021). Presuppositions of determiners are immediately used to disambiguate utterance meaning: A mouse-tracking study on the German language. *Psychological Research*, *85*, 1348–1366.

Schneider, C., Schonard, C., Franke, M., Jäger, G., & Janczyk, M. (2019). Pragmatic processing: An investigation of the (anti-)presuppositions of determiners using mouse-tracking. *Cognition*, *193*, 1–14.

Schriefers, H., Friederici, A. D., & Kuhn, K. (1995). The processing of locally ambiguous relative clauses in German. *Journal of Memory and Language*, *34*(4), 499–520.

Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, *27*(5), 499–520.

Schwanenflugel, P. J., & Shoben, E. J. (1983). Differential context effects in the comprehension of abstract and concrete verbal materials. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *9*(1), 82–102.

Schwarz, F. (2007). Processing presupposed content. *Journal of Semantics*, *4*(24), 373–416.

Schwarz, F. (2009). *Two Types of Definites in Natural Language*.

Schwarz, F., & Tiemann, S. (2012). Presupposition processing - The case of German wieder. In M. Aloni, V. Kimmelmann, F. Roelofsen, G. Sassoon, K. Schulz, & M. Westera (Eds.), *Amsterdam Colloquium* (pp. 200–209). Springer.

Silverstein, M. (1976). Hierarchy of features and ergativity. In R. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112–171). Australian Institute of Aboriginal Studies.

Smith, N., & Levy, R. (2013). The effect of word predictability on reading time is logarithmic. *Cognition*, *128*(3), 302–319.

Spivey-Knowlton, M., & Sedivy, J. (1995). Resolving attachment ambiguities with multiple constraints. *Cognition*, *55*, 227–267.

Staub, A. (2010a). Eye movements and processing difficulty in object relative clauses. *Cognition*, *116*, 71–86. https://doi.org/10.1016/j.cognition.2010.04.002

Staub, A. (2010b). Response time distributional evidence for distinct varieties of number attraction. *Cognition*, *114*, 447–454.

Stoops, A., Luke, S. G., & Christianson, K. (2014). Animacy information outweighs morphological cues in Russian. *Language, Cognition and Neuroscience*, *29*(5), 584–604.

Strawson, P. (1950). On referring. *Mind*, *59*(235), 320–344.

Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, *48*(3), 542–562. https://doi.org/10.1016/S0749-596X(02)00536-3

Sturt, P., & Kwon, N. (2017). Agreement attraction: roles of active dependencies and attractor position. *Paper Presented at the 30th CUNY Conference on Human Sentence Processing.*

Takami, K. (1992). *Preposition stranding: from syntactic to functional analysis.* Mouton de Gruyter.

Tanenhaus, M. K., & Trueswell, J. C. (1995). Sentence comprehension. In J. Miller & P. Eimas (Eds.), *Speech, language, and communication* (Vol. 11, pp. 217–262). Academic Press.

Tarollo, F., & Myhill, J. (1983). Interference and natural language processing in second language acquisition. *Language Learning*, *33*, 55–76.

Tiemann, S. (2014). *The processing of wieder ('again') and other presupposition triggers.*

Tiemann, S., Schmid, M., Bade, N., Rolke, B., Hertrich, I., Ackermann, H., Knapp, J., & Beck, S. (2011). Psycholinguistic evidence for presuppositions: On-line and off-line data. In I. Reich (Ed.), *Proceedings of Sinn & Bedeutung 15* (pp. 581–595).

Torrego, E. (1998). *The dependencies of objects.* MIT Press.

Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object relative clauses: Evidence from eye movements. *Jornal of Memory and Language*, *47*(1), 69–90.

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

Tucker, M. A., & Almeida, D. (2017). The complex structure of agreement errors: evidence from distributional analyses of agreement attraction in Arabic. In A. Lamont & K. Tetzloff (Eds.), *Proceedings of the 47th Meeting of the North East Linguistics Society* (pp. 45–54). GLSA.

Tucker, M. A., Idrissi, A., & Almeida, D. (2015). Representing number in the real-time processing of agreement: self-paced reading evidence from Arabic. *Frontiers in Psychology*, *6*, 347.

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

van Dyke, J. A. (2007). Interference effects from grammatically unavailable constituents during sentence processing. *Journal of Experimental Psychology*, *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 misanalyzed ambiguities. *Journal of Memory and Language*, *49*, 285–316.

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

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

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

van Valin, R. D. Jr. (1995). Toward a functionalist account of so-called extraction constraints. In B. Devriendt, L. Goossens, & J. van der Auwera (Eds.), *Complex structures: A functionalist perspective* (pp. 26–60). Mouton de Gruyter.

Vasishth, S., Chen, Z., Li, Q., & Guo, G. (2013). Processing Chinese Relative Clauses: Evidence for the Subject-Relative Advantage. *PLOS ONE*, *8*(10), 1–15.

Villata, S., & Franck, J. (2020). Similarity-based interference in agreement comprehension and production: Evidence from object agreement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(1), 170–188.

Villata, S., Tabor, W., & Franck, J. (2018). Encoding and retrieval interference in sentence comprehension: evidence from agreement. *Frontiers in Psychology*, *9*(2).

von Heusinger, K. (2007). Accessibility of definite noun phrases. In M. Schwarz-Friesel, M. Consten, & M. Knees (Eds.), *Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference* (pp. 123–144). John Benjamins.

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

Walker, M., Joshi, A., & Prince, E. (1998). *Centering in discourse*. Clarendon Press.

Warren, T., & Gibson, E. (1999). The effects of discourse status on intuitive complexity: Implications for quantifying distance in a locality-based theory of linguistic complexity. *Poster Presented at the 12th CUNT Sentence Processing Conference*.

Warren, T., & Gibson, E. (2002). The influence of referential processing on sentence complexity. *Cognition*, *85*, 79–112.

Warren, T., & Gibson, E. (2005). Effects of NP type in reading cleft sentences in English. *Language and Cognitive Processes*, *20*(6), 751–767. https://doi.org/10.1080/01690960500051055

Waters, G. S., Caplan, D., & Rochcon, E. (1995). Processing capacity and sentence comprehension in patients with Alzheimer's disease. *Cognitive Neuropsychology*, *12*, 1–30.

Xiang, M., Dillon, B., & Phillips, C. (2009). Illusory licensing effects across dependency types: ERP evidence. *Brain and Language*, *108*(1), 40–55. https://doi.org/10.1016/j.bandl.2008.10.002

Xiang, M., Grove, J., & Giannakidow, A. (2013). Dependency dependent interference: NPI interference, agreement attraction, and global pragmatic inferences. *Frontiers in Psychology*, *4*(708).

Appendix A

ITEM SETS FOR EXPERIMENTS

Item sets for NP types in null contexts: Experiment 1

1. It was {the physicist/a physicist} who {the dean/a dean/ I} praised at the meeting.
What type of scientists was praised?
a. physicist   b. biologist   c. chemist

2. It was {the coach/a coach} who {the referee/a referee/you} criticized after the game.
What was the profession of the person who was criticized?
a. sportsperson   b. coach   c. magician

3. It was {the banker/a banker} who {the lobbyist/a lobbyist/you} contacted in the morning.
What was the profession of the person who was contacted?
a. customer   b. manager   c. banker

4. It was {the dancer/a dancer} who {the reporter/a reporter/I} phoned on New Year's Eve.
What was the profession of the person who was phoned?
a. dancer   b. singer   c. athlete

5. It was {the architect/an architect} who {the fireman/a fireman/we} liked before the argument began.
What was the profession of the person who was liked?
a. painter   b. architect   c. director

6. It was {the detective/a detective} who {the secretary/a secretary/I} disliked during card games.
What was the profession of the person who was disliked?
a. fireman   b. officer   c. detective

7. It was {the politician/a politician} who {the mailman/a mailman /you} insulted after reading the newspaper article.
What was the profession of the person who was insulted?
a. politician   b. secretary   c. policeman

8. It was {the actor/an actor} who {the director/a director/we} thanked before the show.
What was the profession of the person who was thanked?
a. director   b. actorc. writer

9. It was {the chef/a chef} who {the cashier/a cashier/I} distrusted after the restaurant closed.
What was the profession of the person who was distrusted?
a. customer   b. client   c. chef

10. It was {the violinist/a violinist} who {the conductor/a conductor/you} complimented at Carnegie Hall.
What type of musicians was complimented?
a. violinist   b. pianist   c. cellist

11. It was {the tutor/a tutor} who {the student/a student/I} questioned during summer vacation.
What was the profession of the person who was questioned?
a. editor   b. tutor   c. writer

12. It was {the editor/an editor} who {the author/an author/we} recommended after a new merger was announced.
What was the profession of the person who was recommended?
a. poet   b. novelist   c. editor

13. It was {the tailor/a tailor} who {the customer/a customer/you} described at the banquet.
What was the profession of the person who was described?
a. tailor   b. lawyer   c. physician

14. It was {the admiral/an admiral} who {the general/a general/I} advised before the trip got underway.
What was the profession of the person who was advised?
a. policeman   b. admiral   c. pilot

15. It was {the lawyer/a lawyer} who {the auditor/an auditor /you} interviewed in the very small office.
What was the profession of the person who was interviewed?
a. accountant   b. inspector   c. lawyer

16. It was {the plumber/a plumber} who {the electrician/an electrician/I} called from the payphone.

What was the profession of the person who was called?
a. plumber   b. analyst   c. engineer

17. It was {the clown/a clown} who {the magician/a magician/you} entertained in the auditorium.
What was the profession of the person who was entertained?
a. counselor   b. clown   c. technician

18. It was {the gardener/a gardener} who {the homeowner/a homeowner/we} envied after the lottery ended.
What was the profession of the person who was envied?
a. dentist   b. plumber   c. gardener

19. It was the poet/a poet that the painter/a painter/you inspired outside the coffeeshop.
What was the profession of the person who was inspired?
a. poet   b. artist   c. novelist

20. It was {the lawyer/a lawyer} who {the businessman/a businessman/we} avoided at the party.
What was the profession of the person who was avoided?
a. accountant   b. lawyer   c. administrator

21. It was {the fisherman/a fisherman} who {the clerk/a clerk/I} saw at the store.
What was the profession of the person who was seen?
a. farmer   b. salesman   c. fisherman

22. It was {the director/a director} who {the actress/an actress/you} impressed at the audition.
What was the profession of the person who was impressed?
a. director   b. actor   c. singer

23. It was {the tenant/a tenant} who {the landlord/a landlord/we} sued for a lot of money.
What was the profession of the person who was sued?
a. agent   b. tenant   c. housemate

24. It was {the secretary/a secretary} who {the editor/an editor/you} hired some time last week.
What was the profession of the person who was hired?
a. author   b. president   c. secretary

Item sets for definites in contexts: Experiment 2

[1]
Felicitous condition
John, an actor, and Matt, a director, were at the coffee shop. They met Andy, a writer, who graciously thanked John for his incredible performance in the latest movie.

It is the actor who the writer graciously thanked for the incredible performance.
Is the last statement true or false?

It is the actor who a writer graciously thanked for the incredible performance.
Is the last statement true or false?

Presupposition failure condition
John, an actor, and Matt, a writer, were at the coffee shop. They met Andy, another writer, who graciously thanked John for his incredible performance in the latest movie.

It is the actor who the writer graciously thanked for the incredible performance.
Is the last statement true or false?

It is the actor who a writer graciously thanked for the incredible performance.
Is the last statement true or false?

[2]
Felicitous condition
Paul, a physicist, and Mary, a psychologist, presented their collaborative grant proposal to Helen, a linguist, who effusively praised Mary for her excellent presentation after the meeting.

It was the psychologist who the linguist effusively praised after the meeting.
Is the last statement true or false?

It was the psychologist who a linguist effusively praised after the meeting.
Is the last statement true or false?

Presupposition failure condition
Paul, a linguist, and Mary, a psychologist, presented their collaborative grant proposal to Helen, another linguist, who effusively praised Mary for her excellent presentation after the meeting.

It was the psychologist who the linguist effusively praised after the meeting.
Is the last statement true or false?

It was the psychologist who a linguist effusively praised after the meeting.
Is the last statement true or false?

[3]
Felicitous condition
Bob, a chef, and Sydney, an accountant, ran a restaurant. They worked with Tom, a manager, who unwillingly distrusted Bob after the restaurant closed.

It was the chef who the manager unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

It was the chef who a manager unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

Presupposition failure condition
Bob, a chef, and Sydney, a manager, ran a restaurant. They worked with Tom, another manager, who unwillingly distrusted Bob after the restaurant closed.

It was the chef who the manager unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

It was the chef who a manager unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

[4]
Felicitous condition
Maria, a janitor, and John, a teacher, were talking in the store. They didn't notice Linda, a student, who immediately recognized John when she entered the store.

It was the teacher who the student immediately recognized at the store.
Is the last statement true or false?

It was the teacher who a student immediately recognized at the store.
Is the last statement true or false?

Presupposition failure condition
Maria, a student, and John, a teacher, were talking in the store. They didn't notice

Linda, another student, who immediately recognized John when she entered the store.

It was the teacher who the student immediately recognized at the store.
Is the last statement true or false?

It was the teacher who a student immediately recognized at the store.
Is the last statement true or false?

[5]
Felicitous condition
Joe, a lawyer, and Mike, a paralegal, took the same elevator in the office building. They arrived at the office right after Ann, an auditor, who unsuspectingly interviewed Joe in the very small office.

It was the lawyer who the auditor unsuspectingly interviewed in the very small office.
Is the last statement true or false?

It was the lawyer who an auditor unsuspectingly interviewed in the very small office.
Is the last statement true or false?

Presupposition failure condition
Joe, a lawyer, and Mike, an auditor, took the same elevator in the office building. They arrived at the office right after Ann, another auditor, who unsuspectingly interviewed Joe in the very small office.

It was the lawyer who the auditor unsuspectingly interviewed in the very small office.
Is the last statement true or false?

It was the lawyer who an auditor unsuspectingly interviewed in the very small office.
Is the last statement true or false?

[6]
Felicitous condition
Tracy, a masseuse, and Emily, a hairdresser, planned to open a salon together. They also invited Kim, a manicurist, who genuinely impressed Emily for her attentiveness.

It was the hairdresser who the manicurist genuinely impressed for her attentiveness.
Is the last statement true or false?

It was the hairdresser who a manicurist genuinely impressed for her attentiveness.
Is the last statement true or false?

Presupposition failure condition
Tracy, a manicurist, and Emily, a hairdresser, and planned to open a salon together. They also invited Kim, another manicurist, who genuinely impressed Emily for her attentiveness.

It was the hairdresser who the manicurist genuinely impressed for her attentiveness.
Is the last statement true or false?

It was the hairdresser who a manicurist genuinely impressed for her attentiveness.
Is the last statement true or false?

[7]
Felicitous condition
Tony, a violinist, and Mike, a pianist, performed their own parts in front of Dorothy, a cellist, who effusively complimented Tony for his excellent form at Carnegie Hall.

It was the violinist who the cellist effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

It was the violinist who a cellist effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

Presupposition failure condition
Tony, a violinist, and Mike, a cellist, performed their own parts in front of Dorothy, also a cellist, who effusively complimented Tony for his excellent form at Carnegie Hall.

It was the violinist who the cellist effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

It was the violinist who a cellist effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

[8]
Felicitous condition
Amanda, a guitarist, and Isabelle, a dancer, performed on a street on Christmas day. They impressed Henry, a singer, who apparently called Isabelle the next day to ask for help with his performance.

It was the dancer who the singer apparently called next day.

Is the last statement true or false?

It was the dancer who a singer apparently called next day.
Is the last statement true or false?

Presupposition failure condition
Amanda, a singer, and Isabelle, a dancer, performed on a street on Christmas day.
They impressed Henry, another singer, who apparently called Isabelle the next day to
ask for help with his performance.

It was the dancer who the singer apparently called next day.
Is the last statement true or false?

It was the dancer who a singer apparently called next day.
Is the last statement true or false?

[9]
Felicitous condition
George, a referee, and Megan, a soccer player, attended a swimming competition.
After the game, they saw Andrew, a coach, who shamefully criticized George for
being unfair during the game.

It was the referee who the coach shamefully criticized after the game.
Is the last statement true or false?

It was the referee who a coach shamefully criticized after the game.
Is the last statement true or false?

Presupposition failure condition
George, a referee, and Megan, a coach, attended a swimming competition. After the
game, they saw Andrew, another coach, who shamefully criticized George for being
unfair during the game.

It was the referee who the coach shamefully criticized after the game.
Is the last statement true or false?

It was the referee who a coach shamefully criticized after the game.
Is the last statement true or false?

[10]
Felicitous condition

Mike, a bodyguard, and Jeff, a detective, were playing cards with Tiffany, a secretary, who openly blamed Jeff for his unfair play during card games.

It was the detective who the secretary openly blamed during card games.
Is the last statement true or false?

It was the detective who a secretary openly blamed during card games.
Is the last statement true or false?

Presupposition failure condition
Sydney, a secretary, and Jeff, a detective, were playing cards with Tiffany, another secretary, who openly blamed Jeff for his unfair play during card games.

It was the detective who the secretary openly blamed during card games.
Is the last statement true or false?

It was the detective who a secretary openly blamed during card games.
Is the last statement true or false?

[11]
Felicitous condition
Chris, a biologist, and Tracy, a chemist, were invited to give lectures at a conference, where they met Kim, a psychologist, who allegedly offended Chris with racist remarks.

It was the biologist who the psychologist allegedly offended with racist remarks.
Is the last statement true or false?

It was the biologist who a psychologist allegedly offended with racist remarks.
Is the last statement true or false?

Presupposition failure condition
Chris, a biologist, and Tracy, a psychologist, were invited to give lectures at a conference, where they met Kim, another psychologist, who allegedly offended Chris with racist remarks.

It was the biologist who the psychologist allegedly offended with racist remarks.
Is the last statement true or false?

It was the biologist who a psychologist allegedly offended with racist remarks.
Is the last statement true or false?

[12]
Felicitous condition
Tom, a stage manager, and Anne, a director, decided to have an audition for a major role in a new film. They finally picked Robert, an actor, who apparently impressed Anne at the audition.

It was the director who the actor apparently *impressed at the audition.*
Is the last statement true or false?

It was the director who an actor apparently *impressed at the audition.*
Is the last statement true or false?

Presupposition failure condition
Tom, an actor, and Anne, a director, decided to have an audition for a major role in a new film. They finally picked Robert, another actor, who apparently impressed Anne at the audition.

It was the director who the actor apparently *impressed at the audition.*
Is the last statement true or false?

It was the director who an actor apparently *impressed at the audition.*
Is the last statement true or false?

[13]
Felicitous condition
Mark, a poet, and William, a musician, were at a coffee shop. They ran into Catherine, a sculptor, who curiously ignored Mark when he visited her exhibition.

It was the poet that the sculptor curiously ignored at the exhibition.
Is the last statement true or false?

It was the poet that a sculptor curiously ignored at the exhibition.
Is the last statement true or false?

Presupposition failure condition
Mark, a poet, and William, a sculptor, were at a coffee shop and met Catherine, another sculptor, who curiously ignored Mark when he visited her exhibition.

It was the poet that the sculptor curiously ignored at the exhibition.
Is the last statement true or false?

It was the poet that a sculptor curiously ignored at the exhibition.
Is the last statement true or false?

[14]
Felicitous condition
Tracy, a poet, and Chris, a composer, were discussing a possible collaboration. They consulted Kim, a novelist, who apparently recommended Chris for a prestigious award.

It was the composer who the novelist apparently recommended for a prestigious award.
Is the last statement true or false?

It was the composer who a novelist apparently recommended for a prestigious award.
Is the last statement true or false?

Presupposition failure condition
Tracy, a novelist, and Chris, a composer, were discussing a possible collaboration. They consulted Kim, another novelist, who apparently recommended him for a prestigious award.

It was the composer who the novelist apparently recommended for a prestigious award.
Is the last statement true or false?

It was the composer who a novelist apparently recommended for a prestigious award.
Is the last statement true or false?

[15]
Felicitous condition
Justin, a lawyer, and Peter, a judge, were drunk at a party. They invited Wendy, a secretary, who tried to carefully avoid Justin at work after the party.

It was the lawyer who the secretary carefully avoided at the party.
Is the last statement true or false?

It was the lawyer who a secretary carefully avoided at the party.
Is the last statement true or false?

Presupposition failure condition

Justin, a lawyer, and Peter, a secretary, were drunk at a party. They invited Wendy, another secretary, who tried to carefully avoid Justin at work after the party.

It was the lawyer who the secretary carefully avoided at the party.
Is the last statement true or false?

It was the lawyer who a secretary carefully avoided at the party.
Is the last statement true or false?

[16]
Felicitous condition
Phillip, a poet, and Lydia, a novelist, were discussing a project, but it did not go well. So they decided to get advice from Rachel, an editor, who finally hired Lydia last week.

It was the novelist who the editor finally hired last week.
Is the last statement true or false?

It was the novelist who an editor finally hired last week.
Is the last statement true or false?

Presupposition failure condition
Phillip, an editor, and Lydia, a novelist, were discussing a project, but it did not go well. So they decided to get advice from Rachel, another editor, who finally hired Lydia last week.

It was the novelist who the editor finally hired last week.
Is the last statement true or false?

It was the novelist who an editor finally hired last week.
Is the last statement true or false?

Item sets for pronouns in contexts: Experiment 3

[1]
High accommodation cost
Jane, a writer, and Alice another writer, were at the coffee shop. One of them graciously thanked Matt, an actor, for his incredible performance in the latest movie.
It is the actor who a writer graciously thanked for the incredible performance.
Is the last statement true or false?

It is the actor who she graciously thanked for the incredible performance.
Is the last statement true or false?

Low accommodation cost
Jane, a writer, and Alice, another writer, were at the coffee shop. Jane graciously thanked Matt, an actor, for his incredible performance in the latest movie.
It is the actor who a writer graciously thanked for the incredible performance.
Is the last statement true or false?

It is the actor who she graciously thanked for the incredible performance.
Is the last statement true or false?

[2]
High accommodation cost
Paul, a psychologist, and Mike, another psychologist, attended the presentation of Helen, a linguist, for a grant proposal. One of the psychologists effusively praised her excellent proposal after the meeting.
It was the linguist who a psychologist effusively praised after the meeting.
Is the last statement true or false?

It was the linguist who he effusively praised after the meeting.
Is the last statement true or false?

Low accommodation cost
Paul, a psychologist, and Mike, another psychologist, attended the presentation of Helen, a linguist, for a grant proposal. Mike effusively praised her excellent proposal after the meeting.
It was the linguist who a psychologist effusively praised after the meeting.

Is the last statement true or false?

It was the linguist who he effusively praised after the meeting.
Is the last statement true or false?

[3]
High accommodation cost
Britney, a manager, and Megan another manager, worked at a restaurant. One of them unwillingly distrusted Tom, an accountant, after the restaurant closed.
It was the accountant who a manager unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

It was the accountant who she unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

Low accommodation cost
Britney, a manager, and Megan, another manager, worked at a restaurant. Britney unwillingly distrusted Tom, an accountant, after the restaurant closed.
It was the accountant who a manager unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

It was the accountant who she unwillingly distrusted after the restaurant closed.
Is the last statement true or false?

[4]
High accommodation cost
Maria, a student, and Linda, another student, were talking in the store. One of them immediately recognized John, a teacher, when he entered the store.
It was the teacher who a student immediately recognized at the store.
Is the last statement true or false?

It was the teacher who she immediately recognized at the store.
Is the last statement true or false?

Low accommodation cost
Maria, a student, and Linda, another student, were talking in the store. Linda immediately recognized John, a teacher, when he entered the store.
It was the teacher who a student immediately recognized at the store.
Is the last statement true or false?

It was the teacher who she immediately recognized at the store.
Is the last statement true or false?

[5]
High accommodation cost
Joe, an auditor, and Mike, another auditor, took the same elevator with Ann, a lawyer, in the office building. One of the auditors unsuspectingly interviewed Ann in the very small office.
It was the lawyer who an auditor unsuspectingly interviewed in the very small office.
Is the last statement true or false?

It was the lawyer who he unsuspectingly interviewed in the very small office.
Is the last statement true or false?

Low accommodation cost
Joe, an auditor, and Mike, another auditor, took the same elevator with Ann, a lawyer, in the office building. Joe unsuspectingly interviewed Ann in the very small office.
It was the lawyer who an auditor unsuspectingly interviewed in the very small office.
Is the last statement true or false?

It was the lawyer who he unsuspectingly interviewed in the very small office.
Is the last statement true or false?

[6]
High accommodation cost
Peter, a hairdresser, and James, another hairdresser, invited Kim, a manicurist, to discuss their plan on opening a salon together. One of the hairdressers genuinely impressed Kim for his attentiveness.
It was the manicurist who a hairdresser genuinely impressed for his attentiveness.
Is the last statement true or false?

It was the hairdresser who he genuinely impressed for his attentiveness.
Is the last statement true or false?

Low accommodation cost
Peter, a hairdresser, and James, another hairdresser, invited Kim, a manicurist, to discuss their plan on opening a salon together. James genuinely impressed Kim for his attentiveness.
It was the manicurist who a hairdresser genuinely impressed for his attentiveness.
Is the last statement true or false?

It was the hairdresser who he genuinely impressed for his attentiveness.
Is the last statement true or false?

[7]
High accommodation cost
Tony, a pianist, and Mike, another pianist, went to Carnegie Hall to see the performance of Dorothy, a cellist. One of the pianists effusively complimented Dorothy for her excellent form.
It was the cellist who a pianist effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

It was the cellist who he effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

Low accommodation cost
Tony, a pianist, and Mike, another pianist, went to Carnegie Hall to see the performance of Dorothy, a cellist. Tony effusively complimented Dorothy for her excellent form.
It was the cellist who a pianist effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

It was the cellist who he effusively *complimented at Carnegie Hall.*
Is the last statement true or false?

Is the last statement true or false?

[8]
High accommodation cost
Amanda, a guitarist, and Isabelle, another guitarist, saw Henry, a singer, performing on a street. One of the guitarists apparently called Henry the next day to ask for help with their performance.
It was the singer who a guitarist apparently called next day.
Is the last statement true or false?

It was the singer who she apparently called next day.
Is the last statement true or false?

Low accommodation cost

Amanda, a guitarist, and Isabelle, another guitarist, saw Henry, a singer, performing on a street. Isabelle apparently called Henry the next day to ask for help with their performance.
It was the singer who a guitarist apparently called next day.
Is the last statement true or false?

It was the singer who she apparently called next day.
Is the last statement true or false?

[9]
High accommodation cost
George, a swimmer, and Matt, another swimmer, attended a swimming competition. After the game, one of them shamefully criticized Jessica, a referee, for being unfair during the game.
It was the referee who a swimmer shamefully criticized after the game.
Is the last statement true or false?

It was the referee who he shamefully criticized after the game.
Is the last statement true or false?

Low accommodation cost
George, a swimmer, and Matt, another swimmer, attended a swimming competition. After the game, George shamefully criticized Jessica, a referee, for being unfair during the game.
It was the referee who a swimmer shamefully criticized after the game.
Is the last statement true or false?

It was the referee who he shamefully criticized after the game.
Is the last statement true or false?

[10]
High accommodation cost
Mary, a detective, and Tiffany, another detective, were playing cards with Jeff, a bodyguard.
One of the detectives openly blamed Jeff for his unfair play during card games.
It was the bodyguard who a detective openly blamed during card games.
Is the last statement true or false?

It was the bodyguard who she openly blamed during card games.
Is the last statement true or false?

Low accommodation cost
Mary, a detective, and Tiffany, another detective, were playing cards with Jeff, a bodyguard.
Tiffany openly blamed Jeff for his unfair play during card games.
It was the bodyguard who a detective openly blamed during card games.
Is the last statement true or false?

It was the bodyguard who she openly blamed during card games.
Is the last statement true or false?

[11]
High accommodation cost
Chris, a biologist, and Tom, another biologist, were invited to give lectures at a conference. One of them allegedly offended Kim, a psychologist, with offensive remarks.
It was the psychologist who a biologist allegedly offended with offensive remarks.
Is the last statement true or false?

It was the psychologist who he allegedly offended with offensive remarks.
Is the last statement true or false?

Low accommodation cost
Chris, a biologist, and Tom, another biologist, were invited to give lectures at a conference. Chris allegedly offended Kim, a psychologist, with offensive remarks.
It was the psychologist who a biologist allegedly offended with offensive remarks.
Is the last statement true or false?

It was the psychologist who he allegedly offended with offensive remarks.
Is the last statement true or false?

[12]
High accommodation cost
Tina, a singer, and Anne, another singer, attended an audition for a main singer in a new girl band. One of them apparently impressed Robert, a producer, at the audition.
It was the producer who a singer apparently *impressed at the audition.*
Is the last statement true or false?

It was the producer who she apparently *impressed at the audition.*
Is the last statement true or false?

Low accommodation cost

Tina, a singer, and Anne, another singer, attended an audition for a main singer in a new girl band. Anne apparently impressed Robert, a producer, at the audition.
It was the producer who a singer apparently *impressed at the audition.*
Is the last statement true or false?

It was the producer who she apparently *impressed at the audition.*
Is the last statement true or false?
[13]
High accommodation cost
Mark, a sculptor, and William, another sculptor, invited Catherine, a painter, to their exhibition this week. One of the sculptors curiously ignored Catherine when she visited their exhibition.
It was the painter who a sculptor curiously ignored at the exhibition.
Is the last statement true or false?

It was the painter who he curiously ignored at the exhibition.
Is the last statement true or false?

Low accommodation cost
Mark, a sculptor, and William, another sculptor, invited Catherine, a painter, to their exhibition this week. Mark curiously ignored Catherine when she visited their exhibition.
It was the painter who a sculptor curiously ignored at the exhibition.
Is the last statement true or false?

It was the painter who he curiously ignored at the exhibition.
Is the last statement true or false?

[14]
High accommodation cost
Anna, a poet, and Kim, another poet, were discussing candidates for New Writer Prize this year. One of them apparently recommended Chris, a novelist, for the prestigious award.
It was the novelist who a poet apparently recommended for a prestigious award.
Is the last statement true or false?

It was the novelist who she apparently recommended for a prestigious award.
Is the last statement true or false?

Low accommodation cost
Anna, a poet, and Kim, another poet, were discussing candidates for New Writer Prize this year. Kim apparently recommended Chris, a novelist, for the prestigious award.

It was the novelist who a poet apparently recommended for a prestigious award.
Is the last statement true or false?

It was the novelist who she apparently recommended for a prestigious award.
Is the last statement true or false?

[15]
High accommodation cost
Lydia, a secretary, and Rachel, another secretary, went to a party and ran into Justin, a lawyer, who work with them. One of the secretaries carefully avoided Justin at work after the party.
It was the lawyer who a secretary carefully avoided at the party.
Is the last statement true or false?

It was the lawyer who she carefully avoided at the party.
Is the last statement true or false?

Low accommodation cost
Lydia, a secretary, and Rachel, another secretary, went to a party and ran into Justin, a lawyer, who work with them. Lydia carefully avoided Justin at work after the party.
It was the lawyer who a secretary carefully avoided at the party.
Is the last statement true or false?

It was the lawyer who she carefully avoided at the party.
Is the last statement true or false?

[16]
High accommodation cost
Phillip, a chemist, and Justin, another chemist, were planning to hire a new member for a new project. One of them finally decided to hire Lydia, a biologist, last week.
It was the biologist who a chemist finally decided to hire last week.
Is the last statement true or false?

It was the biologist who he finally decided to hire last week.
Is the last statement true or false?

Low accommodation cost
Phillip, a chemist, and Justin, another chemist, were planning to hire a new member for a new project. Justin finally decided to hire Lydia, a biologist, last week.
It was the biologist who a chemist finally decided to hire last week.
Is the last statement true or false?

It was the biologist who he finally decided to hire last week.
Is the last statement true or false?

Item sets for suject-verb agreement dependency: Experiments 4, 5
a: Object condition-singular attractor   b: Object condition-plural attractor
c: Oblique condition-singular attractor   d: Oblique condition-plural attractor
e: Long-oblique condition-singular attractor   f: Long-oblique condition-plural attractor

1a. The fan who the musician met certainly {was/were} displeased with the horrible performance.
1b. The fan who the musicians met certainly {was/were} displeased with the horrible performance.
1c. The fan who the musician met with certainly {was/were} displeased with the horrible performance.
1d. The fan who the musicians met with certainly {was/were} displeased with the horrible performance.
1e. The fan who the musician met with on Monday certainly {was/were} displeased with the horrible performance.
1f. The fan who the musicians met with on Monday certainly {was/were} displeased with the horrible performance.

What was the profession of the person who met the fan?
What was the profession of the people who met the fan?

a. journalist   b. dancer   c. musician

2. a. The photographer who the manager sent obviously {was/were} excited to take pictures of the animals.
2b. The photographer who the managers sent obviously {was/were} excited to take pictures of the animals.
2c. The photographer who the manager sent for obviously {was/were} excited to take pictures of the animals.
2d. The photographer who the managers sent for obviously {was/were} excited to take pictures of the animals.
2e. The photographer who the manager sent for on Tuesday obviously {was/were} excited to take pictures of the animals

2f. The photographer who the managers sent for on Tuesday obviously {was/were} excited to take pictures of the animals

What was the profession of the person who organized the photographer?
What was the profession of the people who organized the photographer?

a. director   b. manager   c. model

3a. The doctor who the nurse called clearly {was/were} nervous about the stressful situation.
3b. The doctor who the nurses called clearly {was/were} nervous about the stressful situation.
3c. The doctor who the nurse called for clearly {was/were} nervous about the stressful situation.
3d. The doctor who the nurses called for clearly {was/were} nervous about the stressful situation.
3e. The doctor who the nurse called for on Wednesday clearly {was/were} nervous about the stressful situation.
3f. The doctor who the nurses called for on Wednesday clearly {was/were} nervous about the stressful situation.

What was the profession of the person who requested the doctor?
What was the profession of the people who requested the doctor?
a. lab manager   b. researcher   c. nurse

4a. The thief that the cop chased finally {was/were} caught and arrested in New Jersey.
4b. The thief that the cops chased finally {was/were} caught and arrested in New Jersey.
4c. The thief that the cop chased after finally {was/were} caught and arrested in New Jersey.
4d. The thief that the cops chased after finally {was/were} caught and arrested in New Jersey.
4e. The thief that the cop chased after in New York finally {was/were} caught and arrested in New Jersey.
4f. The thief that the cops chased after in New York finally {was/were} caught and arrested in New Jersey.

What was the profession of the person who chased the thief?
What was the profession of the people who chased the thief?

a. bodyguard   b. soldier   c. cop

5a. The neighborhood which the cop patrolled evidently {was/were} surprised by a stranger.
5b. The neighborhood which the cops patrolled evidently {was/were} surprised by a stranger.
5c. The neighborhood which the cop patrolled through evidently {was/were} surprised by a stranger.
5d. The neighborhood which the cops patrolled through evidently {was/were} surprised by a stranger.
5e. The neighborhood which the cop patrolled through daily evidently {was/were} surprised by a stranger.
5f. The neighborhood which the cops patrolled through daily evidently {was/were} surprised by a stranger.

What was the profession of the person patrolling the village?
What was the profession of the people patrolling the village?

a. cop b. secretary   c. engineer

6a. The dog who the worker walked clearly {was/were} happy to be outside in the sunny weather.
6b. The dog who the workers walked clearly {was/were} happy to be outside in the sunny weather.
6c. The dog who the worker walked toward clearly {was/were} happy to be outside in the sunny weather.
6d. The dog who the workers walked toward clearly {was/were} happy to be outside in the sunny weather.
6e. The dog who the worker walked toward in the park clearly {was/were} happy to be outside in the sunny weather.
6f. The dog who the workers walked toward in the park clearly {was/were} happy to be outside in the sunny weather.

What was the status of the person who interacted with the dog?
What was the status of the people who interacted with the dog?

a. detective   b. student   c. worker

7a. The new cook who the chef picked certainly {was/were} the best with Mexican food.

7b. The new cook who the chefs picked certainly {was/were} the best with Mexican food.
7c. The new cook who the chef picked on certainly {was/were} the best with Mexican food.
7d. The new cook who the chefs picked on certainly {was/were} the best with Mexican food.
7e. The new cook who the chef picked on at work certainly {was/were} the best with Mexican food.
7f. The new cook who the chefs picked on at work certainly {was/were} the best with Mexican food.

What was the job of the person who the cook was working with?
What was the job of the people who the cook was working with?

a. chef   b. doctor   c. staff

8a. The referee who the soccer player kicked obviously {was/were} disappointed about fights during the game.
8b. The referee who the soccer players kicked obviously {was/were} disappointed about fights during the game.
8c. The referee who the soccer player kicked at obviously {was/were} disappointed about fights during the game.
8d. The referee who the soccer players kicked at obviously {was/were} disappointed about fights during the game.
8e. The referee who the soccer player kicked at on Saturday obviously {was/were} disappointed about fights during the game.
8f. The referee who the soccer players kicked at on Saturday obviously {was/were} disappointed about fights during the game.

What type of sportsperson was involved in this incident?

a. soccer player   b. tennis player   c. swimmer

9a. The soldier who the worker fought likely {was/were} surrounded by a team of insurgents.
9b. The soldier who the workers fought likely {was/were} surrounded by a team of insurgents.
9c. The soldier who the worker fought against likely {was/were} surrounded by a team of insurgents.
9d. The soldier who the workers fought against likely {was/were} surrounded by a team of insurgents.

9e. The soldier who the worker fought against on Thursday likely {was/were} surrounded by a team of insurgents.
9f. The soldier who the workers fought against on Thursday likely {was/were} surrounded by a team of insurgents.

What was the status of the person who combatted the soldiers?
What was the status of the people who combatted the soldiers?

a. fighter   b. worker   c. policeman

10a. The protestor who the lawyer met undoubtedly {was/were} upset about the bad news.
10b. The protestor who the lawyers met undoubtedly {was/were} upset about the bad news.
10c. The protestor who the lawyer met with undoubtedly {was/were} upset about the bad news.
10d. The protestor who the lawyers met with undoubtedly {was/were} upset about the bad news.
10e. The protestor who the lawyer met with on Monday undoubtedly {was/were} upset about the bad news.
10f. The protestor who the lawyers met with on Monday undoubtedly {was/were} upset about the bad news.

What was the profession of the person who met the protester?
What was the profession of the people who met the protester?

a. reporter   b. lawyer   c. journalist

11a. The girl who the waitress called unsurprisingly {was/were} unhappy about all the noise.
11b. The girl who the waitresses called unsurprisingly {was/were} unhappy about all the noise.
11c. The girl who the waitress called at unsurprisingly {was/were} unhappy about all the noise.
11d. The girl who the waitresses called at unsurprisingly {was/were} unhappy about all the noise.
11e. The girl who the waitress called at in Wendy's unsurprisingly {was/were} unhappy about all the noise.
11f. The girl who the waitresses called at in Wendy's unsurprisingly {was/were} unhappy about all the noise.

What was the profession of the person who was calling?
What was the profession of the people who were calling?

a. waitress   b. waiter   c. manager

12a. The sibling who the sister pushed definitely {was/were} excited to swing higher and higher.
12b. The sibling who the sisters pushed definitely {was/were} excited to swing higher and higher.
12c. The sibling who the sister pushed back definitely {was/were} excited to swing higher and higher.
12d. The sibling who the sisters pushed back definitely {was/were} excited to swing higher and higher.
12e. The sibling who the sister pushed back on the swing {was/were} excited to swing higher and higher.
12f. The sibling who the sisters pushed back on the swing {was/were} excited to swing higher and higher.

What was the role of the person who pushed the siblings?
What was the role of the people who pushed the siblings?
a. brother   b. cousin   c. sister

13a. The professor who the student called likely {was/were} impressed by the finished work.
13b. The professor who the students called likely {was/were} impressed by the finished work.
13c. The professor who the student called on likely {was/were} impressed by the finished work.
13d. The professor who the students called on likely {was/were} impressed by the finished work.
13e. The professor who the student called on at the weekend likely {was/were} impressed by the finished work
13f. The professor who the students called on at the weekend likely {was/were} impressed by the finished work

What was the status of the person who finished the work?
What was the status of the people who finished the work?
a. student   b. researcher   c. journalist

14a. The preacher who the worship chair arranged somehow {was/were} uncomfortable preaching to empty pews.

14b. The preacher who the worship chairs arranged somehow {was/were} uncomfortable preaching to empty pews.
14c. The preacher who the worship chair arranged for somehow {was/were} uncomfortable preaching to empty pews.
14d. The preacher who the worship chairs arranged for somehow {was/were} uncomfortable preaching to empty pews.
14e. The preacher who the worship chair arranged for on Sunday somehow {was/were} uncomfortable preaching to empty pews.
14f. The preacher who the worship chairs arranged for on Sunday somehow {was/were} uncomfortable preaching to empty pews.

What was the status of the person who arranged the preacher?
What was the status of the people who arranged the preacher?

a. priest   b. worship chair   c. minister

15a. The lamp which the student hid definitely {was/were} very tall compared with others.
15b. The lamp which the students hid definitely {was/were} very tall compared with others.
15c. The lamp which the student hid behind definitely {was/were} very tall compared with others.
15d. The lamp which the students hid behind definitely {was/were} very tall compared with others.
15e. The lamp which the student hid behind in class definitely {was/were} very tall compared with others.
15f. The lamp which the students hid behind in class definitely {was/were} very tall compared with others.

What was the status of the person who hid?
What was the status of the people who hid?

a. soldier   b. student   c. officer

16a. The driver who the fireman pulled thankfully {was/were} clear of the wreckage.
16b. The driver who the firemen pulled thankfully {was/were} clear of the wreckage.
16c. The driver who the fireman pulled out thankfully {was/were} clear of the wreckage.
16d. The driver who the firemen pulled out thankfully {was/were} clear of the wreckage.

16e. The driver who the fireman pulled out in a hurry thankfully {was/were} clear of the wreckage
16f. The driver who the firemen pulled out in a hurry thankfully {was/were} clear of the wreckage

What was the profession of the person who rescued the driver?
What was the profession of the people who rescued the driver?

a. robber   b. fireman   c. policeman

17a. The athlete who the coach trained definitely {was/were} prepared for the championship game.
17b. The athlete who the coaches trained definitely {was/were} prepared for the championship game.
17c. The athlete who the coach trained with definitely {was/were} prepared for the championship game.
17d. The athlete who the coaches trained with definitely {was/were} prepared for the championship game.
17e. The athlete who the coach trained with on Sunday definitely {was/were} prepared for the championship game.
17f. The athlete who the coaches trained with on Sunday definitely {was/were} prepared for the championship game.

What was the profession of the person who trained the athlete?
What was the profession of the people who trained the athlete?

a. referee   b. coach   c. assistant

18a. The kid who the mom drove unsurprisingly {was/were} tired of playing all day.
18b. The kid who the moms drove unsurprisingly {was/were} tired of playing all day.
18c. The kid who the mom drove with unsurprisingly {was/were} tired of playing all day.
18d. The kid who the moms drove with unsurprisingly {was/were} tired of playing all day.
18e. The kid who the mom drove with on Friday unsurprisingly {was/were} tired of playing all day.
18f. The kid who the moms drove with on Friday unsurprisingly {was/were} tired of playing all day.

Which type of family member drove a car?
a. father   b. aunt c. mom

Appendix B

IRB/HUMAN SUBJECTS APPROVAL

DATE: July 1, 2020

TO: Myung Hye Yoo
FROM: University of Delaware IRB

STUDY TITLE: [1619021-1] Different types of noun phrases
SUBMISSION TYPE: New Project

ACTION: DETERMINATION OF EXEMPT STATUS
EFFECTIVE DATE: July 1, 2020

REVIEW CATEGORY: Exemption category # *(2)*

Thank you for your New Project submission to the University of Delaware Institutional Review Board (UD IRB). According to the pertinent regulations, the UD IRB has determined this project is EXEMPT from most federal policy requirements for the protection of human subjects. The privacy of subjects and the confidentiality of participants must be safeguarded as prescribed in the reviewed protocol form.

This exempt determination is valid for the research study as described by the documents in this submission. Proposed revisions to previously approved procedures and documents that may affect this exempt determination must be reviewed and approved by this office prior to initiation. The UD amendment form must be used to request the review of changes that may substantially change the study design or data collected.

In-person research interaction with subjects cannot begin until the UD moratorium in response to the declaration of national emergency related to the COVID-19 pandemic is lifted. Please continue to reference https://research.udel.edu/coronavirus for the most up-to-date recommendations.

Unanticipated problems and serious adverse events involving risk to participants must be reported to this office in a timely fashion according with the UD requirements for reportable events.

A copy of this correspondence will be kept on file by our office. If you have any questions, please contact the UD IRB Office at (302) 831-2137 or via email at hsrb-research@udel.edu. Please include the study title and reference number in all correspondence with this office.

INSTITUTIONAL REVIEW BOARD

www.udel.edu