ABSTRACT

Title of dissertation: THE ACQUISITION AND PROCESSING OF BACKWARDS ANAPHORA

Nina Kazanina, Doctor of Philosophy, 2005

Dissertation directed by: Professor Colin Phillips
Department of Linguistics

This dissertation investigates long-distance backwards pronominal dependencies (*backwards anaphora* or *cataphora*) and constraints on such dependencies from the viewpoint of language development and real-time language processing. Based on the findings from a comprehension experiment with Russian-speaking children and on real-time sentence processing data from English and Russian adults I argue for a position that distinguishes structural and non-structural constraints on backwards anaphora. I show that unlike their non-syntactic counterparts, structural constraints on coreference, in particular Principle C of the Binding Theory (Chomsky 1981), are active at the earliest stage of language development and of real-time processing.

In language acquisition, the results of a truth-value judgment task with 3-6 year old Russian-speaking children reveal a striking developmental asymmetry between Principle C, a cross-linguistically consistent syntactic constraint on coreference, and a Russian-specific discourse constraint on coreference. Whereas Principle C is respected by children already at the age of three, the Russian-specific (discourse) constraint is not operative in child language until the age of five. These findings present a challenge for input-driven accounts of language acquisition and are most naturally explained in theories that admit the existence of innately specified principles that underlie linguistic representations.

In real-time processing, the findings from a series of self-paced reading experiments on English and Russian show that in backwards anaphora contexts the parser initiates an active search for an antecedent for the pronoun which is limited to positions that are not subject to structural constraints on coreference, e.g. Principle C. This grammatically constrained active search mechanism associated observed in the processing of backwards anaphora is similar to the mechanism found in the processing of another type of a long-distance dependency, the *wh*-dependency. I suggest that the early application of structural constraints on long-distance dependencies is due to reasons of parsing efficiency rather than due to their architectural priority, as such constraints aid to restrict the search space of possible representations to be built by the parser. A computational parsing algorithm is developed that combines the constrained active search mechanism with a strict incremental left-to-right structure building procedure.

THE ACQUISITION AND PROCESSING OF BACKWARDS ANAPHORA

by

Nina Kazanina

Dissertation submitted to the faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirement for the degree of
Doctor of Philosophy
2005

Advisory Committee:
Professor Colin Phillips, Chair
Professor Kira Gor
Professor Norbert Hornstein
Professor David Poeppel
Professor Patrick Sturt

ACKNOWLEDGMENTS

Now that graduate school is over, I realize how many people were crucially involved in my life during that period. I am deeply grateful to many people and especially so to:

- Colin Phillips, first, as a graduate director, for not discarding my very late application for graduate school; then, as my first psycholinguistics teacher, for introducing me to the field; later, as my advisor, for his patience and guidance, and for the endless conversations that we had during these years, and for making it possible to email/call/stop by any time for advice and know that you are always going to be helped. It is to him that I owe most of what I know (and also most of the determiners in this dissertation :));
- David Poeppel for his great classes, for broadening our knowledge to areas beyond language, for the cheer-ups and anecdotes and jokes, and for always offering his support whenever it was needed the most;
- Norbert Hornstein for his advice and his ability to explain things so naturally. Also, for often stopping by at our office on his way in and out of 1413 to talk about random things with us;
- my dissertation committee members, Patrick Sturt for starting me think along the lines explored here and for his time and Kira Gor for her insightful comments;
- Masaya Yoshida for the long conversations that we've had, for the late night teas, and for everything he taught me during these years. The rest of the 'Principle C gang': Ellen Lau, for her energy and for always letting me off the hook whenever it came to running a 9 am subject; Moti Lieberman, for his good mood and team;
- Bill Idsardi, my first phonology teacher, for always being unbelievably generous about his time and sharing his knowledge;
- Jeff Lidz and Julien Musolino, the first people to teach me the know-how of running experiments with children and the children, parents and teachers of the Moscow center *Ulybka* for their enthusiastic participation in these experiments;
- Grace Yeni-Komshian for being my mentor and for all she did for me – thank You!

Leticia Pablos, Ana Gouvea and Eli Murguia – for their friendship, for our gatherings and laughs; you are a true gift to me from Linguistics which I will do my best to never lose! ◆ Scott Fults, the best and most considerate classmate and co-organizer one can imagine ◆ Rob Fiorentino, who I learnt a lot from about many things ◆ Tomo Fujii and Jon Sprouse for being my syntax encyclopedias ◆ Evniki Edgar, Baris Kabak, Heidi Altmann, Jason Lilley, Libby Jonczyk, Tom Pellathy at UDel for being such a friendly crowd to meet when you first land in the US and have no clue about what is going on! ◆ Kathi Faulkingham and Rob Magee at UMD and Jane Creswell at UDel for always doing their best to help us ◆ Jeff Walker for MEG help and Brian Dillon for reading and commenting on this thesis ◆ Mitsue, Utako, Ilhan, Usama, Lydia, Takuya, SooMin, Kaori, Phil, Yuval, Clare, Matt, Diogo, Andrea and many-many more UMD Linguistics, NACS (yes, Virginie and Anthony, that means you!) and HESP students, RAs and faculty, past and present, for making the department the place it is.

My family and friends, some mentioned above and some not, you have always been the driving force behind my actions and the pages of this dissertation are all imbued with you! Thank you for being there!

TABLE OF CONTENTS

1 CHAPTER 1. INTRODUCTION

This dissertation is a psycholinguistic study of a particular type of long-distance dependency, *backwards anaphora*. Backwards anaphora is a relation between two nominal elements in which a referentially dependent element, e.g. the pronoun *he* in (1), linearly precedes the expression that it obtains its reference from, *Pooh*, the antecedent.

(1) After *he* read the book, *Pooh* ate an apple.

Although the backwards anaphora reading is rather natural in (1), it is absolutely ruled out in a sentence that bears a lot of surface similarity to (1), such as (2). In (2) *he* must refer to an individual other than *Pooh*, which suggests that the choice of an antecedent by a dependent element is not arbitrary.

(2) **He* read the book after *Pooh* ate an apple.

Whereas based on sentences like (2) it is unquestionable that anaphora in general and backwards anaphora in particular are subject to constraints, almost any other aspect of these dependencies is a subject of a fundamental debate.

For theoretical linguists, the bone of contention has been the *nature of constraints* on anaphora. Generative theories have tended to view constraints on intra-sentential anaphora as having a structural origin. According to the Binding Theory (Chomsky 1981), for example, (2) is illicit because it violates a structural constraint known as Principle C, which disallows coreference between a pronoun and a name that it c-commands. In contrast, adherents of the functionalist approach to language view constraints on anaphora, including Principle C, as resulting from general semantic/pragmatic principles, such as contextual prominence or accessibility of a referent in discourse (e.g., Kuno 1987, van Hoek 1997).

The psychological status of constraints on referential dependencies also remains in doubt. There are at least two possible explanations for why a representation is judged as ill-formed. Is a representation illicit because the grammatical apparatus of the speaker is incapable of generating such a representation? Or, alternatively, is a representation illicit because after being generated by the grammar, it is filtered out by some well-formedness principles that are part of the grammar?

The psycholinguistic experiments reported here are designed with these issues in mind. Understanding how constraints on backwards anaphora emerge in the child's grammar and how they are implemented during real-time processing should make it possible to re-assess various theoretical alternatives in terms of their predictions for those domains. Even more so, findings from our experiments should make it possible to draw inferences regarding the properties of the

acquisition device and the parser themselves. The main issues that our experiments address in each psycholinguistic domain are discussed below.

Language Acquisition

Long-distance dependencies have served as model cases in many debates on whether the child is innately pre-disposed for language (Pinker 1984, Crain 1991, Otsu 1981 vs. Elman 1993, O'Grady 1997, Pullum & Scholz 2002). The reason for this has to do with how observable properties of long-distance dependencies are based on the input that the child receives.

The first step the child needs to make on his way to representing a dependency is to discover that some element is dependent on another element. This by itself is a non-trivial task, but it is overshadowed by the complexity of another task that the child faces, i.e., realizing that a dependency relation is subject to constraints. The absence of a certain representation from the input to the child cannot ever be taken as evidence for the existence of a constraint on such a representation, since that would have led to exclusion of various phenomena that are grammatically licit but infrequent in the input (Chomsky in Piattelli-Palmarini 1980).[1] The presence of a constraint in the grammar can then be due to one of two reasons: either the constraint is innately specified or it gets into the grammar via some circuitous route.

With respect to the Principle C constraint that applies in sentences like (2), considerations of this sort and the universality of Principle C across languages led to claims that it is an innate constraint (Crain & McKee 1986). Crain and McKee suggested that the child need not derive the constraint from the input but rather has it as part of his language endownment. This claim was supported by findings from experiments with young children, who showed their knowledge of Principle C by rejecting backwards anaphora that is illicit due to the constraint, e.g. (2), while accepting coreference in sentences where Principle C does not apply, e.g. (1).

Here we provide a stronger piece of evidence for the innateness of Principle C by showing that it is already active in the child's grammar at a time when other constraints on anaphora are still dormant. This evidence comes from Russian, which in addition to the universally-valid Principle C, illustrated in (3), has an idiosyncratic discourse constraint that rules out backwards anaphora between the embedded and the main subjects in sentences that are introduced by the subordinator *poka* 'while', e.g. (4).

[1] This claim has been recently challenged by Tenenbaum & Griffiths (2001) who show that the child can make inferences based on the missing evidence if he were a Bayesian learner that makes use of a specific maximization principle. The reader is referred to the original paper and to Regier & Gahl (2004) for further details.

(3) **On* s'el jabloko, poka *Vinni-Pux* chital knigu. [Principle C]
He ate an apple while *Pooh* was reading the book.

(4) *Poka *on* chital knigu, Vinni-Pux s'el jabloko. [*poka*-constraint]
While *he* was reading the book, *Pooh* ate an apple.

We show that although even the youngest Russian-speaking children obey Principle C by robustly rejecting coreference in sentences like (3), the *poka*-constraint (4) becomes operative in the child's grammar much later. Russian-speaking children consistently fail to reject (4) on coreference at the age of three, vary on their degree of success with these sentences for another couple of years, and finally exhibit adultlike performance at the age of 5-6. The dissociation in the youngest children's performance on the two types of sentences is surprising if the child's knowledge is solely shaped by the input, since both (3) and (4) are equally missing from the input.

I conclude that the principles motivating Principle C are innate, whereas the *poka*-constraint emerges in the Russian-speaking child's repertoire as a result of tuning of other parts of the grammar, namely, adjustments in the tense-aspect system.

Parsing

It is known that in order to meet the challenge of real-time processing, the parser makes use of various types of information (e.g. lexical frequency, syntactic information, semantic plausibility, recency, etc.) to anticipate the upcoming material. Backwards anaphora represents a convenient case for testing how immediately structural or non-structural information is taken into consideration in real-time processing (see Marslen-Wilson 1975, MacDonald, Pearlmutter & Seidenberg 1994 vs. Frazier 1978, Ferreira & Clifton 1986 for opposite views and Trueswell & Tanenhaus 1995 for a review). In a left-to-right parser model, backwards anaphora is a dependency that can be detected early, as soon as the first member of the dependency, the referentially dependent element, is encountered. Given the properties of the dependent element, the space of all possible antecedents can be restricted based on two types of information. One source is morphological information, such as gender and number: since the dependent element and its antecedent must match in these features, the parser could immediately disregard any incoming nominal that mismatches in gender/number with the dependent element. The other source is structural constraints on dependencies, e.g. Principle C: the parser could immediately disregard candidates that are in positions which violate Principle C.

Based on the results of a series of self-paced reading experiments I argue that a structural constraint on backward anaphora, Principle C, is taken into consideration by the parser at the earliest stage of processing. In doing so, I also show that the dependency formation is active: the parser triggers a search for an antecedent immediately after the dependent element is encountered and attempts

to complete the dependency as quickly as possible. I further argue that our findings conform to classic findings in the literature on the processing of *wh*-dependencies. *Wh*-dependencies, which also are detectable early by the first member of the dependency, the *wh*-word, are known to involve an active search for the gap which abides by grammatical constraints on *wh*-dependencies, the island constraints (Stowe 1986, Crain & Fodor 1985, Traxler & Pickering 1996). I conclude that the parser employs the same mechanism for both *wh*-dependencies and backwards anaphora: an active search that is constrained by structural considerations.

I argue that the parser's application of structural constraints at the earliest stage of processing need not be due to the architectural prominence of such information; rather it can emerge as a result of what is the most optimal way of restricting the search space by the parser. This is because almost any kind of anticipatory prediction necessarily implies structural information, whereas the degree of relevance of non-structural information varies from case to case. This view can therefore be seen as a middle ground between syntax-first and interactive models of parsing: On the one hand, in the spirit of a syntax-first model, structural information never lags behind other types and, on the other hand, in the spirit of an interactive model, it may be considered simultaneously with other types.

Together, our findings from acquisition and sentence processing reveal strong parallels: constraints of structural nature have a lead at the time scale of language development and online processing.

The organization of this dissertation is as follows.

Chapter 2 introduces constraints on anaphora from the perspective of different frameworks. It focuses on Principle C of the Binding Theory (Chomsky 1981) and its subsequent modifications. I discuss objections raised against Principle C in the literature and argue that these objections pertain only to a particular formulation of Principle C and do not hold once modifications proposed in Heim (1982) and Reinhart (1983, 2000) that enhance Principle C when relevant pragmatic notions are adopted. Although these modifications of Principle C are critical in many respects, they do not affect the most essential point behind it, i.e. that a dependent element must not c-command a referential expression that it corefers with.

Chapter 3 discusses the application of binding constraints during real-time sentence processing. I argue that all findings in the existing literature conform to the view that the Binding Principles (mostly Principles A and B in forwards anaphora contexts) are taken into account at the earliest stage of processing. I then present a series of self-paced reading experiments in English aimed at exploring how the parser establishes backwards anaphora dependencies and implements constraints on such dependencies in real time. We show that the parser actively

searches for an antecedent for the pronoun, but only in positions that would not lead to a violation of Principle C. I argue that this represents a direct parallel between the processing of backwards anaphora dependencies and the processing of *wh*-dependencies, which is suggestive of a common mechanism underlying parsing of various types of structural dependencies.

Chapter 4 focuses on Russian, which in addition to Principle C has another constraint on backwards anaphora: the '*poka*-constraint'. I provide various theoretical considerations that all point to the discourse nature of the *poka*-specific constraint. I then show that the syntactic Principle C constraint takes a lead over the Russian-specific constraint both in acquisition and in parsing. Whereas even 3-year old children consistently respect Principle C, the *poka*-constraint does not become fully operative in the child grammar until the age of 5-6. Similarly, in self-paced reading experiments we find that Russian speakers immediately disregard antecedents that would violate Principle C, whereas they do temporarily consider antecedents that are subject to the *poka*-constraint. Cumulatively, the findings from this chapter support the notion that Principle C is a constraint on structure generation that is also part of the child's innate endowment.

In Chapter 5 I develop an algorithm for incremental building of referential dependencies that imitates the mechanisms characteristic of the human parser, i.e. an active search for an antecedent for a dependent element and immediate respect of the Binding Principles. In addition to binding dependencies I consider another type of referential dependency, the control dependency. The algorithm is based on Hornstein's (1999, 2001) theory which considers binding and control as essentially the same type of dependency that is formed via movement.

Chapter 6 concludes.

2 CHAPTER 2. CONSTRAINTS ON REFERENTIAL DEPENDENCIES

Anaphora is a (referential) dependency relation in which one member of the dependency, the dependent element, obtains its reference via the other member, the antecedent. (1)-(2) are examples of intra-sentential anaphora in which the dependent element – the reflexive *herself* or the pronoun *she* – is referentially evaluated by linking (in the most general sense) to *Alice*.

(1) Martin said that Alice$_i$ pinched herself$_i$ (to make sure she was not dreaming).
(2) Alice$_i$ said that she$_i$ was not tired of Martin.

As we noted in the Introduction, anaphoric dependencies are subject to constraints. Although *Alice* is a licit antecedent for the dependent element in (1)-(2), it is no longer licit if it occupies a different position in the sentence, as in (3)-(4).

(3) *Alice$_i$ said that Becky pinched herself$_i$.
(4) *Becky said she$_i$ was not tired of Alice$_i$.

There are two major types of approach to explaining regulations on anaphora, which contrast as to whether ot not structural considerations are admitted among the factors regulating anaphora, or more generally, as part of language architecture.

The Binding Theory, originally proposed in Chomsky (1981), and its subsequent developments exemplify a generative approach to anaphora in which referential dependencies are largely (but not exclusively) regulated by structural constraints. In the following discussion, a lion's share of attention will go to the Binding Theory's Principle C, a long-distance constraint on coreference with an unbounded scope, to lay the groundwork for the psycholinguistic experiments that will explore this principle in acquisition and sentence processing (Chapters 3 and 4). I summarize objections raised against Principle C in the literature and provide arguments, both new and existing, that can deal with the apparent violations.

A competing view to the Binding Theory is exemplified by the Reference Point and Dominions model (van Hoek 1997) developed in the framework of Cognitive Grammar (Langacker 1987, 1991), which considers anaphora to be regulated solely by pragmatic principles. To this end, I will argue that the success of this approach is jeopardized by some aspects of cross-linguistic variation in the use of full nominal elements.

Chapter 2 is organized as follows. Section 2.1 gives a summary of the Binding Theory in its original version (Chomsky 1981). Section 2.2 dwells on reformulations of Principle C as a constraint on pronouns and places it against a broader background that includes a semantic/pragmatic model of the world (Heim 1982, 1993; Reinhart 1983, 2000). Criticisms raised against Principle C are

presented and discussed in section 2.3, where I argue that most objections go away once Principle C is allowed to interface with pragmatics. Section 2.4 provides an overview of van Hoek's (1997) syntax-free approach to anaphora cast in the framework of Cognitive Grammar and discusses potential problems for this approach. Section 2.5 summarizes.

2.1 REFERENTIAL DEPENDENCIES IN GOVERNMENT AND BINDING THEORY

The Binding Theory (Chomsky 1981) is a theory within the generative framework that derives major restrictions on the interpretation of noun phrases (NPs) from syntactic considerations. The Binding Theory categorizes nominal elements into the three types listed in (5).

(5) Anaphors: reflexives, e.g., *myself, himself* and reciprocals, e.g. *each other*
Pronouns, e.g., *I, you, they, he*
R(eferring)-expressions, e.g., *John, the man*

Each type of nominal from (5) is taken to be subject to a structural constraint on coreference, cited in (6)

(6) *Principles A, B & C of the Binding Theory (Chomsky 1981)*
Principle A: An anaphor must be bound within its governing category.
Principle B: A pronoun must be free within its governing category.
Principle C: R-expressions must be free (= not bound).

To illustrate the key points of the Binding Theory here I give simplified descriptions of the terms in (6). (Formal definitions of the notions implicated in (6) are given in Appendix 2-A). *C-command* (Reinhart 1976) is a structural relation between two elements in the tree. By definition, each node c-commands its sister and all the daughters of its sister. This is illustrated in Figure 1.

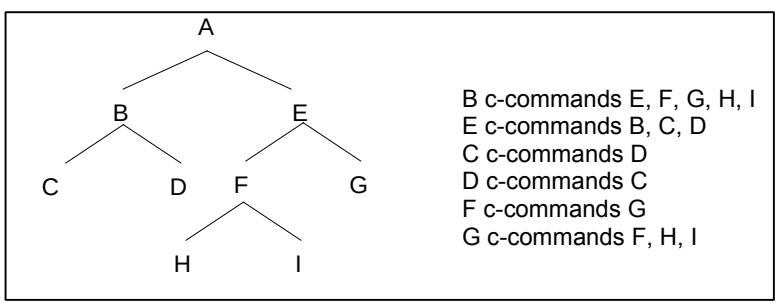


Figure 1. C-command relation. For example, B c-commands its sister E and all the daughter nodes of E (F,G,H,I). An exhaustive list of c-command relations is given to the left of the tree.

A *bound* NP is an NP that is c-commanded by its antecedent. Finally, the *governing category* for a nominal element is the smallest NP or finite clause that properly contains that element.

We are now ready to exemplify the Binding Principles (6) using a few examples. First consider (7), in which the reflexive *herself* obligatorily refers to *Sue*.

(7) Mary_i said that Sue_k kissed $\text{herself}_{*i/k}$.

The choice of an antecedent by the reflexive must conform to Principle A: in order for some NP to be a grammatical antecedent for a reflexive, it must both c-command the reflexive and be inside its governing category. The governing category for the reflexive in (7) is the lower clause. *Sue* indeed satisfies both requirements: it occurs inside the governing category for the reflexive and it c-commands the reflexive. *Mary*, on the other hand, cannot be an antecedent for *herself*, because it is outside the governing category for *herself*.

Let us next look at (8), which was obtained by substituting the reflexive *herself* in (7) with the pronoun *her*. The pattern of coreference in (8) is the opposite of what we had in (7): *her* can refer to *Mary*, but not to *Sue*.

(8) Mary_i said that Sue_k kissed $\text{her}_{i/*k}$.

The pronoun *her* must abide by Principle B whereby an NP is illicit as an antecedent for the pronoun if the NP both c-commands the pronoun and is inside the governing category for the pronoun. Hence, *her* cannot corefer with *Sue*, which both c-commands the pronoun and is inside its governing category (the lower clause). *Mary*, on the other hand, is a licit antecedent for *her*, since it is outside the governing category for the pronoun.

(9) serves as another useful illustration of Principle B. In (9) the pronoun *her* is a possessor inside the NP *her boyfriend*, its governing category (recall that the governing category is the smallest NP or clause that contains the pronoun/reflexive). According to Principle B, every nominal outside that governing category is a licit antecedent for *her*. Hence, in (9) *her* can corefer with either *Mary* or *Sue*.

(9) Mary_i said that Sue_k kissed $\text{her}_{i/k}$ boyfriend.

Note that there in a qualitative difference in how Principle A and Principle B are formulated. Principle A is stated as a 'positive' condition that determines which NP must be an antecedent for the reflexive. Principle B, on the other hand, is formulated as a 'negative' condition that defines which NPs are illicit antecedents for the pronoun. This nicely captures the empirical facts: whereas the reflexive usually has one potential antecedent (or at least a finite number of antecedents), the pronoun may have many more possible antecedents (and potentially an unlimited number of them), as shown in (10) and (11).

(10) John$_i$ told Bill$_j$ that Mark$_l$ read a book about himself$_{*i/*j/k}$.
(11) John$_i$ told Bill$_j$ that Mark$_l$ read a book about his$_{i/j/k}$ hometown.

The last principle, Principle C, is of the most interest to us. In Chomsky (1981) Principle C was formulated as a condition on R-expressions which stated that an R-expression cannot corefer with a nominal that c-commands it. Principle C explains why *John* cannot corefer with *him* in (12)a: in (12)a *John* is subject to Principle C, since it is bound by the pronoun *him*. As expected, coreference becomes allowed if the sentence is modified so that the name is no longer c-commanded by the pronoun, as in (12)b.

(12) a. *He$_i$ knows that Mary likes John$_i$.
b. John$_i$ knows that Mary likes him$_i$.

(12) illustrates Principle C using pronoun-name sequences, i.e. that an R-expression cannot be bound by a pronoun. (13) illustrates Principle C for pairs of R-expressions. In (13) *Ann* and *the teacher* cannot be understood as the same person: had the two NPs coreferred, the R-expression *the teacher* would have been bound and thus would have violated Principle C.

(13) *Ann$_i$ said that the teacher$_i$ was tired.

Principle C elegantly explains the contrast in availability of coreference between *he* and *John* in bi-clausal sentences with adjunct clauses like (14).

(14) a. *He$_i$ entered the room when John$_i$ was eating an ice-cream.
b. When he$_i$ entered the room, John$_i$ was eating an ice-cream.

Figure 2 shows the tree structures for the sentences in (14), based on a standard assumption that sentence-final *when*-clauses are VP-adjoined, whereas sentence-initial *when*-clauses are S-adjoined. In (14)a the R-expression John is c-commanded by the pronoun he, therefore, coreference between these NPs is excluded by Principle C. In (14)b, on the other hand, he does not c-command John (Figure 2b), hence coreference between the two NPs is no longer ruled out by Principle C.

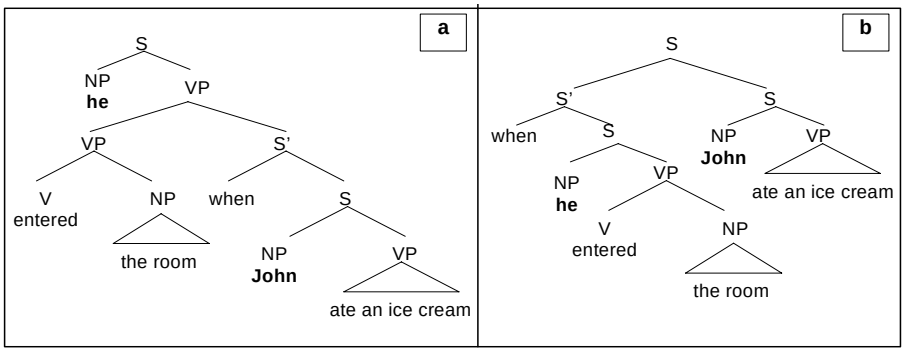


Figure 2. Tree structures for examples (14)a and (14)b.

A note concerning linear word order is worthwhile. As is clear from the definitions in (6), the linear order of the dependent element and its antecedent does not play a decisive role in whether coreference between the elements is allowed. From the perspective of the Binding Principles linear order matters only to the extent it changes the structural relation between the elements. In other words, coreference between a pronoun and a name may be either licit or illicit whether the pronoun follows the R-expression, e.g. (8)ab, or precedes it, e.g., (14)ab.

Note also that not all constraints on referential dependencies are syntactic (Kuno 1987, Lasnik 1989). Linear word order that is irrelevant in syntax does have a prominent role in deciding whether coreference is available at an extra-syntactic interpretive level, usually dubbed as 'discourse', as shown by the comparison of (15)a and (15)b.

(15) a.*He_i opened the door and $John_i$ entered the house.
b. $John_i$ opened the door and he_i entered the house.

Chomsky's (1981) proposal was later modified and refined on various occasions (e.g. Chomsky 1986), some of which were rather substantial. However, for the purpose of this work we can abstract away from the details, since all these proposals share the spirit of the original account in considering constraints on the distribution of nominals to be of a syntactic nature.[2]

2.2 PRINCIPLE C AS A CONSTRAINT ON PRONOUNS

One of the earliest objections against Principle C as formulated by Chomsky (1981) was cases like (16) in which an R-expression is not free. In (16) an R-expression is bound by another R-expression, e.g. in (16)a *Ann* in *Ann's mother* is c-commanded by *Ann* in the main subject position and yet both occurrences of *Ann* may refer to the same person.

(16) a. Ann_i told Mary that Ann_i's mother is a spy.
b. I know what John and Bill have in common. John thinks that Bill is terrific and Bill thinks that Bill is terrific. [Evans 1980]

Moreover, cross-linguistically there are many more exceptions to the generalization that R-expression must be free. According to Lasnik (1986) in Thai and Vietnamese an R-expression may be bound by another R-expression outside its governing category (examples (17) and (21)), and in Thai (but not in Vietnamese) an R-expression can even be bound locally (examples (18) and (22)).

[2] As we said in the previous paragraph, this statement should not be over-generalized to a claim that *all* constraints on referential dependencies are syntactic.

Importantly, in either language an R-expression cannot be bound by a pronoun, either local, e.g. (20) & (24), or non-local, e.g. (19) & (23).

(17) Cɔɔn khit waa Cɔɔn chalaat. [Thai]
John thinks that John is smart
(18) Cɔɔn chɔɔp Cɔɔn.
John likes John
(19) *Khaw khit waa Coon chalaat.
he thinks that John is smart
(20) *Khaw chɔɔp Cɔɔn.
he likes John

(21) John tin John sẽ thắńg. [Vietnamese]
John thinks John will win
(22) *John thương John.
John likes John.
(23) *Nó tin John sẽ thắńg.
he thinks John will win
(24) *Nó thương John.
he likes John.

Based on these facts, Lasnik breaks up Principle C into two subparts given in (25).

(25) Principle C (Lasnik 1991)
(i) An R-expression must be free.
(ii) A pronoun must not bind an R-expression.

Lasnik argues that only part (ii) of Principle C is universal. Part (i), on the other hand, is subject to cross-linguistics parameterization, holding in some languages (e.g. English) but not in others (e.g. Thai).[3] The formulation in (25) limits the domain of the universal subpart of Principle C to pronoun-name sequences and thus obviates problems that the original Principle C had with examples like (16). In what follows, by Principle C I will effectively understand a condition on pronoun-name (or, more broadly, dependent element – name) sequences, as in (25)ii.

Reinhart (1983, 2000) takes an even further departure from the original formulation of the Binding Principles, including Principle C (see Appendix 2-B for details). She argues then that traditional definitions confound different mechanisms of establishing coreference: binding and pragmatic covaluation.

With regard to binding, Reinhart replaces the index-based definition of binding by a definition based on the traditional logical-syntax concept of binding, whereby

[3] Lasnik (1986) suggests that, more generally, part (ii) can be viewed as an instantiation of a general prohibition on the binding of a more referential nominal by a less referential nominal.

pronouns and reflexives are viewed as variables bound by a λ-operator, as in (26). In (26) *he* can corefer with *John* because they are bound by the same λ-operator in accordance with general principles on logical binding. [4]

(26) John thinks he is smart.
LF: John (λx (x thinks x is smart))

The same principles of logical binding have the effect of excluding core Principle C violations where the pronoun binds the name that it c-commands, as in (27)a-c. In (27)a the variable *x* in *x boy* is bound twice, by the pronoun and by the quantifier *every*, which is illegal. In (27)b,c the R-expression *John* cannot be transformed into a variable to be bound by an operator:

(27) a. *He_i likes every boy_i.
LF: every boy (λy (she (λx (x likes y))))
b. *He_i likes $John_i$.
LF: he (λx (x likes John))
c. *He_i said $John_i$ likes wine.
LF: he (λx (x said John likes wine))

Thus, in Reinhart's account there is no need to postulate Principle C separately to explain the lack of binding between a pronoun and a name that it c-commands. However, Reinhart argues that Principle C must be preserved to explain the lack of covaluation in the same cases.

The difference between a bound variable and a covaluation reading is present in every sentence that contains a variable, such as (28), but it can be most straightforwardly demonstrated using sentences with VP-ellipsis such as (29). In (29) binding yields a sloppy-identity reading, whereas covaluation yields a strict-identity reading (also noted in Lasnik 1976).[5]

(28) John likes his dog.

---

[4] A detailed overview of Reinhart (2000) is given in Appendix 2-B.

[5] Sentences with VP ellipsis in which the non-elided conjunct contains a referential dependency usually have two readings: the counterpart of the dependent element in the elided conjunct may either preserve the same reference as the original (the 'strict identity' reading, (a)) or pick up another reference (the 'sloppy identity' reading, (b)). A standard analysis of ellipsis assumes that the strict reading results from substituting the missing VP in the elided clause with a copy of the VP from the non-elided clause that has its variables (pronouns) evaluated, i.e. [likes his_i wife] and *i=John*. The sloppy reading results from copying a non-evaluated version of the VP from the non-elided clause into the elided clause, i.e. [likes his wife], so that the variable in either VP is then evaluated in its own clause.

(i) $John_i$ [$_{VP}$ likes his_i wife], and Bill does too.
a. John likes John's wife and Bill likes John's wife. [strict]
a'. $John_i$ [$_{VP}$ likes his_i wife], and $Bill_k$ [$_{VP}$ likes his_i wife] too.
b. John likes John's wife and Bill likes Bill's wife. [sloppy]
b'. $John_i$ [$_{VP}$ likes his_i wife], and $Bill_k$ [$_{VP}$ likes his_k wife] too.

(29) John likes his dog and Mary does too.
Bound variable: John likes John's dog and Mary likes Mary's dog.
Covaluation: John likes John's dog and Mary likes John's dog.

Reinhart argues that only the bound variable reading is regulated by purely structural principles (principles of logical binding in her case that are similar to Chomsky's (1981) Binding Principles in that they apply only under c-command). The covaluation reading is subject to a broader evaluation procedure that includes pragmatic and economy considerations. According to Reinhart (2000) the mechanism that enables covaluation of a pronoun and an R-expression, for example, applies only if variable binding has failed to apply (e.g. if the elements in the dependency were not in a c-command relation) and if the covaluation mechanism produces a distinct outcome from what the variable-binding outcome could have been.[6] This principle will be illustrated in section 2.3 and will explain a phenomenon that is problematic for earlier versions of Principle C.

I will adopt Reinhart's view of Principle C and henceforth I will use terms coindexation/binding and covaluation in her sense. The term 'coreference' will be used in the most general sense, to refer to cases in which a pronoun/reflexive ends up having the same referent with an a R-expression or a Q-expression, i.e. to indicate either coindexation/binding or covaluation.

2.3 OBJECTIONS AGAINST SYNTACTIC PRINCIPLE C

It has been argued by various researchers that the Binding Theory falls short of explaining certain classes of data. The most problematic cases are represented by sentences in which the pronoun corefers with an R-expression that it c-commands, which are claimed to undermine Principle C. The list below is compiled on the basis of objections raised by various authors, including Bolinger (1979), McCray (1980) and Harris & Bates (2002). I divide the problematic sentences into several categories in (30)-(33).

(30) a. He_i is $Ralf_i$.
b. She_i is the Queen of $Jordan_i$.
(31) a. He_i did what $John_i$ always does…
b. It was rather indelicately pointed out to him_i that $William_i$ would never become a successful accountant.
c. She_i was told that if she wanted to get anywhere in this dog-eat-dog world, $Mary_i$ was going to have to start stepping on some people.
d. The teacher warned him_i that in order to succeed $Walter_i$ was going to have to work a lot harder from now on.
e. He_i had already shot himself before $John_i$ quite knew what he was doing.

[6] See Appendix 2-B for examples of the binding vs. covaluation distinction.

f. He$_i$ would have been like a son to both of us, if my wife and I could have kept Jim$_i$ away from the influence of his family.
g. He$_i$ was just a little boy when I knew John$_i$.

(32) a. I've never been able to talk to him$_i$ about the examples John$_i$ claimed would refute my theory.
b. If you try to tell him$_i$ that the reason why John's$_i$ dog was taken away from him was rabies, he'll get very upset.
c. I've never been able to explain to her$_i$ that Betsy$_i$'s gophers destroyed my lawn each spring.

(33) a. He$_i$ had been staring at the control panel for over an hour when Jack$_i$ received a message from his commander. [attributed to G. Lakoff in Harris & Bates (2002)]
b. He$_i$ was threatening to leave when Jack$_i$ noticed that the computer had died.
c. He$_i$ was about to place a few bets when Mike$_i$ was advised that the cops were in the bar.

Following Reinhart (1983), let us note right away that coreference in (30)-(33) is obtained by non-syntactic means and is covaluation rather than binding. Should the cases in (30)-(33) be authentic cases of binding between the pronoun and the name, the pronoun should also be able to bind a Q-expression in place of the R-expression. However, as shown in (34), coreference is no longer possible if the R-expression is replaced by a quantifier phrase.

(34) a. *He$_i$ is every student$_i$.
b. *He$_i$ did what every man$_i$ always does…
c. *She$_i$ was told that if she wanted to get anywhere in this dog-eat-dog world, every woman$_i$ was going to have to start stepping on some people.
d. *The teacher warned him$_i$ that in order to succeed everyone$_i$ was going to have to work a lot harder from now on.
e. *It was rather indelicately pointed out to him$_i$ that every student$_i$ would never become a successful accountant.
g. *I've never been able to talk to him$_i$ about the examples every physicist$_i$ claimed would refute my theory.
h. *If you try to tell him$_i$ that the reason why every neighbor's$_i$ dog was taken away from him was rabies, he'll get very upset.
i. *I've never been able to explain to her$_i$ that every child$_i$'s gophers destroyed my lawn each spring.
i. *He$_i$ had been staring at the control panel for aver an hour when every captain$_i$ received a message from his commander.
j. *He$_i$ was threatening to leave when every programmer$_i$ noticed that the computer had died.

The remainder of this section will aim at explaining why covaluation is available in sentences (30)-(33). I will show that although the availability of coreference in

these sentences violates Principle C as first formulated in Chomsky (1981), most cases stop being problematic once the Binding Principles are placed into a broader context that establishes a necessary semantic/pragmatic basis for anaphora (i.e. Heim 1982, Reinhart 1983, 2000). Such an enriched Principle C is different from its predecessor in two respects. First, as briefly mentioned in the previous section and explained in Appendix 2-B, this revised Principle C rules out covaluation between an R-expression and a variable that c-commands it only if this covaluation yields a result that is indistinguishable from binding, had binding been possible (Reinhart 1983, 2000; Grodzinsky & Reinhart 1993). This fact will be critical in explaining identity sentences (30). Second is the notion of guises (Heim 1982), according to which the same individual may be represented by more than one index given that the pragmatics of the situation supplies distinct perspectives on the individual. This consideration can help in explaining why coreference is licit in (32).

Let us start with identity sentences like (30), repeated under the same number below. Reinhart (2000) argues that covaluation in these sentences is allowed because it yields a result that is distinguishable from the hypothetical binding result. Putting it simply, the binding evaluation of (30)a, for example, if available, would have involved a tautological statement, i.e. *x likes x*, and its evaluation to a concrete person, i.e. *x = Ralf*. Covaluation, on the other hand, does not involve the same steps and rather is a specific statement about Ralf, i.e. *x* is *Ralf*.

(30) a. He_i is $Ralf_i$.
b. She_i is the Queen of $Jordan_i$.

We conclude that the availability of coreference in identity sentences like (30) is consistent with Principle C.

The availability of a coreference reading in sentences like (31) (repeated below) can be explained by appealing to the notion of *guises* (Heim 1982). Heim convincingly argues that for the purposes of reference an individual entity need not be a minimal indivisible unit. She suggests that nominals pick out mental rather than physical representations and that a minimal unit of reference in fact corresponds to a single guise of an individual. For example, in (31)a *he* and *John*, one and the same person, are viewed from two different perspectives, each providing a separate guise: the perspective of what he is usually like and the perspective of what he is like at the moment. Similarly, in (31)b *William* is presented from the perspective of what he thinks of himself and what others think of him. Accordingly, each guise, i.e. the pronoun and the name in (31), will carry a different index and would not violate Principle C.

(31) a. He_i did what $John_i$ always does…
b. It was rather indelicately pointed out to him_i that $William_i$ would never become a successful accountant.

c. She$_i$ was told that if she wanted to get anywhere in this dog-eat-dog world, Mary$_i$ was going to have to start stepping on some people.
d. The teacher warned him$_i$ that in order to succeed Walter$_i$ was going to have to work a lot harder from now on.

A similar reasoning may explain the rest of sentences in (31). In (31)c, for example, the first guise of *Mary* is that of a receiver of advice, and the second guise is a hypothetical description of what she should or could be like in some extension of the real world. The same applies to *Walter* in (31)d. As we would predict, coreference becomes unavailable if these sentences are changed in a way which would make it harder to construct different guises of the same individual, e.g. (35).[7]

(35) a. *She$_i$ was told that because she wanted to get somewhere in this dog-eat-dog world, Mary$_i$ started stepping on some people.
b. *The teacher told him$_i$ that in order to succeed Walter$_i$ worked long hours for the past six months.

So, sentences like (31) are not counter-examples to Principle C if coreference is understood in terms of guises of individuals, rather than individual entities. In order for this explanation to be valid one must ensure that guises are not overused, which requires developing a detailed semantic-pragmatic theory of conditions under which additional guises are introduced. Such a theory could perhaps borrow many insights from functionalist approaches to language, e.g. Cognitive Grammar theory discussed in section 2.4.

At the moment I have nothing to offer as an explanation for examples in (32). One must note, however, that coreference in (32) is not as easy to obtain as one may think. For example, speakers who easily accept (30), (31) and (33) on coreference are significantly more reluctant to accept (32), and many of them only allow a disjoint reading.

(32) a. I've never been able to talk to him$_i$ about the examples John$_i$ claimed would refute my theory.
b. If you try to tell him$_i$ that the reason why John's$_i$ dog was taken away from him was rabies, he'll get very upset.

[7] One may challenge our argument by saying that the illicitness of (35) on the coreference reading is due to the implausibility of such a reading, since the agent (*Mary* or *William*) is normally informed about her/his own past actions. Although this intuition is plausible, this reasoning is unlikely to be the main reason for the unavailability of coreference in (35). Indeed, the same constrast as in (31)c,d vs. (35)a,b takes place with verbs of communication that can equally felicitously express coreference between the main object and the agent of the embedded event regardless of whether the embedded event is future or past. This can be illustrated with the verb *to remind*:

(i) The teacher reminded him$_i$ that in order to succeed Walter$_i$ was going to have to work a lot harder from now on.
(ii) *The teacher reminded him$_i$ that in order to succeed Walter$_i$ had to work long hours for the past six months.

c. I've never been able to explain to her$_i$ that Betsy$_i$'s gophers destroyed my lawn each spring.

The final set of cases is bi-clausal sentences with *when*-clauses like (33). These sentences were studied by Harris & Bates (2002) in a series of psycholinguistic experiments, where participants had to judge the availability of coreference between the pronoun in the first clause and the subject of the following *when*-clause. Harris & Bates report that participants accepted coreference in discourse-initial sentences like (33) at a very high rate, in approximately 73% of cases, which is unexpected for Principle C since the pronoun in the first clause c-commands the subject of the *when*-clause.

(33) a. He$_i$ had been staring at the control panel for over an hour when Jack$_i$ received a message from his commander.
b. He$_i$ was threatening to leave when Jack$_i$ noticed that the computer had died.
c. He$_i$ was about to place a few bets when Mike$_i$ was advised that the cops were in the bar.

I will argue that the reason for the availability of coreference in (32) is because the pronoun in fact does not c-command the following R-expression. In other words, I will try to justify the claim that *when* in (33) should not be analyzed as a VP-subordinator, but rather as a sentential modifier.[8]

I will start with the observation that sentences like (33) all contain an event in the *when*-clause that is at the very least not controlled by the agent of the first clause, and in many cases is surprising or even shocking for that agent. For example, all 24 sets used in Harris & Bates (2002) (Experiment 1) contained a non-agentive event in the second clause. The most widely used predicates in the *when*-clause were what the authors called 'mental state' verbs (*felt, heard, noticed, saw*) or passives (*was struck, was ordered*). The subject of the *when*-clause therefore was an experiencer, a patient or a goal, but never an agent. If the content of the *when*-clause is changed so that it expresses an agentive event as in (36), the availability of coreference between the two subjects sharply declines.

(36) a. ??He$_i$ had been staring at the control panel for over an hour when Jack$_i$ gave an order to his soldier.
b. *He$_i$ was threatening to leave when Jack$_i$ turned on his computer.
c. *He$_i$ was about to place a few bets when Mike$_i$ started singing a song.

Next, I will argue that the status of the *when*-clause in sentences like (33) is different from other types of *when*-clauses, using questions as a diagnostic.

[8] This idea is discussed in Harris & Bates as a potential explanation for their sentences by a formal linguist, but rejected as incoherent. I will return to their criticism at the end of this section.

Let us first demonstrate the various interpretations that *when* can have. Moens & Steedman (1988) note, and Harris & Bates (2002) agree, that *when* is ambiguous and can be interpreted differently depending on the properties of the events that it links. In (33) it introduces an event for which the main clause serves as a setting (background). On another occasion, e.g. (37), *when* may largely correspond to the subordinator *while* and set the ground for the main event. Moreover, a *when*-clause can introduce an event that is a cause, with the main event expressing the result of that cause. This type of *when* can be approximated by *after* and is exemplified by (38). Note that in both (37) and (38) *when* locates the main event inside the embedded event (which is rendered as 'immediately after' if the embedded event is punctual).[9] In sentences (37)-(38) where *when* has a 'while' or an 'after' reading, coreference between the main subject pronoun and the subject of the *when*-clause is illicit, as shown in (39).

(37) a. Mary was talking on the phone when John was cooking dinner.
b. Sue entered the room when Bill was talking on the phone.
(38) Kate broke the glass when John kicked the door.
(39) a. $He_{*i/*k}$ was talking on the phone when $John_i$ was cooking dinner for $Mark_k$.
b. *He_i entered the room when $Bill_i$ was talking on the phone.
c. *He_i broke the glass when $John_i$ kicked the door.

Interestingly, whereas (37)-(38) can all serve as an answer to a *when*-question (40) about the corresponding main event, the same does not hold of (33). A *when*-question to each of the sentences in (33) is given in (41). Even when coreference is taken out of play, as in (42), these sentences remain infelicitous as an answer to the corresponding *when*-question.

(40) a. When was Mary talking on the phone?
b. When did Sue enter the room?
c. When did Kate break the glass?
(41) When had he been staring at the control panel?
When was he threatening to leave?
When was he about to place a few bets?
(42) Larry had been staring at the control panel for over an hour when Jack received a message from his commander.
Mark was threatening to leave when Jack noticed that the computer had died.
Samuel was about to place a few bets when Mike was advised that the cops were in the bar.

Such a contrast in behavior with respect to *when*-questions is surprising if *when* in (33) or (42) is exactly the same as in (37)-(38). However, the split follows naturally

[9] Note that the temporal relation between the main and embedded events in (33) is different.

if there were two different *when*s: one that adjoins to VP and introduces a VP-modifier clause and the other that adjoins to IP as a sentential modifier.

The ill-formedness of (33) or (42) as an answer to a *when*-question is explained by the fact that the question probes temporal properties of the VP and expects a VP-modifier, which the *when*-clause in (33)/(42) is not, as an answer. In other words, unlike (37)-(38) in which *when* has the primary function of being a subordinator, in (33) *when* is akin to a coordinator with an approximate meaning 'and/but suddenly at that moment'.[10] Sentences like (33) then should be regarded as having the same status as bi-clausal sentences coordinated by *and* or *but*. Coreference in these sentences is not ruled out by Principle C (since the two clauses are sisters in the tree structure) and is moderated by extra-syntactic considerations, e.g. parallelism, discourse prominence, etc. [11]

As I mentioned earlier, Harris and Bates anticipate such a move from a generative linguist, but they are skeptical about the legitimacy of this step. Harris and Bates claim that in such a case, interpretation would dictate the structure, which violates the principles of a generative model with autonomous syntax that claims the opposite directionality of this relation, i.e. that interpretation is dictated by form.[12] I believe that their argument is flawed and I will attempt to show why by comparing *when*-sentences to a somewhat less controversial case of sentences with PP-modifiers.[13]

---

[10] This also explains why when-clauses in (33) are not agentive: in order for the when-event to come as a surprise, it should rather be planned and/or executed by someone else.

[11] As noted by Harris & Bates (2002), it is only the subordinator *when* that yields what they consider to be a violation of Principle C. Similar structures with other subordinators, including temporal subordinators *before*, *after* and *while* behave in accordance with Principle C and disallow coreference between the pronoun and the embedded subject, as exemplified in (i).

(i) *He_i had been staring at the control panel for over an hour while $Jack_i$ received a message from his commander.
*He_i was threatening to leave after $Jack_i$ noticed that the computer had died.
*He_i was about to place a few bets before $Mike_i$ was advised that the cops were in the bar.

This fact strengthens the account of coreference in sentences like (33) advocated here, i.e. that an apparent violation of Principle C is due to the idiosyncratic properties of *when*.

[12] A relevant citation from Harris & Bates (2002) is given below:

> "Mandating that interpretation be dictated by the form (structure) of the phrase-structure tree should have the advantage of allowing theories to be falsi.ed. But in cases like (i) and (ii) interpretation appears to dictate structure. *[NK: (i) and (ii) mentioned in the preceding sentence are given below. I use (i) and (ii) instead of the original numbers for the corresponding sentences from the paper]*. Phrase structure trees become thus a way of describing the factors which give rise to the interpretation." [p. 259]
> (i) She_i was jumping into the car when $Debra_i$ saw a large man lurking in the bushes.
> (ii) He_i hasn't contacted me, but I'm sure $John_i$ is back. [attributed to Mittwoch 1979]

[13] PP-modifier ambiguity is a source of a long and heated debate in the processing literature that tries to determine which factors determine the attachment site for the PP during online processing (e.g. Spivey-Knowlton & Sedivy 1995). Importantly, it is uncontroversial that cases of PP attachment such as (43) below are structurally ambiguous, and it is this part that we need for our demonstration.

Let us look at a well-known case of PP-modifier attachment (43) form the parsing literature.

(43) The cop saw the spy with the binoculars.

a. NP-modifier, low: The cop [$_{VP}$ saw [$_{NP}$ the spy [$_{PP}$ with the binoculars]]].
b. VP-modifier, high: The cop [$_{VP}$ saw [$_{NP}$ the spy] [$_{PP}$ with the binoculars]].

(43) is an ambiguous sentence: the PP *with the binoculars* can be understood as either modifying the NP *the spy* or as modifying the seeing event. These interpretations are instantiated as two different syntactic structures, with the PP either attached low to the NP *the spy* as in (43)a or high to the VP as in (43)b. Syntactically, either low or high attachment of the PP is legitimate, and the final attachment spot is determined by extra-syntactic factors, such as recency, plausibility, frequency, etc. For example, although (43) is globally ambiguous between an NP-modifier and an instrument readings, (44) and (45) that have a similar structure are not: in (44) the PP ends up being an NP-modifier and in (45) it is an instrument.

(44) The cop saw the cat with the kittens. [NP-modifier, low attachment]
(45) The cop saw the fingerprint with the microscope. [VP-modifier, high attachment]

Importantly, once the decision is made and the PP is attached to a certain position in the structure, it starts showing a characteristic behavior for a phrase in that position. For example, forming an identical type of question on the basis of (44) & (45) by fronting the N that immediately follows the verb yields differential results, as shown (46) & (47). This result is unsurprising given that the fronted phrase is a constituent in (45), but not in (44).[14]

(46) *What did the cop see [$_{NP}$ __ [$_{PP}$ with the kittens]] ?
(47) What did the cop see [$_{NP}$ __] [$_{PP}$ with the microscope]?

Now let us return to sentences with *when*-clauses. If *when* is syntactically ambiguous between a VP-modifier and a sentential modifier, then, similar to the PPs above, the placement of the *when*-clause may be decided based on extra-syntactic considerations. For example, if the event in the *when*-clause does not provide an appropriate temporal background for the first event, then the *when*-clause adjoins high, as a sister to the first clause. However, once the *when*-clause is attached to the tree, it behaves in accordance with its position in the tree. Specifically, coreference is ruled out by Principle C if the *when*-clause is subordinated to the first clause as in (39). If the *when*-clause adjoins as a clausal conjunct as in (33), then coreference is

[14] In English fronting part of the NP is prohibited, the so-called Left-Branch Extraction violation (Ross 1967).

not ruled out by structural considerations and is subject to discourse/pragmatic relations. Hence, the availability of coreference in sentences like (33) is not at odds with Principle C.

Summarizing the discussion, most of the cases in (30)-(33) that are claimed to be violations of Principle C are only problematic for Principle C circa 1981, and can be explained once appropriate revisions to Principle C, most notably Heim's (1982) and Reinhart's (1983, 1986, 2000) revisions, are taken into consideration. I conclude therefore that there is no strong evidence against Principle C.

2.4 A FUNCTIONALIST APPROACH TO COREFERENCE

The data from the previous section shows that in order to explain covaluation structural constraints on coreference must be placed in a broader context that includes semantic/pragmatic factors. The following question arises immediately: why not assume that all anaphoric relations are regulated by principles that are fundamentally semantic/pragmatic in nature?

More generally, the assumption that language does not need an independent structural level lies at the core of a range of functionalist theories of language (Chafe 1976, 1987, 1991, 1994; Fauconnier 1985; Fillmore 1977; Lakoff 1988; Langacker 1987, 1991 among others). There have been several studies of anaphora within functionalist framework (Karttunen 1976; Kuno 1987; van Hoek 1997 among others). Here I provide a summary of one of the most extensive of them that is due to van Hoek (1997), which has received significant attention in the recent literature (Tomasello 2004, Harris & Bates 2002). I will then try to show a type of data that is problematic for van Hoek's theory.

2.4.1 A Cognitive Grammar approach to coreference: van Hoek (1997)

Van Hoek (1997) presents an account of anaphora in English set in the framework of Cognitive Grammar (CG, Langacker 1986, 1987, 1991). There are three types of linguistic units in CG: phonological, semantic and symbolic, where a symbolic unit is a pairing of a phonological unit with a semantic unit. Crucially, in contrast to generative theories, CG does not posit an autonomous syntactic component, such as tree structures, or structural relations that are syntactic primitives, such as c-command. Van Hoek's account of anaphora abides by the postulates of CG: it lacks any syntactic principles and is executed in terms of semantic primitives that are already justified in the system, such as prominence, and linear word order and conceptual connectivity.

The core of van Hoek's approach to anaphora is the model of reference points and dominions. In this model nominals may serve as conceptual *reference points*, which in turn are associated with *dominions* – conceptual domains relative to

which the reference point is highly accessible and functions as a central part of the background context. Three factors can influence the choice of a reference point: prominence, linear order and conceptual connectivity (involving cause-effect and other pragmatically grounded factors that define the extent of the reference point's dominion). Another important component is seen in the principles of the Accessibility Theory (Givon 1983, Ariel 1988, 1990 *inter alia*), in which various nominal forms are claimed to differ according to the degree of accessibility of their respective referents in discourse. For example, the use of a full noun phrase indicates that its referent is currently of a low accessibility or salience, whereas the use of a pronoun indicates higher accessibility in the immediate context.

Illegal cases of pronoun-name coreference that are traditionally ruled out by Principles B and C of the Binding Theory are viewed here as resulting from violating semantic and pragmatic principles of organization of reference points and their dominions. The following principles describe regulations on anaphora in language (van Hoek 1997):

(48) (i) a full nominal cannot appear in the dominion of a reference point that it corresponds to;
(ii) an antecedent for a pronoun must be sufficiently salient within the context in which the pronoun appears so that it can plausibly be construed as a reference point with the pronoun in its dominion.

Sentences such as (49) are disallowed on coreference for both of the reasons in (48). First, a nominal form that signals low accessibility (i.e. the full noun phrase) is embedded in the dominion of a coreferential reference point of a higher accessibility (i.e. the pronoun). [The reference point is the subject of the main clause.] At the same time the antecedent for the pronoun is not a reference point in the dominion of the pronoun (the pronoun itself is the reference point).

(49) a. **He* loves *John*.
b. **He* loves *John's* mother.
c. **He* wanted Mary to phone *John*.

Whenever (48) is not violated, as in (50) where the subject is a full nominal which can be construed as a reference point for the pronoun, coreference is available. A major part of the theory, which I leave out from this review for reasons of space, is devoted to explaining why and how nominals other than an unambiguous reference point can serve as antecedents for a pronoun, as in (51). In a nutshell, explanations of these phenomena are derived via relations between heads and their complements or modifiers and via their interaction with the factors of conceptual organization mentioned earlier.

(50) a. *John* likes *his* cat.
b. *John* knows that Mary likes *him*.

(51) a. *John's* mother likes *him.*
b. Before *John* read the letter, *he* locked the door.

Before finishing, we must mention another type of information that is available in the grammar and influences anaphora-related decisions. Consider (52) below:

(52) **John* likes *him*.

According to (48), in (52) coreference between *John* and *him* should be available for the same reasons as in (50)b. Van Hoek then argues that what rules out coreference in (52) is the presence of a highly-entrenched schema in the language (the so-called 'reflexive schema' illustrated by (53)) which is always the first choice for expressing a coreference relation between verb co-arguments English. The reflexive schema trumps all other means of expressing coreference between co-arguments, including (52).

(53) *John* likes *himself*.

To summarize, constraints on coreference are largely explained through a combination of conceptual principles such as (48), principles of organization of complements and modifiers within a sentence and schemas that are entrenched in the language.

2.4.2 A Cross-linguistic Challenge for CG Approaches to Anaphora

Arguing against functionalist grammars is challenging in light of the gradedness of the principles that they rest on. Even so, however, I believe that van Hoek's theory can be shown to have problems in light of some cross-linguistic data. The specific argument that I will make concerns van Hoek's use of the Accessibility scale for explaining anaphoric regularities.

The Accessibility Scale (Ariel 1988, 1990) that ranks nominals in terms of the degree of accessibility of their referents in discourse is an integral part of the principles (48) that reflect the fundamentals of semantic-pragmatic organization. According to this scale a full nominal cross-linguistically marks lower accessibility of its referent in the discourse than an overt pronoun, which in turn signals lower accessibility than a null pronoun. The precise degree of accessibility associated with different nominals may vary across languages, but their relative positioning on the scale must remain constant.

With this in mind, let us take a look at the data from Hmong (Mortensen 2004). In Hmong (54) and (55) are a common way of expressing coreference or disjointness between the verb co-arguments. [(54) is similar to (18) and (22) from Thai and Vietnamese.]

(54) Pov_i yeej qhuas Pov_i.
Pao always praise Pao
Meaning: 'Pao always praises himself.'

(55) Pov_i yeej qhuas $nwg_{*i/j}$.
Pao always praise 3Sg
Meaning: 'Pao always praises him.'
*'Pao always praises himself.'

Hmong sentences show a pattern directly opposite from English: a nominal of lower accessibility, an R-expression, can be part of a dominion with a coreferent R-expression as a reference point, whereas a nominal of a higher accessibility, e.g. a pronoun, cannot. This challenges van Hoek's approach in one of two ways: either the hypothesis that there is a universally-valid Accessibility scale has to be dropped, or the way in which it is used in van Hoek's theory to explain the distribution of nominals across languages is inadequate. Note, that adjusting the Accessibility hierarchy would not give the desired result: commonsensically full nominals have to be at the lowest end of the scale given the richness of lexical content that they convey.[15] Thus, the Hmong facts in (54)-(55) are problematic for van Hoek's account.

Perhaps even more problematic is that Hmong has proper reflexives in its repertoire, as exemplified in (56).

(56) Pov_i yeej qhuas nwg tug $kheel_i$.
Pao always praise 3Sg CLF self
Pao always praises himself.

Given the availability of (56), van Hoek's theory would predict that Hmong has an entrenched 'reflexive schema' that should block the occurrence of a nominal of any other type in place of the reflexive in constructions like (56). Yet, as (54) shows, the reflexive can be substituted by a full nominal without affecting the interpretation.[16]

To conclude, Hmong poses two challenges for van Hoek's proposal. First, pronouns seem to be of lower accessibility than R-expressions, which is the direct opposite of what we have seen earlier for English. Second, the presence of a reflexive schema does not exclude the appearance of full nominals in the position of a reflexive.

[15] This does not mean that no other part of the scale may be paramererized cross-linguistically without undermining its pragmatic validity, e.g. it is imaginable that the relative positioning of overt vs. null pronouns may vary depending on how strongly pro-drop a language is.

[16] A similar point is made by Boeckx, Hornstein & Nunes (ms.), who note that data from Hmong and similar languages is problematic for generative theories of anaphora based on morpheme-specific operations that rely on a hierarchy of anaphoric dependence roughly of the form *reflexive>pronoun>R-expression* (e.g. Burzio 1991, Safir 2004).

I believe that the Hmong example demonstrates a general problem that van Hoek's approach faces in light of cross-linguistic variation. The source of the problem lies in the fact that nominals are taken at face value, i.e., the appearance of a particular nominal form (e.g. a reflexive, a null or overt pronoun, an R-expression) is regarded as a direct reflection of how accessible the respective referent is in discourse, i.e. the surface form must be in a one-to-one correspondence with pragmatics. Should a structural level be available in the system, the form of the nominal could be divorced from pragmatic factors such as the accessibility of referents. For example, in some cases the form may be regarded as (at least partially) defined by the structural position of the element, rather than a direct reflection of the accessibility of the reference.[17, 18]

One may reply to my criticism by saying that the same effect as above can be achieved via individual constructions. Traditionally, CG constructions have placeholders for nouns, verbs, etc. which are filled by individual forms from the lexicon. This is not sufficient in our cases, where a placeholder obtains its value as a function of another placeholder in the construction. Having placeholders as variables dependent on other placeholders would be a notational variant of an autonomous structure that would violate the Content Requirement (Langacker 1987) whereby no rules or structures which do not correspond to conceptual or phonological forms that actually occur in the language can be posited.

Summarizing, some aspects of cross-linguistic variation seem problematic for van Hoek's theory. Until these problems are resolved, I conclude that structural representations are indispensable in accounting for the distribution of anaphora.

2.5 CONCLUSION

I have reviewed two distinct ways in which anaphoric regulations may be conceptualized, either with (the Binding Theory) or without (the Cognitive Grammar approach to anaphora) appeal to structural notions. I argued that referential dependencies are regulated by a combination of constraints of a syntactic and discourse nature. With regards to Principle C, I showed that most objections pertain only to earlier formulations of Principle C and are no longer problematic once it is revised following suggestions by Heim (1982) and Reinhart (1983, 1986, 2000).

[17] This approach would actually allow to preserve the insights of the Accessibility scale, e.g. by allowing it to take into account positions of nominals.

[18] See Mortensen (2004) for an account of anaphora in Hmong, including an idea that full nominals in construction like (54) are treated as grammatical formatives rather than lexical elements in the spirit of Hornstein (1999).

The next chapter, Chapter 3, concerns the application of Principle C in real-time sentence processing. For the purposes of Chapter 3 it will suffice to formulate Principle C as in (57):

(57) Principle C: a dependent element (usually, a pronoun) cannot corefer with an R-expression that it c-commands.

The aim of the experiments presented in the next chapter is to investigate how speakers arrive at the conclusion that a given NP is an illicit antecedent for the pronoun. Are the Binding Principles, in particular, Principle C part of the parser's initial structure building procedure or they are applied at a later stage? I will argue that the parser is extremely efficient and accurate in implementing Principle C: it rules out a dependency between a pronoun and a following NP as soon as it can predict that the NP belongs to the domain that is c-commanded by the pronoun.

Appendix 2-A: Main definitions from the Binding Theory (Chomsky 1981)

The binding theory distinguishes three types of noun phrases listed in (5).

(1) Anaphors: reflexives, e.g., *myself, himself* and reciprocals, e.g. *each other*
Pronouns, e.g., *I, you, they, he*
R(eferring)-expressions, e.g., *John, the man*

Each of the nominal types is subject to one of three principles (Principles A, B and C) of the Binding Theory, summarized in (6).

(2) *Principles A, B & C of the Binding Theory (Chomsky 1981)*
Principle A: An anaphor must be bound within its governing category.
Principle B: A pronoun must be free within its governing category.
Principle C: R-expressions must be free (= not bound).

The formal definitions of the notions implicated in (6) are given in (3)-(6).

(3) Binding
α binds β if and only if
(i) α c-commands β, and
(ii) α and β are coindexed.

(4) C-command
α c-commands β if and only if
(i) neither α nor β dominates the other, and
(ii) the lowest branching node that dominates α also dominates β.

(5) Governing Category
α is the governing category for β if and only if α is the minimal category containing β and a governor of β, where α = NP or S.

(6) Government
α governs β iff (i) $\alpha = X^0$ and (ii) α c-commands β and if γ c-commands β then γ either c-commands α or is c-commanded by β.

Appendix 2-B. Reinhart (2000)

Reinhart's (2000) theory incorporates and extends ideas from her previous work (Reinhart 1983, Grodzinsky & Reinhart 1993, Reinhart & Reuland 1994) with the aim of broadening its empirical coverage to facts that were previously unexplained.

Reinhart rejects the definition of binding in terms of coindexation (or identity of variables) used in the standard theory both for conceptual reasons and because that view faces empirical problems and makes it difficult to state generalizations about binding. Reinhart replaces the coindexation-based definition of binding with a definition based on the traditional logical-syntax concept of binding. Importantly, Reinhart draws a clear distinction between binding and coreference (Reinhart, 1983) or later, between binding and covaluation (Reinhart 2000), which will be discussed in greater detail below.[19] Covaluation is not governed by considerations of the computational system but by an interface strategy, which requires reference-set computation (although not necessarily due to economy considerations).

(A)-Binding vs. Covaluation

Reinhart takes a view of pronouns as free variables. Thus, (5) corresponds to (6), with a pronoun substituted by a free variable *z*. Such a pronoun variable can be resolved either via *binding* or via *covaluation*. In a nutshell, (A-)binding is the logical relation - a relation between an operator and a variable, whereas covaluation is a relation between arguments – variables, or the indices of discourse entities. In the case of binding, the pronoun is bound by a λ-operator as in (6)b. In the case of covaluation, the pronoun variable receives a value from the discourse (e.g. from a file card that was created earlier) as in (6)c. The distinction between binding and covaluation holds regardless of the referential status of the relevant expressions.

(5) Lili thinks she's got the flu.
(6) a. Lili (λx (x thinks z has got the flu))
b. *Binding*: Lili (λx (x thinks x has got the flu))
c. *Covaluation*: Lili (λx (x thinks z has got the flu) & z =Lucie))

In (6), and in most other cases binding and covaluation yield an identical truth-conditional output, which was the reason why the two mechanisms were not separated in the standard approach. The necessity to distinguish binding from covaluation can be illustrated with (7) for which the two construals yield truth-conditionally-distinct outputs:

[19] The difference between coreference and covaluation for Reinhart is that covaluation, unlike coreference, is available regardless of the referential status of the antecedent, and shows up also in quantified contexts (following Heim 1993). In this work we use the term coreference is the most general sense to refer to cases in which the pronoun and noun end up having the same referent, either through coindexation or through covaluation.

(7) a. Only Lucie respects her husband.
b. *Binding*: Only Lucie (λx (x respects x's husband))
c. *Covaluation*:Only Lucie (λx (x respects her husband) & her = Lucie)

(7)b entails that unlike Lucie, all other women do not respect their husbands; (7)c entails that the only person who respect's Lucie's husband is Lucie herself. The same ambiguity as found in (7) can be found in sentences such as (8) with a non-referential antecedent (observation due to Heim 1993). (8) can be construed as entailing that every wife thinks that other wives do not respect their husbands (binding, (8)b), or that every wife thinks that other wives do not respect her husband (covaluation, (8)c). So, Reinhart (2000) no longer restricts covaluation to cases where the antecedent is a referential NP (cf. Renhart 1983).

(8) a. Every wife thinks that only she respects her husband.
b. *Binding*: Every wife (λx (x thinks that [only x respects x's husband]))
c. *Covaluation*: Every wife_i thinks that only she_i respects her_i husband.

A(rgument)-Binding

Reinhart defines binding differently from the traditional definitions by Chomsky (1981), as suggested by the comparison of (9) and (10). We will follow Reinhart in referring to her definition of binding as 'A(rgument)-binding' to distinguish from Chomsky's binding. Her definition of binding (10) rests on logical-syntax and does not appeal to the notion of c-command per se. Note, however, that if α A-binds β in accordance with (10), then α necessarily c-commands β, (since it is a sister of a node containing β). Thus, A-binding effectively recreates the configuration under which c-command obtains; moreover, c-command is relevant for the syntactic conditions under which λ-predicates can be formed (compositionality).

(9) Traditional definition of binding (Chomsky 1981):
α binds β iff α and β are coindexed, and α c-commands β.
(10) A-Binding (logical-syntax based definition, Reinhart 2000):
α A-binds β iff α is the sister of a λ-predicate whose operator binds β.

Next we will illustrate how Reinhart (2000) rules out configurations that are traditionally ruled out by Principles B and C of the Binding Theory (Chomsky 1981).

Reinhart argues that A-binding is subject to the general restrictions imposed by logical syntax (e.g. the same variable cannot be bound twice by different λ-operators) and also to the Chain condition (Reinhart & Reuland 1993, roughly condition B).

The properties of logical syntax explain the ungrammaticality of the coreference reading in well-known 'strong crossover' configurations, e.g. (11)a:

(11) a. Who did he say we should invite t?
b. who (λx (he said we should invite x))
b'. who (λx (he (λy (y said we should invite x))))
c. *Binding*: *who (λx (he (λx (x said we should invite x))))

In (11)b the trace is bound by the *wh*-operator, so it cannot be A-bound again by the pronoun ((11)b' is the full representation including also the VP λ-predicate). In (11)b', λy cannot bind x since x is already bound. If the pronoun were to A-bind the trace, that would lead to a nonsensical logical representation like (11)c where the same variable is bound by two operators. Note that this by itself does not exclude an alternative anaphora construal for (11) via covaluation (see later).

Principle B. The core cases such as (12) that are traditionally ruled out by Principle B are now subject to the Chain Condition (an equivalent of Condition B, sometimes referred to as Condition B, Reinhart & Reuland 1993). Unlike the condition on non-duplicate variable binding, the Chain condition is specific to the Computational System.

(12) a. Max touched him.
b. *Binding*: Max (λx (x touched x))

The construal in (12)b satisfies all requirements by the logical syntax. Rather, it is ruled out by the Chain condition. We will return to the issue why covaluation is illicit in (12) later.

Principle C. Reinhart claims that cases that are traditionally viewed as violations of Principle C straightforwardly follow from the conditions on logical syntax. An additional condition is needed for the purposes of excluding covaluation in a subset of configurations that are traditionally subject to Principle C.

Most straightforward cases that are subject to the traditional Principle C are now accounted for by the fact that NPs that cannot be transformed into variables cannot be A-bound (since by definition A-binding is variable-binding). This explains the impossibility of the coreference reading in sentences like (13)a: in (13)b *Lucie* is not the type of NP that can be bound by the λ-operator. Binding is ruled out even if the NP *Lucie* undergoes QR and therefore can be substituted by a variable as in (13)c: if so the trace of the NP *Lucie* is bound by the upper λ-operator and cannot be bound again by the pronoun.

(13) a. ***She*** said we should invite **Lucie**.
b. *Binding, no QR*: She (λx (x said we should invite Lucie))
c. *Binding, with QR*: Lucie (λy (she (λx (x said we should invite y))))
(14) a. ***She*** likes **every girl**.
b. *Binding*: Every girl (λy (she (λx (x likes y))))

Exactly the same logic as in (13)c is used in sentences like (14) in which the pronoun c-commands a Q-expression to correctly rule out coindexation between *she* and *every girl* (the same mechanism also accounted for the Strong Crossover effect in (11)). As before, we will consider why covaluation is not an option in (13)-(14) later. To conclude, the impossibility of binding in cases that are traditionally ruled out by Principle C is excluded in Reinhart's view by the rules of the logical syntax.

Covaluation

Now we are ready to discuss the other strategy of pronoun resolution – covaluation. Recall that covaluation was argued to be a separate strategy from binding and a necessary one based on sentences like (7) that are truth-conditionally ambiguous. Similarly, covaluation is permitted in the configuration of A-binding in (15) and gives rise to the 'strict' reading of these sentences.

(15) Covaluation permitted in a configuration of A-binding:
a. **Lili** thinks **she** has got the flu, and Max does too.
b. Only **Lucie** respects **her** husband.
c. Every wife thinks that only **she** respects **her** husband.

In Reinhart's theory, covaluation is subject to an interface strategy known as Rule I in (16).

(16) **Rule I** (an interface rule): - evaluation should proceed condition-by-condition
α and β cannot be covalued in a derivation D, if
a. α is in a configuration to A-bind β, and
b. α cannot A-bind β in D, and
c. the covaluation interpretation is indistinguishable from what would be obtained if α A-binds β. [To check this option, construct a comparison-representation by replacing β, with a variable A-bound by α.]

The conditions in Rule I are ordered as shown in (16) and must be evaluated one after the other. Rule I tells us that covaluation is unavailable if all three of the conditions (a)-(c) are met, and is available if at least one of the conditions is violated. We diagram the algorithm underlying Rule I in Figure 3.

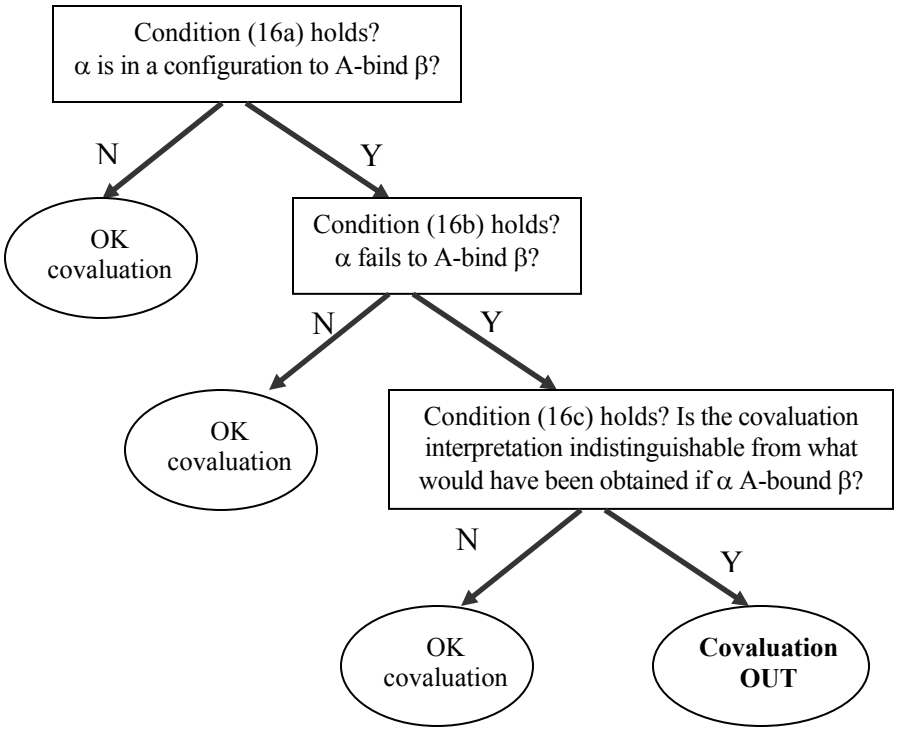


Figure 3. Diagrammed representation of Rule I from Reinhart (2000).

We are now ready to see how Rule I works at each step.

Covaluation as a result of failing at condition (a) of Rule I. (17) and (18) can be understood by the speakers as expressing coreference between *him* and *Max*. Yet, the coreference reading of (17) and (18) cannot be due to the binding, since the configuration for A-binding (see (10)) does not obtain. Covaluation is available as a result of not passing step (a) of Rule I, whereby covaluation is allowed when the pronoun and the name are not in an A-binding configuration.

(17) a. The woman next to him touched Max.
b. The woman next to **him** (λx (x touched **Max**))
(18) Max's mother loves him (he = Max).

Covaluation as a result of failing at condition (b) of Rule I. In (19) condition (a) holds, but condition (b) does not, because *Max* can and does A-bind *his*. Thus, covaluation is available as a result of not meeting the requirements of condition (b) of Rule I.

(19) Max loves his mother.

Covaluation as a result of failing at condition (c) of Rule I. Rule I also allows covaluation in sentences like (20), but in this case the verdict can only be made by

the third condition of the rule. This is because in (20) the first two conditions of Rule I hold. (16)a is true because *he* is in a configuration to A-bind *Max*. That (16)b holds – i.e. that *he* fails to A-bind *Max* – is illustrated in (20)b-b' for both QR and non-QR version of the sentence. Without QR, Max is not a suitable type of NP to be A-bound; under QR, binding is excluded since it violates the logical requirement of non-duplication of binders.

(20) a. Only he (himself) still thinks that Max is a genius.
b. *Binding*, no QR: Only he (λy (y thinks Max is a genius))
b'. *Binding, with QR*: Max (λy (Only he λx(x still thinks that y is a genius)).
c. *Covaluation*: Only he (λy (y thinks **Max** is a genius) & (he =Max))
d. *Hypothetical Binding*: Only he (λy (y thinks **y** is a genius) & (he =Max))

However, condition (c) is false since covaluation yields a truth-conditionally distinct output from binding, if the latter did obtain. (20)c is not equivalent to (20)d, in which binding hypothetically takes place: (20)c is true if everyone considers himself a genius, as long as no one but Max considers Max a genius, whereas (20)d is false in this situation.

Impossibility of coreference as a result of meeting all three conditions of Rule I. (21)a-(22)a (repeated from (11)a-(13)a) are cases in which coreference is excluded altogether. We have shown earlier why binding is illicit in these cases. Covaluation is ruled out by Rule I as a result of meeting all three conditions.

As shown above, the pronoun and the name in (21)-(22) are in a configuration that is necessary for A-binding (i.e. condition (a) of Rule I is met), but binding is excluded by a restriction on binding (i.e. condition (b) of Rule I is met). According to condition (c), we should next evaluate whether covaluation yields an indistinguishable output from (hypothetical) binding. Covaluation construals are given in (21)b-(22)b, and the corresponding binding comparisons are in (21)c-(22)c (the element β that is replaced by a variable bound by the element α according to condition (c) of Rule I is highlighted in bold in these construals) .

(21) 'Strong crossover' ((11)a)
a. Who did he say we should invite t?
b. Covaluation: who (λx (x said we should invite **x**)) *or* who (λx (x (λz (z said we should invite **x**))))
c. Binding-comparison: who (λx (x (λz (z said we should invite **z**))))

(22) Sentences traditionally excluded by Principle C ((13)a)
a. She said we should invite Lucie.
b. Covaluation: she (λx (x said we should invite **Lucie**) & (she = Lucie))
c. Binding-comparison: She (λx (x said we should invite **x**) & (**x** = Lucie))

(23) Sentences traditionally excluded by Principle B ((12)a)
a. Max saw him.

b. Covaluation: Max (λx (x saw **him**) & (him = Max))
c. Binding-comparison: Max (λx (x admires **x**)) & (**x** = Max)

In each case covaluation produces an indistinguishable output from the hypothetical binding construal used for comparison. Thus covaluation in (21)a-(23)a is ruled out.

3 CHAPTER 3. REAL-TIME PROCESSING OF LONG-DISTANCE DEPENDENCIES[20]

3.1 USING GRAMMATICAL CONSTRAINTS TO LIMIT THE SEARCH SPACE

One of the most fundamental properties of language is its ability to establish dependencies between its elements. By a dependency we will understand a relation that involves two (or more) elements in which one of the elements is licensed by another element. As shown below, such licensing may concern the appearance of the dependent element as a whole, the morphological make-up (e.g. number, gender, case) of the dependent element or its interpretation.

Some dependencies are strictly local (i.e. within the maximal projection of a given head) and are instantiated via a head-specifier or a head-complement relation. An example of a head-complement relation is case assignment from a verb to its object: different verbs may choose to assign different cases to their object, as in (1). Head-specifier relations can be exemplified by gender and/or number agreement between a determiner or an adjective and the noun inside NPs or DPs, as in (2).

(1) a. Ivan uvidel devochku. [Russian]
Ivan saw girl.Gen
b. Ivan pomog devochke.
Ivan helped girl.Dat

(2) a. une grande pomme [French]
a.fem big.fem apple.fem
b. un grand auto
a.masc big.masc car.masc

Most morpho-syntactic phenomena, similarly to those illustrated above, are based on relations between elements that are strictly local and can, depending on the theory, either be stored as part of the lexical entry of the relevant head or be derived in L(exical)-syntax. In addition to such cases languages also have phenomena that are based on long-distance relations. Elements that are involved in a long-distance dependency need not be in a specifier-head or specifier-complement configuration. For example, a *wh*-dependency between the *wh*-element, also sometimes referred to as the (*wh*-)filler, and its gap position in (3)a involves the Specifier of the CP and the complement position of the verb.[21] Even

[20] The experiments reported in this chapter were conducted in collaboration with Ellen Lau, Moti Lieberman, Colin Phillips and Masaya Yoshida.

[21] Whether *wh*-movement involves gaps or should be represented by directly associating the moved wh-filler with the verb is subject of an ongoing debate in the literature [see Pickering 1993 for a review]. In this work although we will use the term filler-gap dependency in reference to *wh*-dependencies, it should not be taken as committal to either theoretical position. For our current purposes the differences between these theories will not be significant.

further, members of the dependency need not even be located in the same clause and can potentially be separated by an indefinite amount of intervening material. For example, in (3)b,c the *wh*-element is licensed by a gap position in another clause. (3)d is added as evidence that a *wh*-element must have a licensor –a *wh*-element is illicit in the absence of a gap.

(3) a. What did John eat __?
b. What did Mary see that John ate __?
c. What did Bill say that Mary saw that John ate __?
d. *What did John eat an apple?

Another type of long-distance dependency involves negative-polarity items (NPIs) such as *anything*. NPIs must be licensed by a negative element as shown in (4)a (compare with (4)b). In English the negative element may belong to a different clause, as in (4)b. (4)c demonstrates that an NPI cannot be used unless it is properly licensed.[22]

(4) a. John did**n't** steal **anything**.
b. Mary did**n't** say that John stole **anything**.
Meaning: 'It is not true that Mary said that there was something that John stole.'
c. *John stole anything.

(5)a is an example of a referential dependency between the pronoun *he* and its antecedent *John*, in which the pronoun picks up its reference through the main subject NP.

(5) a. $John_i$ says that he_i wants to take a few days off.
b. Every $employee_i$ says that he_i wants to take a few days off.
c. John/every $employee_i$ says that the administration knows that he_i wants to take a few days off.
d. John/every $employee_i$ says that his wife believes that the administration knows that he_i wants to take a few days off.

That pronouns indeed may enter into a dependency relation with their antecedents rather than accidentally corefer with them becomes unambiguous when an antecedent is a quantifier phrase. (5)b is different from (5)a in that the main subject *John* is replaced by a quantifier phrase *every employee*. As a result, not only does *he* in (5)b range over all members of the set of employees, but also *he* cannot pick an arbitrary referent from that set: (5)b can only mean that each of the employees expressed his own wish to take a few days off, as diagrammed in Figure 4a, and it cannot mean that every employee expressed a wish of some

[22] The NPI reading of *anything* that is in question here should not be confused with its free-choice reading.

other employee to take a vacation (Figure 4b). This shows a clear binding relation between a pronoun and its antecedent.

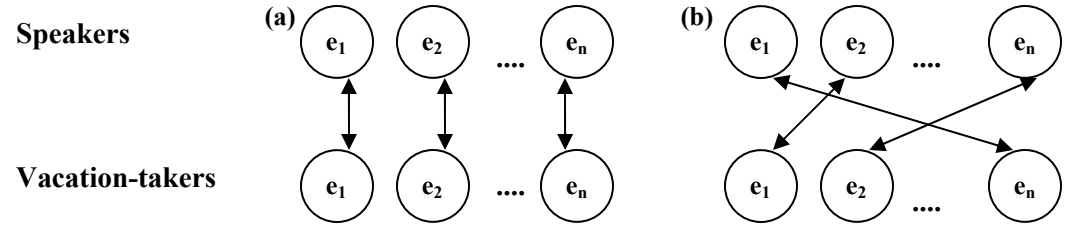


Figure 4. The logically possible relations between speakers (upper row) and vacation-takers (lower row). e_1, e_2, .. e_n represent individual employees. Although both (a) and (b) are logically possible, a sentence like (5)b can only have the interpretation that corresponds to the diagram (a).

Referential dependencies differ from dependencies that involve *wh*-elements or NPIs in that what is licensed through the dependency is not the presence of the dependent element (the pronoun), but rather an interpretation that that pronoun may receive.[23] (5)c,d illustrate that referential dependencies can span multiple clauses. Other examples of long-distance dependencies attested in human languages include at least scrambling, control and raising.

Having shown that long-distance dependencies exist, i.e. that there are classes of elements that require licensing and can be licensed non-locally, we are ready to discuss another fact about long-distance dependencies. This fact is that no matter how long or unbounded these dependencies may be, they all are subject to conditions on well-formedness. A sentence may contain both members from a dependency and yet not be licit. For example, (6) contains a *wh*-element and a gap position, yet the sentence is unacceptable. In (7) the NPI *anything* fails to be licensed by the negation. In (8) *he* cannot receive its referent by relating to the quantifier expression *every employee* and must instead refer to some other (unspecified) male referent.

(6) ***What** did the boy who bought __ likes Mary?
intended meaning: 'What did the boy who likes Mary buy?'

(7) *The girl that he did**n't** like said that John stole **anything**.
intended meaning: 'The girl that he didn't like say that John stole something.'

(8) *The letter signed by **every employee$_i$** said that **he$_i$** was unhappy about the situation in the office.
intended meaning: 'The letter that was signed by every employee said that that employee was unhappy.'

[23] This is a description, rather than a claim about the nature of the relation between bound and non-bound pronouns (i.e. whether the two types of pronouns are the same or different formatives).

(9) ***He$_i$** said that **John$_i$** is tired.
intended meaning: ‘ John said that he (=John) was tired’.

In the last forty years or so, linguists have made enormous headway in discovering, formulating and formalizing conditions on the well-formedness of long-distance dependencies. How successful such formalizations are and whether they rest on principles that are independently known to be operative in languages varies for various types of dependencies. A significant body of cross-linguistic research has revealed which constraints on dependencies are attested universally and which are subject to variation (e.g., Rizzi 1982; Kayne 1984; Travis 1984).

Long-distance dependencies also pose interesting questions from a psychological perspective. One might wonder about the psychological status of illicit structures such as (6)-(8) and the constraints that rule them out. Are mental representations like these ever generated? How are constraints on long-distance dependencies represented? And more generally, what makes a given structure ungrammatical?

One way of defining an ungrammatical representation is to assume that it was generated at some stage and subsequently ruled out by a grammatical filter. This should be reminiscent of the move made by Optimality Theory (OT, Prince & Smolensky 1993). In this theory the generative component generates an unlimited number of candidates (‘Richness of the Base’), which are subsequently evaluated against a set of (ranked) grammatical constraints. As a result of constraint application one candidate that violated the least number of constraints on the top of the constraint hierarchy is the winner, i.e. the well-formed output. The rest of the candidates that were less optimal than the winning candidate can be classified as ungrammatical (possibly, to a various degree of ungrammaticality). Similarly, in Government and Binding theory (GB theory, Chomsky 1981) most illicit configurations are ruled out in by various filters that apply to outputs. For example, filters known as the Binding Principles exclude certain patterns of coreference that would otherwise be logically possible.[24] Thus, both in OT and GB the generative component overgenerates (to a varying degree) and some of the generated candidates are filtered out at a later stage.

A qualitatively different perspective on illicit representations is that ill-formedness is the result of an inability to generate a structure in the generative component. In this view there is no need for an independent system of grammatical filters that apply to outputs. Generalized Phrase Structure Grammar (GPSG, Gazdar 1982) may serve as an example of a theoretical model that shares a similar assumption: filters are applied all at once during the generation stage, hence no illicit structure is ever generated. Another example of this sort can be some versions of Minimalism (Chomsky 1995): illicit candidates never survive through the generation stage as they are non-compliant with the requirements of

[24] Theoretical models are given purely to exemplify different intuitions, and are not treated as models of sentence generation.

the model (e.g., failure to check a feature, violations of Shortest Move, Merge over Move, etc.).[25]

In this chapter we will investigate the representation of constraints on one particular type of long distance dependency – a referential dependency between a pronoun and a name. The issue of how illicit referential dependencies are ruled out will be explored from a parsing perspective. We will attempt to track how and when non-local referential dependencies are established in real time and how constraints on such dependencies apply during online processing. The focus of the discussion will be on backwards anaphora, a referential dependency in which the dependent element precedes its licensor. Before we turn to the discussion of backwards anaphora, however, we will discuss what is known about online processing of a different type of long-distance dependency, *wh*-dependencies (section 3.2). The reason for looking at *wh*-dependencies is that there are a number of properties they share with backwards anaphora (section 3.3). Section 3.4 will review the existing literature on the processing of referential dependencies. Section 3.5 is a presentation of a series of experiments aimed at establishing which mechanisms underlie the formation of a backwards anaphora dependency during online processing, and whether grammatical constraints on backwards anaphora, in particular Principle C, are immediately respected by the parser.

3.2 REAL TIME PROCESSING OF WH-DEPENDENCIES

Fodor (1978) lays out two possibilities with respect to how filler-gap dependencies might be parsed. She contrasts a 'filler-driven' view of *wh*-dependency formation whereby the parser starts building a dependency upon encountering a *wh*-filler with a 'gap-driven' view whereby dependency formation is initiated by a gap. Crain & Fodor (1985) and Stowe (1986) were among the first sentence-processing studies of *wh*-dependencies that explored this issue. In particular, they showed that *wh*-dependency formation is 'filler-driven', i.e. that upon encountering a filler, the parser starts an active search for a gap position in an attempt to complete a filler-gap dependency as soon as possible (also known as the 'Active Filler Strategy', Frazier & Clifton 1989). Moreover, Stowe (1986) argued that active dependency formation is immediately constrained by grammatical constraints on *wh*-dependencies, i.e. the parser does not generate representations that would incur island constraints violations.

25 One must note that there is overgeneration in Minimalism, in a slightly different form. One of the most crucial notions is use of economy principles, whereby a licit configuration is established as a result of comparing several derivations, some of which are less economical, hence illicit.
I will use this opportunity to once again emphasize a purely illustrative status of the comparisons used to describe opposite psychological models.

Stowe (1986) compared reading times for sentences that contain a licit *wh*-dependency ((10)a) with their non-*wh* controls (10)b using a word-by-word self-paced reading technique.

(10) a. My brother wanted to know who Ruth will bring us home to __ at Christmas.
b. My brother wanted to know if Ruth will bring us home to Mom at Christmas.

Stowe found a slowdown in reading times at the NP *us*, the object position of the transitive verb *bring*, in (10)a compared to (10)b. Stowe took the slowdown to reflect the parser's surprise at encountering the overt NP *us* in the position where the parser postulated a gap associated with the *wh*-filler ('Filled Gap Effect'). She argued that the postulation of a gap in this position is driven by the parser's desire to complete an incomplete *wh*-dependency as soon as possible and, crucially, before the parser has any bottom-up evidence that the verb does have a missing argument. In other words, the *wh*-filler initiates *an active search* for a licensor by predicting and expecting a licensor in every relevant upcoming position (i.e. an A-position). The same effect was found by Crain & Fodor (1985).[26]

The second question addressed by Stowe is whether active *wh*-dependency formation is sensitive to island constraints on movement. If the process of online dependency formation immediately adheres to grammatical constraints, the parser should not attempt to postulate a gap inside an island, even despite its strong intention to complete the dependency as soon as possible. On the contrary, if initial dependency formation is not regulated by grammatical constraints, we should find a Filled Gap Effect at any NP occupying a potential gap position either outside or inside an island.

Stowe's Experiment 2 tested sentences such as (11)-(12). The position of interest is the first argument position that follows the *wh*-filler, i.e. the complement of the preposition *about.* In (11)a this position is inside a subject NP, an island for *wh*-movement, whereas in (12)a the same position is not inside an island.[27] Hence, we

[26] If a *wh*-filler searches for an antecedent in every licit position, as claimed by an active search mechanism, the main subject position, the first position on the parser's left-to-right search path, should yield a filled-gap effect whenever it is filled, i.e. everywhere but in subject *wh*-questions. However, the filled-gap effect in the subject position was not observed in Stowe's (1986) and other studies, which cast doubts on the notion of an active search. However, Lee (2004) found a subject filled-gap effect in sentences where the *wh*-filler and the subject position are separated by intervening material, thus suggesting that in previous studies in which the subject position directly followed the *wh*-filler the effect at the subject was too weak to notice (e.g. due to ease of reanalysis).

[27] Movement out of subject NPs is generally not allowed, as exemplified in (i). (ii) shows that movement of the same NP is allowed when it is an object.

(i) [The book about polar bears] was sitting on the desk.
*What was [the book about _] sitting on the desk?

expect a Filled Gap Effect in both (11) and (12) if the parser initially disregards island constraints. On the other hand, if the parser respects island constraints during online processing, we should find a Filled Gap Effect in (12), but not in (11):

(11) a. The teacher asked what [$_{NP}$ the silly story about Greg's older brother] was supposed to mean __.
b. The teacher asked if the silly story about Greg's older brother was supposed to mean anything.

(12) a. The teacher asked what the team laughed about Greg's older brother fumbling__.
b. The teacher asked if the team laughed about Greg's older brother fumbling the ball.

Stowe found a Filled Gap Effect at *Greg's* in (12), but not in (11). The absence of the Filled Gap Effect inside an island suggests that the parser respects constraints on *wh*-movement even at the earliest stages of filler-gap dependency formation and avoids generating representations that are ungrammatical due to island constraints.

Stowe's results were confirmed in several subsequent studies, including an eyetracking study by Traxler & Pickering (1996) that used a different type of experimental paradigm based on manipulation of semantic plausibility of *wh*-filler gap relation. Traxler and Pickering argue for island sensitivity of the parser by showing that the plausibility effect that is present with well-formed dependencies disappears when a dependency is ungrammatical due to an island constraint. They investigated pairs of sentences like (13)-(14) with relative clauses headed by a noun (e.g., *the book* or *the city*) which is either a plausible or an implausible object for the verb *wrote* which is the first possible head that in the input following the noun.

(13) *No-island conditions, plausible/implausible*
We like **the book/ the city** that the author **wrote** unceasingly and with great dedication about __ while waiting for a contract.

(14) *Island conditions, plausible/implausible*
We like **the book/ the city** that the author [who **wrote** unceasingly and with great dedication] saw __ while waiting for a contract.

The plausibility effect, i.e. longer reading times for the implausible candidate than for its plausible counterpart, that is present in (13), in which the dependency between the head of the relative clause and *wrote* is grammatical, disappears in (14), in which such a dependency is ruled out by a grammatical constraint (Ross 1967, *wrote* is inside a subject relative clause).

(ii) John read [the book about polar bears].
What did John read [the book about _]?

Summarizing Stowe (1986) and Traxler & Pickering (1996), upon encountering a *wh*-filler, the parser initiates an active search for a gap position and this process is grammatically constrained.[28]

3.3 WH-DEPENDENCIES VS. BACKWARDS ANAPHORA

Let us start by reviewing the main properties of w*h*- and referential dependencies.[29] As we said earlier, a dependency can be viewed as a relation between a dependent element (a *wh*-word or a pronoun) and its licensor (a gap or an antecedent). The *wh*-word needs to find a gap to receive a thematic role. The pronoun needs an antecedent to obtain a referential interpretation. Neither *wh*- nor referential dependencies are clause-bound. In either case, the dependent and the licensor need not belong to the same clause and may be separated from each other by an indefinite number of other clauses, as in (15)-(16). Both types of dependencies are subject to grammatical constrains on well-formedness, which are island constraints for *wh*-dependencies and binding principles for referential dependencies.

(15) a. Who does John know that Mary kissed __ ?
b. Who does John know that Bill said that Mary kissed __ ?
(16) a. John$_i$/every boy$_i$ knows that Mary likes him$_i$.
b. John$_i$/every boy$_i$ knows that Bill said that Mary likes him$_i$.

One difference between *wh*- and referential dependencies which is worth pointing out concerns the order of the dependent and the licensor. Cross-linguistically, if a *wh*-element is displaced it moves leftward so that on the surface it linearly precedes its gap position (obviously in wh-in-situ languages the *wh*-element and its gap position overlap).[30] In contrast, referential dependencies show more variation and allow both orders. The most typical case is *forwards anaphora* as in (16) in which the pronoun refers back to a previously mentioned referent, i.e. the dependent element linearly follows its licensor. The other possibility is *backwards*

[28] To say that every study of *wh*-dependencies showed immediate respect for island constraints is perhaps an overstatement. Phillips (submitted) reviews the literature on the subject of island constraints in online processing. He concludes that generally the parser never postulates a gap inside an island unless that gap my turn out to be parasitic and be saved later by a licit gap, with the possible exception of some complex NP islands that do not clearly support parasitic gaps yet seem to be violated.

[29] We will concentrate on referential dependencies that involve pronouns and leave out dependencies involving reflexives which are generally shorter and more restricted.

[30] A potential counterexample to this generalization is American Sign Language which has sometimes been argued to have rightward *wh*-movement (Neidle, Kegl, Bahan, Aarons & MacLaughlin 1997). Another potential counterexample is from Dhivehi, an Indo-Aryan language of the Maldives, an SOV language in which a *wh*-word may occur either *in situ* or in a special postverbal position (Cain & Gair 2000). I thank Brian Dillon for bringing the Dhivehi case to my attention.

anaphora in which the pronoun linearly precedes its antecedent. Despite being used less frequently than forwards anaphora, backwards anaphora nevertheless is used to a certain extent in most languages. In many of the environments where backwards anaphora is used it is not at all marginal, e.g. (17):

(17) a. The book that she$_i$ read upset Mary$_i$.
b. Before he$_i$ dialed the number, John$_i$/every boy$_i$ re-read the letter.

Given the various similarities between *wh*- and referential dependencies we would like to investigate whether their real-time processing is handled by similar parsing mechanisms. Since linear order is of crucial importance in processing, we will concentrate on comparing *wh*-dependencies with backwards anaphora, in which the dependent element precedes its licensor. The advantage of backwards anaphora for the investigation of online dependency formation is that it is possible to isolate the parser's treatment of each NP that follows the dependent element. I will return to this point later in section 3.4.2.

One question that can be asked of referential dependencies is whether their processing involves the same active search mechanism that applies during the processing of *wh*-dependencies. As shown by Stowe (1986) and Crain & Fodor (1985), the parser starts an active search for the gap position as soon as it encounters a *wh*-filler. We would like to see whether cataphoric pronouns also trigger an active search for an antecedent.[31] Such a conjecture is definitely plausible given that the primary function of the pronoun is to refer to an entity in the discourse context. However, unlike *wh*-fillers that require a gap position, the necessity to have an antecedent is not a grammatical requirement for the pronoun. Sentences or discourse fragments where a pronoun is ultimately left unresolved as in (18) are pragmatically infelicitous and perhaps not as informative as they could be, yet grammatically well-formed (compare with an unacceptable (19) which lacks a gap for a *wh*-filler). In cases where the pronoun lacks a licit overtly mentioned antecedent, it is taken to refer to an unspecified discourse referent.

(18) He said that it was going to rain.
(19) *What did you eat an apple?

If pronouns do pattern with *wh*-fillers and trigger an active search for an antecedent in spite of the fact that it is not a grammatical requirement, we will be in a position to ask a second, more interesting question that concerns the real-time application of grammatical constraints on long-distance dependencies. It has been shown for *wh*-dependencies that despite the parser's strong desire to find a gap as soon as possible, it avoids even temporarily postulating a gap inside an island. We would like to know whether the binding principles also have the status of

[31] For current purposes we define a cataphoric pronoun as a pronoun that occurs in the context without any previous referents and that can only be referentially resolved by an NP that linearly follows the pronoun.

constraints on structure generation. If they do, we would expect that during an active search for an antecedent triggered by a cataphoric pronoun the parser never includes positions that would violate the binding constraints in its list of candidate antecedents. Our experiments will specifically concern Principle C, which applies in sentences like (20) and rules out coreference between the pronoun and the name (sentences in (20) could have been instances of licit backwards anaphora if it were not for Principle C).

(20) a. *He$_i$ likes John$_i$.
b. *He$_i$ said that John$_i$ likes wine.
a. *He$_i$ drank beer while John$_i$ watched a soccer game.

Before going on, let us clarify our terminology. Following the tradition established in the parsing literature, our use of the term 'antecedent' in relation to an NP will not necessarily entail that that NP is a licit licensor for the dependent element in question. Instead it will be used as a shortcut for 'a candidate antecedent' which could become a licit antecedent provided that it (i) occurs in a licit structural position, and (ii) matches the pronoun in its morphological features (gender, number). An NP is a 'binding-accessible antecedent' for the dependent element if it occurs in a position that is an appropriate antecedent position from the perspective of the binding theory. An NP is a 'binding-inaccessible antecedent' if its structural position is ruled out by a grammatical constraint on coreference. Similarly, the terms forwards anaphora or backwards anaphora pertain to the linear order of the dependent element in relation to other R-expressions in the sentence and/or context and do not entail that the sentence contains a well-formed referential dependency.
In section 3.4 I provide a review of the existing literature on the time-course of application of the binding principles during processing of referential dependencies. In section 3.5 I discuss our experiments that were designed to test the time-course of application of Principle C in the real-time processing of sentences with backwards anaphora.

3.4 PROCESSING OF REFERENTIAL DEPENDENCIES

3.4.1 Previous research on Principles A & B in online processing of forwards anaphora constructions

The studies reviewed in this section concern the application of Principles A and B during the processing of forwards anaphora. They lay the groundwork for testing the issue concerning the time-course of application of Principle C in backwards anaphora in at least two ways. First, they can serve as an introduction to psycholinguistic methods that are used for testing anaphora-related issues and

demonstrate a range of effects that can be obtained as a result of anaphora resolution. Second, they provide information on the time-course of application of Principles A and B in forwards anaphora, which is relevant for a broader discussion on how backwards and forwards anaphora compare in terms of parsing or whether different principles of the Binding Theory have the same status for the parser.

All studies discussed in this section, with the exception of Nicol & Swinney (1989), used reading time measures (either eye movements during reading or self-paced reading times) to investigate the time-course of application of Principles A and B during online processing. Although the experiments differ in various ways, they all share a key methodological insight that is employed for testing the issue. This insight has to do with an implicit assumption which is present in every theory of referential dependencies: a distinction between the position of a certain NP in the structure (either syntactic or prominence-based) and the 'content' of that position in terms of its morphological and/or lexical properties. In order to know whether the parser considers a certain accessible or inaccessible antecedent as a licensor for the dependent element, the lexical properties of this antecedent are manipulated so that it fits or does not fit the dependent element. For example, a morphological feature (usually gender or number) of an accessible or inaccessible antecedent can be manipulated such that it either matches or mismatches the features of the dependent element. An effect of such a manipulation on the processing of the dependent element (in the form of an increase or decrease in reading times) is taken to signal that the parser at least temporarily considers a dependency between the antecedent and the dependent element; the lack of such an effect is considered as evidence for the parser's lack of interest in the respective position.

We now turn to a review of individual studies.

3.4.1.1 Sturt (2003)

Sturt (2003) used eye-tracking to investigate online processing of sentences with reflexives. The aim of the study was to determine at which stage of processing Principle A applies. According to the *binding-as-early-filter* hypothesis Principle A applies immediately and supervises formation of the initial set of antecedent candidates for the reflexive. This view contrasts with the *binding-as-late-filter* hypothesis whereby Principle A applies at a delay to a set of pre-chosen candidate antecedents, and with the *defeasible filter* hypothesis whereby Principle A applies at the initial stage, but may later be reconsidered and violated due to other, e.g. discourse-related, factors.

In Experiment 1 participants read sentences with forwards anaphora where the reflexive was preceded by an NP that was a binding-accessible antecedent for the reflexive (*the surgeon* in (21)), and another NP that was an inaccessible

antecedent for the reflexive (*he* or *she* in (21)) and that was the most salient entity in the discourse context. Experimental sets such as (21) were formed by crossing the antecedent type (accessible vs. inaccessible) and gender congruency with respect to the reflexive (match vs. mismatch).[32]

(21) a. *Accessible-match/Inaccessible-match*
Jonathan was pretty worried at the City Hospital. **He** remembered that **the surgeon** had pricked **himself** with a used syringe needle. There should be an investigation soon.
b. *Accessible-match/Inaccessible-mismatch*
Jennifer was pretty worried at the City Hospital. **She** remembered that **the surgeon** had pricked **himself** with a used syringe needle. There should be an investigation soon.
c. *Accessible-mismatch/Inaccessible-match*
Jennifer was pretty worried at the City Hospital. **She** remembered that **the surgeon** had pricked **herself** with a used syringe needle. There should be an investigation soon.
d. *Accessible-mismatch/Inaccessible-mismatch*
Jonathan was pretty worried at the City Hospital. **He** remembered that **the surgeon** had pricked **herself** with a used syringe needle. There should be an investigation soon.

Sturt found a **gender-mismatch effect (GMME)** at the reflexive as a result of gender-manipulation in the accessible antecedent: the first-fixation and first-pass reading times were longer when the gender of the anaphor mismatched the stereotypical gender of the accessible antecedent (e.g., *surgeon ... herself*) than when they matched (e.g., *surgeon ... himself*). However, reading times at the same region did not differ reliably as a function of whether the inaccessible antecedent matched the anaphor. The early effect of the inaccessible antecedent was absent despite the fact that that antecedent was the most salient entity in the discourse context. However, there was a significant effect of the inaccessible antecedent (in addition to the effect of the accessible antecedent) in measures indicative of later processing, namely, second-pass reading times at a later region (*used syringe*).

In Experiment 1 the inaccessible antecedent was put into focus by introducing it as the subject of the first sentence and re-introducing it by the pronoun subject of the 2nd sentence. This was done to maximize the chance of the inaccessible antecedent having an early effect on the processing of the reflexive. Nevertheless,

[32] Note that due to using an occupation NP as the accessible antecedent, in all conditions, including the 'accessible mismatch' (21)c,d, the reflexive has a licit antecedent. Sentences (21)c,d may originally seem ill-formed due to a mismatch between the feminine gender of the reflexive *herself* and the stereotypical masculine gender of the accessible antecedent *the surgeon* (the so-called 'stereotypical gender effect'). Eventually though, after the speaker overcomes the stereotypical gender bias and allows both male and female candidates into the set of surgeons, (21)c,d become licit.

it was the accessible antecedent that was always linearly closest to the reflexive, which could be the reason for the effect. Experiment 2 was designed to test for this possibility. In Experiment 2 the antecedent that was inaccessible due to Principle A was linearly closest to the pronoun, as shown in (22).

(22) a. *Accessible-match/Inaccessible-match*
Jonathan was pretty worried at the City Hospital. **The surgeon** who treated **Jonathan** had pricked **himself** with a used syringe needle. There should be an investigation soon.
b. *Accessible-match/Inaccessible-mismatch*
Jennifer was pretty worried at the City Hospital. **The surgeon** who treated **Jennifer** had pricked **himself** with a used syringe needle. There should be an investigation soon.
c. *Accessible-mismatch/Inaccessible-match*
Jennifer was pretty worried at the City Hospital. **The surgeon** who treated **Jennifer** had pricked **herself** with a used syringe needle. There should be an investigation soon.
d. *Accessible-mismatch/Inaccessible-mismatch*
Jonathan was pretty worried at the City Hospital. **The surgeon** who treated **Jonathan** had pricked **herself** with a used syringe needle. There should be an investigation soon.

Nevertheless, there were no significant effects of the inaccessible antecedent in either early or late measures in Experiment 2. Similar to Experiment 1, there was a reliable effect of the gender congruency of the accessible antecedent.

Summarizing the results, early measures from both experiments suggest that the initial set of antecedents for the reflexive excluded candidates that are violators of Principle A. In Experiment 1 those candidates seemed to be subject to consideration at a subsequent point. Based on these combined results Sturt concluded that Principle A acts as an early but defeasible filter.

A somewhat stronger interpretation of the results is to say that they are largely compatible with a stronger view of Principle A, i.e. that it is an early filter. Sturt's conclusion on the defeasibility of Principle A implicitly weighs the effect of the inaccessible antecedent on processing at the same level as the effects from the accessible antecedent. It seems, however, that the influence of the inaccessible antecedent was much weaker than that of the accessible antecedent. Unlike the effect of the accessible antecedent that was found in a number of early measures (first fixations, first-pass reading times, regression-path times) at the reflexive and the following spill-over region, the effect of the inaccessible antecedent was much more localized (pre-final region only) and appeared only in one measure (second-pass reading times). Moreover, there was no effect of the inaccessible antecedent in Experiment 2 although, arguably, the degree of discourse prominence of that antecedent in Experiment 2 was lower than in Experiment 1. Furthermore, somewhat surprisingly, the GMME from the inaccessible antecedent was entirely

driven by the accessible-match pair, in which the inaccessible-mismatch condition took longer than the inaccessible-match condition. This result is counter-intuitive (although not impossible to explain) even if we admit that an inaccessible antecedent does influence later processing, since it suggests that the influence is stronger when a licit antecedent for the reflexive could be found right away than when it required overcoming a stereotypical gender bias.

3.4.1.2 Nicol & Swinney (1989)

Nicol & Swinney (1989) (also Nicol 1988) studied the time-course of application of syntactic constraints on coreference using a cross-modal priming technique.

While listening to sentences with forwards anaphora such as (23), participants had to perform a lexical decision task to a visual probe that appeared immediately following the reflexive or the pronoun (marked by ^ in (23)). Target probes were semantically related to one of the three NP's preceding the pronoun, e.g. *glove* for *the boxer*, *snow* for *the skier* or *nurse* for *the doctor*. Each of these NP's was either an accessible or an inaccessible antecedent for the pronoun or reflexive. In (23)a *the doctor* is the only binding-accessible antecedent for the reflexive, in accordance with Principle A of the Binding Theory. In (23)b, on the other hand, *the doctor* is an inaccessible antecedent for the pronoun *him* due to Principle B, whereas *the boxer* and *the skier* are accessible antecedents.

The authors reasoned that if grammatical constraints are applied at a delay, then all licit and illicit antecedents should be considered by the parser as candidates upon encountering a referentially dependent NP. On the other hand, if binding principles apply early to restrict the initial set of candidate antecedents, then inaccessible candidates will be inactivated and 'invisible' to the parser.

(23) a. The boxer$_i$ told the skier$_j$ that the doctor$_k$ for the team would blame himself$_{*i/*j/k}$ ^ for the recent injury.
b. The boxer$_i$ told the skier$_j$ that the doctor$_k$ for the team would blame him$_{i/j/*k}$ ^ for the recent injury.

The authors found a significant priming effect for semantic associates of the NP *the doctor*, but not for those of *the boxer* or *the skier* in (23)a. On the contrary, in (23)b there was no priming effect for targets associated with *the doctor,* but there was a significant priming effect for targets associated with both *the boxer* and *the skier*.

The priming pattern obtained in the experiment cannot be explained as a recency or residual activation effect, and must be attributed to the fact that activation of a candidate set of antecedents for a pronoun or reflexive takes place in compliance with syntactic constraints on coreference. In other words, the initial set of candidate antecedents contains all and only those referents that stand in an appropriate syntactic relation to the referentially dependent NP.

3.4.1.3 Badecker & Straub (2002)

Badecker & Straub (2002) (henceforth B&S) investigated the time-course of application of Principles A and B using a different technique, a self-paced reading task. Similarly to Sturt's experiment, they studied whether the processing of the pronoun/reflexive was affected by gender manipulation at preceding accessible or inaccessible antecedent position. Following now familiar logic, B&S hypothesized that there should be no effect as a result of the gender manipulation in the binding-inaccessible position if the binding constraints acted as inviolable constraints on structure-building which exclude binding-inaccessible positions from consideration at the earliest possible stage.

The dependent element in Experiment 1, a pronoun, was preceded by an accessible antecedent (the main subject) and an antecedent that was inaccessible due to Principle B (the embedded subject). The gender of either antecedent was manipulated so that it either matched or mismatched the gender of the pronoun. (24) is a sample stimulus set from that experiment.

(24) Sample stimulus set from Experiment 1, Badecker & Straub (2002).
Accessible-match, inaccessible-match
John thought that **Bill** owed **him** another chance to solve the problem.
Accessible-match, inaccessible-mismatch
John thought that **Beth** owed **him** another chance to solve the problem.
Accessible-mismatch, inaccessible-match
Jane thought that **Bill** owed **him** another chance to solve the problem.
Accessible-mismatch, inaccessible-mismatch
Jane thought that **Beth** owed **him** another chance to solve the problem.

The most important result of Experiment 1 (also found in Experiments 2 & 3) showed a **gender match effect (GME)** for the accessible-match pair: mean reading times in the inaccessible-match condition were significantly longer than in the inaccessible-mismatch condition for the region directly following the pronoun ('another') or the next region ('chance'), or both. This result is potentially problematic for theories that claim that inaccessible antecedents are invisible to the parser from the earlier stages of processing (e.g. 'binding-as-early-filter' model described in Sturt (2003)).

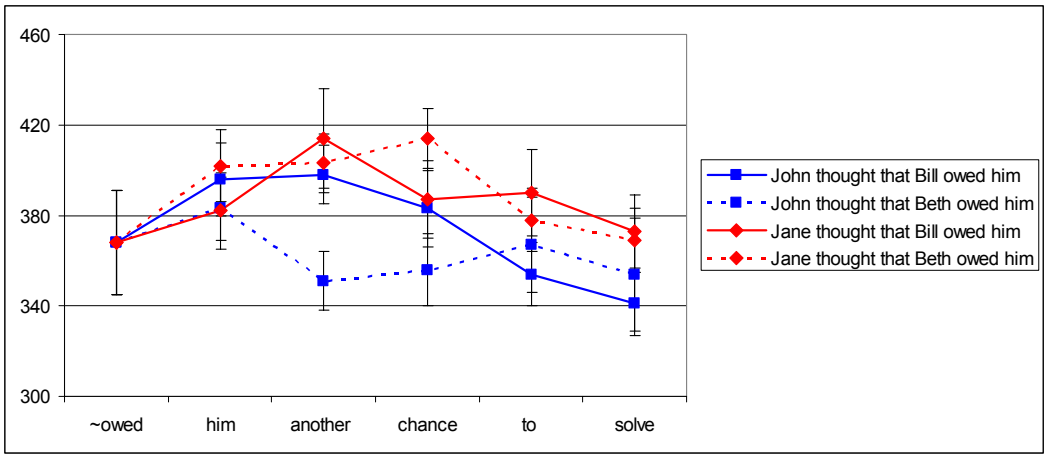


Figure 5. Results of Experiment1, Badecker & Straub (2002). The diagram is based on the values cited in the appendix to the paper.

In Experiment 3 B&S also tested whether the candidate set for a reflexive initially includes candidates that are subject to Principle A. They compared sentences in which the inaccessible antecedent for the reflexive (the main subject) matched or mismatched in gender with the reflexive, as shown in (25):

(25) Reflexive conditions from Experiment 5, Badecker & Straub (2002), *gender manipulation in the **inaccessible** antecedent*
Inaccessible-match
John thought that Bill owed **himself** another opportunity to solve the problem.
Inaccessible-mismatch
Jane thought that Bill owed **himself** another opportunity to solve the problem.

Once again, the authors found that the reading times for the regions following the pronoun were longer when the pronoun matched in gender with the inaccessible antecedent.
Having concluded that inclusion into the initial set of antecedent for a pronoun or reflexive is not influenced by the status of the NP in terms of the binding theory, B&S conducted additional experiments to investigate which factors do influence the composition of the initial set. In Experiment 5 they manipulated the gender of a genitive NP that is of a relatively low discourse-prominence. This genitive NP was an inaccessible antecedent in the reflexive conditions, but an accessible antecedent in the pronoun conditions.

(26) Sample stimulus set from Experiment 5, Badecker & Straub (2002)
*a. Reflexive-condition: gender manipulation in the **inaccessible** antecedent*
Jane thought that **Bill's** brother owed himself another opportunity to solve the problem.

Jane thought that **Beth's** brother owed himself another opportunity to solve the problem.

*b. Pronoun-condition: gender manipulation in the **accessible** antecedent*
Jane thought that **Bill's** brother owed him another opportunity to solve the problem.
Jane thought that **Beth's** brother owed him another opportunity to solve the problem.

The authors found a GMME at the regions following the pronoun (*another opportunity*) in the pronoun condition, i.e. as a result of manipulation in accessible antecedent. No significant effects were found in the reflexive condition. Moreover, manipulation of the gender of an inaccessible antecedent did not affect the processing of the reflexive in another experiment (Experiment 6), where the inaccessible antecedent was also of low discourse prominence (see (27)).

(27) Sample stimuli from Experiment 6, Badecker & Straub (2002)
a. It appeared **to John** that Bill owed **himself** another opportunity to solve the problem.
b. It appeared **to Jane** that Bill owed **himself** another opportunity to solve the problem.

The authors conclude based on these results that NP's that are of a low-prominence in the discourse context are not included in the initial set of candidates.

Thus, B&S concluded that Binding Principles could not be guiding structure-building and that antecedents that are subject to those constraints are nevertheless initially considered in the set of possible antecedent candidates. In their own words, "… the initial candidate set is created in a manner analogous to the process of activating the set of interpretations associated with an ambiguous lexical form: the salient entities of the local discourse that are consistent with the number and gender specification of the pronoun or anaphor make up the set of available interpretations for the referentially dependent expression, and the grammatical constraints on interpretation operate quickly and effectively in the process of selecting from among these options." [B&S 2002: 767][33]

[33] At the time when this section was initially written, I understood the model advocated by Badecker & Straub (2002) as making its initial selection based on gender and discourse-prominence and explicitly excluding the binding constraints should not be classified as an interactive model, as in:

> These results indicate that the binding-theory principles do not function as initial filters on the input to all stages of coreference processing. Instead, the data presented here support the interactive-parallel-constraint model. The initial candidate set is composed of the focused discourse entities (or sentence constituents) that are compatible with the lexical properties of the referentially dependent expression..." [B&S 2002: 764-5].

We argue that B&S's conclusions may be too strong in light of some of their own findings. If the initial set is formulated according to factors such as discourse prominence and gender/number specifications, inclusion of a given NP in the initial set of candidate antecedents should not be affected by whether that NP is accessible in binding terms. Yet in Experiment 5 gender manipulation of the same NP was sufficient to cause an effect only if that NP was an accessible antecedent for the dependent element (i.e. in the pronoun but not in the reflexive condition). We will return to this point in section 3.4.1.6.

Moreover, a detailed look at the results of Experiment 1 also reveals results that are unexpected under B&S's explanation. Figure 5 shows that in the regions following the pronoun mean reading times for the [accessible-match, inaccessible-mismatch] condition are significantly shorter than for the remaining three conditions. These effects must be explained by at least two processes. Increases in reading times for both accessible-mismatch conditions are explained by the 'no-antecedent surprise', i.e. the parser's surprise at not being able to find a licit antecedent for the pronoun (which by itself implies that the pronoun is seeking a grammatically appropriate antecedent). The increased reading time for the multiple-match condition is due to interference from the inaccessible antecedent. There, both the accessible and inaccessible antecedents match in gender with the pronoun and comprise the initial set of candidates; the longer reading times reflect an 'ambiguity' and the parser's increased effort in evaluating both antecedents and excluding one of them. However, if the number of candidates in the initial set yields a noticeable increase in reading times, we should expect the same effect in the comparison of the two accessible-mismatch conditions. Indeed, although eventually in either of them the parser fails to find a licit antecedent for the pronoun, in the no-match condition this outcome should be more immediate than in the inaccessible-match condition. This is because in the no-match condition the set of candidate antecedents is empty since no NP matches the pronoun in gender, whereas in the inaccessible-match condition the set has a candidate that is later ruled out for structural reasons. Yet the results show no difference in the processing of these two cases, thus undermining a claim whereby only gender information and discourse prominence, but not the structural accessibility of an NP, can affect formation of an initial candidate set for a pronoun/reflexive.

If so, the model proposed in B&S can be classified as a morphology-first & discourse-first model, but crucially not as an interactive model. This is because although an interactive model does not require that effects of various constraints showed up simultaneously, it does assume that information from all possible sources is taken into account at once, and does not favor any one of the sources, unless there is an inherent ordering in determining types of information.

Later, in a personal communication in June 2005, Bill Badecker explained that the results reported in the paper which suggest late application of the binding constraints could but need not be taken as due to an architectural inferiority of the binding constraints. The view that is proposed here is similar to such a view in that the binding constraints are argued not to be delayed for any architectural reasons.

It should be clear that despite our claim that some of the findings provide less counterevidence to the 'binding-as-early-filter' model than may initially be thought, it nevertheless is the case that some findings do stay as counterevidence. Those are the GME as a result of the manipulation of the gender of an inaccessible antecedent for both the pronoun and reflexive. Most puzzlingly for the 'binding-as-early-filter' model, there was an effect of an inaccessible antecedent on the processing of the reflexive, even though a licit accessible antecedent was always available and moreover was linearly the closest NP to the pronoun. However, even those effects are explicable if the effects that we see are late and reflect stages of processing that go beyond the initial set of candidate antecedents (see the end of the next section).

3.4.1.4 Kennison (2003)

Kennison (2003) reports results from three self-paced reading experiments. Experiments 1 and 2 had an identical design, with the difference that Experiment 1 tested the feminine pronoun *her*, whereas Experiment 2 tested the masculine *his*. The pronoun *her* is ambiguous between the possessor and the direct object which may crucially affect the logic and results of the experiment. Hence, we will not discuss Experiment 1 and concentrate instead on Experiment 2.[34]

Experiment 2 tested simple monoclausal sentences with forwards anaphora in which the pronoun was either the direct object (*him* in the NP conditions) or the possessor for the direct object (*his* in the Specifier-conditions). Thus, the subject was an accessible antecedent for the pronoun in the Specifier-condition, but due to Principle B it was an inaccessible antecedent in the NP-condition. Kennison reasoned that if Principle B was a strong constraint on structure building, there should be no effect of gender manipulation in the main subject position on the reading times at the pronoun when the subject is inaccessible due to Principle B (i.e. in the NP condition).

(28) Stimuli from Experiment 2, Kennison (2003)
Specifier conditions (gender manipulation in the accessible antecedent)
- *gender-match*
Carl watched **his** classmate during the open rehearsals of the school play.
- *gender-mismatch*
Susan watched **his** classmate during the open rehearsals of the school play.

NP conditions (gender manipulation in the inaccessible antecedent)
- *gender-match*

[34] I will only discuss the results of the manipulation of the gender of the subject, but not number. The number feature may be a weaker factor for manipulation in this type of experiment because of the split-anteedent phenomenon and because English allows the plural *they* to be used as a gender-neutral 3rd person singular pronoun..

Carl watched **him** yesterday during the open rehearsals of the school play.
- *gender-mismatch*
Susan watched **him** yesterday during the open rehearsals of the school play.

Kennison found a GMME in the Specifier conditions: in the regions directly following the pronoun (*classmate* and *during*) mean reading times were longer in the gender-mismatch sentence than in its gender matched counterpart. In the NP conditions a reverse effect was found: a GME due to longer reading times at the regions following the pronoun in the gender-match sentences than in their gender-mismatch counterparts.[35]

In Experiment 3 NP conditions from the previous experiment were placed in the context of a lead-in sentence, the subject of which was always a valid and discourse-prominent antecedent for the pronoun, as demonstrated in (29). As a result, the effect of the inaccessible antecedent (2nd sentence subject) on pronoun processing was no longer found.

(29) Stimuli from Experiment 3, Kennison (2003)

Lead in:	**Billy** complained about having a stomach ache.
Critical sentence:	
gender-match	**Carl** watched **him** closely throughout the day.
gender-mismatch	**Susan** watched **him** closely throughout the day.

Kennison takes this result to be supporting "the view that the extent to which structurally unavailable entities will influence coreference resolution can be determined by whether coreference is achieved and how quickly it is achieved. Finding a good match between an antecedent and the pronoun being processed presumably terminates the process of search for an antecedent. When no antecedent is available or when the fit between a structurally available antecedent and the pronoun is less strong, the process of searching for an antecedent is allowed to continue, during which time structurally unavailable antecedents can exert an influence on processing time." [Kennison 2003: 348]. She concludes therefore that both structurally accessible and inaccessible antecedents are included in the initial set of candidates during real-time coreference resolution. A contrasting position is that inasmuch as the effects that were found do indeed reflect properties of the initial candidate set (as claimed by the author), they cannot be explained without recognizing the effect of the Binding theory on the earliest stage of processing. For example, the opposite direction of the effect in the Specifier and the NP condition are unexpected if all that matters for formation of the initial candidate set is the discourse prominence of the antecedent and whether it matches the pronoun in its morphological features, since both of this

[35] The GME was only significant in the analysis that was based on cumulative reading times from two regions following the pronoun.

parameters are identical between conditions (see section 3.4.1.6 for a further discussion).

Overall, Kennison's (2003) view is similar to that of Badecker & Straub (2002): binding-inaccessible antecedents are included in the initial set of candidates during the resolution of pronouns/reflexives. This conclusion may be challenged based on a couple of reasons. For example, in B&S's Experiments 1 and 2 and Kennison's Experiments 1 and 2 a significant portion of the experimental conditions (75% in Kennison and 50% in B&S) lacked a licit antecedent for the pronoun.[36] This could have resulted in participants adopting an artificial strategy during the experiment. This fact could be especially problematic since the effects of the inaccessible antecedent found in these experiments constituted the core of the argument against the binding-as-early-filter model.[37]

Another consideration is whether the effects found in Kennison's and Badecker & Straub's studies which were taken to characterize the initial stage of linguistic processing, in fact reflect a later stage of processing (e.g. the parser's conscious self-check) or even a combination of behaviors specific to the parser and those that are cognition-general. It is well known that there is a cognition-general 'similarity effect' – a higher cost of processing of two or more items when they are similar to each other along some dimension compared to when they are not (Lewis 1996, Gordon, Hendrick & Levine 2002). If so, the GME from B&S's Experiment 3 may result from a load of this sort, rather than reflecting the state of the candidate set for the pronoun/reflexive formed by the parser.[38] The GME in the NP-condition in Kennison's Experiment 2 could be a second-pass effect: in the absence of an accessible antecedent the parser could have made a conscious check of all NPs that were compatible along the morphological dimension.

Similar concerns were raised by Lee & Williams (2005), whose experiments are presented next.

3.4.1.5 Lee & Williams (2005)

Lee & Williams (2005) tested the same configurations as Badecker & Straub (2003), but they managed to avoid having stimuli in which the pronoun does not get resolved intra-sententially. This was attained by choosing the gender-accessible antecedent, the main subject, to be a noun (typically an occupational

[36] In the other studies that we reviewed (Sturt 2003, Nicol & Swinney 1989) a reflexive or a pronoun always had a licit intra-sentential antecedent.

[37] This criticism does not apply to Experiment 3 in B&S, in which a pronoun always had a licit intra-sentential antecedent.

[38] The interference effect only emerges during coreference resolution and is not found otherwise when no dependent elements are present (Experiment 2, B&S 2002). This is expected, since the feature gender becomes more salient when it is actively considered for the reason of anaphor resolution and, therefore, can cause a stronger interference effect.

term) that has a widely-accepted stereotypical gender, rather than an unambiguous male or female name. This manipulation made it possible to avoid stimuli in which the cataphoric pronoun does not get resolved intra-sententially. Stimuli from Experiment 1 are given in (30).

(30) Stimuli from Experiment 1, Lee & Williams (2005).
Accessible-match, inaccessible-match
The **surgeon** confirmed that **Jonathan** warned **him** about the hostile demonstrations outside the hospital.
Accessible-match, inaccessible-mismatch
The **surgeon** confirmed that **Jennifer** warned **him** about the hostile demonstrations outside the hospital.
Accessible-mismatch, inaccessible-match
The **midwife** confirmed that **Jonathan** warned **him** about the hostile demonstrations outside the hospital.
Accessible-mismatch, inaccessible-mismatch
The **midwife** confirmed that **Jennifer** warned **him** about the hostile demonstrations outside the hospital.

In the main ANOVAs L&W found a main effect of gender of the main subject (accessible antecedent) in the three regions following the pronoun, but no main effect or interaction involving the embedded subject (inaccessible antecedent).

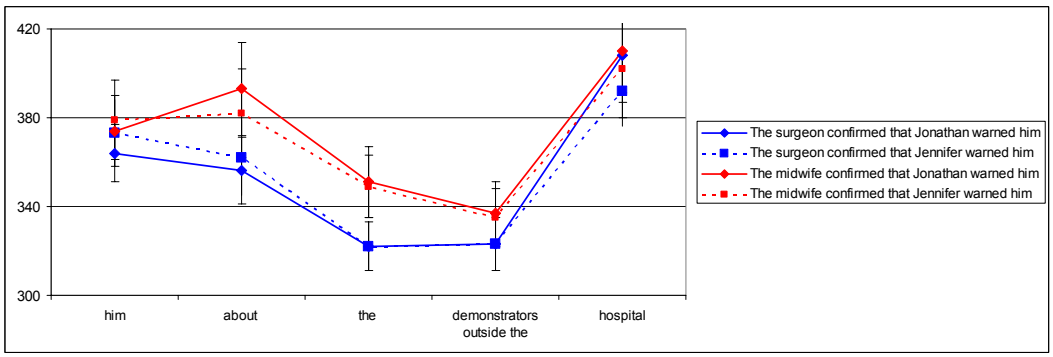


Figure 6. Results of Experiment 1, Lee & Williams (2005). The figure is constructed based on the values cited in the poster.

In order to confidently conclude that the reason for the lack of influence from the embedded subject lies in its inaccessibility in terms of the Binding Theory, the authors needed to rule out an alternative explanation in terms of lower prominence of the embedded subject compared to the main subject. To do so they embedded the pronoun deeper inside the object NP, which made both the main and the embedded subject an accessible antecedent for the pronoun ((31)).

(31) Stimuli from Experiment 2, Lee & Williams (2005).
Main-match, embedded-match
The **plumber** realized that **Jonathan** paid **his** kids well for odd jobs in the

house.
Main-match, embedded-mismatch
The **plumber** realized that **Jennifer** paid **his** kids well for odd jobs in the house.
Main-mismatch, embedded-match
The **cleaner** realized that **Jonathan** paid **his** kids well for odd jobs in the house.
Main-mismatch, embedded-mismatch
The **cleaner** realized that **Jennifer** paid **his** kids well for odd jobs in the house.

The authors reasoned that if the embedded subject was disfavored in Experiment 1 for a reason other than Principle B, it would continue to be neglected in Experiment 2 as well. However, the results showed a consistent effect of the gender congruency of the embedded subject in regions following the pronoun, either in the form of a main effect, or an interaction with the main subject. Thus, the absence of influence from the embedded subject in Experiment 1 must be attributed to its structural inaccessibility in terms of the Binding Theory.

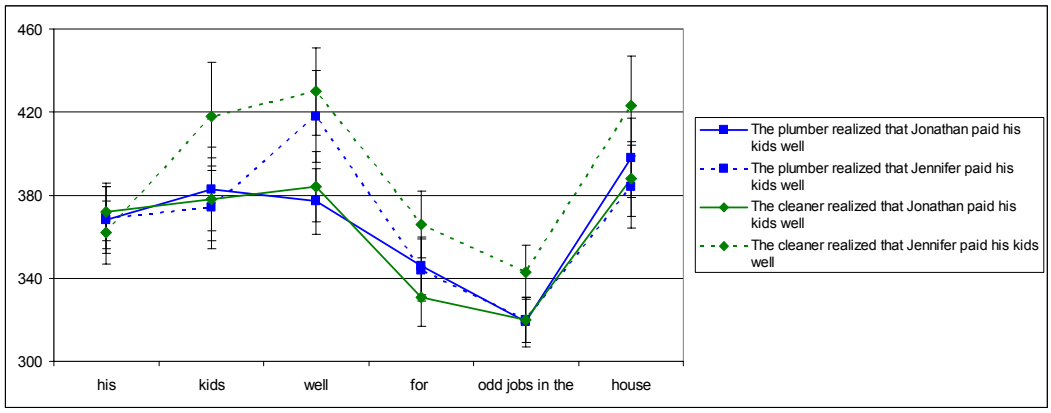


Figure 7. Results of Experiment 2, Lee & Williams (2005). The diagram is built based on the values cited in the poster.

Lee & Williams conclude that Principle B constrains the initial stage of pronoun resolution. Their conclusion contrasts with the conclusion by Badecker & Straub (2003), despite the fact that the two studies tested similar structures and used the same technique. The most plausible reason for the difference concerns availability of a licit antecedent for the pronoun in all conditions in Lee &Williams's experiments, whereas in Badecker & Straub (2003) the pronoun was always left unresolved in the accessible-mismatch condition.

3.4.1.6 Summary

There is no unanimous conclusion in the literature regarding whether antecedents that are inaccessible due to Principle A or B are ruled out immediately from the

set of candidate antecedents for the reflexive/pronoun. Below I will make an attempt to show that there are good reasons to believe that these principles influence formation of the initial set of candidate antecedents for a reflexive or pronoun.

The review of the literature on processing of forwards anaphora demonstrates that gender manipulation in the antecedent so that it matches or mismatches the gender of the pronoun/reflexive can yield all three logically possible effects: an increased reading time at or after the dependent element (GMME), a decrease in reading times (GME), or a null-effect. Table 1 summarizes and classifies all reviewed studies according to the type of effect that was found, and specifies the type of antecedent that caused the effect.

I believe that strong conclusions can be made based on the distribution of the effects. It takes no more than a quick look at Table 1 to realize that the direction of the effect strongly correlates with the structural accessibility of the relevant antecedent. In all but one case, GMME was found as a result of manipulation in gender of the accessible antecedent.[39] Conversely, a GME was found as a result of a feature-manipulation in the inaccessible antecedent. All cases in which the manipulation of morphological features did not yield any effect on processing also involved binding-inaccessible antecedents. Interestingly, other logical possibilities have not been attested. For example, the manipulation in the accessible antecedent never results in a GME or no-effect, whereas the inaccessible antecedent (almost) never yields a surprise GMME even when it is the most salient discourse referent. These gaps remain mysterious unless one accepts that the initial set of candidate antecedents is formed in accordance with binding constraints. However, once such a view is adopted most correlations fall out naturally. Particularly, the parser that forms its initial antecedent set in accordance with the binding constraints should not find a gender-incongruous NP in a binding-inaccessible position any more surprising than a gender-congruous NP; this explains why manipulation in the inaccessible antecedent never results in a GMME. That manipulation of gender in the accessible antecedent always results in a GMME is also natural provided that the parser aims at finding an antecedent for the dependent element as soon as possible: it is a 'surprise', if not a 'shock' reaction after realizing that the antecedent that made into the initial candidate fails on morphological grounds.[40]

[39] The only exception to the pattern was a late GMME to an inaccessible antecedent in Sturt (2003). This effect, however, was found only in one of many measures (second-pass reading times) and only in one of two experiments. Further investigation is necessary to make a definite conclusion about this case.

[40] In most studies with the exception of Lee & Williams (2005) there was only one accessible antecedent for the pronoun/reflexive. Thus, we cannot distinguish whether GMME indicates the parser's disappointment regarding the failure of a specific candidate, or a more general disappointment related to failure to find an antecedent for the dependent element among referents from the discourse context.

Study	*Sample Stimuli*	*Binding Acc vs Inacc*
GMME		
Sturt (2003)	… **the surgeon** pricked **himself/herself**…	Accessible
Sturt (2003)	**Jonathan/Jennifer** remembered that X pricked **himself** …	Accessible
Badecker & Straub (2002)	**John/Jane** thought that X owed **him** another chance …	Accessible
Lee & Williams (2005)	**The midwife** confirmed that X warned **her/him** …	Accessible
Kennison (2003), Exp. 2	**Carl/Susan** watched **his** classmate ….	Accessible
GME		
Badecker & Straub (2002), Exp. 1-3	John thought that **Bill/Beth** owed **him** another chance …	Inaccessible
Badecker & Straub (2002), Exp. 3	**John/Jane** thought that Bill owed **himself** another chance …	Inaccessible
Kennison (2003), Exp. 1	**Carl/Susan** watched **him** yesterday ….	Inaccessible
Null-effect		
Sturt (2003)	**Jonathan/Jennifer** remembered that X pricked **himself**…	Inaccessible
Badecker & Straub (2002), Exp. 5	Jane thought that **Bill's/Beth's** brother owed himself …	Inaccessible
Badecker & Straub (2002), Exp. 6	It appeared to **John/Jane** that Bill owed **himself** …	Inaccessible
Kennison (2003), Exp. 3	**Billy** complained about having a stomach ache. **Carl/Susan** watched **him** yesterday ….	Inaccessible
Lee & Williams (2005)	The midwife confirmed that **Jonathan/Jennifer** warned **her/him** …	Inaccessible

Table 1. Summary of findings (GMME, GME, null-effects) in sentence-processing studies of forwards anaphora in English.

The only remaining question concerns inaccessible antecedents. According to our claim inaccessible antecedents are not part of the initial set of candidates and therefore should not have any effect on the processing of the dependency; yet they occasionally yield a GME. We have already mentioned one possible reason for this effect, namely that it is a cognition-general interference effect due to the presence in working memory of several items that share some feature (Lewis & Nakayama 2002). If so, the inaccessible antecedent interferes with processing at a different level, as the bearer of gender value that is in the spotlight due to the dependency formation. This explains the direction of the effect: GME is a result of additional load from a morphologically-congruent inaccessible antecedent.

There could be an alternative account for why certain (usually discourse-prominent) inaccessible antecedents cause a GME, according to which a GME should be found only with antecedents that are inaccessible due to Principle B, but not due to Principle A. I mention this as a possibility despite the fact that this account is at odds with the results of Badecker & Straub (2003) who found a GME caused by antecedents that are subject to Principle B, and to Principle A. The proposal highlights the fact that reflexives, but not pronouns, have a very restricted range of licit antecedent positions. Accordingly, in most theories regulations on reflexives are formulated as a condition which appoints a grammatical antecedent for the reflexive (either through locality requirement in GB, or via a categorizing reflexive schema in CG, etc), whereas regulations on pronouns are stated by specifying inaccessible antecedents.[41] If these conditions are psychologically valid, the parser should define a set of accessible antecedents for the pronoun as a complement set to the set of inaccessible antecedents. If so, interference from an inaccessible antecedent could be explained by its explicit mention in the process of deriving licit candidates (this should be reminiscent of "Don't look at the elephant!" situations where even most compliant and cooperative subjects cannot help but to look at the elephant).

[41] This also holds for Reinhart & Reuland's theory of anaphora, in which Principle A & B are stated as follows:

Condition A: A reflexive-marked (syntactic) predicate is reflexive
Condition B: A reflexive (semantic) predicate is reflexive-marked

Definitions relevant for the conditions are:

- a predicate is reflexive iff two of its arguments are coindexed;
- a predicate is reflexive-marked iff it is lexically reflexive or one of its arguments is a SELF anaphor;
- a SELF-anaphor is a morphologically complex anaphor.

Thus, according to Condition A a reflexive item like *himself* must be part of a reflexive-marked predicate, which must be semantically reflexive. Thus the antecedent for *himself* must be the other argument of the predicate that the reflexive is an argument of. Condition B is also formulated in terms of reflexive marking (roughly, it refers to anaphors). Condition B states that a pronoun cannot be coindexed with another argument of its predicate, since that would require the predicate to be reflexively marked, and a pronoun cannot provide that reflexive-marking.

I conclude based on examination of individual studies and based on broader considerations that findings on Principle A and B in forwards anaphora contexts generally conform to the view of these constraints as 'early filters'.

3.4.2 Active search mechanisms in backwards anaphora and the application of Principle C

We are now ready to take up the main question of our study: how backwards anaphora relations are established during online processing and whether antecedents that are subject to Principle C are ever considered by the parser as a candidate licensor for the pronoun.

There are several reasons for our interest in backwards anaphora. First, backwards anaphora may be convenient for methodological reasons. In a typical study of forwards anaphora the discourse that precedes the dependent element contains more than one candidate antecedent. Typically, an effect of the manipulation of the properties of one of the antecedents is attributed entirely to the relation between the antecedent and the dependent element. However, manipulating some antecedent not only changes its relation with the dependent element, but also its relation with other antecedents in the input, which may in turn lead to short-term memory interference effects similar to those mentioned in sections 3.4.1.4 and 3.4.1.6. Hence, the effect observed at the dependent element may confound various sources, and it may be hard to separate out the influence of each individual antecedent on the processing of the dependent element. This problem does not arise in backwards anaphora, in which each candidate licensor for the pronoun appears at a different time, which makes it possible to confidently assess an effect from an individual candidate (at least for the first candidate antecedent that follows the pronoun).

An ongoing debate in theoretical and experimental linguistics concerns whether all structural relations can be reduced to the lexical entry of individual items. If this is possible, then an independent syntactic level should be unnecessary. The answer based on Principles A & B is encouraging for the lexicalist view: these principles may be formulated as restrictions on the interpretations of clause-mates (GB) or on the interpretations of co-arguments of a predicate (e.g., Reinhart & Reuland 1993), and therefore could be (or even, according to Reinhart & Reuland, at least partly, must be) coded in the lexical entry for each verb. Reducing Principle C to the lexicon is more problematic, since it regulates relations between elements across clause boundaries and can range over virtually unbounded dependencies. If Principle C is respected immediately during online processing, this would call for an explanation from the strongly lexicalist theories of parsing (e.g. MacDonald et al. 1994). I will return to this discussion in section 3.5.4.4.

3.4.3 Previous research on backwards anaphora and Principle C

Hirst & Brill (1980) were the first authors who studied the effect of Principle C on processing. They used a whole-sentence self-paced reading task using items such as (32). In these items the first name (the subject of the first sentence, *John*) is a binding-accessible antecedent for the pronoun, whereas the second name (*Henry*) is a binding-inaccessible antecedent.

(32) **John** stood watching. **He** ran for a doctor after **Henry** fell down some stairs.

The authors report that reading times for the second sentence differed as a function of the plausibility of the inaccessible antecedent as the referent for the pronoun. Hirst & Brill conclude based on this result that people use contextual information during the process of coreference resolution even when it is not decisive for determining the correct antecedent. However, the nature of the task and its coarse temporal resolution do not make it possible to make inferences regarding the time-course and the source of the effect. Specifically, the effect could be due to a late stage of processing, at which information from multiple sources is put together and the sentence is evaluated for its overall plausibility.

The most substantial research on Principle C and backwards anaphora is due to Cowart & Cairns (1987). The study aimed to establish whether the parser actively attempts to find an antecedent for a cataphoric pronoun and how various types of semantic and syntactic information are used by the parser during pronoun resolution.

Cowart & Cairns investigated the processing of pairs of sentence fragments such as (33)-(35). Each pair contained a *they*-condition with the pronoun *they* in the first clause and a *you*-condition with the pronoun *you* in the same position. Unlike *you*, *they* calls for an antecedent from the discourse context to be interpreted; a potential position for such antecedent is the subject position of the 2nd clause. In the absence of preceding context the first phrase of the 2nd clause (*charming babies*, *frying eggs*, *visiting uncles*) is syntactically ambiguous between an NP and a gerund. If the option of finding an antecedent for the pronoun is ranked high on the parser's list, we expect that each of the ambiguous phrases is initially treated as an NP, rather than a gerund. Note though that even if this phrase were an NP, eventually it would not be a suitable antecedent for the pronoun either because of a semantic anomaly ((33): babies normally don't lecture) or selectional restrictions ((34): *they* cannot be *eggs*, since they eat, and eggs don't eat) or a grammatical constraint on coreference ((35): *they* c-commands *visiting uncles*, thus coreference is ruled out by Principle C).

(33) a. Whenever they lecture during the procedure, charming babies…
b. Whenever you lecture during the procedure, charming babies…

(34) a. Even though they eat very little oil, frying eggs…
b. Even though you eat very little oil, frying eggs…
(35) a. If they want to believe that visiting uncles…
b. If you want to believe that visiting uncles…

The main verb in each sentence was singular *is* that was compatible with the gerund interpretation of the ambiguous phrase, but not with the NP interpretation. In a reading task, longer reading times at the verb *is* in the *you*-condition vs. *they*-condition would indicate the parser's initial treatment of the ambiguous phrase as an NP, presumably for the abovementioned reason of satisfying the pronoun. Rather than measuring the reading times at the main predicate, the authors measured the naming times for *is*, which was presented immediately after the participant heard the sentence fragment. They found longer response latencies for *is* in the *they*-condition as compared to the *you*-condition in (33) and (34), but there were no significant differences in the response latencies in (35). The authors attributed these results to the fact that before getting to the verb, the parser attempted to link the pronoun *they* with the first available antecedent, but only if that antecedent was in a binding-accessible position (i.e. not subject to Principle C). Unlike syntactic restrictions, semantic and selectional restrictions did not have an immediate effect of excluding an incongruent NP from the set of candidate antecedents from the pronoun.

The experiment enables us to make two important conclusions. First, the cataphoric pronoun *they* initiates an active search for an antecedent in the main subject position. This preference may result from various reasons or their combinations, such as assignment of the pronoun to the subject of the main clause, the first noun phrase of the clause or the first NP that has the same grammatical role as the pronoun. Second, the antecedent is not projected in positions that are ruled out by Principle C. In other words, the parser does not ever consider a syntactically inappropriate antecedent, although it does consider antecedents that are semantically inappropriate.

However, some objections can be raised with respect to the design of the study and the interpretation of the results. The authors' claim that the parser immediately respects Principle C is based on the contrast that the pronoun-type effect is present in the 'semantic' pairs (33) & (34), but not in the 'syntactic' pair (35). This argument requires the implicit assumption that different semantic and syntactic violations produce an effect of a comparable magnitude at the same point in time. In light of this, the conclusions drawn from these findings should be more cautious, perhaps that the syntactic considerations are taken into account by the parser earlier than the semantic considerations. However, the claim that the parser never violates a syntactic constraint needs stronger evidence, e.g. using structures which are not subject to any constraints as controls. Moreover, as the authors themselves point out, the results were only significant in the analysis by participants but not in the analysis by items, which raises the possibility that the results are dependent on the peculiarities of the materials. Finally, the naming task

that was employed did not provide information on positions beyond the position that was probed, which could have concealed important effects at an earlier or later point in processing.[42]

A study by van Gompel & Liversedge (2003) provides information on various stages in the processing of structures with licit backwards anaphora. They tested bi-clausal sentences as in (36) (i.e. similar to (33)-(34) but without any semantic anomalies) using eyetracking. The gender of the main subject was manipulated such that it either matched (36)a or mismatched (36)b with the preceding pronoun.[43]

(36) Stimuli from Experiment 1, van Gompel & Liversedge (2003).
a. gender-match
When **he** was at the party, **the boy** cruelly teased the girl during the party games.
b. gender-mismatch
When **he** was at the party, **the girl** cruelly teased the boy during the party games.
c. control
When **I** was at the party, **the boy** cruelly teased the girl during the party games.

The authors found a GMME in early measures at the region immediately following the subject: the first-pass reading times were slower at *cruelly* in (36)a than in (36)b.

Importantly, longer reading times following the 2nd subject in (36)b vs. (36)a were unlikely to have been due to the introduction of a new discourse entity in the main subject position, since reading times in the critical regions in the control condition (36)c, where the main subject also introduced a new discourse referent, did not differ from (36)a (i.e. the 2nd subject in the incongruous condition (36)b was read more slowly than both in the congruous condition (36)a and the control condition (36)c). Thus slower reading times in (36)b must be attributed to difficulty related to referential dependency formation between the main subject and the preceding pronoun.

Thus, van Gompel & Liversedge's findings support Cowart & Cairns's (1987) claim that in structures like (36) the parser expects the subject of the 2nd clause to be an antecedent for the preceding pronoun. Appearance of a gender-mismatching NP in that position yields a 'surprise' GMME on the part of the parser. The most natural interpretation for the GMME is that upon encountering the pronoun the parser made a forward prediction for an antecedent in the 2nd subject position and

[42] The task could also have been quite unnatural for participants since all stimuli in the experiment were incomplete sentence fragments.
[43] The number of the main subject was manipulated in a separate experiment and yielded the same results as the gender manipulation.

was surprised to find out that this prediction was not borne out due to a gender mismatch.

The significance of van Gompel & Liversedge's finding is that it provides us with a valid method to track the time-course of dependency formation in real-time processing. They showed that if the parser expects an antecedent for a cataphoric pronoun in a specific upcoming position, the appearance of a gender-mismatching NP in that position yields an identifiable GMME on the part of the parser. Taking this finding as a starting point, we can further investigate why and when the parser has an expectation for an antecedent in specific positions during backwards anaphora resolution, and whether positions that are subject to a constraint such as Principle C are initially considered by the parser as candidate antecedent positions and ruled out at a subsequent interpretive stage, or are avoided from the earliest stages of processing.

3.5 SELF-PACED READING EXPERIMENTS ON BACKWARDS ANAPHORA

3.5.1 Experiment 1

The goal of Experiment 1 was to test whether grammatical constraints on coreference, in particular Principle C, are immediately taken into consideration by the parser during online processing.

To address this question we compared the processing of nouns that followed a cataphoric pronoun when that noun was an accessible antecedent for the pronoun (no-constraint conditions) or an inaccessible antecedent due to Principle C (Principle C conditions).

The no-constraint conditions (37)b were modeled after the structures used in van Gompel & Liversedge (2003), who found a GMME as a result of gender manipulation in the embedded subject position in such sentences, most likely due to the parser's expectation of an antecedent in that position.[44] The Principle C conditions (44)a were formed by moving the conjunction *while* to the beginning of the second clause thereby changing the embedding relation between the two clauses. In the Principle C conditions the sentence-initial pronoun c-commands the following embedded *while*-clause, including its subject position;[45] hence, coreference between the pronoun and the 2nd clause subject is ruled out by Principle C. We reasoned that unless Principle C immediately rules out the 2nd subject position from consideration as a potential antecedent for the preceding

[44] The conditions in (37) are simplified stimuli from the experiment. The actual stimuli are listed in (39).

[45] In terms of the tree structure, the embedded while-clause in the Principle C conditions modifies the main verb phrase, i.e. it adjoins to the main VP and, hence, is c-commanded by the main subject.

pronoun, a GMME similar to that in the no-constraint condition should be obtained in the Principle C conditions.

(37) *a. Principle C condition*
She was taking classes full-time while **Kathryn/Russell** was working two jobs to pay the bills.
b. No-constraint condition
While **she** was taking classes full-time **Kathryn/Russell** was working two jobs to pay the bills.

The experiment consisted of an offline questionnaire and an online self-paced reading task, which we discuss as Experiments 1a and 1b respectively.

3.5.1.1 Experiment 1a: Off-line rating experiment

The offline experiment was a rating task. We used a methodology similar to that used by Gordon & Hendrick (1997) to obtain coreference ratings for various types of structures. Participants were asked to rate the acceptability of coreference between two NPs highlighted in bold on a 1 (impossible)-to-5 (absolutely natural) scale. We sought to confirm that, provided matching gender/number values, participants indeed accept coreference between the cataphoric pronoun and the 2nd subject in the no-constraint environments, but reject it in the Principle C environments.

(38) Stimuli from an offline rating task, Experiment 1a.
a. *Principle C*
Because last semester **she** was taking classes full-time while **Kathryn** was working two jobs to pay the bills, Erica felt guilty.
b. *No-constraint*
Because last semester while **she** was taking classes full-time **Kathryn** was working two jobs to pay the bills, Russell never got to see her.

Twelve sets were chosen from the stimuli of the online part of Experiment 1 and reduced to the two conditions shown in (38) (see Experiment 1b for an explanation why the experimental stimuli had another level of embedding in the uppermost *because*-sentence). Two different lists were constructed based on these 12 sets using a Latin Square design, and each list had two different stimulus order randomizations. 40 participants judged 6 instances of each condition in (38). The questionnaire also contained 20 additional sentences, twelve of which tested configurations used in Experiment 2 (see Experiment 2a). All participants were undergraduate students at the University of Maryland, enrolled in an introductory class in Linguistics and naïve to the purpose of the experiment.

Results

The mean coreference rating score from Experiment 1a are given in Table 2.

Table 2. Mean rating scores from Experiment 1a.

Condition	*Mean score (standard error)*
Principle C	1.4 (.12)
No-constraint	4.1 (.13)

The mean rating score for the Principle C condition was significantly lower than in the no-constraint condition both in the participants analysis and items analysis (two-tailed paired *t*-test, *ps*<.001).

The results fully agree with the predictions of the Binding theory: coreference between the pronoun in the main subject position and the name in the 2nd subject position is allowed only when the pronoun does not c-command the name, i.e. in the no-constraint condition but not in the Principle C condition.

3.5.1.2 Experiment 1b: Online Sentence Reading Experiment

Participants

60 native speakers of English from the University of Maryland undergraduate population with normal or corrected-to-normal vision and no history of language disorders. Each participant gave informed consent and was paid $10 for participation in the experiment.

Materials & Design

There were 30 sets of 5 conditions, four of which were organized in a 2x2 design with the factors *constraint* (Principle C vs. no-constraint) and *gender congruency* (gender-match vs. gender-mismatch between the pronoun and the 2nd clause subject). The gender of the cataphoric pronoun was balanced across stimulus sets: half of the sets were built on the basis of the masculine pronoun *he*, and the other half on the basis of the feminine pronoun *she*. Within each condition, the gender-match and gender-mismatch sentences differed only in the gender of the subject of the 2nd clause. The 2nd clause subject was always a gender-unambiguous proper name, matched for the number of letters and syllables in the gender-match and gender-mismatch variants within each set.

We wanted to ensure that if the coreference interpretation is ruled out by the parser, the exclusion cannot be due to pragmatic infelicity, but rather to a grammatical constraint on coreference. In creating the stimuli we therefore chose events in each clause of the sentence such that they could be performed either by different agents or, crucially, by the same agent.

In addition, in order to ensure that the cataphoric pronoun does receive an inter-sentential grammatical antecedent in every case, we embedded our sentences in a further sentence introduced by the conjunctions *although* or *since*. The choice of

the gender of the third clause subject was such that each sentence had a unique antecedent for the pronoun. So, in the Principle C conditions the subject of the third clause always matched in gender with the first pronoun and served as a grammatical antecedent for that pronoun. In the no-constraint conditions, due to the possibility of coreference between the pronoun and the 2nd clause subject in the gender-match condition, the gender of the third clause subject mismatched that of the pronoun. Conversely, in the gender-mismatch condition of the no-constraint pair, the third clause subject matched the gender of the pronoun.

(39) A full set of stimuli from Experiment 1b.
a. Principle C, gender-match
Because last semester **she** was taking classes full-time while **Kathryn** was working two jobs to pay the bills, **Erica** felt guilty.
b. Principle C, gender-mismatch
Because last semester **she** was taking classes full-time while **Russell** was working two jobs to pay the bills, **Erica** felt guilty.
c. No constraint, gender-match
Because last semester while **she** was taking classes full-time **Kathryn** was working two jobs to pay the bills, **Russell** never got to see her.
d. No constraint, gender-mismatch
Because last semester while **she** was taking classes full-time **Russell** was working two jobs to pay the bills, **Erica** promised to work part-time in the future.
e. No constraint, name
Because last semester while **Erica** was taking classes full-time **Russell** was working two jobs to pay the bills, **she** promised to work part-time in the future.

Following van Gompel & Liversedge (2003), we added a fifth 'name' condition (39)e to each set. This condition was identical to condition (39)d in the number of referents introduced in each clause. It was needed to ensure that any observed GMME at the 2nd subject in the no-constraint pair cannot merely be due to the introduction of a new discourse referent. If the number of referents is indeed the reason for a GMME in the no-constraint pair, we should expect the same increase in reading times in the name condition (i.e. (39)d & (39)e should not be different from each other and both slower than (39)c at the 2nd subject).

The online experiment was a self-paced reading task with word-by-word presentation of stimuli. The regions used for the data analysis corresponded either to a single word or to a phrase. The critical 2nd subject position in which the gender manipulation occurred corresponded to region 7 in all conditions.

In addition to the target sentences, the experiment included 90 filler sentences of various lengths and complexity (target-to-filler ratio of 1:3). In order to mask experimental sentences, the fillers bore a number of similarities with the target items, including length and average clause number and were designed in several

subgroups, each built around some salient feature of the targets, e.g. use of proper names and pronouns, a subordinator followed by a temporal modifier at the beginning of the sentence, etc. There were no instances of unresolved anaphora in filler sentences. This was done in order to encourage readers to always expect a pronoun to find an antecedent within the sentence.

Procedure

Participants were tested using a desktop PC, running the Linger software (Doug Rohde, MIT). Stimuli were distributed among five presentation lists in a Latin Square design. Each list contained 30 experimental sentences (6 per condition) and 90 filler sentences. Filler sentences were identical across all five lists. Each participant was assigned to one of the lists according to his participant number in the experiment; the order of the stimuli within the presentation list was randomized for each participant.

The experiment employed a standard word-by-word moving window paradigm (Just et al. 1982). Each trial started with a blank screen. Upon pressing the space-bar, a sentence masked by dashes appeared on the screen. The masks extended to all letters and punctuation marks, but left spaces unmarked. As the participant pressed the spacebar, a new word appeared on the screen as the previous one was re-masked by dashes.

A comprehension question appeared after the end of each sentence all at once and unmasked (e.g. *Was Kathryn/Russell working two jobs?*). Participants were instructed to read sentences at a natural pace and to respond the comprehension questions as accurately as possible. To answer the question the subject pressed the *f*-key for 'yes' and the *j*-key for 'no'. If the question was answered incorrectly the word 'Incorrect' appeared briefly in the center of the screen.

Analysis

Only sentences for which the corresponding comprehension question was answered correctly were included in the analysis. Reading times that exceeded the threshold of 2.5 standard deviations above the subject's mean reading rate for each region were replaced by the threshold value. This procedure affected 2.4% of the data (range 1.9-2.6% for individual conditions).

The data from the first four conditions were entered into a 2x2 ANOVA with the factors *constraint* (Principle C, no constraint) and *congruency* (gender-match, gender-mismatch). Reading times from the name condition (39)e were compared pairwise to either of the variants from the no-constraint condition in a one-way ANOVA. ANOVAs were computed on the participant mean raw reading times collapsing over items (F1), and on item means collapsing over participants (F2). All significant main effects and interactions with $p < .05$ are reported.

Results

Comprehension question accuracy

Two out of 60 participants showed comprehension question accuracy that was more than 2.5 standard deviations below the mean accuracy. As a result, the lowest scoring participant from each list was removed from the analysis to equalize the number of subjects remaining on each list. For the remaining 55 participants, the mean question response accuracy was 91%, with a range of 90-93% for individual conditions. Differences in accuracy between individual conditions were not significant.

Self-paced reading

We will first present the results of the 2x2 ANOVA on the Principle C and no-constraint conditions (39)a-d, and then discuss the name condition. The results from the Principle C conditions are presented in Figure 8 and from the no-constraint and name conditions in Figure 9.

Principle C and no-constraint conditions. There were no reliable main effects or interactions in any region in the first clause.

In the main ANOVA there was a main effect of *constraint* at the 2nd subject NP (region 8) both in the participant and in the items analysis ($F1(1,54) = 9.7$, $p < .01$, $F2(1,29)=11.1$, $p< . 01$) due to longer mean reading times in the Principle C conditions than in the no-constraint conditions. This effect can be attributed to the differences in the content of the preceding region between the two pairs of conditions. Critically, there also was a significant *constraint* x *congruency* interaction ($F1(1,54) = 5.5$, $p < .05$, $F2(1,29) = 3.9$, $p < .05$) at the 2nd subject NP. Pairwise comparisons within each level of the factor *constraint* showed no effect of gender congruency for the Principle C conditions (all Fs < 1.1, ps>.1), and a significant effect of gender congruency for the no-constraint conditions that was reliable in the participants analysis ($F1(1,54) = 4.7$, $p < .05$) and marginally significant in the items analysis ($F2(1,29) = 2.9$, $p = .09$) due to a slowdown of 34 ms in average reading times in the gender-mismatch condition (i.e. a GMME). The main effect of *constraint* was also found at the region immediately following the 2nd subject (region 9, $F1(1,54) = 8.9$, $p < .01$, $F2(1,29) = 11.7$, $p < .01$), but its interaction with *congruency* was not significant($Fs \leq 1$) and there were no significant differences in the pairwise comparisons within each level of factor constraint (all Fs$\leq$1).

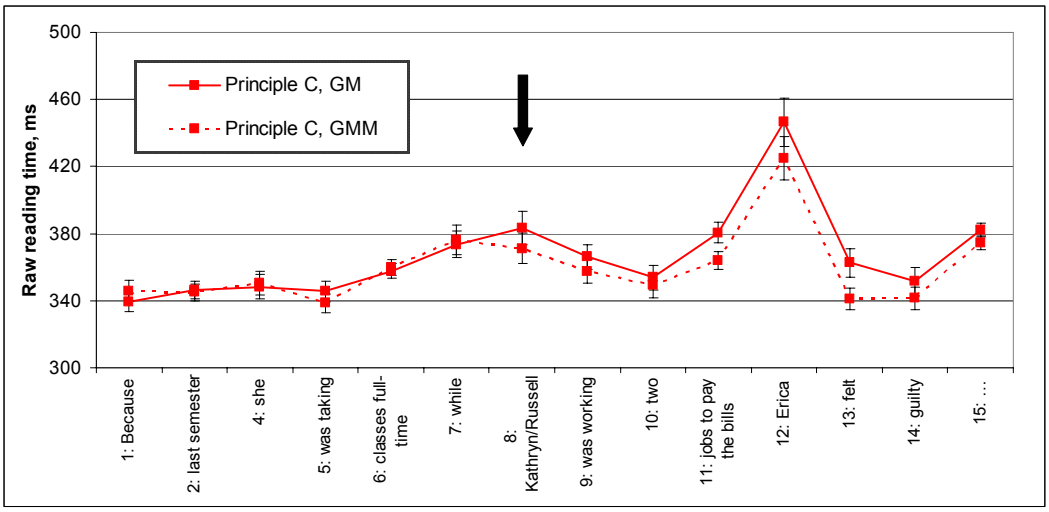


Figure 8. Mean reading times in milliseconds for the Principle C conditions from Experiment 1b.

The breakdown into regions was as follows: Because$_1$ (last semester)$_2$ she$_4$ (was taking)$_5$ (classes full-time)$_6$ while$_7$ Kathryn/Russell$_8$ (was working)$_9$ two$_{10}$ (jobs to pay the bills,)$_{11}$ Erica$_{12}$ felt$_{13}$ guilty$_{14}$. The arrow marks the critical 2nd subject.

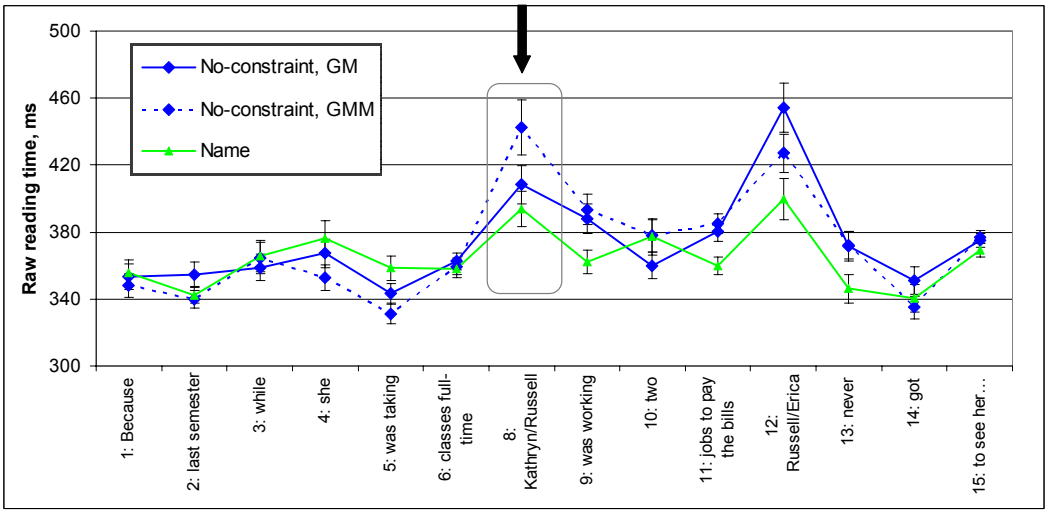


Figure 9. Mean reading times in milliseconds for the no-constraint and the name conditions from Experiment 1b.

The breakdown into regions was as follows: Because$_1$ (last semester)$_2$ while$_3$ she/she/Erica$_4$ (was taking)$_5$ (classes full-time)$_6$ Kathryn/Russell/Russell$_8$ (was working)$_9$ two$_{10}$ (jobs to pay the bills,)$_{11}$ Russell/Erica/she$_{12}$ never$_{13}$ got$_{14}$ (to see her.)$_{15}$ The arrow marks the critical 2nd subject.

Other reliable effects included a main effect of *congruency* at the 3rd subject (region 12) that was significant in the participants analysis ($F1(1,54) = 4.2$, $p < .05$) and marginally significant in the items analysis ($F2(1,29) = 3.0$, $p = .09$).[46] The effect was due to longer reading times for the 3rd subject in the gender-matching conditions than in the gender-mismatching conditions. There also was a main effect of *constraint* at the following word (region 13, $F1(1,54) = 6.2$, $p < .05$, $F2(1,29) = 4.6$, $p < .05$) due to slower reading times in the Principle C conditions. Pairwise comparisons in the same region showed a significant effect of *congruency* for the Principle C pair ($F1(1,54) = 7.0$, $p < .01$, $F2(1,29) = 4.2$, $p < .05$) due to longer reading times in the gender-match condition (22 ms on average). There was no effect of congruency in region 13 in the no-constraint pair (both Fs < 1). All other differences were not statistically reliable.

Name condition. In the first clause there was a significant difference in reading times between the name and the gender-mismatch condition of the no-constraint pair at the subject (region 4, $F(1,54)=4.7$, $p<.05$, $F2(1,29)=3.3$, $p=.07$) and the next region (region 5, $F(1,54)=11.1$, $p<.001$, $F2(1,29)=8.0$, $p<.01$) and between the name and the gender-match condition in region 5 ($F(1,54)=3.9$, $p<.05$, $F2(1,29)=2.7$, $p<.1$), all due to longer reading times in the name condition. These effects are expected given differences in the length of the 1st subject, which was the pronoun *he*/*she* in the no-constraint conditions and a proper name in the name condition.

At the critical 2nd subject NP reading times in the gender-mismatch condition of the no-constraint pair were significantly longer than in the name condition ($F1(1,54) = 7.6$, $p < .01$, $F2(1,29) = 6.3$, $p < .05$). However, reading times between the name and the gender-match condition did not differ significantly in the same region (both Fs ≤ 1). There were various significant differences between the name condition and either of the no-constraint conditions following the critical region. These differences are expected in light of the parser's attempt to resolve coreference at the 2nd subject in the no-constraint conditions, but not in the name condition and are not reported here as uninteresting.

Discussion

Experiment 1 compared sentences where a cataphoric pronoun was followed by a name (2nd subject) that was either an accessible antecedent for the preceding pronoun (no-constraint conditions (39)c,d), or an inaccessible antecedent due to Principle C (Principle C conditions (39)a,b). Comparison of the parsing profiles for the two conditions allows us to determine at which point during sentence processing NPs that are subject to Principle C are excluded from consideration by the parser as potential antecedents for the preceding pronoun.

[46] Recall that the factor *congruency* refers to a (mis)match in gender between the pronoun and the 2nd subject, rather than the 3rd subject.

The offline ratings showed that participants judged the 2nd subject to be a licit antecedent for the preceding pronoun in the no-constraint condition, but not in the Principle C condition. The results of the offline task thus agree well with the predictions of the Binding Theory.

In the self-paced reading task, we replicated the gender-mismatch effect (GMME) found in an eyetracking study by van Gompel & Liversedge (2003). In the no-constraint condition, there was a slowdown at the 2nd subject NP if that NP mismatched in gender with the preceding pronoun (i.e. longer RTs in (39)d vs. (39)c). As mentioned earlier, the GMME may be interpreted by appealing to the existence of an expectation on the part of the parser for a well-formed antecedent in the 2nd subject position, and a surprise reaction whenever this expectation did not come true. In contrast, in the Principle C conditions there was not a significant effect of gender congruency at the critical 2nd subject position, or in any other region in the 2nd clause. The null-effect in the Principle C condition suggests that in this condition the parser never considered the subject of the 2nd clause as a potential antecedent, thus making the gender-congruency between that subject and the preceding pronoun irrelevant. The lack of an effect of the manipulation of the gender of the 2nd subject NP on the reading times is explained by the parser's immediate exclusion of this position from the set of candidate antecedent positions in accordance with Principle C.

Furthermore, we can rule out an alternative explanation for the GMME in the no-constraint condition that attributes the effect to an additional discourse referent at the 2nd subject position in the gender-mismatch condition compared to the gender-match sentence. The results from the control 'name' condition rules out this possibility. If the effect were due to the introduction of a new referent, we would expect an increase in reading times at the 2nd subject in the name condition (39)e similar to that in (39)d. Instead, the reading times at the 2nd subject in (39)e patterned together with the gender-matching (39)c, which casts doubt upon the notion that longer reading times in (39)d reflect a cost for an extra discourse referent (see van Gompel & Liversedge 2003 for a similar result). Hence, we can be confident in asserting that the GMME in the no-constraint pair reflects processes related to coreference resolution.

A justified question is whether the GMME at the critical subject in the no-constraint condition necessarily indicates that a prediction was generated about that position before the NP was processed. In other words, does a cataphoric pronoun indeed initiate an active and predictive search for an antecedent in the following material? Instead, the parser may perform a 'passive' search whereby each NP encountered in the bottom-up input is 'tried' as an antecedent, but crucially, no forward predictions about an upcoming antecedent position are made. Importantly for us, two points remain valid. First, the sentence-initial pronoun is not taken to refer to an unspecified discourse referent: the parser attempts to link it to an overtly mentioned referent in the following discourse. Second, even under the passive-search hypothesis, in order to explain our pattern of results, one must

accept that the parser considers the availability of coreference in terms of the relative structural positions of the pronoun and the current NP before it considers whether the two items match in gender. Otherwise, the same GMME as in the no-constraint conditions should have been observed in the Principle C conditions. Thus, regardless of whether the correct model must be formulated in terms of an active or passive search, the structural constraints must be as applying early during the online process of dependency formation.

In light of the discussion of the exact source of the GMME and the 'active' search mechanism that is likely to underlie this effect, another property of the stimuli that we used in Experiment 1 becomes relevant. Since the argument for the early on-line effects of Condition C involves the absence of a GMME, the argument depends on the assumption that there is no other reason for the absence of a GMME in the Condition C conditions. Along these lines, however, there may be an inherent asymmetry between the Principle C and the no-constraint conditions related to the parser's ability to predict the 2nd subject position. Specifically, it could be the case that the pronoun starts an active search for an antecedent if and only if there is an independent prediction for an upcoming NP position at the time when the pronoun is processed. If so, there is a qualitative difference between the conditions in Experiment 1 concerning the existence of an independent prediction for the critical 2nd subject position at the time when the pronoun is processed. Recall that the actual stimuli used in the experiment were 3-clause structures, diagramed in (40).

(40) Diagram of conditions from Experiment 1b.
No-constraint: [Because [while *pronoun* [2nd subject…....]], 3rd subject…]
Principle C: [Because [*pronoun*….[while 2nd subject …..]], 3rd subject…]

In the no-constraint conditions the subordinator *while* that appears before the pronoun sets up an expectation for an embedded clause and the following main clause. If upon encountering the pronoun the parser checks whether there is a predicted antecedent position, it can find one, namely the subject of the main clause projected by *while* (i.e. 2nd subject). On the other hand, if we repeat the same procedure step-by-step in the Principle C conditions, the parser's choice would fall on the subject position of the main clause of the *because*-sentence (the 3rd subject). Thus, the lack of a GMME in the Principle C conditions may be an artifact of an expectation for an antecedent in a different position, namely the main subject of the outermost *because*-sentence.

The claim above was based on the assumption that (i) GMME emerges in a given position only if there is a previously established prediction for an antecedent in that position, and (ii) the parser projects an antecedent for a cataphoric pronoun only if there is an independently predicted structural position at the time when the pronoun is encountered. A somewhat different version of this claim is to modify the second condition to say that the prediction need not be present right away at the time when the pronoun is processed, but may be supplied by an element later

in the sentence. Under such a modified view, there is a prediction for the 2nd subject position in both conditions, but it appears much later in the Principle C conditions, then in the no-constraint conditions (depending on where in the sentence the subordinator *while* appears). Then the lack of a GMME in the Principle C conditions may be an artifact of the closeness of the subordinator *while* and the critical 2nd subject: the parser may not have had sufficient time to fully develop a prediction for the critical 2nd subject. In other words, the absence of a GMME may be due to the lack of time, but crucially not the lack of motivation, to project an antecedent in the 2nd subject position.

Under both of these hypotheses all differences in the processing of the Principle C and no-constraint conditions could have arisen due to properties of the stimuli that are unrelated to constraints on coreference. In order to reject these alternatives and to strengthen our claim regarding the immediate application of Principle C during online processing, we need to ensure that the lack of GMME at the 2nd subject in the Principle C condition was not due to the lack of prediction for that position at the time when the pronoun was encountered, and/or to a late prediction of this position. We will attempt to resolve these issues in Experiments 2 & 3 by testing configurations that are matched in their ability to predict the crucial NP position across all conditions.

3.5.2 Experiment 2

The goal of Experiment 2 was to test the effect of Principle C without confounding the presence of a grammatical constraint with the presence of a prediction for the critical subject position. In Experiment 2 both the Principle C and the no-constraint conditions were identical with respect to the prediction of the second clause at all times, including the moment when the pronoun was encountered. Thus, if we obtained the same pattern of results using these materials as in Experiment 1, the crucial difference in the processing of the Principle C and no-constraint conditions could no longer be attributed to differences in prediction for the critical subject position, and must be attributed to the presence or absence of a grammatical constraint on coreference.

In Experiment 2 we used sentences that started with an expletive *it*-clause. In the Principle C condition (41)a the cataphoric pronoun was the direct complement of the dative preposition *to* or *for* and referred to the experiencer of the predicate. In the no-constraint condition (41) c,d the pronoun was embedded more deeply as the possessor inside the prepositional complement.[47]

[47] The conditions in (41) are simplified compared to the actual stimuli used in the online experiment. A sample set of original materials used for the online experiment is given in (46).

(41) a/b. *Principle C conditions*
It seemed worrisome to **him** that **John/Ruth** was gaining so much weight.
c/d. *No constraint conditions*
It seemed worrisome to **his** family that **John/Ruth** was gaining so much weight.

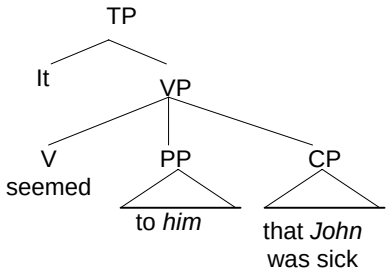
We reasoned that a sentence fragment like *It seemed worrisome to ...* would create a strong expectation for another clause and hence for another subject position. Hence, the 2nd subject position is expected at the time the pronoun is processed in both conditions. If we were wrong in attributing the lack of GMME in the Principle C condition in Experiment 1 to the presence of a constraint and the null effect was instead due to the lack of prediction for the critical subject at the time when the cataphoric pronoun was encountered, we should find a GMME in the Principle C condition in this experiment.

3.5.2.1 Transparency of the *to*-phrase

The reader may note that under a traditional definition of c-command, the pronoun does not c-command the subject of the main clause even in our Principle C condition (41)a,b. If so, these sentences should not fall under the scope of Principle C which applies only if the name is c-commanded by the pronoun. Nevertheless, the facts are that English speakers consistently reject (41)a on the coreference interpretation, although they overwhelmingly accept this interpretation in (41)c (see Experiment 2a).

The unavailability of coreference in (41)a,b could be explained by appealing to an additional constraint on coreference that applies to such configurations. Alternatively, this case could fall under Principle C if the pronoun in (41)a,b did c-command the embedded subject position. A tree structure for the Principle C conditions is shown in (42).[48]

(42)



[48] The tree (42) is drawn using ternary branching for simplicity. The reader is free to use other ways of representing the structure for (41), e.g. using Larson's (1989) VP-shell hypothesis. This choice does not influence the logic of the argument.

Works by Reinhart (1983) and Brody (1994) point towards the latter option, as the assumption that the complement of *to* does c-command the main clause (either because *to* is a case marker, or because it is somehow 'transparent') helps to explain a range of phenomena, including binding regularities, negative polarity item (NPI) licensing and scope assignment in such configurations. Two representative arguments are reproduced below.

The first argument for the existence of c-command from the complement of the preposition *to* to the main subject comes from the domain of negative-polarity item (NPI) licensing. NPIs are licensed in structures where they are c-commanded by a negative element (Klima 1965). For example, the NPI *anybody* is licensed in (43)a and (43)b by a negative subject *nobody* or negation *not*, which c-command the NPI. The use of *anybody* is illegal whenever a licensor is not present, as in (43)c.

(43) a. **Nobody** noticed that **anybody** had left.
b. **John** didn't notice that **anybody** had left.
c.***John/someone** noticed that **anybody** had left.

Turning to our cases, (44)a shows that a negative NP as the complement of *to* licenses the use of an NPI as the subject of the main clause. (44)b confirms that the NPI *anything* is indeed licensed by *few of his colleagues* and is illicit if *few of his colleagues* is replaced by a positive expression, such as *many of his colleagues*.

(44) a. It was apparent to **few of his colleagues** that **anything** had changed about John.
b.*It was apparent to **many of his colleagues** that **anything** had changed about John.

The second argument comes from the availability of the bound variable reading of a pronoun in the embedded clause in examples where the NP inside the PP is quantificational, e.g. (45).

(45) It seemed obvious to every $girl_i$ that she_i was the best dancer on the team.

Although various accounts may differ in how exactly the c-command relation is obtained, there is notable consensus in the literature concerning the 'transparent' status of the preposition *to* (see Baltin & Postal 1996 for a summary). We thus conclude that the unacceptability of coreference in (41)a is due to the fact that the pronoun c-commands the main subject NP, i.e. Principle C.

3.5.2.2 Experiment 2a: Offline Rating Experiment

The aim of this experiment was to test whether coreference between the pronoun and the 2nd subject is available in the no-constraint condition, but not in the Principle C condition in (41). Experiment 2a was run together with Experiment 1a

and tested 40 participants. The reader is referred to the Design section in Experiment 1a for additional details on methods and design.

Results

The mean coreference rating scores from Experiment 2a are given in Table 3.

Table 3. Mean rating scores for Experiment 2a.

Condition	*Mean score(standard error)*
Principle C	1.5 (.12)
No-constraint	4.2 (.13)

The mean rating score for the Principle C conditions was significantly lower than in the no-constraint conditions in the participants and items analyses (two-tailed paired *t*-test, both *p*s<.001).

3.5.2.3 Experiment 2b: Online Sentence Reading Experiment

Participants

Participants were 60 native-speakers of English from the University of Maryland undergraduate population who did not participate in Experiment 1. All participants had normal or corrected-to-normal vision and no history of language disorders. They gave informed consent and were paid $10 for participation in the experiment.

Materials &Design

Experiment 2 followed a 2 x 2 design with the factors *constraint* (Principle C vs. no constraint) and *congruency* (gender-match vs. gender-mismatch). Sentences were presented in a word-by-word non-cumulative moving window paradigm.

As in Experiment 1, we aimed at providing an accessible intra-sentential antecedent for the cataphoric pronouns in all cases. This was achieved by adding a *but*-clause to the end of the sentence, the subject of which was a licit antecedent for the pronoun, as shown in (46). The gender of the 3rd subject was chosen such that there was a unique antecedent for the pronoun in each condition.

(46) A full set of stimuli from Experiment 2b.
a. Principle C, gender-match
It seemed worrisome to **him** that **John** was gaining so much weight, but **Matt** didn't have the nerve to comment on it.
b. Principle C, gender-mismatch
It seemed worrisome to **him** that **Ruth** was gaining so much weight, but **Matt** didn't have the nerve to comment on it.

c. No constraint, gender-match
It seemed worrisome to **his** family that **John** was gaining so much weight, but **Ruth** thought it was just a result of aging.
d. No constraint, gender-mismatch
It seemed worrisome to **his** family that **Ruth** was gaining so much weight, but **Matt** thought it was just a result of aging.

In creating the stimuli we ensured that the scenarios were chosen such that both Principle C and no-constraint conditions are plausible on the coreference reading (should this reading be available), e.g. that John and his family are equally likely to be worried about something negative involving John (e.g. gaining weight).

The full list of materials is given in Appendix 3-B.

Analysis

The analysis was performed on the data from all 60 subjects who participated in the experiment, distributed equally among the four presentation lists. The procedure was identical to that used in Experiment 1: sentences for which the comprehension question was answered incorrectly were excluded from the analysis, and reading times that exceeded the threshold of 2.5 standard deviations above the mean reading rate for each region per subject were replaced by the threshold value (winsorized). Winsorizing affected 2.0% of trials (1.9 –2.2% for individual conditions).

The regions used for the data analysis corresponded to single words, except for regions corresponding to the end of the clause, for which several words were averaged together due to variation in the clause length between items (see the legend for Figure 10 for regions). Raw reading times for each region were entered into a 2-way ANOVA with the factors *constraint* and *gender congruency*. Planned pairwise comparisons within each level of the factor *constraint* are also reported.

Results

Comprehension question accuracy

The mean question accuracy was 95.7% and ranged between 94.4% and 97.3 % for individual conditions. Differences between conditions were not significant (Fs<1).

Self-paced reading

The results from Experiment 2b are plotted in Figure 10.

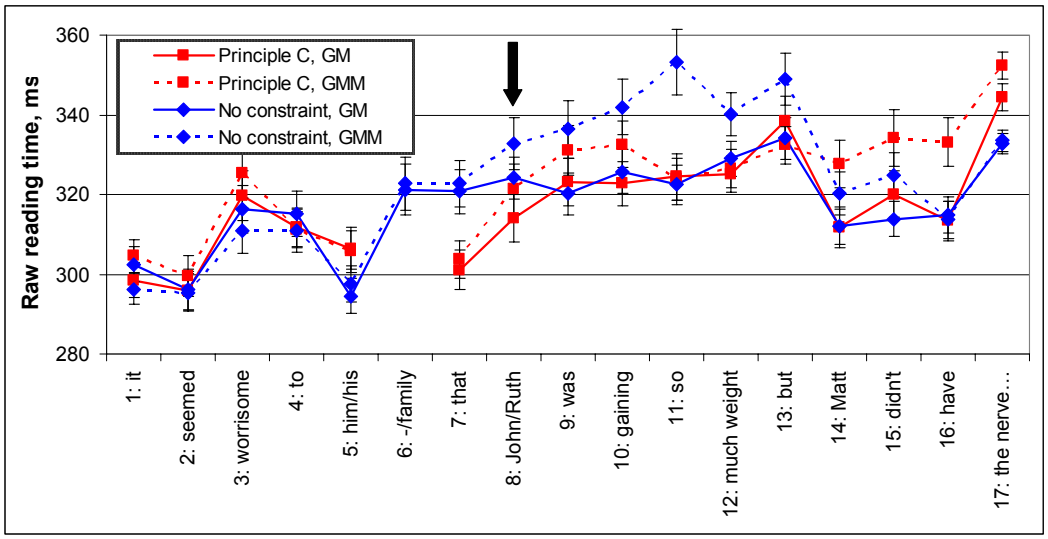


Figure 10. Mean reading times in milliseconds from Experiment 2b. The regions were as follows: It_1 $seemed_2$ $worrisome_3$ to_4 him_5/ his_5 $family_6$ $that_7$ $John/Ruth_8$ was_9 $gaining_{10}$ so_{11} (much weight)$_{12}$, but_{13} $Matt_{14}$ $didn't_{15}$ $have_{16}$ (the nerve to comment on it)$_{17}$.

In region 1 in the general ANOVA the interaction of the factors *constraint* and *congruency* was significant in the participants analysis and marginally significant in the items analysis ($F1(1,59)=4.7$, $p<.05$, $F2(1.31)=2.9$, $p=.096$). However, this interaction was not supported in the pairwise comparisons within each level of the *constraint* factor (all *ps*>1). There were no other significant effects or interactions anywhere in the first clause (all *p*s > .1).

In region 7, the complementizer *that*, there was a main effect of *constraint* ($F1(1,59)=16.3$, $p < .001$, $F2(1.31)=10.6$, $p < .01$). The effect was due to longer reading times for the no-constraint conditions than for the Principle C conditions, most likely due to the lexical difference in the preceding regions in the two pairs of conditions (*to her* vs. *to her professors*). A main effect of *constraint* in the same direction was also found in the following region, the subject of the 2nd clause (region 8, $F1(1,59)=4.7$, $p < .05$, $F2(1.31)=4.3$, $p < .05$). However, the interaction of the *constraint* factor with *congruency* did not reach significance (both *Fs* <<1), and planned pairwise comparisons revealed no effect of *congruency* in either condition (all $Fs \leq 1.6$, all *ps*> .2).

The analyses of variance in the three regions immediately following the subject (regions 9-11 in Figure 10) showed a significant or a marginally significant main effect of *congruency*,ith gender-mismatching conditions read more slowly than their gender-matching counterparts (region 9: $F1(1,59)=3.7$, $p=.06$, $F2(1.31)=3.9$, $p=.06$, region 10: $F1(1,59)=8.8$, $p<.01$, $F2(1.31)=5.0$, $p < .05$); region 11: $F1(1,59)=4.0$, $p<.05$, $F2(1.31)=3.1$, $p = .09$). The interaction of *constraint* and *congruency* was significant in region 11 ($F1(1,59)=6.7$, $p<.05$, $F2(1.31)=9.0$, $p < .01$), but not in earlier regions 9 or 10 (both *Fs*<1). However, planned pairwise comparisons in regions 9 and 10 showed a significant or marginally significant

effect of gender congruency (a gender-mismatch effect, the incongruent condition read more slowly than the congreuent condition) in the no-constraint conditions (region 9: F1(1,59)=3.2, p = .08, F2(1.31)=3.6, p = .07; region 10: F1(1,59)=4.1, p <.05, F2(1.31)=3.3, p = .08; region 11: F1(1,59)=10.0, p < .01, F2(1.31)=10.9, p < .01). One must note, however, that the GMME was numerically small until region 11 (region 9: 336 ms in the gender-mismatch condition vs. 320 ms in the gender-match condition; region 10: 342 vs. 326 ms; region 11: 353 vs. 322 ms). The effect of gender congruency did not reach significance in the Principle C conditions in any region in the 2nd clause. Nevertheless, the reading times were numerically longer following the incongruent 2nd subject in comparison to the congruent 2nd subject in region 9 (331 vs. 323 ms) and region 10 (333 vs. 323 ms). In region 10 the difference in reading times in the Principle C conditions was marginally significant (F1(1,59)=2.8, p =.1, F2(1.31)=3.1, p = .09).

In the 3rd clause there was a significant main effect of *congruency* at the subject of the 3rd clause (region 14) in the overall ANOVA (F1(1,59)=5.8, p <.05, F2(1.31)=8.5, p < .01), that was due to longer reading times in the gender-mismatch conditions than in the gender-match conditions. In the pairwise comparisons, the effect of gender congruency was significant in the Principle C conditions (F1(1,59)=6.5, p <.05, F2(1.31)=6.7, p < .05), but not in the no-constraint conditions (both *ps*>.1).

Discussion

In Experiment 2b we found a significant GMME in the no-constraint conditions 3 words down from the 2nd subject, and not in the Principle C conditions. However, we must admit that in the GMME in the no-constraint conditions in Experiment 2 was both weaker and delayed in comparison to Experiment 1.

One of the assumptions that we had about the stimuli in Experiment 2 is that the main predicate such as *It seemed clear* creates a strong prediction for an upcoming finite clause. Consequently, the subject of this clause is an antecedent position that is independently predicted, similar to the 2nd subject position in the no-constraint conditions in Experiment 1. If the two experiments are similar with respect to the properties of the critical subject position, it is unclear why in Experiment 2 the parser did not have the same strong reaction to manipulations in the 2nd subject position as in Experiment 1.

This consideration forced us to explore the issue of whether the complement clause is unambiguously predicted by the predicate in more detail. This question can be tested using a sentence completion task. If we were right in our intuitions, a sentence fragment like (47) would have to be completed by providing both an NP complement for the preposition *to* and a complement clause that would satisfy the content role of the predicate (a sample continuation that we predicted is given in (48)).

(47) It seemed clear to ...
(48) ... John that Mary was tired.

Experiment 2c was a completion task that was designed to quantify how strong the expectation for an embedded clause was.

3.5.2.4 Experiment 2c: Offline Completion Experiment

Participants

Thirteen undergraduate students from the University of Maryland took part in the experiment.

Materials & Design

Participants were given a pen-and-paper questionnaire which contained an initial sentence fragment such as (49). They were instructed to provide a completion to each sentence that seemed natural to them.

(49) It seemed clear to his….

We present completion data for fragments like (49). The main goal is to quantify how often the completion for the fragment contains a complement clause. Moreover, as (49) shows, the last word of the fragment was a possessor pronoun, e.g. *his*. This was done to estimate which proportion of the completions that contained an embedded clause also contained a referent for the pronoun as part of the embedded clause. All target fragments contained the masculine 3rd person singular pronoun *his* or plural *their*. The feminine pronoun *her* was not used in target sentences to avoid an ambiguity between the pronoun and the possessor homophones.
There were 6 target items involving *his* and an equal number if targets involving *their*, interspersed with 60 fragments of other types.

Results

Participants provided a clausal continuation to a fragment like (49) in 69% of all cases (71% in the 'his' condition and 67% in the 'their' condition). In 70% of the cases that had a second clause the complement clause contained a potential antecedent for the cataphoric pronoun *his* or *their* (82% for ‘his’ and 56% for ‘their’). This referent was the subject of the 2nd clause in 87% of all cases when the referent was present (81% for ‘him’, 96% for ‘their’).

Discussion

The results of Experiment 2c suggest that a sentence fragment such as (49) did not establish an unambiguous expectation for an upcoming embedded clause. In

almost 30% of all cases the completed sentence was a simple sentence. In part, this result might be explained by the subjects' desire to give the shortest possible continuation. Nevertheless such monoclausal completions are grammatical, given that the sentence-initial *it* can be taken for a 3rd person neuter pronoun with an unspecified referent, rather than an expletive.

The results of the completion task undermine our original assumption that the embedded clause is reliably predicted based on the initial predicate. If so, there was not a strong prediction for an upcoming NP position at the time when the pronoun was encountered in either the Principle C or the no-constraint conditions in the online Experiment 2b.

3.5.2.5 Experiment 2: General Discussion

The goal of Experiment 2 was to test the claim that Principle C immediately restricts dependency formation during online processing of backwards anaphora using more rigorous stimuli.

As in Experiment 1, in Experiment 2 we found a significant GMME in the no-constraint condition, and no significant effect in the Principle C condition. These findings support the claim that the parser does not search for an antecedent for the pronoun in positions that are subject to a grammatical constraint on coreference.

In Experiment 1 the absence or presence of the coreference constraint was confounded with variation in when the critical subject position could be projected in an anticipatory fashion. Experiment 2 was an improvement over the previous experiment in that it kept the availability of a prediction for the critical 2nd subject position constant across both conditions. Thus the presence or absence of a GMME in no-constraint vs. Principle C conditions in Experiment 2 cannot be an artifact of the presence or absence of a prediction for the critical 2nd subject position.

Overall, the results of Experiment 2 support our claim that Principle C immediately restricts the set of candidates considered by the parser. The GMME was significant following the crucial subject in the no-constraint condition, but not in the Principle C condition. However, we must admit that the results were not as clear as expected. For example, the 2-way interaction *constraint* x *congruency* did not become significant until region 11 (three regions downstream from the critical subject position), and the differences between the congruent and the incongruent conditions in the no-constraint pair were also weaker and later than in the previous experiment. Additionally, in the Principle C conditions there was a marginally significant trend in the direction of a GMME.

Moreover, the results of the completion experiment showed that our stimuli were ambiguous with regard to whether the beginning of the sentence sets up an independent prediction for the embedded clause. Our original belief was that the

embedded clause is strongly predicted in either condition on the basis of the main clause predicate (hence, before the pronoun is encountered). According to the completion task, however, continuations for a fragment like *It seemed worrisome to ...* contained a second clause only in 61% of all cases.

The unexpected results of the completion task do not create a challenge for our claim that grammatical constraints on coreference are immediately respected by the parser. Although lower than we had predicted, the degree of expectation for an embedded clause should have been the same across the no-constraint conditions and the Principle C conditions in Experiment 2 and thus could not have been the reason for the differences in the effects between the two pairs of conditions. However, the ambivalent status of stimuli in Experiment 2 prevents us from drawing clear conclusions regarding the nature of the search mechanism in backwards anaphora. Does the parser initiate a search for an antecedent for the cataphoric pronoun only when there is an independent prediction for an antecedent position at the time when the pronoun is processed, or is such prediction not required?

Let us compare the no-constraint conditions from Experiments 1 and 2, repeated for convenience in (50).

(50) *a. No-constraint conditions from Experiment 1*
Because last semester while **she** was taking classes full-time
Kathryn/Russell was working two jobs to pay the bills,…
b. No-constraint conditions from Experiment 2
It seemed worrisome to **his** family that **John/Ruth** was gaining so much weight, but …

Although both types of structures yielded a GMME at the 2nd subject in the no-constraint conditions, the effect was weaker in Experiment 2. This difference could be explained by stating that a search for an antecedent is initiated only if an independent prediction for an upcoming antecedent position is available at the time when the pronoun is encountered. The strength of the GMME would then correlate with the degree of availability of such prediction for a given structure. According to speakers' completions in Experiment 2c, the prediction is available and hence the search for an antecedent is initiated with a 100% likelihood in Experiment 1, but only with a 69% likelihood in Experiment 2. Then, putting it simply, the surprise effect to a mismatching antecedent should be observed in every trial in Experiment 1, whereas in Experiment 2 it should be present only roughly in 6/10 trials in which the parser had a previous expectation for an antecedent.

There is another explanation for why the GMME was weaker in Experiment 2 which does not make an assumption that the parser's active search for an antecedent is restricted to specific configurations. The parser might attempt to resolve anaphora regardless of the availability of a prediction for an antecedent

position at the time when the pronoun is encountered; thus, the no-constraint conditions in Experiments 1 and 2 did not differ in this respect. However, although a prediction for an antecedent position is not necessary to initiate an antecedent search, it could be necessary to obtain an observable GMME. This effect may be sufficiently strong to measure only if there is some previously existing top-down prediction for an antecedent that is not fulfilled bottom-up. In Experiment 1, the parser had ample time to establish a firm top-down prediction for an antecedent in the critical 2nd subject position (based on the subordinator *while* that made the main clause inevitable) before that subject was encountered bottom-up. In Experiment 2, on the other hand, due to the closeness between the complementizer *that* and the critical subject, the parser may not have had sufficient time to establish a prediction for that subject, which resulted in a weaker surprise effect for the mismatching condition.

Thus, under the first hypothesis whether the cataphoric pronoun triggers a search for an antecedent depends on the types of structural predictions that are available at the time when the pronoun is processed. If this is true, then resolution of backwards anaphora is different from *wh*-dependency processing: the parser always looks for a dependent element that is a *wh*-filler, but only does so under certain conditions when the dependent element is a cataphoric pronoun. Under the second hypothesis the pronoun initiates a search for an antecedent regardless of whether there is a prediction for an NP position. If so, the processing of *wh*-dependencies and backwards anaphora can be viewed as sharing a common 'active search' mechanism.

The aim of Experiment 3 was to distinguish between these hypotheses. To do so we tested whether the parser initiates a search for an antecedent if there is absolutely no syntactic prediction for a potential antecedent position available at the time when the cataphoric pronoun is encountered. The prediction appears later on, through a subordinator which is separated from the following critical gender-marked head noun by other lexical material, to supply ample time for the parser to process the prediction and thus increase the strength of the surprise effect. If GMME is found in such configuration, that would be evidence for the 'activeness' of the parser's search for an antecedent.

3.5.3 Experiment 3

Experiment 3 tested configurations with licit backwards anaphora (no-constraint conditions, (51) c,d) and configurations in which backwards anaphora is prohibited by Principle C (Principle C conditions, (51)a,b). All conditions in Experiment 3 unambiguously lacked a prediction for a potential antecedent position at the point when the cataphoric pronoun first appeared in the sentence. The first possible cue for such a position appeared later in the sentence when the

subordinator *while* was encountered. Importantly, the subordinator *while* did not immediately precede the critical gender-marked head noun of the 2nd subject NP as in Experiment 2 and was separated from the head noun by the determiner *the* followed by two adjectives. This was done to ensure that the parser had every opportunity to fully process the information associated with the subordinator before encountering the critical noun, i.e. that (i) there is a new clause and, hence, a subject position upcoming, and (ii) in the Principle C condition this clause is c-commanded by the cataphoric pronoun and thus every NP inside that clause is subject to Principle C.

(51) *a. Principle C, gender-match*
He chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids.
b. Principle C, gender-mismatch
She chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids.
c. No constraint, gender-match
His managers chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids.
d. No constraint, gender-mismatch
Her managers chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids.

Following familiar logic, if Principle C indeed immediately restricts the space of candidate antecedents, we expect no effect of gender congruency of the inaccessible 2nd subject in the Principle C condition. As usual, this null-effect will only be informative if there is an effect of gender congruency in the no-constraint condition, where the critical subject is a valid antecedent for the preceding pronoun.

3.5.3.1 Experiment 3a: Offline Rating Experiment

Participants

Participants were 60 students from the undergraduate population of the University of Maryland. They took part in Experiment 3a after completing the online task (Experiment 3b).

Materials & Design

Experiment 3a was a pen-and-paper questionnaire. The subject's task was to rate the plausibility of coreference between a pronoun and a noun highlighted in bold using a 1-to-5 scale, with 1 = absolutely impossible, 5 = absolutely natural.

The offline task contained the four conditions shown in (52) that were based on the congruent variant of the Principle C condition (53)a used for the online experiment (see Materials for Experiment 3b).

(52) A sample set of stimuli from an offline Experiment 3a
a. Principle C
He chatted amiably with some fans while **the talented, young quarterback** signed autographs for the kids.
b. No-constraint
His managers chatted amiably with some fans while **the talented, young quarterback** signed autographs for the kids.
c. Forwards Anaphora
The talented, young quarterback chatted amiably with some fans while **he** signed autographs for the kids.
d. But-condition
He chatted amiably with some fans while the talented, young quarterback signed autographs for the kids, but **Steve** wished the children's charity event would end soon so he could go home.

We expected our participants to reject coreference between the subject pronoun of the main clause and the embedded subject NP in (52)a that was subject to Principle C and to allow it in (52)b where the constraint does not apply. The purpose of including a forwards anaphora condition, formed by switching the order of the pronoun and the name in (52)a, was to ensure that with Principle C taken out of the picture the main and embedded events can be plausibly interpreted as being simultaneously performed by the same agent. The final condition (52)d was added to confirm that our effort to provide a licit intra-sentential antecedent for the pronoun through the subject of an additional *but*-clause reached the desired goal.

Twenty-four sets of stimuli were constructed based on 24 target stimulus sets from Experiment 3b. Four experimental lists were constructed using a Latin-square design, such that each subject saw only one condition from each set. In addition to the 24 experimental sentences, each questionnaire also contained 36 filler sentences that were identical across the four lists.

Results

The mean coreference rating scores from Experiment 3a are summarized in Table 4.

Table 4. Mean rating scores from Experiment 3a.

Condition	*Mean score (standard error)*
Principle C	1.7 (.09)
No-constraint	3.4 (.13)
Forwards anaphora	4.3 (.08)
But-condition	3.9 (.09)

The Principle C condition received a mean rating score of 1.7 that was significantly lower than the rating score in the other three conditions (2-tailed paired t-test, all *p*s ≤ .001).[49]

3.5.3.2 Experiment 3b: Online Sentence Reading Experiment

Participants

Participants were 60 students from the undergraduate population of the University of Maryland.

Materials & Design

Twenty-four sets of experimental stimuli such as (53) were created with an equal number of sets that had a masculine or a feminine noun as the subject of the 2nd clause.

(53) A full set of stimuli from Experiment 3b.
a. Principle C, gender-match
He chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids, but **Steve** wished the children's charity event would end soon so he could go home.
b. Principle C, gender-mismatch
She chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids, but **Carol** wished the children's charity event would end soon so she could go home.
c. No constraint, gender-match
His managers chatted amiably with some fans while the talented, young

[49] One may note that the coreference rating score in the no-constraint and the *but*-conditions was significantly lower than in the forwards anaphora condition. This is expected given that forwards anaphora is the preferred way of expressing coreference in the given contexts.

quarterback signed autographs for the kids, but **Carol** wished the children's charity event would end soon so she could go home.
d. No constraint, gender-mismatch
Her managers chatted amiably with some fans while the talented, young **quarterback** signed autographs for the kids, but **Carol** wished the children's charity event would end soon so she could go home.

In the Principle C condition the pronoun was the subject of a main clause that was followed by an embedded *while*-clause. In the no-constraint condition the pronoun was the possessor inside the main subject. Gender congruent and incongruent sentences were formed by manipulating the gender of the sentence-initial pronoun such that it either matched or mismatched the gender of the head noun of the subject NP of the 2nd clause, which was the critical position. The 2nd subject NP always consisted of the definite determiner *the*, followed by two adjectives and a gender-unambiguous head noun. The nouns used were either lexically (e.g. *king, queen*) or conventionally (e.g. *quarterback*) strongly gender-specific. Because in English neither determiners nor adjectives are gender-marked, and because the adjectives were chosen such that they were semantically plausible with both masculine and feminine head nouns (i.e. we avoided using adjectives that are biased towards one of the genders, such as *handsome* or *pregnant*), it was not until the head noun that the gender of the subject NP became clear.

As in the previous experiments, in order to ensure the possibility of intra-sentential coreference, the sentences had an additional clause that contained a licit antecedent for the cataphoric pronoun whenever no suitable antecedent was available in the 2nd clause.

The experiment followed a Latin Square design, with 24 experimental sentences interspersed with 72 fillers. The experimental procedure was identical to the procedure in Experiments 1 and 2.

Analysis

There were technical problems with the data from 2 out of 60 subjects (lists B & D) – the data were not saved by the software. To balance the number of subjects across each list, we excluded an additional subject from lists A and C. Those were also the only two subjects whose overall mean reading rate was 2.5 standard deviations above the average. Thus, analysis was performed on 56 subjects distributed equally between presentation lists. The procedure was identical to that used in Experiment 1: sentences that were answered incorrectly were excluded, and values that exceeded a threshold of 2.5 standard deviations above the mean reading rate for each region were replaced by that threshold value (winsorizing). The mean question accuracy was 92.6% (91.6-93.8% for individual conditions). Winsorizing affected 2.2% of all data (2.1-2.3% for individual conditions). The data were entered into a 2x2 ANOVA with the factors *constraint* (Principle C, no constraint) and *congruency* (gender-match, gender-mismatch).

Results

The results from all conditions in Experiment 3b are presented in Figure 11.

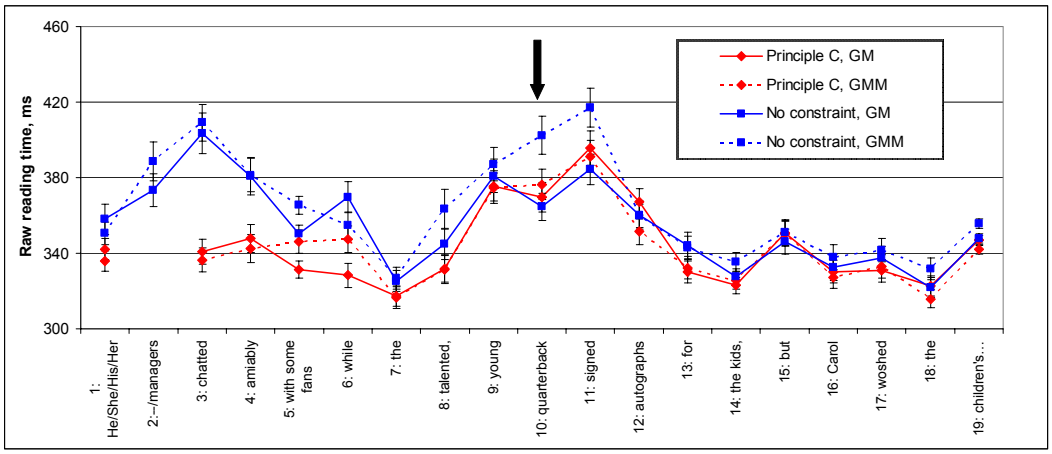


Figure 11. Mean reading times for the Principle C and no-constraint conditions from Experiment 3b. The arrow marks the position of the critical gender-marked head-noun. The regions were as follows:
He/she$_1$ [Principle C] … *or* His/her$_1$ managers$_2$ [no-constraint]…
… chatted$_3$ amiably$_4$ (with some fans)$_5$ while$_6$ the$_7$ talented,$_8$ young$_9$ quarterback$_{10}$ signed$_{11}$ autographs$_{12}$ for$_{13}$ (the kids,)$_{14}$ but$_{15}$ Steve/Carol$_{16}$ wished$_{17}$ the$_{18}$ (children's charity event would end soon so he could go home.)$_{19}$

In the first clause, there was a significant main effect of *constraint* in regions 3 & 4 (all *Fs* >12.5, all *ps* <.01) due to longer reading times in the no-constraint conditions than in the Principle C conditions. This effect can be explained by appealing to lexical differences in the preceding regions between the Principle C and no-constraint conditions. At the conjunction *while* (region 6) there was a significant main effect of *constraint* (F(1,55)=13.7, p<.001, F2(1,23)=14.9, p<.001) and a significant interaction *constraint* x *congruency* (F(1,55)=7.6, p<.01, F2(1,23)=11.2, p<.01). Pairwise comparisons revealed that the interaction was determined by an unexpected effect of *congruency* in the Principle C conditions in which region 6 was read significantly more slowly in the congruent sentence than in its incounguent counterpart (F(1,55)=4.1, p<.05, F2(1,23)=4.5, p<.05). The effect of *congruency* was not significant in the no-constraint conditions (both *Fs*<2.9, *ps*≥.1). The effect of *constraint* was marginally significant in region 7 (determiner *the*) (F(1,55)=3.6, p=.06, F2(1,23)=4.2, p=.05), and significant at the following adjective (region 8, *talented*) (F(1,55)=6.2, p<.05, F2(1,23)=11.4, p<.01).

In the critical region, at the subject noun of the 2nd clause (region 10) there was a main effect of congruency (F(1,55)=8.7, p<.01, F2(1,23)=4.9, p<.05) and a significant interaction *constraint* x *congruency*: (F(1,55)= 8.1, p<.01,

F2(1,23)=6.2, p<.05). Separate pairwise comparisons of the Principle C and no-constraint conditions revealed a strong effect of *congruency* in the no-constraint pair (F(1,55)=18.4, p<.001, F2(1,23)=14.0, p=.001), but no corresponding effect in the Principle C pair (both *Fs*<1). In the no-constraint pair the effect was in the predicted direction, with reading times in the incongruent condition on average 37 ms longer than in the congruent condition (i.e. a GMME).

The interaction *constraint* x *congruency* was also significant in the main ANOVA at the word following the subject noun (region 11): F(1,55)=4.2, p<.05, F2(1,23)=5.7, p<.05. Once again, pairwise comparisons within each level of the factor *constraint* showed that the interaction was due to the presence of a significant effect of congruency in the no-constraint conditions (F(1,55)=6.5, p=.01, F2(1,23)=8.6, p<.01), and its absence in the Principle C conditions (both *Fs*<1). There were no other significant effects in the remainder of the 2nd clause (in anticipation of potential concerns, the effect of congruency in region 12 in the Principle C conditions was significant only in the by-subject analysis: F(1,55)=7.1, p=.01, F2(1,23)=1.9, p>.1 and showed slower reading times in the gender matching condition, which is the opposite of a gender mismatch effect). There we no significant effects anywhere in the final clause.

Discussion

The main finding of Experiment 3 was the presence of a GMME in the no-constraint condition at the critical gender-marked head noun. The effect was present despite the fact that there was no independent prediction for an antecedent position at the time when the cataphoric pronoun in (53) was encountered. This confirms the hypothesis that the parser always initiates an active search for an antecedent for the cataphoric pronoun.

The GMME was not found at the critical 2nd subject or in any other region in the 2nd clause in the Principle C condition. These results confirm the claim based on our previous experiments that the initial set of antecedents for the cataphoric pronoun does not contain NP candidates that violate Principle C. Note that this in itself does not exclude the possibility that the 'illicit' NPs are considered at some later point, before they are ultimately rejected.

3.5.4 General Discussion of Experiments 1, 2 & 3

The logic and design of all three experiments was such that a null effect was expected in the Principle C condition under the hypothesis that binding constraints apply during the earliest stage of processing. The results of Experiment 1 showed a GMME in the no-constraint conditions, but the lack of such an effect in the Principle C conditions, and thus supported the view of the early application of the binding constraint.

However, the null effect in the Principle C conditions should be taken as informative only if there is evidence that a GMME could have been possible if it were not for the constraint. To this end there was an alternative possible interpretation of the results of Experiment 1, since there was an undesirable difference between the Principle C and no-constraint conditions related to the predictability of the critical 2nd subject position. This position was predictable at the time when the pronoun was processed in the no-constraint condition but not in the Principle C conditions, which could explain the respective presence vs. absence of an effect in that position. Consequently, Experiments 2 and 3 were designed such that the Principle C and no-constraint conditions were carefully matched on all parameters except the grammatical constraint. Experiment 2 was originally conceived as a test of the Principle C and no-constraint conditions with an equal and high predictive power for the critical subject position (i.e. prediction was originally thought to be available at the cataphoric pronoun, although the results of the completion experiment did not support that idea). In contrast, in Experiment 3 the critical antecedent position was equally unpredictable at the pronoun across the two conditions. In both Experiment 2, and especially Experiment 3, we obtained essentially the same results as in Experiment 1, i.e. the parser considered the critical subject as a potential antecedent, but only if it was not subject to a grammatical constraint on coreference.

Re-evaluating the results of Experiment 2, we can conclude that a weaker GMME is likely to be due to the closeness of the 2nd subject position to the complementizer *that*, its structural predictor. This is supported by the results of Experiment 3, in which the critical subject position was entirely unpredictable at the time when the pronoun was processed, yet was separated from its precursor by three words. This increased distance enabled the parser to fully process the subordinator and the accompanying structural prediction for a subject position (also supported by the appearance of the determiner and adjectives).

As for Experiment 1, we said earlier that the absence of a GMME in the Principle C condition could have been due to one of three reasons: (i) a grammatical constraint on coreference, (ii) an inability to make an independent structural prediction for the 2nd subject position at the time when the pronoun is processed, or (iii) the relative closeness between the subordinator *while*, which introduces a clause that is c-commanded by the cataphoric pronoun, and the critical subject position. Experiment 3 showed that the ability to predict a potential antecedent position at the time of processing the pronoun is not required to obtain a GMME, and that introducing significant separation between the predictor (*while*) and the gender-marked noun failed to elicit a GMME in the Principle C condition, thus ruling out options (ii) and (iii). Hence, we can now confidently attribute the lack of GMME in the Principle C condition in Experiment 1 to the parser's immediate respect for Principle C.

3.5.4.1 Binding Theory as an Early Filter during Coreference Resolution

We can now extend Table 1 that summarized the effects of Principles A & B on the processing of forwards anaphora to include the findings on backwards anaphora. Table 5 is a summary of the finding; new additions are italicized.

	Study	*Sample Stimuli*	*Binding Accessible vs Inaccessible*
	GMME		
FA	Sturt (2003)	… **the surgeon** pricked **himself/herself**…	Accessible
FA	Sturt (2003)	**Jonathan/Jennifer** remembered that X pricked **himself** …	Accessible
FA	Badecker & Straub (2002)	**John/Jane** thought that X owed **him** another chance …	Accessible
FA	Lee & Williams (2005)	**The midwife** confirmed that X warned **her/him** …	Accessible
BA	*Sturt et al. (2005)*[50]	After making **herself**/**himself** a cup of tea, **the nurse**…	Accessible
FA	Kennison (2003), Exp. 2	**Carl/Susan** watched **his** classmate ….	Accessible
BA	Van Gompel & Liversedge (2003)	When **he** was at the party, the **boy/girl** cruelly …	Accessible
BA	*this chapter, Exp. 1*	While **she** was taking classes full-time **Kathryn**/**Russell** was working two jobs …	Accessible
BA	*this chapter, Exp. 2*	It seemed worrisome to **his** family that **John/Ruth** was gaining weight …	Accessible
BA	*this chapter, Exp. 3*	**His/her** managers chatted amiably with some fans while the talented young **quarterback** signed autographs…	Accessible
	GME		
FA	Badecker & Straub (2002), Exp. 1-3	John thought that **Bill/Beth** owed **him** another chance …	Inaccessible
FA	Badecker & Straub (2002), Exp. 3	**John/Jane** thought that Bill owed **himself** another chance …	Inaccessible
FA	Kennison (2003), Exp. 1	**Carl/Susan** watched **him** yesterday ….	Inaccessible

[50] The study will be discussed in chapter 5.

	Null-effect		
FA	Sturt (2003)	**Jonathan/Jennifer** remembered that X pricked **himself**…	Inaccessible
FA	Badecker & Straub (2002), Exp. 5	Jane thought that **Bill's/Beth's** brother owed himself …	Inaccessible
FA	Badecker & Straub (2002), Exp. 6	It appeared to **John/Jane** that Bill owed **himself** …	Inaccessible
FA	Kennison (2003), Exp. 3	**Billy** complained about having a stomach ache. **Carl/Susan** watched **him** yesterday ….	Inaccessible
FA	Lee & Williams (2005)	The midwife confirmed that **Jonathan/Jennifer** warned **her/him** …	Inaccessible
BA	*this chapter, Exp. 1*	**She** was taking classes full-time while **Kathryn**/**Russell** was working two jobs …	Inaccessible
BA	*this chapter, Exp. 2*	It seemed worrisome to **him** that **John**/**Ruth** was gaining weight …	Inaccessible
BA	*this chapter, Exp. 3*	**He/she** chatted amiably with some fans while the talented young **quarterback** signed autographs…	Inaccessible

Table 5. Summary of findings (GMME, GME, null-effects) in sentence-processing studies of forwards and backwards anaphora.

We can see that the new additions to Table 5 corroborate our previous observations. In backwards anaphora GMME is found as a result of manipulation of the accessible antecedent. Manipulation in the antecedent that was inaccessible due to Principle C did not yield any effect on processing.[51] A sharp contrast in the treatment of binding-accessible and binding-inaccessible antecedents that holds in both forwards and backwards anaphora strengthens the claim that syntactic constraints on coreference are taken into consideration by the parser at the earliest stage of dependency formation. Moreover, this suggests that the parser uses a

[51] A relevant study that is not mentioned in Table 5 is by Aoshima, Phillips & Weinberg (2003) on referential dependencies in Japanese involving the personal pronouns *kare/kanojo* 'he/she' and a pronoun *soko* which allows both a bound and referential interpretation. Aoshima et al. argued that Japanese speakers actively search for antecedents of these pronouns only in grammatically sanctioned positions: there was a congruity effect at the antecedent in a grammatically licit position, but not in grammatically illicit position. The congruous accessible antecedent was read faster than its incongruous counterpart in the experiment involving personal pronouns, thus supporting our prediction. The effect was reversed in the experiment with bound pronouns, but the authors claimed that the reversal was due to an independent reason (the parser's preferential treatment of *soko* as referential in the absence of discourse cues and with the necessity to reanalyze this preference once an accessible quantificational antecedent is encountered). Assuming that this explanation is correct, Aoshima et al.'s finding can be taken as further support for the claim presented in this section.

common search mechanism in the processing of forwards and backwards anaphora, a mechanism that actively searches for an antecedent for the dependent element, but never looks in a binding-inaccessible position.

3.5.4.2 The Role of structural information in parsing

In explaining the results of our experiments we proposed that the parser makes an anticipatory top-down prediction regarding the accessibility of an upcoming antecedent position based on the cues that have been encountered bottom-up. An alternative to this is a view whereby the parser makes no anticipatory predictions for an antecedent and, instead, checks whether each NP could serve as an antecedent for the cataphoric pronoun as that NP becomes available from the input. Crucially, regardless of whether the parser makes a forward prediction, we must agree that it initially considers the accessibility of the position of an NP before checking whether that NP and the pronoun match in gender. Otherwise, as has been pointed out on several occasions, we would expect the same reaction by the parser to every morphologically incongruent NP independent of whether it is in a binding-accessible or inaccessible position[52].

In the remainder of this section I would like to put forward considerations which could explain why structural information provides expedient means for making forward predictions regarding candidate antecedents. The core reason is that this is often the only type of information about the antecedent that is reliably derivable before the NP is encountered bottom up. Other factors that affect whether an NP is a licit antecedent for the cataphoric pronoun are morphological features and semantic plausibility. However, unlike the structural position of an NP which may be independently predicted top-down, these types of information cannot be known until the NP is encountered bottom-up. In other words, the parser can confidently restrict the search space of candidate antecedents in an anticipatory fashion based on the structural position of a yet-to-be-seen NP. However, the parser cannot rule out an upcoming NP as a candidate antecedent based on a morphological mismatch with the dependent element in an anticipatory fashion, since the morphological features of an NP cannot be confidently predicted (that would

[52] Although the results are somewhat equivocal, at least some aspects point towards the view based on a forward prediction. In particular, if the parser does not make any forward prediction regarding potential antecedents, we would expect, contrary to fact, that the effect in Experiments 2 and 3 should be identical. These experiments look identical from the perspective of a parser that waits until an NP is available in the input before making any judgments about it. However, we have seen that increased distance between the predictor of an accessible antecedent and the antecedent itself results in a more immediate effect of that antecedent on processing (Experiment 3).

require clairvoyance, rather than just look-ahead, on the part of the parser) and can be known only when the NP is available bottom-up.[53]

Note therefore that there are natural reasons for why syntactic information takes precedence over other types of information in processing of backwards anaphora, or for this reason, any other case where the dependent element precedes its licensor. The argument that syntactic considerations apply first follows directly from considering what would constitute an optimal procedure if the parser is driven to establish a referential dependency with a cataphoric pronoun as soon as possible.[54] Syntactic information turns out to have the crucial role because it enables the parser to make predictions about the upcoming material earlier than any other type of information. This reasoning is similar in spirit to interactive approaches to parsing in that there are no architectural constraints that force certain information types to have priority.

If so, it is the phenomenon of forwards anaphora that provides the necessary conditions for exploring the question of whether syntactic information has an architectural priority over other types of information. In forwards anaphora the search for the antecedent is retrospective and all information about the antecedent is available simultaneously. Yet, as shown by the review of the literature in section 3.4.1, despite various types of information being equally accessible, structural considerations seem to never lag behind morphological considerations. Why should syntactic information be taken into consideration immediately even in forwards anaphora?

I suggest that this preference may be due the role of structural vs. morphological considerations in validating a dependency. Whereas structural considerations are relevant for every type of dependency (e.g. every dependency is subject to structural constraints on well-formedness), morphological information may at times be less useful. A vivid example is a backwards anaphora dependency with null elements, such as obligatory control dependencies involving null PRO in English. The dependent element PRO must be controlled by an antecedent in a specific position in the structure. For example, in sentences like (54), the PRO subject of the adjunct clause must be controlled by the subject of the main clause, i.e. the lemonade drinker must be *Marie*. That PRO must have an antecedent and is illicit unless there is one is illustrated by (55) & (56). Both sentences are ill-formed, because the main subject is not an appropriate antecedent for the PRO, either because it is an expletive *there* and thus not an NP at all, or because it is an inappropriate agent of drinking (*the weather*).

[53] This statement primarily pertains to English. In languages with richer morphological agreement gender/number of an NP may be known before the NP is encountered bottom-up. This point is addressed later in the section.

[54] Here optimality means finding a licit antecedent as soon as possible and with the least possible amount of false alarms.

(54) a. After drinking some lemonade, Marie sat down to watch soccer with Alik.
b. After PRO drinking some lemonade, Marie sat down to watch soccer.
(55) *After drinking some lemonade, there were no drinks left.
(56) *After drinking some lemonade, the weather started to improve.

Thus, a PRO subject of the adjunct clause must have an antecedent in the main subject position. The parser can use (and, in fact, does use according to Sturt et al. 2005) this structural information to unambiguously determine the appropriate antecedent, (most likely) in an anticipatory manner. However, since PRO is phonologically null (and in the absence of overt agreement on the predicate in English), the parser cannot possibly make predictions about any of the morphological features of the antecedent and restrict the space of possible antecedents based on their gender/number. Interestingly, there seem to be no reverse cases in which morphological information by itself is fully sufficient for defining which licensors are licit for some dependent element and structural constraints do not apply.[55] Even in languages with rich morphological agreement where the morphological properties of a yet-unseen NP can be predicted from, say, agreement morphemes ona previously encountered verb, the prediction is only available for nominals in certain structural positions (e.g., subject, object, indirect object). Thus, once again, morphological information is coupled together with structural information, similar to what we have seen in forwards anaphora. If this is correct and if the parser is indeed geared towards an optimal solution, proposals where a certain type of morphological information takes architectural precedence over structural information (Kennison 2003; Badecker & Straub 2002) seem unproductive.

To summarize, we have argued that if there is a common basic mechanism of active dependency formation that applies to various types of dependencies, it should strongly rely on structural considerations for reasons of efficiency and optimality. In any dependency the structural relation between its elements is always relevant and critical, whereas the relevance of other factors may vary. These general considerations, in addition to the empirical evidence that the parser relies on structural information to avoid false alarms, provide justification and support for the claim that structural constraints on dependency formation are respected immediately by the parser.

---

[55] This statement should not be confused with the claim that morphology never helps disambiguating between several candidate antecedents. The pair of examples below shows that it definitely can. For example in (i) below the fact that the pronoun *her* must unambiguously refer to *Ana* is due to gender considerations (compare to (ii) where morphology cannot disambiguate *her* between *Eli* and *Ana*). Note however, that morphology can never be the sole factor for determining the licit antecedent, as otherwise we would expect *Ana* to be the antecedent of *her* in (iii), contrary to fact.

(i) Ana told John that everyone in Bilbao missed her.
(ii) (ii) Ana told Eli that everyone in Bilbao missed her.
(iii) (iii) John told David that Ana missed her.

3.5.4.3 Wh-dependency vs. Backwards Anaphora

We can also make a more informed conclusion on whether backwards anaphora dependencies can be viewed as involving the same active dependency completion mechanisms as *wh*-dependencies. On the one hand, given various similarities between these types of long-distance dependencies (see section 3.3), similar active mechanisms could plausibly be implicated in processing of these dependencies: the dependent element initiates an active search for the licensor in the following material. On the other hand, despite the various similarities, there are good reasons to believe that the processing mechanism need not be identical. Unlike a *wh*-filler, the pronoun is not required to have an overt licensor in the discourse and may in principle refer to an unspecified discourse referent. So, if an online search mechanism is sensitive to how mandatory a dependency is, we may see less pressure for dependency formation in the case of a cataphoric pronoun than with a *wh*-filler. Unlike the *wh*-filler which would always trigger a search for a gap, the pronoun may do so only under some special conditions. For example, it is conceivable that the search for an antecedent is initiated only if there is an independent prediction for an upcoming NP position at the time when the parser encounters the cataphoric pronoun.

We can get a fair amount of mileage regarding the nature of the search mechanisms from the results of Experiment 3. In this experiment in both the Principle C and no-constraint conditions the critical 2nd subject position could not be predicted at the time when the pronoun was processed. A top-down prediction for this subject emerged as soon as the parser encountered the subordinator *while* later in the sentence. Nevertheless, we observed a GMME in the no-constraint condition, suggesting that the parser actively considers the possibility of an antecedent in this position. Thus, the existence of a reliable structural prediction for an NP position at the time when the pronoun is processed is not a necessary requirement for predicting an antecedent position in future material. Such a prediction can be made later as some new information projects an upcoming NP position. To put it in simply, the cataphoric pronoun's need to find an antecedent remains active even when that pronoun cannot be immediately linked to some projected antecedent in the upcoming input. Hence, the processing of *wh*-dependencies and backwards anaphora may use the same underlying active mechanism that involves look-ahead to predict the position of dependency completion as soon as possible, yet it is selective in that it actively avoids postulating a dependency that would cause a violation of a grammatical constraint.[56]

[56] In the most straightforward case the earliest possible completion of a dependency refers to satisfying all needs of the dependent element by a licensor as early as possible. In reality, however, a dependent element may have more than one feature that need to be satisfied, e.g. scope and thematic-role for a *wh*-word, and if so, the earliest possible licensing position may not be the same

It is worth mentioning that an active search for an antecedent is triggered even in the case where the dependent element is of a relatively low discourse prominence. Badecker & Straub (2003) discuss a prominence hierarchy for R-expressions (previously discussed in Gordon, Hendrick, Ledoux & Yang 1999 and primarily based on Ariel's (1990) Accessibility Theory) according to which genitive NP possessors are less prominent than nominative NP subjects. One could conceivably apply the same ranking to pronouns in terms of their search strength. If so, pronouns that are more prominent on the hierarchy would presumably be more active 'seekers' for an antecedent than the less prominent ones. If so, in Experiment 3 the cataphoric pronoun *he* or *she* in the Principle C conditions should have triggered a more active search for an antecedent than a possessive pronoun *his* or *her* in the no-constraint condition. Nevertheless, we obtained the GMME in the no-constraint condition, but not in the Principle C condition. Our results thus suggests that (i) even less prominent pronouns set an active search for an antecedent, and (ii) regardless of how pressing is the need for an antecedent, the parser obeys grammatical constraints online during formation of referential dependencies.

A note of caution is appropriate here, in light of the narrow range of positions that we probed in our experiments. In all our experiments the critical position was the subject position. According to our view the same effects should obtain in non-subject positions, e.g. an object position. In SVO languages the prediction for the object position comes from the verb, e.g. transitive *kissed* in (57) or ditransitive *sent* in (57):

(57) a. **His** mother kissed **John/ Mary** before leaving.
b. **His** mother sent a parcel to **John/Mary** on December 20.

In (57), the parser that is actively expecting an antecedent for the pronoun in the earliest accessible position should take the verb complement to be such antecedent. Hence, a surprise GMME should obtain whenever there is a gender mismatch between the pronoun and the object. However, the same position should not be under consideration in (58), in which coreference between the pronoun subject and the object violates Principle B.

(58) a. **He** kissed **John/Mary** before leaving.
b. **He** sent a parcel to **John/Mary** on December 20.

---

for different features. In such cases it needs to be seen which of the features is the major driving force behind an active dependency formation. Based on their research on Japanese referential dependencies, Aoshima, Phillips & Yoshida (2005) propose that the search for an antecedent is first driven by considerations of scope when the dependent element is scope-taking (*wh*-word) and by a thematic role when it is not scope taking (a referential NPs, e.g. *that boy*) (although, needless to say, that all features need to be ultimately satisfied).

Configurations such as (57)&(58) are potentially good candidates for future tests of our model.

3.5.4.4 A Note On Fully Lexicalized Parsing

MacDonald, Pearlmutter & Seidenberg (1994) propose a reductionist model whereby all principles of sentence processing are represented as part of a lexical entry for individual words in the lexicon. I briefly assess this model in terms of binding dependencies and claim that the parser's immediate respect of Principle C poses a challenge for the model.

In MacDonald et al.'s model sentences arise as chains of X-bar schemas from the lexicon, which are put together in accordance with constraints on how X-bar structures can be linked together. The bulk of the job is performed by argument structures of lexical items, but MacDonald et al. (1994) recognize that there may be other constraints (e.g. island constraints or binding constraints), which in the strong version of their theory should also all be statable in the lexicon.[57] It is this strong version of the theory that I would like to argue against using evidence on processing of referential dependencies.

One of the most challenging tests for the claim that all syntax is reducible to local X-bar schemas involves long-distance dependencies. To their credit, MacDonald et al. (1994) explicitly discuss this point, although the issue is much less prominent, if raised at all, in later developments of the theory. MacDonald et al. (1994) agree that the parser's ability to construct authentic long-distance dependencies would pose a challenge for their account. However, the authors express some doubt whether what is considered to be a long-distance dependency is more than a sequence of local relations (e.g. successive cyclicity of *wh*-dependencies, Chomsky 1986).

The status of referential dependencies and constraints on them in this model should be considered along the lines of wh-dependencies. In order to stay loyal to the spirit of the proposal, one must say that binding constraints operate locally, in terms of specific lexical items or X-bar templates. Such a formulation is available for Principles A and B which have a roughly clause-size scope; moreover, as mentioned in section 3.4.2, Reinhart & Reuland (1994) claim that these principles *must be* stated in terms of properties of predicates (see also fn. 41, p. 58). On the contrary, Principle C (or Reinhart's (1983, 2000) rule on logical binding), regulates relations between elements across clause boundaries and can scope over

[57] The strong view that all syntax can be lexicalized which is stated cautiously in MacDonald, Pearlmutter & Seidenberg (1994) becomes prevalent in later work (Seidenberg & MacDonald 1999, Allen & Seidenberg 1999).

virtually unbounded dependencies. The existence of a constraint that is not reducible to local notions is problematic for a lexicalized account.[58]

Arguably, though, this fact per se is insufficient, since MacDonald et al.'s claims pertain to parsing, rather than static grammatical representations. If Principle C were violated by the parser, it could signal a special status of Principle C which, unlike locally-defined constraints, is 'unhandlable' in real-time by the parser. If so, this would support a claim whereby parsing principle are considered exclusively in terms of local lexical choices. Yet, we found that the parser respects Principle C immediately, just like in the case of other binding principles that are local, which is problematic for the strongly lexicalist view of parsing.

3.6 CONCLUSION

This chapter discussed the real-time processing of sentences with backwards anaphora in English. We showed that upon encountering a cataphoric pronoun the parser starts an active search for its antecedent in the following material. Importantly, during this search the parser does not consider positions which are inaccessible antecedent positions due to a structural constraint on coreference, Principle C. These results support a view whereby binding constraints are considered to apply immediately during the process of structure building.

[58] More precisely, the issue is not so much whether Principle C can be represented in terms of local relations, but rather whether the outcome is anything other than a notational variant of a syntactic constraint. For examples, Principle C can be instantiated as a restriction on coreference that percolates down from the X-bar structure that contains a pronoun 'link-by-link' to each subsequent X-bar structure that it c-commands. However, such formulation is based on structural considerations and therefore foreign to at least a strong version of MacDonald et al.'s theory.

Appendix 3-A. Full list of stimuli from Experiment 1B

The experiment contained 30 sets of five conditions. In each set the conditions are as follows:

Condition a: Principle C, gender-matched
Condition b: Principle C, gender-misma ched
Condition c: no-constraint, gender-matched
Condition d: no-constraint, gender-mismatched
Condition e: name control

1a. Although every Sunday she ate breakfast while Sue worked on the crossword, Jessica never offered to help.
? Was it a crossword that Sue was working on? Y
1b. Although every Sunday she ate breakfast while Ben worked on the crossword, Jessica never offered to help.
? Was it a crossword that Ben was working on? Y
1c. Although every Sunday while she ate breakfast Sue worked on the crossword, Ben never offered to help.
? Was it a crossword that Sue was working on? Y
1d. Although every Sunday while she ate breakfast Ben worked on the crossword, Jessica never wanted to help.
? Was it a crossword that Ben was working on? Y
1e. Although every Sunday while Jessica ate breakfast Ben worked on the crossword, she never wanted to help.
? Was it a crossword that Ben was working on? Y

2a. Although after breakfast she vacuumed the hotel's lobby while Meg waited for the repairman to call, Alicia didn't have time to do the hallway.
? Was it the hotel's lobby that got vacuumed? Y
2b. Although after breakfast she vacuumed the hotel's lobby while Jim waited for the repairman to call, Alicia didn't have time to do the hallway.
? Was it the hotel's lobby that got vacuumed? Y
2c. Although after breakfast while she vacuumed the hotel's lobby Meg waited for the repairman to call, Jim got tired of waiting and contacted them himself.
? Was it the hotel's lobby that got vacuumed? Y
2d. Although after breakfast while she vacuumed the hotel's lobby Jim waited for the repairman to call, Meg ended up having to answer the phone herself.
? Was it the hotel's lobby that got vacuumed? Y
2e. Although after breakfast while Meg vacuumed the hotel's lobby Jim waited for the repairman to call, she ended up having to answer the phone herself.
? Was it the hotel's lobby that got vacuumed? Y

3a. Although yesterday morning he wandered along the park's trails while Carl investigated the local wildlife, Brad didn't complain of boredom.
? Did Brad complain of being bored? N
3b. Although yesterday morning he wandered along the park's trails while Anne investigated the local wildlife, Brad didn't complain of boredom.
? Did Brad complain of being bored? N
3c. Although yesterday morning while he wandered along the park's trails Carl investigated the local wildlife, Anne was the one who discovered the rare lizard specimen.
? Was it a rare insect that Anne discovered? N
3d. Although yesterday morning while he wandered along the park's trails Anne investigated the local wildlife, Brad was the one who discovered the rare lizard specimen.
? Was it a rare insect that Brad discovered? N
3e. Although yesterday morning while Brad wandered along the park's trails Anne investigated the local wildlife, he was the one who discovered the rare lizard specimen.
? Was it a rare insect that Brad discovered? N

4a. Although during the fire he was searching frantically for the apartment's fire extinguisher while Jacob was screaming for help, Bill didn't have much hope that they would escape.
? Did Bill find the fire extinguisher? N
4b. Although during the fire he was searching frantically for the apartment's fire extinguisher while Laura was screaming for help, Bill didn't have much hope that they would escape.
? Did Bill find the fire extinguisher? N

4c. Although during the fire while he was searching frantically for the apartment's fire extinguisher Jacob was screaming for help, Laura could hear only the sound of the smoke alarm.
? Did Laura hear the screaming? N
4d. Although during the fire while he was searching frantically for the apartment's fire extinguisher Laura was screaming for help, Jacob could hear only the sound of the smoke alarm.
? Did Bill hear the screaming? N
4e. Although during the fire while Bill was searching frantically for the apartment's fire extinguisher Laura was screaming for help, he could hear only the sound of the smoke alarm.
? Did Bill hear the screaming? N

5a. Although last fall he was traveling in Egypt while Andrew was digging for ancient artifacts, Doug never got a chance to visit the excavation site.
? Was it dinosaur bones that Andrew was digging for? N
5b. Although last fall he was traveling in Egypt while Carrie was digging for ancient artifacts, Doug never got a chance to visit the excavation site.
? Was it dinosaur bones that Carrie was digging for? N
5c. Although last fall while he was traveling in Egypt Andrew was digging for ancient artifacts, Carrie never got a chance to visit him at the excavation site.
? Was it dinosaur bones that Andrew was digging for? N
5d. Although last time fall while he was traveling in Egypt Carrie was digging for ancient artifacts, Doug never got a chance to visit her at the excavation site.
? Was it dinosaur bones that Carrie was digging for? N
5e. Although last time fall while Doug was traveling in Egypt Carrie was digging for ancient artifacts, he never got a chance to visit her at the excavation site.
? Was it dinosaur bones that Carrie was digging for? N

6a. Although last weekend he planted daisies in the church's garden while Dennis pulled up weeds, Chad decided to switch jobs for this weekend.
? Was it roses that were being planted? N
6b. Although last weekend he planted daisies in the church's garden while Rachel pulled up weeds, Chad decided to switch jobs for this weekend.
? Was it roses that were being planted? N
6c. Although last weekend while he planted daisies in the church's garden Dennis pulled up weeds, Rachel accidentally replanted the flowering ones later.
? Was it roses that were being planted? N
6d. Although last weekend while he planted daisies in the church's garden Rachel pulled up weeds, Chad still saw some left in the flowerbed.
? Was it roses that were being planted? N
6e. Although last weekend while Chad planted daisies in the church's garden Rachel pulled up weeds, he still saw some left in the flowerbed.
? Was it roses that were being planted? N

7a. Although last year he was training for the Olympics while Aaron was writing a book, Frank made sure they met for lunch at least once a week.
? Was it a book that Aaron was working on? Y
7b. Although last year he was training for the Olympics while Wendy was writing a book, Frank made sure they met for lunch at least once a week.
? Was it a book that Wendy was working on? Y
7c. Although last year while he was training for the Olympics Aaron was writing a book, Wendy made him still set aside time for her.
? Was it a book that Aaron was working on? Y
7d. Although last year while he was training for the Olympics Wendy was writing a book, Frank made sure they met for lunch at least once a week.
? Was it a book that Wendy was working on? Y
7e. Although last year while Frank was training for the Olympics Wendy was writing a book, he made sure they met for lunch at least once a week.
? Was it the Olympics that Frank was training for? Y

8a. Although Monday morning he filled the car's tank while Adam asked the employee for directions, Dave ran out of gas halfway there anyway.
? Did Dave get enough gas for the whole trip? N

8b. Although Monday morning he filled the car's tank while Judy asked the employee for directions, Dave ran out of gas halfway there anyway.
? Did Dave get enough gas for the whole trip? N
8c. Although Monday morning while he filled the car's tank Adam asked the employee for directions, Judy thought they should just buy a map.
? Was it a flashlight that Judy thought they should buy? N
8d. Although Monday morning while he filled the car's tank Judy asked the employee for directions, Dave thought they should just buy a map.
? Was it a flashlight that Dave thought they should buy? N
8e. Although Monday morning while Dave filled the car's tank Judy asked the employee for directions, he thought they should just buy a map.
? Was it a flashlight that Judy thought they should buy? N

9a. Although every night she pored over the course's textbook while Amber watched the kids, Cindy continually got poor grades on all the papers.
? Did Cindy get bad grades on her papers? Y
9b. Although every night she pored over the course's textbook while Brian watched the kids, Cindy continually got poor grades on all the papers.
? Did Cindy get bad grades on her papers? Y
9c. Although every night while she pored over the course's textbook Amber watched the kids, Brian still put them to bed himself.
? Did Brian still put the kids to bed? Y
9d. Although every night while she pored over the course's textbook Brian watched the kids, Cindy still put them to bed herself.
? Did Cindy still put the kids to bed? Y
9e. Although every night while Cindy pored over the course's textbook Brian watched the kids, she still put them to bed herself.
? Did Cindy still put the kids to bed? Y

10a. Although Sunday night he slept soundly while John was attacked by mosquitoes, Matt was awakened the following night by gnats biting his arms.
? Was it Sunday night that Ruth was attacked? Y
10b. Although Sunday night he slept soundly while Ruth was attacked by mosquitoes, Matt was awakened the following night by gnats biting his arms.
? Was it Sunday night that Ruth was attacked? Y
10c. Although Sunday night while he slept soundly John was attacked by mosquitoes, Ruth didn't get bitten at all.
? Was it Sunday night that John was attacked? Y
10d. Although Sunday night while he slept soundly Ruth was attacked by mosquitoes, Matt was awakened the following night when the bugs came after him.
? Was it Sunday night that Ruth was attacked? Y
10e. Although Sunday night while Matt slept soundly Ruth was attacked by mosquitoes, he was awakened the following night when the bugs came after him.
? Was it Sunday night that Ruth was attacked? Y

11a. Although this week she was beginning a season on Broadway while Jen was recovering from the flu, Audrey made a point of visiting her after performances.
? Was it pneumonia that Jen was recovering from? N
11b. Although this week she was beginning a season on Broadway while Rob was recovering from the flu, Audrey made a point of visiting him after performances.
? Was it pneumonia that Rob was recovering from? N
11c. Although this week while she was beginning a season on Broadway Jen was recovering from the flu, Rob couldn't hear any weakness in her voice.
? Was it pneumonia that Jen was recovering from? N
11d. Although this week while she was beginning a season on Broadway Rob was recovering from the flu, Audrey made him come anyway to provide moral support.
? Was it pneumonia that Rob was recovering from? N
11e. Although this week while Jen was beginning a season on Broadway Rob was recovering from the flu, she made him come anyway to provide moral support.
? Was it pneumonia that Rob was recovering from? N

12a. Although after the concert he was scribbling a positive review while Heather was waiting in line for an autograph, Jeff changed it when the band blew the fans off.
? Was the band gracious to the fans? N
12b. Although after the concert he was scribbling a positive review while Heather was waiting in line for an autograph, Jeff changed it when the band blew the fans off.
? Was the band gracious to the fans? N
12c. Although after the concert while he was scribbling a positive review Brandon was waiting in line for an autograph, Heather didn't see him because his head was down.
? Did Heather see Brandon? N
12d. Although after the concert while he was scribbling a positive review Heather was waiting in line for an autograph, Jeff didn't see her because his head was down.
? Did Jeff see Heather? N
12e. Although after the concert while Jeff was scribbling a positive review Heather was waiting in line for an autograph, he didn't see her because his head was down.
? Did Jeff see Heather? N

13a. Although in December he was studying for finals while Jason was vacationing in the Caribbean, Max didn't waste much time feeling jealous.
? Was it in December that Jason was vacationing in the Caribbean? Y
13b. Although in December he was studying for finals while Cindy was vacationing in the Caribbean, Max didn't waste much time feeling jealous.
? Was it in December that Cindy was vacationing in the Caribbean? Y
13c. Although in December while he was studying for finals Jason was vacationing in the Caribbean, Cindy hoped that he would be able to get at least a little work done.
? Was it in December that Jason was vacationing in the Caribbean? Y
13d. Although in December while he was studying for finals Cindy was vacationing in the Caribbean, Max didn't waste much time wishing he had gone along.
? Was it in December that Cindy was vacationing in the Caribbean? Y
13e. Although in December while Max was studying for finals Cindy was vacationing in the Caribbean, he didn't waste much time wishing he had gone along.
? Was it in December that Cindy was vacationing in the Caribbean? Y

14a. Although after the midterm he was hanging out in the club's loft while Mark was playing pool, Nick wasn't in the mood to join the game.
? Was Nick in the mood to play pool? N
14b. Although after the midterm he was hanging out in the club's loft while Jill was playing pool, Nick wasn't in the mood to join the game.
? Was Nick in the mood to play pool? N
14c. Although after the midterm while he was hanging out in the club's loft Mark was playing pool, Jill wasn't in the mood to join the game.
? Was Jill in the mood to play pool? N
14d. Although after the midterm while he was hanging out in the club's loft Jill was playing pool, Nick wasn't in the mood to join the game.
? Was Nick in the mood to play pool? N
14e. Although after the midterm while Nick was hanging out in the club's loft Jill was playing pool, he wasn't in the mood to join the game.
? Was Nick in the mood to play pool? N

15a. Although during the hurricane she made oatmeal cookies while Barbara listened attentively to the public service updates, Michelle was only pretending to be calm.
? Was it music that Barbara was listening to? N
15b. Although during the hurricane she made oatmeal cookies while Barbara listened attentively to the public service updates, Michelle was only pretending to be calm.
? Was it music that Charlie was listening to? N
15c. Although during the hurricane while she made oatmeal cookies Barbara listened attentively to the public service updates, Charlie kept trying to have a conversation with her.
? Was it music that Barbara was listening to? N
15d. Although during the hurricane while she made oatmeal cookies Charlie listened attentively to the public service updates, Michelle kept trying to have a conversation with him.
? Was it music that Charlie was listening to? N

15e. Although during the hurricane while Michelle made oatmeal cookies Charlie listened attentively to the public service updates, she kept trying to have a conversation with him.
? Was it music that Charlie was listening to? N

16a. Because last semester she was taking classes full-time while Kathryn was working two jobs to pay the bills, Erica felt guilty.
? Was Kathryn working two jobs? Y
16b. Because last semester she was taking classes full-time while Russell was working two jobs to pay the bills, Erica felt guilty.
? Was Russell working two jobs? Y
16c. Because last semester while she was taking classes full-time Kathryn was working two jobs to pay the bills, Russell never got to see her.
? Was Kathryn working two jobs? Y
16d. Because last semester while she was taking classes full-time Russell was working two jobs to pay the bills, Erica promised to work part-time in the future.
? Was Russell working two jobs? Y
16e. Because last semester while Erica was taking classes full-time Russell was working two jobs to pay the bills, she promised to work part-time in the future.
? Was Russell working two jobs? Y

17a. Because in May he was auctioning off the museum's jewels while Danny was trying to arrange tickets to South America, Carl was afraid they had aroused suspicion.
? Was it the museum's jewels that were auctioned off? Y
17b. Because in May he was auctioning off the museum's jewels while Nancy was trying to arrange tickets to South America, Carl was afraid they had aroused suspicion.
? Was it the museum's jewels that were auctioned off? Y
17c. Because in May while he was auctioning off the museum's jewels Danny was trying to arrange tickets to South America, Nancy started to get suspicious.
? Was it the museum's jewels that were auctioned off? Y
17d. Because in May while he was auctioning off the museum's jewels Nancy was trying to arrange tickets to South America, Carl was afraid they would arouse suspicion.
? Was it the museum's jewels that were auctioned off? Y
17e. Because in May while Carl was auctioning off the museum's jewels Nancy was trying to arrange tickets to South America, he was afraid they would arouse suspicion.
? Was it the museum's jewels that were auctioned off? Y

18a. Because in early 2003 he was editing the film's last scenes while Justin was finishing up the soundtrack, Brad expected the movie to hit theatres by summer.
? Did Brad think the movie would be out by Christmas? Y
18b. Because in early 2003 he was editing the film's last scenes while Denise was finishing up the soundtrack, Brad expected the movie to hit theatres by summer.
? Did Brad think the movie would be out by Christmas? Y
18c. Because in early 2003 while he was editing the film's last scenes Justin was finishing up the soundtrack, Denise expected the movie to hit theatres by summer.
? Did Denise think the movie would be out by Christmas? Y
18d. Because in early 2003 while he was editing the film's last scenes Denise was finishing up the soundtrack, Brad expected the movie to hit theatres by summer.
? Did Brad think the movie would be out by Christmas? Y
18e. Because in early 2003 while Brad was editing the film's last scenes Denise was finishing up the soundtrack, he expected the movie to hit theatres by summer.
? Did Brad think the movie would be out by Christmas? Y

19a. Because last month he was heading the party's congressional campaign while Peter came under investigation by the IRS, Jack had to cut ties with him.
? Was it the FBI who was investigating Peter? N
19b. Because last month he was heading the party's congressional campaign while Karen came under investigation by the IRS, Jack had to cut ties with her.
? Was it the FBI who was investigating Karen? N
19c. Because last month while he was heading the party's congressional campaign Peter came under investigation by the IRS, Karen replaced him as chief advisor.
? Was it the FBI who was investigating Peter? N

19d. Because last month while he was heading the party's congressional campaign Karen came under investigation by the IRS, Jack removed her from the advisory staff.
? Was it the FBI who was investigating Karen? N
19e. Because last month while Jack was heading the party's congressional campaign Karen came under investigation by the IRS, he removed her from the advisory staff.
? Was it the FBI who was investigating Karen? N

20a. Because last summer she lifeguarded at the city's public pool while Pam practiced diving, Molly decided to try the sport herself.
? Did Molly decide to try diving? Y
20b. Because last summer she lifeguarded at the city's public pool while Tom practiced diving, Molly decided to try the sport herself.
? Did Molly decide to try diving? Y
20c. Because last summer while she lifeguarded at the city's public pool Pam practiced diving, Tom decided to ask her for lessons.
? Did Tom ask for diving lessons? Y
20d. Because last summer while she lifeguarded at the city's public pool Tom practiced diving, Molly came to know his trademark dives by heart.
? Did Molly learn Tom's trademark dives? Y
20e. Because last summer while Molly lifeguarded at the city's public pool Tom practiced diving, she came to know his trademark dives by heart.
? Did Molly learn Tom's trademark dives? Y

21a. Because last fall she was running for re-election while Sandra was being treated for cancer, Julie made health care a core campaign issue.
? Was education Julie's core issue? N
21b. Because last fall she was running for re-election while Walter was being treated for cancer, Julie made health care a core campaign issue.
? Was education Julie's core issue? N
21c. Because last fall while she was running for re-election Sandra was being treated for cancer, Walter advised her to drop out of the race.
? Was Walter's advice to stay in the race? N
21d. Because last fall while she was running for re-election Walter was being treated for cancer, Julie made health care her core campaign issue.
? Was education Julie's core issue? N
21e. Because last fall while Julie was running for re-election Walter was being treated for cancer, she made health care her core campaign issue.
? Was education Julie's core issue? N

22a. Because on Monday he was making reservations while Steven was still adding to the guest list, Kyle was told to call back later.
? Was the guest list complete yet? N
22b. Because on Monday he was making reservations while Debbie was still adding to the guest list, Kyle was told to call back later.
? Was the guest list complete yet? N
22c. Because on Monday while he was making reservations Steven was still adding to the guest list, Debbie told him to call back later.
? Was the guest list complete yet? N
22d. Because on Monday while he was making reservations Debbie was still adding to the guest list, Kyle decided to call back later.
? Was the guest list complete yet? N
22e. Because on Monday while Debbie was making reservations Steven was still adding to the guest list, he decided to call back later.
? Was the guest list complete yet? N

23a. Because last year she was in France while Erin was working under a famous Parisian chef, Melissa ate gourmet meals almost every night.
? Was it an Italian chef that Erin was working for? N
23b. Because last year she was in France while Alan was working under a famous Parisian chef, Melissa ate gourmet meals almost every night.
? Was it an Italian chef that Alan was working for? N

23c. Because last year while she was in France Erin was working under a famous Parisian chef, Alan went to eat at that restaurant all the time.
? Was it an Italian chef that Erin was working for? N
23d. Because last year while she was in France Alan was working under a famous Parisian chef, Melissa was able to get reservations at the restaurant whenever she wanted.
? Was it an Italian chef that Alan was working for? N
23e. Because last year while Melissa was in France Alan was working under a famous Parisian chef, she was able to get reservations at the restaurant whenever she wanted.
? Was it an Italian chef that Alan was working for? N

24a. Because during intermission she was testing the stage's microphone while Sarah was preparing for the big scene, Elaine got several dirty looks from the other actors.
? Did Elaine get dirty looks from people? Y
24b. Because during intermission she was testing the stage's microphone while Roger was preparing for the big scene, Elaine got several dirty looks from the other actors.
? Did Elaine get dirty looks from people? Y
24c. Because during intermission while she was testing the stage's microphone Sarah was preparing for the big scene, Roger kept having to repeat her cues.
? Did Roger have to repeat Sarah's cues? Y
24d. Because during intermission while she was testing the stage's microphone Roger was preparing for the big scene, Elaine tried to keep her voice down.
? Did Elaine try to keep her voice down? Y
24e. Because during intermission while Elaine was testing the stage's microphone Roger was preparing for the big scene, she tried to keep her voice down.
? Did Elaine try to keep her voice down? Y

25a. Because in April she was training for the women's World Cup while Mary was still recovering from a hamstring injury, Sharon couldn't visit her often at the hospital.
? Was it a hamstring injury that Mary was recovering from? Y
25b. Because in April she was training for the women's World Cup while Ryan was still recovering from a hamstring injury, Sharon couldn't visit her often at the hospital.
? Was it a hamstring injury that Ryan was recovering from? Y
25c. Because in April while she was training for the women's World Cup Mary was still recovering from a hamstring injury, Ryan advised her to run less than usual.
? Was it a hamstring injury that Mary was recovering from? Y
25d. Because in April while she was training for the women's World Cup Ryan was still recovering from a hamstring injury, Sharon usually ran twice as much as he did.
? Was it a hamstring injury that Ryan was recovering from? Y
25e. Because in April while Sharon was training for the women's World Cup Ryan was still recovering from a hamstring injury, she usually ran twice as much as he did.
? Was it a hamstring injury that Ryan was recovering from? Y

26a. Because Sunday afternoon he was singing along with Eminem's new single while Richard was watching the football game, Josh was told in no uncertain terms to shut up.
? Was it Friday night that Richard was watching the game? N
26b. Because Sunday afternoon he was singing along with Eminem's new single while Theresa was watching the football game, Josh was told in no uncertain terms to shut up.
? Was it Friday night that Theresa was watching the game? N
26c. Because Sunday afternoon while he was singing along with Eminem's new single Richard was watching the football game, Theresa asked him if he could just do one or the other.
? Was it Friday night that Richard was watching the game? N
26d. Because Sunday afternoon while he was singing along with Eminem's new single Theresa was watching the football game, Josh was told in no uncertain terms to shut up.
? Was it Friday night that Theresa was watching the game? N
26e. Because Sunday afternoon while Josh was singing along with Eminem's new single Theresa was watching the football game, he was told in no uncertain terms to shut up.
? Was it Friday night that Theresa was watching the game? N

27a. Because Thursday night she was in the kitchen cooking dinner while Tina was watching the TV drama, Cassie could hear frequent gasps of surprise and horror from the other room.
? Could Cassie hear Tina gasping at the show? Y

27b. Because Thursday night she was in the kitchen cooking dinner while Eric was watching the TV drama, Cassie could hear frequent gasps of surprise and horror from the other room.
? Could Cassie hear Eric gasping at the show? Y
27c. Because Thursday night while she was in the kitchen cooking dinner Tina was watching the TV drama, Eric could hear her frequent gasps of surprise and horror.
? Could Eric hear Tina gasping at the show? Y
27d. Because Thursday night while she was in the kitchen cooking dinner Eric was watching the TV drama, Cassie could hear his frequent gasps of surprise and horror.
? Could Cassie hear Eric gasping at the show? Y
27e. Because Thursday night while Cassie was in the kitchen cooking dinner Eric was watching the TV drama, she could hear his frequent gasps of surprise and horror.
? Could Cassie hear Eric gasping at the show? Y

28a. Because last Friday she was bartending while Crystal was checking IDs, Linda heard lots of stories about local kids with fake licenses.
? Was it checking for weapons that was Crystal's job? N
28b. Because last Friday she was bartending while Michael was checking IDs, Linda heard lots of stories about local kids with fake licenses.
? Was it checking for weapons that was Michael's job? N
28c. Because last Friday while she was bartending Crystal was checking IDs, Michael told her to be tough on all the local kids with fake licenses.
? Was it checking for weapons that was Crystal's job? N
28d. Because last Friday while she was bartending Michael was checking IDs, Linda grew to know him very well.
? Was it checking for weapons that was Michael's job? N
28e. Because last Friday while Linda was bartending Michael was checking IDs, she grew to know him very well.
? Was it checking for weapons that was Michael's job? N

29a. Because in 1998 she was a Chief Researcher at the Pentagon while Dawn was working on a major space project, Joyce offered to help with obtaining grant money.
? Was it in 1998 that Dawn was working on the space project? Y
29b. Because in 1998 she was a Chief Researcher at the Pentagon while Paul was working on a major space project, Joyce offered to help with obtaining grant money.
? Was it in 1998 that Paul was working on the space project? Y
29c. Because in 1998 while she was a Chief Researcher at the Pentagon Dawn was working on a major space project, Paul saw her regularly at government workshops.
? Was it in 1998 that Dawn was working on the space project? Y
29d. Because in 1998 while she was a Chief Researcher at the Pentagon Paul was working on a major space project, Joyce saw him regularly at government workshops.
? Was it in 1998 that Paul was working on the space project? Y
29e. Because in 1998 while Joyce was a Chief Researcher at the Pentagon Paul was working on a major space project, she saw him regularly at government workshops.
? Was it in 1998 that Paul was working on the space project? Y

30a. Because last night she was studying physics while Lisa was listening to the heavy metal station, Maria had a hard time concentrating on the assignment.
? Did Maria find it hard to concentrate because of the music? Y
30b. Because last night she was studying physics while Tony was listening to the heavy metal station, Maria had a hard time concentrating on the assignment.
? Did Maria find it hard to concentrate because of the music? Y
30c. Because last night while she was studying physics Lisa was listening to the heavy metal station, Tony decided to work in a different part of the house.
? Did Maria find it hard to concentrate because of the music? Y
30d. Because last night while she was studying physics Tony was listening to the heavy metal station, Maria had a hard time concentrating on the assignment.
? Did Tony decide to work in another part of the house? Y
30e. Because last night while Maria was studying physics Tony was listening to the heavy metal station, she had a hard time concentrating on the assignment.
? Did Maria find it hard to concentrate because of the music? Y

Appendix 3-B. Full list of stimuli from Experiment 2B

The experiment contained 30 sets of four conditions. In each set the conditions are as follows:

Condition a: Principle C, gender-matched
Condition b: Principle C, gender-mismatched
Condition c: no-constraint, gender-matched
Condition d: no-constraint, gender-mismatched

1a. It was clear to her that Sue would win the scholarship, but Jessica decided to apply for it anyway.
? Was it clear that Sue would win the scholarship? Y
1b. It was clear to her that Ben would win the scholarship, but Jessica decided to apply for it anyway.
? Was it clear that Ben would win the scholarship? Y
1c. It was clear to her professors that Sue would win the scholarship, but Ben decided to apply for it anyway.
? Was it clear that Sue would win the scholarship? Y
1d. It was clear to her professors that Ben would win the scholarship, but Jessica decided to apply for it anyway.
? Was it clear that Ben would win the scholarship? Y

2a. It seemed suspicious to her that Meg always kept the curtains closed, but Alicia was afraid to investigate further.
? Did Meg always keep the curtains closed? Y
2b. It seemed suspicious to her that Jim always kept the curtains closed, but Alicia was afraid to investigate further.
? Did Jim always keep the curtains closed? Y
2c. It seemed suspicious to her neighbors that Meg always kept the curtains closed, but Jim didn't think anything of it.
? Did Meg always keep the curtains closed? Y
2d. It seemed suspicious to her neighbors that Jim always kept the curtains closed, but Alicia didn't think anything of it.
? Did Jim always keep the curtains closed? Y

3a. It was surprising to him that Carl hadn't finished reading the book, although Brad too had found it quite difficult.
? Had Brad found the book easy? N
3b. It was surprising to him that Anne hadn't finished reading the book, although Brad too had found it quite difficult.
? Had Brad found the book easy? N
3c. It was surprising to his classmates that Carl hadn't finished reading the book, although Anne had also found it quite difficult.
? Had Anne found the book easy? N
3d. It was surprising to his classmates that Anne hadn't finished reading the book, although Brad had also found it quite difficult.
? Had Brad found the book easy? N

4a. It seemed bizarre to him that Jacob was running for class president, but Mark tried to be supportive.
? Was Mark trying to be supportive? Y
4b. It seemed bizarre to him that Linda was running for class president, but Mark tried to be supportive.
? Was Mark trying to be supportive? Y
4c. It seemed bizarre to his friends that Jacob was running for class president, but Linda encouraged them to be supportive.
? Did Linda encourage them to be supportive? Y
4d. It seemed bizarre to his friends that Linda was running for class president, but Mark encouraged them to be supportive.
? Did Mark encourage them to be supportive? Y

5a. It seemed strange to him that Andrew wore dark glasses indoors, but Doug figured it was probably some kind of fashion statement.
? Did it seem normal that Andrew was wearing dark glasses indoors? N
5b. It seemed strange to him that Carrie wore dark glasses indoors, but Doug figured it was probably some kind of fashion statement.
? Did it seem normal that Carrie was wearing dark glasses indoors? N
5c. It seemed strange to his students that Andrew wore dark glasses indoors, but Carrie knew that the professor had a serious health condition.
? Did it seem normal that Andrew was wearing dark glasses indoors? N
5d. It seemed strange to his students that Carrie wore dark glasses indoors, but Doug knew that the professor had a serious health condition.
? Did it seem normal that Carrie was wearing dark glasses indoors? N

6a. It was amazing to him that Dennis had landed a job at Microsoft, even though Chad had heard they were hiring a lot of people.
? Was it at McDonald's that Dennis had landed a job? N
6b. It was amazing to him that Rachel had landed a job at Microsoft, even though Chad had heard they were hiring a lot of people.
? Was it at McDonald's that Rachel had landed a job? N
6c. It was amazing to his colleagues that Dennis had landed a job at Microsoft, even though Rachel had said they were hiring a lot of people.
? Was it at McDonald's that Dennis had landed a job? N
6d. It was amazing to his colleagues that Rachel had landed a job at Microsoft, even though Dennis had said they were hiring a lot of people.
? Was it at McDonald's that Rachel had landed a job? N

7a. It was irritating to him that Aaron was always late for work, but Frank hadn't decided whether to issue an official reprimand.
? Had Frank decided whether to issue a reprimand? N
7b. It was irritating to him that Wendy was always late for work, but Frank hadn't decided whether to issue an official reprimand.
? Had Frank decided whether to issue a reprimand? N
7c. It was irritating to his staff that Aaron was always late for work, but Wendy thought that an official reprimand would be too harsh.
? Did Wendy think a reprimand was a good idea? N
7d. It was irritating to his staff that Wendy was always late for work, but Frank thought that an official reprimand would be too harsh.
? Did Frank think a reprimand was a good idea? N

8a. It seemed funny to him that Adam wasn't more nervous about the interview, but Dave thought it was probably a good thing.
? Was Adam very nervous about the interview? N
8b. It seemed funny to him that Judy wasn't more nervous about the interview, but Dave thought it was probably a good thing.
? Was Judy very nervous about the interview? N
8c. It seemed funny to his friends that Adam wasn't more nervous about the interview, but Judy thought it was probably a good thing.
? Was Adam very nervous about the interview? N
8d. It seemed funny to his friends that Judy wasn't more nervous about the interview, but Dave thought it was probably a good thing.
? Was Judy very nervous about the interview? N

9a. It was obvious to her that Amber was slacking off at work, but Cindy didn't want to say anything about it.
? Was it obvious that Amber was slacking off at work? Y
9b. It was obvious to her that Brian was slacking off at work, but Cindy didn't want to say anything about it.
? Was it obvious that Brian was slacking off at work? Y
9c. It was obvious to her coworkers that Amber was slacking off at work, but Brian wanted to give her another chance.
? Was it obvious that Amber was slacking off at work? Y
9d. It was obvious to her coworkers that Brian was slacking off at work, but Cindy wanted to give him another chance.
? Was it obvious that Brian was slacking off at work? Y

10a. It seemed worrisome to him that John was gaining so much weight, but Matt didn't have the nerve to comment on it.
? Was John gaining a lot of weight? Y
10b. It seemed worrisome to him that Ruth was gaining so much weight, but Matt didn't have the nerve to comment on it.
? Was Ruth gaining a lot of weight? Y
10c. It seemed worrisome to his family that John was gaining so much weight, but Ruth thought it was just a result of aging.
? Was John gaining a lot of weight? Y
10d. It seemed worrisome to his family that Ruth was gaining so much weight, but Matt thought it was just a result of aging.
? Was Ruth gaining a lot of weight? Y

11a. It seemed encouraging to her that Jen was spending more time on homework, but Audrey still thought a tutor might be needed.
? Did Audrey think a tutor might be needed? Y

11b. It seemed encouraging to her that Rob was spending more time on homework, but Audrey still thought a tutor might be needed.
? Did Audrey think a tutor might be needed? Y
11c. It seemed encouraging to her parents that Jen was spending more time on homework, but Rob still thought a tutor might be needed.
? Did Rob think a tutor might be needed? Y
11d. It seemed encouraging to her parents that Rob was spending more time on homework, but Audrey still thought a tutor might be needed.
? Did Audrey think a tutor might be needed? Y

12a. It seemed unusual to him that Brandon had packed a coat for the trip to Florida, but Jeff didn't question it.
? Did it seem natural that Brandon had packed a coat for the trip to Florida? N
12b. It seemed unusual to him that Heather had packed a coat for the trip to Florida, but Jeff didn't question it.
? Did it seem natural that Heather had packed a coat for the trip to Florida? N
12c. It seemed unusual to his roommates that Brandon had packed a coat for the trip to Florida, but Heather didn't question it.
? Did it seem natural that Brandon had packed a coat for the trip to Florida? N
12d. It seemed unusual to his roommates that Heather had packed a coat for the trip to Florida, but Jeff didn't question it.
? Did it seem natural that Heather had packed a coat for the trip to Florida? N

13a. It seemed alarming to her that Barbara was stationed in the war zone, but Michelle tried not to think about it.
? Did it seem alarming that Barbara was in the war zone? Y
13b. It seemed alarming to her that Charlie was stationed in the war zone, but Michelle tried not to think about it.
? Did it seem alarming that Charlie was in the war zone? Y
13c. It seemed alarming to her family that Barbara was stationed in the war zone, but Charlie told them there was nothing to worry about.
? Did it seem alarming that Barbara was in the war zone? Y
13d. It seemed alarming to her family that Charlie was stationed in the war zone, but Michelle told them there was nothing to worry about.
? Did it seem alarming that Charlie was in the war zone? Y

14a. It seemed natural to her that Sandra would want to move away from Alaska, but Julie advised her to finish out the year.
? Was it Hawaii that Sandra wanted to move away from? N
14b. It seemed natural to her that Walter would want to move away from Alaska, but Julie advised her to finish out the year.
? Was it Hawaii that Walter wanted to move away from? N
14c. It seemed natural to her boss that Sandra would want to move away from Alaska, but Walter thought she should wait for her kids to finish school.
? Was it Hawaii that Sandra wanted to move away from? N
14d. It seemed natural to her boss that Walter would want to move away from Alaska, but Julie thought she should wait for his kids to finish school.
? Was it Hawaii that Walter wanted to move away from? N

15a. It seemed appropriate to her that Erin step down after the scandal, even though Melissa believed her to be innocent.
? Did Melissa believe her to be guilty? N
15b. It seemed appropriate to her that Alan step down after the scandal, even though Melissa believed him to be innocent.
? Did Melissa believe him to be guilty? N
15c. It seemed appropriate to her employees that Erin step down after the scandal, even though Alan had proof of her innocence.
? Was it proof of her guilt that Alan had? N
15d. It seemed appropriate to her employees that Alan step down after the scandal, even though Melissa had proof of his innocence.
? Was it proof of her guilt that Melissa had? N

16a. It seemed logical to her that Sarah finish school before getting married, even though Elaine thought it was a really good match.
? Did it seem logical that Sarah finish school before getting married? Y
16b. It seemed logical to her that Roger finish school before getting married, even though Elaine thought it was a really good match.
? Did it seem logical that Roger finish school before getting married? Y

16c. It seemed logical to her sister that Sarah finish school before getting married, even though Roger didn't want to wait.
? Did it seem logical that Sarah finish school before getting married? Y
16d. It seemed logical to her sister that Roger finish school before getting married, even though Elaine didn't want to wait.
? Did it seem logical that Roger finish school before getting married? Y

17a. It was offensive to him that Jacob insulted the diplomat, but David decided to ignore the incident.
? Was it the secretary that Jacob insulted? N
17b. It was offensive to him that Debra insulted the diplomat, but David decided to ignore the incident.
? Was it the secretary that Debra insulted? N
17c. It was offensive to his advisor that Jacob insulted the diplomat, but Debra decided to ignore the incident.
? Was it the secretary that Jacob insulted? N
17d. It was offensive to his advisor that Debra insulted the diplomat, but David decided to ignore the incident.
? Was it the secretary that Debra insulted? N

18a. It was unfortunate for him that Kyle had misplaced the house keys, but Greg managed to get in through a window.
? Did Greg manage to get in? Y
18b. It was unfortunate for him that Rita had misplaced the house keys, but Greg managed to get in through a window.
? Did Greg manage to get in? Y
18c. It was unfortunate for his roommates that Kyle had misplaced the house keys, but Rita managed to get in through a window.
? Did Rita manage to get in? Y
18d. It was unfortunate for his roommates that Rita had misplaced the house keys, but Kyle managed to get in through a window.
? Did Kyle manage to get in? Y

19a. It was surprising to him that Thomas had requested the money, but Daniel felt few qualms about sending it.
? Was it surprising that Thomas had requested the money? Y
19b. It was surprising to him that Allison had requested the money, but Daniel felt few qualms about sending it.
? Was it surprising that Alison had requested the money? Y
19c. It was surprising to his accountant that Thomas had requested the money, but Alison was not shocked at all.
? Was it surprising that Thomas had requested the money? Y
19d. It was surprising to his accountant that Alison had requested the money, but Thomas was not shocked at all.
? Was it surprising that Alison had requested the money? Y

20a. It was acceptable to him that Nick would drive the car through the mountains, but Mike hoped there would be no need for it.
? Was it in the city that Nick would drive the car? N
20b. It was acceptable to him that Dana would drive the car through the mountains, but Mike hoped there would be no need for it.
? Was it in the city that Dana would drive the car? N
20c. It was acceptable to his companions that Nick would drive the car through the mountains, but Dana hoped there would be no need for it.
? Was it in the city that Nick would drive the car? N
20d. It was acceptable to his companions that Dana would drive the car through the mountains, but Nick hoped there would be no need for it.
? Was it in the city that Dana would drive the car? N

21a. It was fine with him that Harry had cancelled the magazine subscription, but Keith wished that Harry had asked first.
? Did Keith wish Harry had asked first? Y
21b. It was fine with him that Laura had cancelled the magazine subscription, but Keith wished that Laura had asked first.
? Did Keith wish Laura had asked first? Y
21c. It was fine with his patients that Harry had cancelled the magazine subscription, but Laura was worried that the nurses would miss it.
? Was Laura worried the nurses would miss the magazine? Y
21d. It was fine with his patients that Laura had cancelled the magazine subscription, but Harry was worried that the nurses would miss it.
? Was Harry worried the nurses would miss the magazine? Y

22a. It was shocking to him that Eric was pleading guilty to the charges, but Will came to terms with it quickly.
? Was it shocking that Eric was pleading guilty to the charges? Y
22b. It was shocking to him that Gail was pleading guilty to the charges, but Will came to terms with it quickly.
? Was it shocking that Gail was pleading guilty to the charges? Y

22c. It was shocking to his lawyers that Eric was pleading guilty to the charges, but Gail understood that he did it to save the familyÆs honor.
? Was it shocking that Eric was pleading guilty to the charges? Y
22d. It was shocking to his lawyers that Gail was pleading guilty to the charges, but Eric understood that he did it to save the familyÆs honor.
? Was it shocking that Gail was pleading guilty to the charges? Y

23a. It was laughable to him that Frank intended to record rap albums, but Barry didnÆt want to discourage him from trying.
? Was it bluegrass albums that Frank intended to record? N
23b. It was laughable to him that Sarah intended to record rap albums, but Barry didnÆt want to discourage her from trying.
? Was it bluegrass albums that Sarah intended to record? N
23c. It was laughable to his siblings that Frank intended to record rap albums, but Sarah saw it as something exciting and new.
? Was it bluegrass albums that Frank intended to record? N
23d. It was laughable to his siblings that Sarah intended to record rap albums, but Frank saw it as something exciting and new.
? Was it bluegrass albums that Sarah intended to record? N

24a. It was comforting to her that Renee had landed safely, but Karen still felt nervous about her daughter being so far away.
? Did Karen feel nervous about her daughter being far away? Y
24b. It was comforting to her that Kevin had landed safely, but Karen still felt nervous about her son being so far away.
? Did Karen feel nervous about her son being far away? Y
24c. It was comforting to her manager that Renee had landed safely, but Kevin was still worried about the rest of the trip.
? Was Kevin still worried about the rest of the trip? Y
24d. It was comforting to her manager that Kevin had landed safely, but Renee was still worried about the rest of the trip.
? Was Renee still worried about the rest of the trip? Y

25a. It seemed unwise to her that Pamela remarry weeks after the divorce, but Marcia didnÆt want to voice any misgivings.
? Did it seem unwise that Pamela remarry so soon? Y
25b. It seemed unwise to her that Steven remarry weeks after the divorce, but Marcia didnÆt want to voice any misgivings.
? Did it seem unwise that Steven remarry so soon? Y
25c. It seemed unwise to her therapist that Pamela remarry weeks after the divorce, but Steven thought it was the right choice.
? Did it seem unwise that Pamela remarry so soon? Y
25d. It seemed unwise to her therapist that Steven remarry weeks after the divorce, but Marcia thought it was the right choice.
? Did it seem unwise that Steven remarry so soon? Y

26a. It seemed wrong to her that Paula took credit for the project, but Anita swallowed her pride and continued working hard.
? Did it seem right that Paula took credit for the project? N
26b. It seemed wrong to her that Gavin took credit for the project, but Anita swallowed her pride and continued working hard.
? Did it seem right that Gavin took credit for the project? N
26c. It seemed wrong to her colleagues that Paula took credit for the project, but Gavin claimed not to have given it a second thought.
? Did it seem right that Paula took credit for the project? N
26d. It seemed wrong to her colleagues that Gavin took credit for the project, but Anita claimed not to have given it a second thought.
? Did it seem right that Gavin took credit for the project? N

27a. It seemed unimportant to her that Beverly had taken the car keys, but Crystal tried to sympathize with her housematesÆ frustrations.
? Was it the office keys that Beverly had taken? N
27b. It seemed unimportant to her that Charlie had taken the car keys, but Crystal tried to sympathize with her housematesÆ frustrations.
? Was it the office keys that Charlie had taken? N

27c. It seemed unimportant to her housemates that Beverly had taken the car keys, but Charlie asked her to leave a note in the future.
? Was it the office keys that Beverly had taken? N
27d. It seemed unimportant to her housemates that Charlie had taken the car keys, but Beverly asked him to leave a note in the future.
? Was it the office keys that Charlie had taken? N

28a. It was evident to her that Liz had eaten all the donuts, but Amy decided to pretend not to notice.
? Did Amy pretend not to notice? Y
28b. It was evident to her that Bob had eaten all the donuts, but Amy decided to pretend not to notice.
? Did Amy pretend not to notice? Y
28c. It was evident to her partner that Liz had eaten all the donuts, but Bob decided to pretend not to notice.
? Did Bob pretend not to notice? Y
28d. It was evident to her partner that Bob had eaten all the donuts, but Liz decided to pretend not to notice.
? Did Liz pretend not to notice? Y

29a. It seemed troubling to her that Polly listened to such loud music, but Irene eventually passed it off as a custom of youth.
? Did Polly listen to loud music? Y
29b. It seemed troubling to her that Craig listened to such loud music, but Irene eventually passed it off as a custom of youth.
? Did Craig listen to loud music? Y
29c. It seemed troubling to her relatives that Polly listened to such loud music, but Craig was fine with it as long as the chores got done.
? Did Polly listen to loud music? Y
29d. It seemed troubling to her relatives that Craig listened to such loud music, but Polly was fine with it as long as the chores got done.
? Did Craig listen to loud music? Y

30a. It seemed admirable to her that Miriam volunteered to feed the homeless, but Louise didnÆt have time to help out.
? Did Louise have time to help out? N
30b. It seemed admirable to her that Nathan volunteered to feed the homeless, but Louise didnÆt have time to help out.
? Did Louise have time to help out? N
30c. It seemed admirable to her father that Miriam volunteered to feed the homeless, but Nathan thought it was a waste of time.
? Did Nathan think volunteering was worthwhile? N
30d. It seemed admirable to her father that Nathan volunteered to feed the homeless, but Miriam thought it was a waste of time.
? Did Miriam think volunteering was worthwhile? N

31a. It was frightening to him that Trevor was skydiving regularly, but Paul still thought about joining in occasionally.
? Was it frightening that Trevor was skydiving? Y
31b. It was frightening to him that Amanda was skydiving regularly, but Paul still thought about joining in occasionally.
? Was it frightening that Amanda was skydiving? Y
31c. It was frightening to his friends that Trevor was skydiving regularly, but Amanda knew it must be a big rush.
? Was it frightening that Trevor was skydiving? Y
31d. It was frightening to his friends that Amanda was skydiving regularly, but Paul knew it must be a big rush.
? Was it frightening that Amanda was skydiving? Y

32a. It was puzzling to her that Katie was not interested in dancing, but Nancy chalked it up to different interests.
? Was Katie interested in dancing? N
32b. It was puzzling to her that Tommy was not interested in dancing, but Nancy chalked it up to different interests.
? Was Tommy interested in dancing? N
32c. It was puzzling to her teammates that Katie was not interested in dancing, but Tommy had asked about it and now understood why.
? Was Katie interested in dancing? N
32d. It was puzzling to her teammates that Tommy was not interested in dancing, but Katie had asked about it and now understood why.
? Was Tommy interested in dancing? N

Appendix 3-C. Full list of stimuli from Experiment 3B

The experiment contained 24 sets of four conditions. In each set the conditions are as follows:

Condition a: Principle C, gender-matched
Condition b: Principle C, gender-mismatched
Condition c: no-constraint, gender-matched
Condition d: no-constraint, gender-mismatched

1a. He announced the war strategy while the brave, noble king gathered up the troops, but James thought their roles should have been reversed.
? Was it a war strategy that was announced? Y
1b. She announced the war strategy while the brave, noble king gathered up the troops, but Diane thought their roles should have been reversed.
? Was it a war strategy that was announced? Y
1c. His advisors announced the war strategy while the brave, noble king gathered up the troops, but Diane felt their plans lacked cunning.
? Was it a war strategy that was announced? Y
1d. Her advisors announced the war strategy while the brave, noble king gathered up the troops, but Diane felt their plans lacked cunning.
? Was it a war strategy that was announced? Y

2a. He kept grading papers in the teachers' lounge while the tall, imposing headmaster talked loudly to parents on the phone, but Walter tried his best not to listen in.
? Was it the children the headmaster was talking to? N
2b. She kept grading papers in the teachers' lounge while the tall, imposing headmaster talked loudly to parents on the phone, but Sandra tried her best not to listen in.
? Was it the children the headmaster was talking to? N
2c. His aides kept grading papers in the teachers' lounge while the tall, imposing headmaster talked loudly to parents on the phone, but Sandra thought it wasn't proper and asked them to leave.
? Was it the children the headmaster was talking to? N
2d. Her aides kept grading papers in the teachers' lounge while the tall, imposing headmaster talked loudly to parents on the phone, but Sandra thought it wasn't proper and asked them to leave.
? Was it the children the headmaster was talking to? N

3a. He cleared off the table while the exhausted, over-worked waiter took the drink orders, but Matt had so many dirty plates that he left the cups behind.
? Did Matt get the cups? N
3b. She cleared off the table while the exhausted, over-worked waiter took the drink orders, but Jane had so many dirty plates that she left the cups behind.
? Did Jane get the cups? N
3c. His coworkers cleared off the table while the exhausted, over-worked waiter took the drink orders, but Jane was stuck at the cash register and couldn't help anyone.
? Was Jane able to help anyone? N
3d. Her coworkers cleared off the table while the exhausted, over-worked waiter took the drink orders, but Jane was stuck at the cash register and couldn't help anyone.

4a. He waited impatiently in the foyer while the austere, stuffy butler hung up the coats, but George managed not to say anything insulting about going faster.
? Was it in the foyer that he was waiting? Y
4b. She waited impatiently in the foyer while the austere, stuffy butler hung up the coats, but Andrea managed not to say anything insulting about going faster.
? Was it in the foyer that she was waiting? Y
4c. His employers waited impatiently in the foyer while the austere, stuffy butler hung up the coats, but Andrea was sure they wouldn't say anything insulting about going faster.
? Was it in the foyer that they were waiting? Y
4d. Her employers waited impatiently in the foyer while the austere, stuffy butler hung up the coats, but Andrea was sure they wouldn't say anything insulting about going faster.
? Was it in the foyer that they were waiting? Y

5a. He distributed fliers for the campaign while the friendly, personable congressman shook the voters' hands, but Paul wished he had a more important role to play.
? Did Paul wish he had a more important role? Y
5b. She distributed fliers for the campaign while the friendly, personable congressman shook the voters' hands, but Liza wished she had a more important role to play.
? Did Liza wish she had a more important role? Y
5c. His volunteers distributed fliers for the campaign while the friendly, personable congressman shook the voters' hands, but Liza thought they should have put up signs as well.
? Did Liza think they should have put up signs? Y
5d. Her volunteers distributed fliers for the campaign while the friendly, personable congressman shook the voters' hands, but Liza thought they should have put up signs as well.
? Did Liza think they should have put up signs? Y

6a. He continued drinking cheap American beer while the creepy, old bachelor hit on every woman in the bar, but Jake didn't think it was his place to step in.
? Was it whiskey that he was drinking? N
6b. She continued drinking cheap American beer while the creepy, old bachelor hit on every woman in the bar, but Kate didn't think it was her place to step in.
? Was it whiskey that she was drinking? N
6c. His coworkers continued drinking cheap American beer while the creepy, old bachelor hit on every woman in the bar, but Kate asked if one of them could stop drinking and say something.
? Was it whiskey that they were drinking? N
6d. Her coworkers continued drinking cheap American beer while the creepy, old bachelor hit on every woman in the bar, but Kate asked if one of them could stop drinking and say something.
? Was it whiskey that they were drinking? N

7a. He continuously scouted around the ranch while the tanned, weather-beaten cowboy checked on the herd, so Frank felt that they were doing their utmost to keep the area safe.
? Was it the cowboy who checked on the herd? Y
7b. She continuously scouted around the ranch while the tanned, weather-beaten cowboy checked on the herd, so Sarah felt that they were doing their utmost to keep the area safe.
? Was it the cowboy who checked on the herd? Y
7c. His riders continuously scouted around the ranch while the tanned, weather-beaten cowboy checked on the herd, so Sarah felt that they were doing their utmost to keep the area safe.
? Was it the cowboy who checked on the herd? Y
7d. Her riders continuously scouted around the ranch while the tanned, weather-beaten cowboy checked on the herd, so Sarah felt that they were doing their utmost to keep the area safe.
? Was it the cowboy who checked on the herd? Y

8a. He quickly unscrewed the fire hydrant while the fearless, cool-headed fireman prepared to enter the building, but Jeffrey was afraid they wouldn't be fast enough.
? Was Jeffrey afraid they wouldn't be fast enough? Y
8b. She quickly unscrewed the fire hydrant while the fearless, cool-headed fireman prepared to enter the building, but Allison was afraid they wouldn't be fast enough.
? Was Alison afraid they wouldn't be fast enough? Y
8c. His squadmates quickly unscrewed the fire hydrant while the fearless, cool-headed fireman prepared to enter the building, but Allison had already put out most of the fire by herself.
? Had Alison already put out most of the fire? Y
8d. Her squadmates quickly unscrewed the fire hydrant while the fearless, cool-headed fireman prepared to enter the building, but Allison had already put out most of the fire by herself.
? Had Alison already put out most of the fire? Y

9a. He finalized the news reports while the honest, no-nonsense anchorman made his way onto the set, but Keith signaled that the shooting wouldn’t start for another half hour.
? Was shooting going to start within ten minutes? N
9b. She finalized the news reports while the honest, no-nonsense anchorman made his way onto the set, but Laura signaled that the shooting wouldn’t start for another half hour.
? Was shooting going to start within ten minutes? N
9c. His interns finalized the news reports while the honest, no-nonsense anchorman made his way onto the set, but Laura signaled that the shooting wouldn’t start for another half hour.

? Was shooting going to start within ten minutes? N
9d. Her interns finalized the news reports while the honest, no-nonsense anchorman made his way onto the set, but Laura signaled that the shooting wouldn’t start for another half hour.
? Was shooting going to start within ten minutes? N

10a. He continued adjusting the sails while the old, experienced fisherman started telling anecdotes to the crew, but Ron made his way over to listen as soon as he was done.
? Was it emergency instructions that the fisherman was telling the crew? N
10b. She continued adjusting the sails while the old, experienced fisherman started telling anecdotes to the crew, but Amy made her way over to listen as soon as she was done.
? Was it emergency instructions that the fisherman was telling the crew? N
10c. His deckhands continued adjusting the sails while the old, experienced fisherman started telling anecdotes to the crew, but Amy told them to come and listen, too.
? Was it emergency instruction3 4hat the fisherman was telling the crew? N
10d. Her deckhands continued adjusting the sails while the old, experienced fisherman started telling anecdotes to the crew, but Amy told them to come and listen, too.
? Was it emergency instructions that the fisherman was telling the crew? N

11a. He chatted amiably with some fans while the talented, young quarterback signed autographs for the kids, but Steve wished the children's charity event would end soon so he could go home.
? Was it a quarterback who was signing autographs? Y
11b. She chatted amiably with some fans while the talented, young quarterback signed autographs for the kids, but Carol wished the children's charity event would end soon so she could go home.
? Was it a quarterback who was signing autographs? Y
11c. His managers chatted amiably with some fans while the talented, young quarterback signed autographs for the kids, but Carol wished the children's charity event would end soon so she could go home.
? Was it a quarterback who was signing autographs? Y
11d. Her managers chatted amiably with some fans while the talented, young quarterback signed autographs for the kids, but Carol wished the children's charity event would end soon so she could go home.
? Was it a quarterback who was signing autographs? Y

12a. He held the horses back while the friendly, young prince stopped to chat with someone, but Will wanted the conversation to finish so they could ride further on.
? Was it the dogs that he held back? N
12b. She held the horses back while the friendly, young prince stopped to chat with someone, but Mary wanted the conversation to finish so they could ride further on.
? Was it the dogs that she held back? N
12c. His guardsmen held the horses back while the friendly, young prince stopped to chat with someone, but Mary wanted the conversation to finish so they could ride further on.
? Was it the dogs that they held back? N
12d. Her guardsmen held the horses back while the friendly, young prince stopped to chat with someone, but Mary wanted the conversation to finish so they could ride further on.
? Was it the dogs that they held back? N

13a. She checked the placement of the props while the tiny, nervous ballerina struggled to lace up her corset, but Hannah didn't offer to help.
? Did Hannah offer to help? N
13b. He checked the placement of the props while the tiny, nervous ballerina struggled to lace up her corset, but Trevor didn't offer to help.
? Did Trevor offer to help? N
13c. Her friends checked the placement of the props while the tiny, nervous ballerina struggled to lace up her corset, but Trevor didn't offer to help.
? Did Trevor offer to help? N
13d. His friends checked the placement of the props while the tiny, nervous ballerina struggled to lace up her corset, but Trevor didn't offer to help.
? Did Trevor offer to help? N

14a. She slowly drank coffee while the pleasant, helpful stewardess made the rounds of the cabin, but Carrie decided to ask for a soda as well.
? Was it tea that she was drinking? N

14b. He slowly drank coffee while the pleasant, helpful stewardess made the rounds of the cabin, but Andrew decided to ask for a soda as well.
? Was it tea that he was drinking? N
14c. Her coworkers slowly drank coffee while the pleasant, helpful stewardess made the rounds of the cabin, but Andrew was surprised they needed caffeine so early in the flight.
? Was it tea that they were drinking? N
14d. His coworkers slowly drank coffee while the pleasant, helpful stewardess made the rounds of the cabin, but Andrew was surprised they needed caffeine so early in the flight.
? Was it tea that they were drinking? N

15a. She prepared to attack while the mean, wicked witch began casting another spell, but Rachel suddenly tripped over a magic toadstool.
? Was the witch beginning to cast a spell? Y
15b. He prepared to attack while the mean, wicked witch began casting another spell, but Daniel suddenly tripped over a magic toadstool.
? Was the witch beginning to cast a spell? Y
15c. Her servants prepared to attack while the mean, wicked witch began casting another spell, but Daniel had a secret plan up his sleeve.
? Was the witch beginning to cast a spell? Y
15d. His servants prepared to attack while the mean, wicked witch began casting another spell, but Daniel had a secret plan up his sleeve.
? Was the witch beginning to cast a spell? Y

16a. She thought about what to serve for dessert while the cheerful, kind-hearted widow entertained the kids, but Wendy eventually decided they should just go out for ice cream.
? Was it the widow who was entertaining the kids? Y
16b. He thought about what to serve for dessert while the cheerful, kind-hearted widow entertained the kids, but Frank eventually decided they should just go out for ice cream.
? Was it the widow who was entertaining the kids? Y
16c. Her cousins thought about what to serve for dessert while the cheerful, kind-hearted widow entertained the kids, but Frank eventually decided they should just go out for ice cream.
? Was it the widow who was entertaining the kids? Y
16d. His cousins thought about what to serve for dessert while the cheerful, kind-hearted widow entertained the kids, but Frank eventually decided they should just go out for ice cream.
? Was it the widow who was entertaining the kids? Y

17a. She continued battling the wolves while the young, brave heroine swooped in to rescue the captives, but Maria failed to see the goblins approaching from behind.
? Was it tigers that she was battling? N
17b. He continued battling the wolves while the young, brave heroine swooped in to rescue the captives, but Aaron failed to see the goblins approaching from behind.
? Was it tigers that he was battling? N
17c. Her friends continued battling the wolves while the young, brave heroine swooped in to rescue the captives, but Aaron couldn't help them because he was busy dealing with some goblins.
? Was it tigers that they were battling? N
17d. His friends continued battling the wolves while the young, brave heroine swooped in to rescue the captives, but Aaron couldn't help them because he was busy dealing with some goblins.
? Was it tigers that they were battling? N

18a. She kept listening to the breaking news while the tired, overworked maid washed the dishes, but Kathy still heard the plate shatter on the floor.
? Did Kathy hear the plate shatter? Y
18b. He kept listening to the breaking news while the tired, overworked maid washed the dishes, but Roger still heard the plate shatter on the floor.
? Did Roger hear the plate shatter? Y
18c. Her employers kept listening to the breaking news while the tired, overworked maid washed the dishes, but Roger kindly offered to help.
? Did Roger offer to help? Y
18d. His employers kept listening to the breaking news while the tired, overworked maid washed the dishes, but Roger kindly offered to help.
? Did Roger offer to help? Y

19a. She kept fooling around with the costumes while the quiet, unassuming actress recited lines from the play, but Karen knew she was supposed to be practicing her own part.
? Was she fooling around with the costumes? Y
19b. He kept fooling around with the costumes while the quiet, unassuming actress recited lines from the play, but Aaron knew he was supposed to be practicing his own part.
? Was he fooling around with the costumes? Y
19c. Her co-stars kept fooling around with the costumes while the quiet, unassuming actress recited lines from the play, but Aaron asked all of them to finish the rehearsal and leave the room.
? Were her co-stars fooling around with the costumes? Y
19d. His co-stars kept fooling around with the costumes while the quiet, unassuming actress recited lines from the play, but Aaron asked all of them to finish the rehearsal and leave the room.
? Were his co-stars fooling around with the costumes? Y

20a. She gossiped endlessly about scandals while the rich, famous heiress tried on hats at the boutique, but Sophia would have stopped if she had realized it annoyed the people around them.
? Was it shoes that the heiress was trying on? N
20b. He gossiped endlessly about scandals while the rich, famous heiress tried on hats at the boutique, but Thomas would have stopped if he had realized it annoyed the people around them.
? Was it shoes that the heiress was trying on? N
20c. Her chaperones gossiped endlessly about scandals while the rich, famous heiress tried on hats at the boutique, but Thomas tried to engage her in a serious religious discussion instead.
? Was it shoes that the heiress was trying on? N
20d. His chaperones gossiped endlessly about scandals while the rich, famous heiress tried on hats at the boutique, but Thomas tried to engage her in a serious religious discussion instead.
? Was it shoes that the heiress was trying on? N

21a. She set out the gaudy centerpieces obediently while the composed, affable hostess arranged the balloons, but Barbara couldn't help wishing that the bride had chosen more tasteful decorations.
? Did Barbara wish the decorations were more tasteful? Y
21b. He set out the gaudy centerpieces obediently while the composed, affable hostess arranged the balloons, but Brandon couldn't help wishing that the bride had chosen more tasteful decorations.
? Did Brandon wish the decorations were more tasteful? Y
21c. Her helpers set out the gaudy centerpieces obediently while the composed, affable hostess arranged the balloons, but Brandon couldn't help wishing that the bride had chosen more tasteful decorations.
? Did Brandon wish the decorations were more tasteful? Y
21d. His helpers set out the gaudy centerpieces obediently while the composed, affable hostess arranged the balloons, but Brandon couldn't help wishing that the bride had chosen more tasteful decorations.
? Did Brandon wish the decorations were more tasteful? Y

22a. She prayed loudly and with feeling while the vibrant, beloved priestess strode through the main congregation, so Tina was rewarded with a smile.
? Was she praying quietly? N
22b. He prayed loudly and with feeling while the vibrant, beloved priestess strode through the main congregation, so Zack was rewarded with a smile.
? Was he praying quietly? N
22c. Her companions prayed loudly and with feeling while the vibrant, beloved priestess strode through the main congregation, so Zack felt compelled to join in.
? Were her companions praying quietly? N
22d. His companions prayed loudly and with feeling while the vibrant, beloved priestess strode through the main congregation, so Zack felt compelled to join in.
? Were his companions praying quietly? N

23a. She put together a preliminary seating chart while the anxious, jittery bride talked with the caterers, but Lauren insisted that the couple themselves should discuss it in detail before it was finalized.
? Did Lauren insist that the couple discuss the seating chart? Y
23b. He put together a preliminary seating chart while the anxious, jittery bride talked with the caterers, but Justin insisted that the couple themselves should discuss it in detail before it was finalized.
? Did Justin insist that the couple discuss the seating chart? Y
23c. Her assistants put together a preliminary seating chart while the anxious, jittery bride talked with the caterers, but Justin insisted that the couple themselves should discuss it in detail before it was finalized.

? Did Justin insist that the couple discuss the seating chart? Y
23d. His assistants put together a preliminary seating chart while the anxious, jittery bride talked with the caterers, but Justin insisted that the couple themselves should discuss it in detail before it was finalized.
? Did Justin insist that the couple discuss the seating chart? Y

24a. She passed notes with practiced skill while the strict, pious nun spoke at length about discipline, but Jessica really did listen intently to every word in the lecture.
? Was it a priest who was lecturing about discipline? N
24b. He passed notes with practiced skill while the strict, pious nun spoke at length about discipline, but Matthew really did listen intently to every word in the lecture.
? Was it a priest who was lecturing about discipline? N
24c. Her students passed notes with practiced skill while the strict, pious nun spoke at length about discipline, but Matthew thought they had actually paid attention to most of her guest lecture.
? Was it a priest who was lecturing about discipline? N
24d. His students passed notes with practiced skill while the strict, pious nun spoke at length about discipline, but Matthew thought they had actually paid attention to most of her guest lecture.
? Was it a priest who was lecturing about discipline? N

4 CHAPTER 4: CONSTRAINTS ON COREFERENCE IN RUSSIAN[59]

As we have seen in Chapter 3, in English, as in many other languages, sentences like (1), (2), and (4) allow coreference between the subject NPs *Pooh* and *he*, whereas (3) does not. Coreference in (3) is blocked by Principle C of the Binding Theory, which prevents a pronoun from c-commanding its antecedent.

(1) Pooh$_i$ ate the apple while he$_i$ was reading a book.
(2) While Pooh$_i$ was reading a book, he$_i$ ate the apple.
(3) *He$_i$ ate the apple while Pooh$_i$ was reading a book.
(4) While he$_i$ was reading a book, Pooh$_i$ ate the apple.

As noted by Kazanina & Phillips (2001), the Russian counterparts (4)-(8) of the English sentences above exhibit the same coreference possibilities as in English, except that coreference is *disallowed* in (8): the *while*-sentence with backwards anaphora is possible in English, but impossible in Russian.

(5) Pux$_i$ s"el jabloko, poka on$_i$ čital knigu.
Pooh ate.perf apple while he read.imp book
Pooh ate an apple while he was reading a book.
(6) Poka Pux$_i$ čital knigu, on$_i$ s"el jabloko.
while Pooh read.imp book he ate.perf apple
While Pooh was reading a book, he ate an apple.
(7) *On$_i$ s"el jabloko, poka Pux$_i$ čital knigu.
he ate.perf apple while Pooh read.imp book
He ate an apple while Pooh was reading a book.
(8) *Poka on$_i$ čital knigu, Pux$_i$ s"el jabloko.
while he read.imp book Pooh ate.perf apple
While he was reading a book, Pooh ate an apple.

In what follows we will refer to sentences like (8) that are unacceptable on coreference interpretation as *poka*-sentences, since these sentences start with the conjunction *poka* 'while'. We will informally refer to the principle that rules out coreference in (8) as the *poka*-constraint. The name is given purely for convenience and should not be taken to reflect any commintments regarding the underlying nature of the constraint.

I will compare the constraints on backwards anaphora in Russian that apply in (7) and (8) and I will claim that they have different origins. Principle C is a universal and syntactic constraint. The *poka*-constraint, on the other hand, is Russian-specific and is sensitive to discourse factors. In addition to theoretical considerations, this conclusion will be supported by the results of an experiment with Russian-speaking children that explored how these constraints are acquired

[59] The experiments reported in this chapter were conducted in collaboration with Colin Phillips.

and by the results of a self-paced reading experiment with Russian-speaking adults that reveals how these constraints on backwards anaphora apply during real-time processing.

The chapter therefore has three major parts. In the first part (section 4.1) I will present more data that are relevant for pinpointing the exact nature of the *poka*-constraint. Section 4.2 discusses findings from an acquisition experiment that investigated how Principle C and the *poka*-constraint are acquired by Russian-speaking children. Section 4.3 presents a self-paced reading experiment with Russian-speaking adults that was designed to establish how the same constraints apply in real-time processing. Section 4.4 concludes.

4.1 THE POKA-CONSTRAINT: MAIN FACTS

In this section we will present the main facts about the *poka*-constraint in Russian.

As shown above, complex sentences that are introduced by the conjunction *poka* 'while' are subject to an additional Russian-specific constraint that rules out coreference between the pronoun subject of the *while*-clause and the subject NP of the following matrix clause (Kazanina & Phillips, 2001). This is illustrated by the example in (9) (same as (8)).

(9) *Poka on$_i$ smotrel novosti, Ivan$_i$ el jabloko.
while he.Nom watched.Imp news Ivan.Nom ate.Imp apple
While he watched news, Ivan was eating an apple.

Let me point out that that the unavailability of the coreference reading in sentences like (9) is an extremely robust phenomenon across Russian speakers, as confirmed both in a truth-value judgment task (section 4.2.2) and in an acceptability judgment task (section 4.3.1).

In the remainder of this section we will make an attempt to define the range of cases in which the *poka*-constraint applies. We will modify various properties of (9) in order to understand which of them may be critically involved in making the sentence unacceptable on the coreference reading.

Before we turn to more specific issues, it is important to make it clear that Russian does not have a general prohibition against backwards anaphora. A pronoun, including a nominative pronoun, may corefer with a name that linearly follows it. In (10) and (12) the pronoun is the subject of a relative clause that linearly precedes the antecedent of the pronoun.

(10) Sosed, kotorogo ona$_i$ vstretila v metro, rasskazal Mashe$_i$ utrennjuju novost'.
neighbor.Nom which she.Nom met in metro told Masha.Dat morning news.Acc
The neighbor that she met in the metro told Masha the morning news.

(11) Sosed, kotorogo ona$_i$ vstretila v metro, napugal Mashu$_i$.
neighbor.Nom which she.Nom met in metro frightened Masha.Acc
The neighbor that she met in the metro frightened Masha.

(12) Rasskaz, kotoryj ona$_i$ pročitala, rasstroil devočku$_i$.
story.Nom which she.Nom read.Perf upset girl.Gen
The story which she read upset the girl.

Thus, (10) & (12) suggest that the *poka*-constraint is not a sub-case of a general prohibition against backwards anaphora in Russian. This conclusion is further supported by the availability of coreference in bi-clausal sentences in which the embedded clause is introduced by a subordinator other than *poka* 'while'. For example, in (13) & (14) the temporal clause is headed by the subordinator *do togo kak* 'before' or *posle togo kak* 'after' respectively.

(13) Do togo kak on$_i$ posmotrel novosti, Ivan$_i$ s"el jabloko.
before he.Nom watched.perf news Ivan ate.Perf apple
Before he watched the news, Ivan ate an apple.

(14) Posle togo kak on$_i$ posmotrel novosti, Ivan$_i$ s"el jabloko.
after he.Nom watched.perf news Ivan ate.Perf apple
After he watched the news, Ivan ate an apple.

(13) & (14) show that backwards anaphora is allowed in sentences that have a syntactic configuration that is similar to (9), but introduced by a different temporal subordinator. Note that unlike *poka* 'while' which calls for an imperfective predicate in the embedded clause, the subordinators in (13)-(14) require a perfective predicate in the embedded clause.[60]

An interesting case is the subordinator *kogda* 'when' that can be followed by either a perfective or an imperfective predicate in the embedded clause. Similar to English, *kogda* conveys a temporal overlap between the embedded and the main events if the aspect of the embedded predicate is imperfective as in (15) and a successive interpretation of the events if the main predicate is perfective as in (16).[61] Whereas speakers generally allow coreference if *when* is interpreted as *after*, they find the coreference interpretation harder to obtain in sentences where *when* is interpreted as *while*.

---

[60] In addition to the meaning equivalent to the English 'while', the Russian subordinator *poka* can convey the meaning 'until'. In those cases *poka* is usually followed by a negated perfective predicate, as illustrated in (i). In this reading of *poka*, coreference between the embedded subject and the following main subject is not blocked, as shown in (ii).

(i) Masha ne vyključila televizor, poka ona ne dosmotrela vystuplenie Iriny Slutskoj.
Masha not turn-off TV until she not finish-watching performance Irina-Slutskaja.Gen.
Masha didn't turn off the TV until she finished watching Irina Slutskaja's performance.

(ii) Poka ona$_i$ ne dosmotrela vystuplenie Iriny Slutskoj, Masha$_i$ ne vyključila televizor.

Throughout this chapter we will concentrate on the first meaning of *poka*, 'while', which expresses a temporal overlap between the embedded and the main events.

[61] The aspect of the main clause is perfective in our examples, although the remaining two combinations are also available.

(15) Kogda on$_i$ voshel na kuxnju, Ivan$_i$ vkljuchil radio.
when he entered.Perf in kitchen Ivan switched.Perf radio
When (≈ after) he entered the kitchen Ivan turned on the radio.

(16) ??Kogda on$_i$ shel na kuxnju/po koridoru, Ivan$_i$ vkljuchil radio.
when he passed.Imp to kitchen/along hallway Ivan switched.Perf radio
When (≈ while) he was walking to the kitchen/in the hallway Ivan turned on the radio.

Russian has another subordinator, *v to vremja kak*, that can be literally translated as 'in/during the time when' and, hence, semantically rather similar to the subordinator *poka* 'while'.[62] Similarly to *poka* 'while', *v to vremja kak* also requires an imperfective embedded clause, as shown in (17); coreference between the main and the embedded subjects is ruled out.

(17) *V to vremja kak on$_i$ smotrel novosti, Ivan$_i$ el jabloko.
in time when he.Nom watched.Imp news Ivan.Nom ate.Imp apple
While he watched news, Ivan was eating an apple.

To summarize what we have seen so far, sentences that are bad on coreference (i.e. (9) & (16)) are introduced by the subordinator *poka* 'while', *v to vremja kak* 'during' or *kogda* 'when' and have an imperfective predicate in the embedded clause.[63]

Let us now look into more detail at sentences that are introduced by the subordinator *poka* 'while'.[64] Our aim is to see whether every pronoun in the embedded *poka*-clause is illicit on coreference with any R-expression in the main clause, or whether the *poka*-constraint targets only specific pronoun-name combinations. To this end, we will consider pronouns in various positions in the embedded clause. In (9) the pronoun in the *poka*-clause is a Nominative subject and could not have the main subject as an antecedent. In (18) the pronoun is a Dative subject, and the sentence is licit on coreference. The same is true of (19) in which the pronoun in the *poka*-clause is an object.

[62] I would like to thank Kira Gor for bringing this fact to my attention.
[63] As it becomes clear from (16)-(17) which involve suborsinators other than *poka* and from some other examples later in this section, '*poka*-constraint' is somewhat a misnomer for the constraint, since (i) it may apply to sentences that are introduced by a subordinator different from *poka* (although also conveying overlap in two actions, i.e. with the meaning 'while') and, (ii) does not apply to every sentence introduced by the subordinator *poka*'while'. Nevertheless, following Kazanina & Phillips (2001), I chose to keep the original name here.
[64] Note that the unavailability of coreference in the *poka*-sentence (8) cannot be explained via reconstruction of the adjunct *poka*-clause into its base position. If the adjunct clause in the *poka*-sentence were to undergo reconstruction, the resulting sentence would be (5), which is grammatical on coreference.

(18) Poka emu$_i$ bylo žarko, Ivan$_i$ obmaxivalsja gazetoj.
while he was.Imp hot Ivan.Nom waved.Imp newspaper.Instr
While he was hot, Ivan was fanning himself with a newspaper.

(19) Poka Maria rasskazyvala emu$_i$ novosti, Ivan$_i$ el jabloko.
while Maria told.Imp him.Dat news Ivan.Nom are.Imp apple
While Maria told him the news, Ivan was eating an apple.

(18) & (19) suggest that non-nominative pronouns do not induce the *poka*-constraint. (20) & (21) below further suggest that not all nominative pronouns fall under the constraint: coreference between a nominative pronoun and the following matrix subject is available if the sentence has a habitual reading.

(20) Poka on$_i$ učilsja v škole, Ivan$_i$ mečtal stat' kosmonavtom.
while he.Nom studied in school, Ivan.Nom dreamt.Imp become astronaut
While he was a schoolboy, Ivan dreamt of becoming an astronaut.

(21) Poka on$_i$ zanimalsja sportom, Ivan$_i$ ne pil i ne kuril.
while he.Nom did sports Ivan.Nom not drink.Imp and not smoke.Imp
While he did sports, Ivan didn't drink alcohol or smoke.

All nominative subject pronouns that we have seen so far were Agents. (22) shows that an Experiencer nominative subject pronoun in the *poka*-clause can corefer with the main subject. In (22) & (23) the pronoun subject of the *poka*-clause is an Experiencer and the main subject is an Agent.

(22) Poka on$_i$ čuvstvoval bol', Ivan$_i$ deržal ruku v vode.
while he.Nom felt.Imp pain Ivan.Nom kept.Imp hand in water
While he felt pain, Ivan held his hand in the water.

(23) Poka on$_i$ zhdal avtobusa, Ivan$_i$ prochital gazetu vdol' i poperek.
while he waited.Imp bus.Gen Ivan.Nom read.Perf newspaper lengthwise and crosswise
While he was waiting for a bus, Ivan read all of the paper inside out/far and wide.

And finally, an Agent pronoun in the *poka*-clause can corefer with a main subject that is an Experiencer, as in (24) & (25).

(24) Poka on$_i$ čital gazetu, Ivan$_i$ ne čuvstvoval boli v spine.
while he.Nom read.Imp paper Ivan.Nom not felt.Imp pain in back
While he was reading a newspaper, Ivan did not feel any back pain.

(25) Poka ona$_i$ razgovarivala po telefonu, Masha$_i$ vspomnila, chto zabyla kupit' xleba.
while she.Nom talked.Imp on phone Masha.Nom remembered.Perf that forgot buy bread
While she was talking on the phone Masha remembered that she forgot to buy some bread.

To summarize, the *poka*-constraint applies to sentences introduced by the subordinator *poka* 'while' where the subject of the *poka*-clause is a nominative pronoun that bears an agent role and disallows coreference between the embedded and the main subjects. Moreover, the possibility of coreference in *poka*-sentences can be affected by manipulations in the matrix clause. Coreference between the pronoun and the main subject improves significantly if the main subject is non-agentive.

4.1.1 Null subjects

The unavailability of coreference in (9) cannot be explained by a blocking effect from a null pronoun that is available in the position of the overt pronoun.

It has been observed that in languages that allow *pro*-drop, the coreference reading with an overt pronoun may be blocked whenever a null pronoun *pro* is available in the same position. The reason is that overt pronouns in these languages tend to be stressed or focal pronouns, which must be deictic. (26) is an example of this phenomenon from Italian (the example is due to Carminatti 2002).

(26) Maria$_i$ scriviera frequentemente a Ivano quando *pro$_i$* era negli Stati Uniti.
Maria wrote frequently to Piera when she was in the United States.
(27) Maria$_i$ scriviera frequentemente a Ivano quando lei$_{*i}$ era negli Stati Uniti.
Maria wrote frequently to Piera when she$_k$ was in the United States.

Given that Russian allows *pro* in certain contexts, it is possible that the unavailability of coreference in *poka*-sentences might be due to the same reason as in Italian. We tested this possibility by replacing the overt pronoun in (5)-(8) with a null *pro*, as shown in (28). The result is the following: three out of four cases, (28)b-d, are ill-formed sentences on any interpretation, i.e. an overt subject is required in these structures. The only well-formed sentence is (28)a, in which *pro* is the subject of the embedded clause that linearly follows the matrix clause. In (28)a the main and the embedded subjects must be interpreted as coreferent.[65]

(28) a. Ivan el jabloko, poka *pro* čital knigu.
Ivan was eating an apple while *pro* was reading a book.
b. *Poka Ivan čital knigu, *pro* el jabloko.
While Ivan was reading a book, *pro* was eating an apple.
c. **pro* el jabloko, poka Ivan čital knigu.
pro was eating an apple while Ivan was reading a book.

[65] The fact that Russian allows a coreference interpretation of an overt pronoun even in structures where *pro* is licit suggests that in Russian overt pronouns are not always stressed/focused elements, in contrast to prototypical *pro*-drop languages.

d. *Poka *pro* čital knigu, Ivan el jabloko.
While *pro* was reading a book, Ivan was eating an apple.

The example that is critical for us is (28)d, which is a *pro* counterpart of the *poka*-sentence (8). Since (28)d is ill-formed, we can conclude that the reason for the unacceptability of coreference in (8) is not related to competition with its null-subject counterpart. [The same can be said about the Principle C sentence (7) based on the unacceptability of (28)c]. So, the unacceptability of the *poka*-sentences cannot be explained by the fact that coreference is blocked due to the reconstruction of the *poka*-clause into its base position.

We conclude that the unavailability of coreference in (9) in Russian cannot be explained by appealing to the overt pronoun vs. null pronoun distinction.[66]

4.1.2 The *Poka*-constraint is not reducible to Principle C

Let us compare the *poka*-sentence (8) to its Principle C counterpart (7), repeated below for convenience.

(7) *On$_i$ s"el jabloko, poka Pux$_i$ čital knigu.
he ate.perf apple while Pooh was reading.imp book
(8) *Poka on$_i$ čital knigu, Pux$_i$ s"el jabloko.
while he was reading.imp book Pooh ate.perf apple

On the surface, these two sentences share a certain amount of similarity. Both sentences are bi-clausal structures with an embedded clause introduced by the subordinator *while* and have the same truth-conditions. In either sentence the pronoun linearly precedes the noun, but cannot corefer with it. Despite these similarities, we believe that the reasons for the unavailability of backwards anaphora in (7) & (8) are different and cannot be reduced to each other. I will claim that the constraints that apply in (7) & (8) are independent and that unlike Principle C the *poka*-constraint is not syntactic in nature.

First, I would like to argue against the idea that coreference in sentences like (8) is unavailable because the pronoun does have structural priority over the name and thus triggers a violation of Principle C. For example, it might be proposed that the notion of command should be generalized to encompass cases like (8) or that the pronoun undergoes a movement to a position that c-commands the name.

In order to show why such a scenario is unlikely, let us imagine for a moment that the unacceptability of coreference in (8) is due to the fact that the embedded subject pronoun does c-command the matrix subject, and thus gives rise to a

[66] Russian contrasts with other Slavic languages, e.g. Polish and Serbo-Croatian, in which unavailability of the coreference reading in the *poka*-sentence (8) is most likely to be due to the grammaticality of a null-pronoun counterpart of this sentence.

Principle C violation. We would then predict that for the same reason the embedded pronoun should not be able to corefer with any other nominal in the matrix clause. Yet, as shown in (29), coreference between the embedded subject pronoun *he* and the indirect object *Ivan* is allowed. Hence, the original assumption that in (9) the embedded pronoun c-commands the matrix clause and thus gives rise to a Principle C violation is not borne out.

(29) Poka on$_i$ smotrel televizor, Maria gotovila Ivanu$_i$ užin.
while he.Nom watched.Imp TV Maria cooked.Imp Ivan.Dat dinner.Acc
While he was watching TV, Maria cooked dinner for Ivan.

(30) Poka ona$_i$ lovila taksi, k Mashe podoshel bomž$_i$ i poporosil rubl'.
while she.Nom catch.Imp cab to Maria.Dat approached.Perf homeless and asked.Perf ruble
While she was catching a cab, a homeless man approached Maria and asked a ruble.

We conclude that unlike Principle C which is crucially based on the c-command relation, the *poka*-constraint cannot be explained in terms of c-command between the relevant elements.

Second, the *poka*-constraint and Principle C show different sensitivity to certain factors. As illustrated in (31)-(32), the *poka*-constraint is affected by the thematic role of the subject, whereas Principle C is not (presuming that changes in the thematic role do not affect the c-command relation). The *poka*-constraint applies when the subject of the embedded clause is an agent as in (31)a, but not when it is an experiencer, as in (31)b. On the other hand, the Principle C counterparts of these *poka*-sentences are both equally bad on coreference, as illustrated in (32). These facts are unexpected if the impossibility of coreference in *poka*-sentences were due to their structural configuration.

(31) a. *Poka on$_i$ smotrel televizor, Ivan$_i$ el jabloko.
While he.Nom was watching TV, Ivan ate.Imp an apple.
While he was watching TV, Ivan ate an apple.

b. Poka on$_i$ čuvstvoval bol', Ivan$_i$ deržal ruku v vode.
while he.Nom felt.Imp pain Ivan kept hand in water
While he felt pain, Ivan held his hand in the water.

(32) a. *On$_i$ smotrel televizor, poka Ivan$_i$ el jabloko.
he.Nom was watching TV, while Ivan ate.Imp an apple.
He was watching TV while Ivan ate an apple.

b. *On$_i$ čuvstvoval bol', poka Ivan$_i$ deržal ruku v vode.
he.Nom felt.Imp pain while Ivan kept hand in water
He felt pain while Ivan held his hand in the water.

Moreover, as (33)-(34) show, Dative subjects obviate the *poka*-constraint, although they clearly induce Principle C violations.

(33) Poka emu$_i$ bylo xolodno, Ivan$_i$ prygal na odnoj nozhke.
while he.Dat was.neut cold Ivan.Nom jumped.Imp on one foot
While he was cold, Ivan kept jumping on one foot.

(34) Emu$_{*i}$ bylo xolodno, poka Ivan$_i$ derzhal okno zakrytym.
he.Dat was hot while Ivan.Nom kept.Imp window shut
He was hot, while Ivan kept the window shut.

Third, Principle C has been shown to be valid cross-linguistically, whereas the *poka*-constraint is significantly less common across languages. In fact, we could not find a single language which shows the same pattern as Russian, although more investigation is definitely needed in the future.[67]

The fourth and fifth pieces of evidence come from our acquisition experiment with Russian-speaking children and from an online sentence-processing experiment with Russian-speaking adults. The results of the experiments reveal clear-cut differences in the treatment of sentences that are subject to Principle C and the *poka*-constraint.

4.1.3 What IS the *Poka*-constraint?

In section 4.1.3.1 I review Avrutin & Reuland's approach to backwards anaphora (Avrutin & Reuland 2003, Reuland & Avrutin 2005) whereby the *poka*-constraint emerges from a requirement that clauses must be grounded in discourse. In section 4.1.3.2 I present a modification of their account that makes it possible to account for some of the previously unexplained cases.

4.1.3.1 Avrutin & Reuland (2003)

Avrutin & Reuland (2003) and Reuland & Avrutin (2005) investigate ways in which referential dependences can be established in different languages. A somewhat simplified version of their idea is that such dependencies may be formed at three levels: 'narrow' syntax, 'broad' syntax or discourse. These levels are ordered and respect a strict division of labor between themselves, so that a failure to form a dependency at a lower level cannot be bypassed by invoking an alternative strategy to form the same dependency at a higher level.

[67] Unfortunately, other Slavic languages that are obvious candidates for having the *poka*-constraint all allow *pro*-drop in the embedded clause in sentences like (8). Thus although in the *poka*-sentences coreference with an overt pronoun is unavailable, it is unavailable for reasons far more general than the *poka*-constraint.

Dependencies formed in narrow syntax must be encoded through a process of chain formation (Reuland 2001). They are restricted to operations on morpho-syntactic features and are characterized by locality, respect for islands, etc. (see also Hornstein 1999, 2001 for a similar approach). An example of a dependency formed in narrow syntax is a dependency between a simplex anaphor such as Dutch *zich* (e.g. (35)) or Icelandic *sig* and its antecedent, which is instantiated as a syntactic chain and later interpreted in terms of variable binding.[68]

(35) Willem$_i$ schaamt zich$_i$. [Dutch]
Willem shames SE

Dependencies formed in broad syntax, although not instantiated as syntactic chains, are also structure-based. These dependencies still obey structural considerations such as c-command, but they are not encoded as chains and, therefore, do not obey restrictions on chains, e.g. island constraints. An example of a dependency encoded in broad syntax is given in (36).

(36) a. Every lawyer wondered whether Mary would leave after seeing the mess in his office.
b. Every lawyer (λx (x wondered [whether Mary would leave [after seeing the mess in x's office]]))

Finally, there are cases in which a dependency that the speakers establish is clearly not driven by purely syntactic considerations. Grammatically it is possible for the pronoun *his* in (37) to refer to *the diplomat, the tennis player* or *John*. That speakers overwhelmingly opt for *the tennis player* to be the antecedent of the pronoun is purely a result of pragmatic considerations. Dependencies such as in (37) are regarded to belong to and be finalized at the level of discourse.

(37) The diplomat told the tennis player that John had watched his last game.

Since discourse processes directly concern the relation between the current element and the previous context, they are sensitive to factors such as linear precedence. Other factors that discourse processes are sensitive to are notions like pragmatic/semantic plausibility, coherence, world knowledge, etc.

Having reviewed the three levels at which a dependency can be established according to Avrutin & Reuland's proposal, we are ready to show why backwards anaphora is possible in English (4) but not in its Russian counterpart (8).

The core idea is that in English, but not in Russian, there is a temporal dependency between the matrix and embedded T^0 available in broad syntax which serves as a host for a coreference dependency between the main and embedded subjects. The evidence for this claim comes from the phenomenon of tense

[68] Hornstein (1999, 2001) uses a similar mechanism to derive the distribution of English complex anaphors *himself* or *herself*. See Chapter 5 for further description of Hornstein's theory.

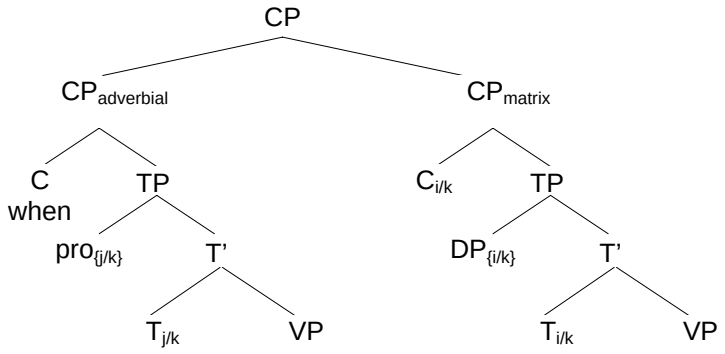
concord, which Avrutin & Reuland consider to be manifestation of a dependency between the tense heads of the main and embedded clause. It is well-known that tense-concord is present in English, but not in Russian, as shown in (38)-(40).

(38) John said that Mary was eating an apple.
(39) John said that Mary is eating an apple.
(40) Ivan skazal, čto Maša est jabloko.
Ivan said that Masha eats.Present an apple
Ivan said that Masha was eating an apple.

In English, in order to express that an embedded event took place simultaneously with a past main event the embedded predicate must also be past tense, as in (38). If the embedded predicate has a present tense as in (39), the sentence can only be interpreted as stating that the embedded event took place at the speech time, but not necessarily at the time of the main event (i.e. (39) means that Mary was eating an apple at the time when (39) was uttered). In Russian, on the other hand, an unmarked way of expressing the ongoingness of the embedded event (*apple-eating*) with a past main event (*saying*) is to use a present tense in the embedded clause, i.e. (40) is the unmarked way of conveying the meaning that Masha was eating an apple at the time of Ivan's report.

Omitting the details, here is how a morpho-syntactic dependency between T_{matr} and T_{emb} is established in English, which in turn may (but, presumably, does not need to) serve as a basis for a referential dependency between the embedded subject pronoun and the main subject. C_{matrix} shares features with T_{matr} (e.g. Haegeman 1992), which in turn shares features with T_{emb} (index *k* in the diagram (41)). The nominative subject of either clause in the [Spec TP] position shares features with the respective T^0 (represented by indexes *i* and *k*; this part follows the proposal by Pesetsky & Torrego (2001)). The resulting configuration is such that the embedded and main subjects are fully identical, which enables them to be coindexed (without entailing a Weak Crossover violation).

(41)



Importantly, the morpho-syntactic dependency between T_{matr} and T_{emb} serves as a basis for establishing a licit syntactic dependency between the main and embedded subjects in sentences like (4) (an English counterpart of the Russian *poka*-sentence (8)), repeated below for convenience.

(4) While hei was reading a book, Poohi ate the apple.
(8) *Poka oni čital knigu, Puxi s"el jabloko.
while he read.imp book Pooh ate.perf apple
While he was reading a book, Pooh ate an apple.

Russian, on the other hand, in the absence of a syntactic dependency between T_{matr} and T_{emb} has no other choice but to attempt establishing coreference using discourse means. Next I review the details of Avrutin & Reuland's (2003) proposal on how referential dependencies can be established in discourse. The key insight of their account is that Russian *poka*-sentences like (8) are ruled out on coreference because of the combination of two factors: the pronoun is the subject of an imperfective clause and it linearly precedes the R-expression.

Avrutin & Reuland's account rests on Enç's (1987) idea that clauses, specifically TPs, must be linked to discourse. This idea is enhanced by a requirement that discourse linking (or *anchoring*) should take place as soon as possible. Discourse linking can be performed in three possible ways, listed in (42).

(42) Means of linking a clause to discourse (Avrutin & Reuland 2003):
(i) by directly valuing the reference time of TP;
(ii) by directly valuing the [Spec, TP], i.e. the subject;
(iii) by valuing with reference to the matrix clause (as last resort).

These options are ordered as shown in (42), i.e. option (ii) cannot be attempted before (i), and (iii) cannot be attempted before (ii). Anticipating the description of each option below, let us notice that option (ii) refers to the [Spec TP] position in which nominative subjects reside and will be crucial in Avrutin & Reuland's explanation of why pronouns in this position cannot corefer with a linearly following R-expression.
Let us exemplify how the three options in (42) apply in Russian. Option (i) can be demonstrated using (43) (from Avrutin & Reuland 2003):

(43) Kogda/kak tol'ko on_i polnostju razvalil stranu, Boris ušol na pensiju.
when/as soon he completely destroyed.Perf country Boris.Nom went on pension
When/as soon as he completely destroyed the country, Boris retired.

The temporal clause in (43) is perfective and, hence, can be independently assigned a temporal interpretation and subsequently linked to discourse. The subject pronoun in [Spec TP] need not immediately be used for anchoring and can be temporarily stored to later enter a dependency relation with an antecedent, e.g.

the subject of the matrix clause. Hence, *he* and *Boris* can legitimately enter into a referential dependency.

Importantly, option (i) would be unavailable if the temporal clause were imperfective, since the imperfective aspect is assumed to be 'weak' and requiring to be evaluated relative to the main clause (Borik 2002). Thus, the embedded clause in (44) cannot be linked to discourse via direct valuation of reference time, since no temporal interval can be assigned to the clause until the matrix clause is processed.

(44) *Poka/*kogda on$_i$ el jabloko, Ivan$_i$ smotrel televizor.
while/when he ate.Imp apple Ivan watched.Imp TV
While/when he was eating an apple, Ivan was watching TV.

Option (ii) is then attempted; the clause is linked to discourse through the subject pronoun *he* in [Spec TP]. This requires that the pronoun is assigned a file card with an unspecified but fixed discourse referent right away. Since the pronoun cannot be put in temporary storage, it cannot be coindexed with the main subject that appears later in the sentence. By the same token the pronoun must be disjoint from any other NP in the following main clause. (This is contrary to (29) & (30), but Avrutin (p.c.) judges them to be illicit on coreference. We will return to this case later, after we illustrate how the final option can be used).[69]

Option (iii) is available in sentences like (45) for which the first two options did not succeed.

(45) Poka emu$_i$ bylo teplo, Ivan$_i$ xodil golym.
while he.Dat was.Neut warm Ivan walked naked.Instr
While he was warm, Ivan walked naked.

The temporal clause in (45) is imperfective, thus option (i) is unavailable. Option (ii) is also unavailable because Dative subjects in Russian are not located in [Spec, TP] (e.g. King 1995). Since the subject pronoun *emu* does not stand in a Specifier-Head relationship with the Tense head, it cannot be used to anchor the TP. In this case option (iii) is used as a last resort. Information from the main clause must be utilized to provide the temporal anchoring. The dative pronoun is temporarily put in storage (just as in the case of the nominative pronoun in (43)) and therefore can later take the main subject for its antecedent.

Avrutin & Reuland's proposal gives an account of the core facts concerning the *poka*-constraint in Russian. A clear advantage of their account is that it derives differences between the availability of backwards anaphora in English and Russian from independently known facts on tense-concord. There are some remaining problems, however.

[69] Avrutin & Reuland (2003) classify this as an Accessibility condition violation (Ariel 1990), whereby in cases of a repeated mention a nominal of higher accessibility (e.g., a pronoun) usually follows a nominal of lower accessibility (an R-expression).

First, some of the data remains unexplained. Recall that Avrutin & Reuland (2003) derive the availability of coreference in all complex sentences with the temporal embedded clause preceding the main clause from the properties of the embedded clause. However, we have seen earlier that some NPs in the main clause may corefer with the pronoun subject of the *poka*-clause (e.g. with the indirect object in (29) & (30) that we alluded to above). Other examples included cases like (46) & (47) in which the embedded subject pronoun can corefer with a matrix subject that is an experiencer. Moreover, for most speakers coreference becomes improved if the main predicate is perfective, e.g. (48) vs. (49).

(46) Poka ona$_i$ slušala novosti, Masha$_i$ vspomnila, čto zabyla vyključit' čajnik.
while she.Nom listened.Imp news Masha.Nom remembered.Imp that forgot.Perf turn.off kettle
While she was listening to the news Masha remember that she forgot to turn off the kettle.

(47) Poka on$_i$ razgovarival po telefonu, Ivanu$_i$ stalo ploxo s serdcem.
while he.Nom talked.Imp on phone Ivan.Dat became.Perf bad wth heart
While he was talking on the phone Ivan had a heart attack.

(48) ?Poka ona$_i$ gotovila užin, Masha$_i$ pozvonila druz'jam.
while she.Nom cooked.Imp dinner Masha.Nom called.Perf frends
While she was cooking dinner Masha called (up) her friends.

(49) *Poka ona$_i$ gotovila užin, Masha$_i$ zvonila druz'jam. [similar to (8)]
while she.Nom cooked.Imp dinner Masha.Nom called.Imp friends
While she was cooking dinner Masha was calling her friends.

Furthermore, the Russian *poka*-sentence (8) is ruled out via option (ii) because in order to anchor the embedded clause the pronoun must be evaluated immediately and cannot wait until the parser encounters the main subject. Whereas this indeed explains why binding between the pronoun and the main subject is excluded, it is unclear what precludes (accidental) covaluation between the main subject and the unspecified referent that the pronoun was assigned.

In the next section I will attempt to explain some of the problematic facts.

4.1.3.2 New account

The proposal presented here adopts some of the important insights of Avrutin & Reuland's (2003) account. The modifications that we propose here are in order to resolve some of the theoretical and empirical problems that the original proposal has.

This section has three subsections. First, I will argue that temporal clauses differ from other embedded clauses in that they are used to evaluate the main event in terms of its tense-aspectual properties (section 4.1.3.2.1). Second, I will modify

Avrutin & Reuland's procedure on how clauses are anchored to discourse (section 4.1.3.2.2). In place of structural notions such as T-head and [Spec TP] as in (42), I will use the temporal endpoint of an event and its agent as potential 'anchors'. Finally I will demonstrate how our proposal accounts for the (un)availability of coreference in *poka*-sentences and in various other sentences discussed earlier (section 4.1.3.2.3).

4.1.3.2.1 Temporal embedded clauses

I will show that in complex sentences with spatio-temporal embedded clauses, such as *poka*-sentences, there is an obligatory relation between the embedded and the main event times.

The reason why such a relation is required relates to the necessity to evaluate the main event at the interval defined by the embedded clause in order to obtain the truth value of the sentence. For example, (50) is false in a situation where the girl finished building a toy house after the boy was done watering the flowers. Although in this situation the proposition of the main clause holds – the girl did build a toy house – it does not hold unless evaluated at the interval delimited by the *while*-clause.

(50) Poka malchik polival cvety, devochka postroila domik.
while boy watered.IMP flowers girl built.PERF house
While the boy was watering the flowers, the girl built a (toy) house.

Thus, in sentences involving temporal embedded clauses the main event must be assessed with respect to the embedded event. As Reuland & Avrutin (2005) argue, in some languages, like English or Dutch, such a relation may be formed in broad syntax, based on a morpho-syntactic dependency between T_{main} and T_{emb}. Russian, on the other hand, fails to encode this relation in broad syntax; therefore, a temporal relation between the main and embedded events must be formed in discourse.

4.1.3.2.2 How clauses are linked to discourse

Next we will modify Avrutin and Reuland's proposal on how a clause can be anchored to discourse. The original proposal by the authors is repeated in (51) below.

(51) Means of linking a clause to discourse (Avrutin & Reuland 2003):
(i) by directly valuing the reference time of TP, i.e. T^0;
(ii) by directly valuing the [Spec, TP], i.e. the subject;
(iii) by valuing with reference to the matrix clause (as last resort).

(52) Means of linking a clause to discourse (our proposal):
(i) by directly valuing the event endpoint;
(ii) by directly valuing the agent of the event.

(52) is our variant of the original proposal. As can be seen, option (iii), which is problematic on conceptual grounds, is taken away completely, whereas the other two options are modified. In the remainder of this section I will explain the formulations in (52) and provide justifications for them.

Let us start by comparing option (i) in (51) vs. (52). Recall that in Avrutin and Reuland's account this option can be employed to anchor a perfective clause, but not an imperfective clause. The T^0 of an imperfective predicate was considered to be 'weak' and not usable for discourse anchoring.

The imperfective in Russian is a broad category that embraces various readings, and can describe progressive, conative, habitual, generic, iterative, etc. events. Rather than considering all imperfectives as weak, we propose that only a subset of them, i.e. those that are associated with events that may lack an endpoint fail to provide a temporal anchor.[70] These are cases when the imperfective is used in its conative or progressive reading, e.g. (53) & (54). In its progressive reading the imperfective presents an event as ongoing at some time interval, as in (i). In its conative reading the imperfective is used to refer to a permanently incomplete event, as in (53). In either case the imperfective zooms into the middle portion of the event and excludes the endpoint of the event from view.

(53) Ivan stroil.IMP dom v derevne(, no tak i ne dostroil ego).
Ivan was building a house in the village (, but never built it completely).
(54) Ivan smotrel.IMP televizor s vos'mi do desjati utra.
Ivan was watching TV from 8 to 10 am.

Importantly, the conative reading of the imperfective is only available with agentive events. This reading is unavailable with events that have an experiencer subject that does not exert control over the event. The contrast is illustrated by licitness of (53) vs. illicitness of (55).

(55) #Ivan chuvstvoval bol' v spine, no tak i ne pochuvstvoval ee.
Ivan.Nom felt.Imp pain in back but so not felt.Perf it
Ivan was feeling pain in his back, but didn't feel it completely.

If the only reading of the imperfective in which an event endpoint is missing is the conative reading, and if that reading is unavailable for events with experiencer subjects, then all events with experiencer subjects provide a temporal endpoint which can be used to anchor the event to discourse. The same can be said of habitual events which intuitively are super-events composed of many local events

[70] Relevance of the endpoint to the location of an event in absolute time comes is due to the fact that such a temporal point can allow to locate the event relative to other events in the context.

with individual endpoints and are conceptualized as recurrent events spread in time.[71]

Now let us turn to option (ii). Capitalizing on a specifier-head relation between the subject of the clause and its tense head T^0, Avrutin & Reuland's suggest that subjects located in [Spec TP] can be used as discourse anchors for clauses with a weak T^0. However, this formulation does not capture the contrast between nominative agents and experiencers, which both reside in [Spec TP], and, nevertheless, behave differently with respect to coreference (cf. (9) vs. (22)). We re-formulate option (ii) in terms of thematic roles and suggest that only agents can be used to ground a clause in discourse. This idea is rather intuitive given that the agent of the event exerts most power over temporal properties of the event, more than any other event role. For example, as we have shown earlier, only agentive events can be conative (i.e. events may end up being incomplete only due to the agent's cessation of action, although the latter can be caused either internally to the agent or by an external force).[72]

Furthermore, the idea that event roles rather than positions in the syntactic tree may be relevant for discourse anchoring has been independently proposed in the literature. Ritter & Wiltschko (2005) argue based on the data from a tenseless language (Blackfoot, Algonquian) that discourse anchoring can take place through either an agent or a patient role, whichever is higher on the discourse-based person/animacy hierarchy.

Summarizing the discussion in this section, option (i) can be used to anchor any clause that contains an event with a defined endpoint. These are all events that are described by a perfective predicate, habitual events (described by an imperfective) or imperfective events that have non-agentive, e.g. experiencer, subjects. The only events that require grounding through option (ii) are imperfective events with agentive subjects which may be conative and thus lack an event endpoint.

---

[71] This is supported by the fact that even agentive habitual events cannot be viewed as conative, as shown in (i).

(i) *Masha delala zarjadku každoe utro, no tak i (ni razu) ne sdelala ee.
Masha.Nom do.Imp exercise every morning but so (not once) not made.Perf it
Masha did exercise every morning but did not do it even once.

[72] More evidence for the special status of agents comes from pairs of sentences like (i) & (ii) below. The main clause subject *Ivan* is an agent in (i) and a theme in (ii). The contrast in grammaticality of the two sentences suggests that whereas an agent is a licit controller for the PRO in the adjunct clause, a theme is not. This shows that, in addition to occupying an appropriate structural position, the controller of PRO must be agentive.

(i) Razgovarivaja po telefonu, Ivan stroil glazki stojavsh'ej poblizosti devushke.
talking on phone Ivan.Nom made.Imp eyes.Acc standing nearby girl.Dat
While talking on the phone Ivan was flirting with a girl who was standing nearby.

(ii) *Razgovarivaja po telefonu, Ivan ponravilsja stojavsh'ej poblizosti devushke.
talking on phone Ivan.Nom appealed.Perf standing nearby girl.Dat
While talking on the phone Ivan appealed to a girl who was standing nearby.

4.1.3.2.3 Explaining backwards anaphora in Russian

We are now ready to explain how our proposal explains the facts related to backwards anaphora in Russian. The (un)availability of backwards anaphora is a result of an interplay of the factors in (56).

(56) Summary of our proposal:
- Temporal embedded clauses are used to evaluate the main clause;
- Clauses must be linked to discourse as soon as possible:
(i) by directly valuing the endpoint of the event;
(ii) by directly valuing the agent of the event.

As we said in the conclusion to the previous section, any perfective clause or imperfective clause that describes an event that has a temporal endpoint will be anchored to discourse through its event time (option i). Option (ii) will be used to ground imperfective clauses with progressive/conative agentive events.

Whenever a clause is anchored to discourse via option (i), the cataphoric pronoun in the fronted embedded clause will not be used for discourse linking. It therefore can be temporarily stored as unevaluated and can be assigned a referent once an appropriate NP is encountered later in the sentence. Sentence types in which the embedded clause is grounded to discourse via option (i) are listed below.

- Sentences with perfective temporal clauses, e.g. introduced by subordinators *before* or *after*

(57) Do togo kak on$_i$ posmotrel novosti, Ivan$_i$ s"el jabloko. [= (13)]
before he.Nom watched.perf news Ivan ate.Perf apple
Before he watched the news, Ivan ate an apple.

(58) Posle togo kak on$_i$ posmotrel novosti, Ivan$_i$ s"el jabloko. [= (14)]
after he.Nom watched.perf news Ivan ate.Perf apple
After he watched the news, Ivan ate an apple.

- Sentences with nominative and dative experiencer subjects. As explained in section 4.1.3.2.2, the relevant events always contain an event endpoint.

(59) Poka on$_i$ čuvstvoval bol', Ivan$_i$ deržal ruku v vode. [= (31)b]
while he.Nom felt.Imp pain Ivan kept.Imp hand in water
While he felt pain, Ivan held his hand in the water.

(60) Poka emu$_i$ bylo žarko, Ivan$_i$ obmaxivalsja gazetoj. [= (18)]
while he was.Imp hot Ivan.Nom waved.Imp newspaper.Instr
While he was hot, Ivan was fanning himself with a newspaper.

- Sentences with habitual predicates, e.g. (20)-(21) cited above [(21) is repeated below as (61)]. Despite the fact that the embedded predicate is imperfective, the event has an endpoint and can be anchored to discourse via its event time.

(61) Poka on$_i$ zanimalsja sportom, Ivan$_i$ ne pil i ne kuril. [= (21)]
while he.Nom did sports Ivan.Nom not drink.Imp and not smoke.Imp
While he did sports, Ivan didn't drink alcohol or smoke.

Let us now go through various sentences in which the event in the *poka*-clause lacks an endpoint and thus has to be linked to discourse via option (ii).

- In (62) the embedded clause is grounded through its agent *Maria*. This leaves the possibility for the object pronoun *him* to wait for a reference till later on. Hence *him* can corefer with the main subject *Ivan*.

(62) Poka Maria rasskazyvala emu$_i$ novosti, Ivan$_i$ el jabloko. [= (19)]
while Maria told.Imp him.Dat news Ivan.Nom are.Imp apple
While Maria told him the news, Ivan was eating an apple.

- In (63)-(65) in order to ground the *poka*-clause into the discourse the agent pronoun *(s)he* must be used. What remains puzzling is why the pronoun can corefer with the main subject that is an experiencer (as in (63)) or with some other NP from the main clause (as in (64)) , but not with the main subject that is an agent (in (65)).

(63) Poka ona$_i$ razgovarivala po telefonu, Masha$_i$ vspomnila, chto zabyla kupit' xleba. [= (25)]
while she.Nom talked.Imp on phone Masha.Nom remembered.Perf that forgot.Perf to-buy bread
While she was talking on the phone Masha remembered that she forgot to buy some bread.

(64) Poka ona$_i$ lovila taksi, k Mashe$_i$ podoshel bomž i poporosil rubl'. [= (30)]
while she.Nom catch.Imp cab to Maria.Dat approached.Perf homeless and asked.Perf ruble
While she was catching a cab, a homeless man approached Maria and asked a ruble.

(65) *Poka on$_i$ smotrel novosti, Ivan$_i$ el jabloko. [= (9)]
while he.Nom watched.Imp news Ivan.Nom ate.Imp apple
While he watched news, Ivan was eating an apple.

In all three sentences above the embedded clause must be grounded via option (ii). This requires that the agent of the *poka*-clause, the pronoun, be assigned a discourse referent. At the time when the pronoun is encountered, the file card stack is empty, since there are no preceding NPs. Consequently, the pronoun is assigned to a new file card with an unspecified discourse reference.[73] What we

[73] The original 'File Change Semantics' model was developed by Heim (1982). According to this model, discourse participants keep a file (similar to a clerk's file) which contains information relevant to the discourse, e.g cards that correspond to the entities mentioned in the discourse, their properties and actions. Cards that correspond to entities are usually introduced by indefinite NPs

suggest next is different from Avrutin & Reuland's proposal: we assume that such a file card can be updated later by information from an NP that follows the pronoun (in Avrutin & Reuland's case that was assumed to always lead to an Accessibility condition violation).

This explains (63) and (64), i.e. why coreference between the embedded pronoun subject and an NP within the main clause is available. Why cannot the same mechanism be used for (65)? We believe that the reason lies in how evaluation of the main event takes place. In (65) *Ivan* is the agent of the main event and thus is evaluated whenever the temporal properties of the event are assessed. The agent subject of the embedded clause, the pronoun, is used to evaluate the main event with its agent and cannot in turn be dependent on that agent. Using the File Change Semantics metaphor, the information from the main subject cannot be used to update the pronoun's file card and coreference between the embedded and main subjects is unavailable.[74, 75]

If coreference between the embedded and main subjects in (65) is indeed excluded as a result of temporal evaluation of the main event and its agent by the embedded event and its agent, we should expect coreference between two subjects to be available when no temporal assessment is involved. This is exactly what we find in sentences like (66), in which the pronoun inside an imperfective relative clause that modifies the main subject corefers with a subject of another clause later in the sentence (*Mikhail*).

(66) Sosedka, kotoroj on$_i$ stroil dom, raznesla po derevne slux, chto Mixail$_i$ iščet sebe nevestu.
neighbor.fem who he.Nom built.Imp house spread.Perf around village rumor that Mikhail.Nom looks.Imp himself fiancée
The neighbor (fem.) who he was building a house for spread the rumor that Mikhail was looking for a fiancée all around the village.

---

and updated by information from subsequent discourse in accordance with syntactic and discourse regulations. For further details on the File Card Semantics the reader is referred to Heim (1982, 1983), as well as Avrutin (1999).

[74] Alternatively, the reason why the embedded and main subjects cannot be coindexed is because that would lead to a pattern that is indistinguishable from what narrow syntax failed to obtain.

[75] Norbert Hornstein (p.c.) pointed out that referential 'cross-dependencies' seem to be dispreferred more generally by the grammar. For example, it is hard-to-impossible to undertand (i) as describing an event in which the mother was kissed by her own son which corresponds to the indexing pattern in (ii). Illiitness of (i) on coreference interpretation, where which both pronouns are referentially dependent on each other, on a par with availability of coreference in (iii) or (iv), where only one of the elements is dependent on the other one, can be easily explained if cross-dependencies were disallowed.

(i) Her son kissed his mother.
(ii) *[Her$_k$ son]$_i$ kissed [his$_i$ mother]$_k$.
(iii) The son$_i$ kissed his$_i$ mother$_k$.
(iv) Her$_k$ son$_i$ kissed the mother$_k$.

The relative clause in (66) does not introduce a temporal endpoint for the corresponding event and thus needs to be linked to discourse via its pronoun subject. The pronoun gets assigned a file card with some unspecified referent; however, this file card can be updated by the material from any NP in the following clause, including its subject *Michael*, because the relative clause is not used for temporal assessment of that clause. Thus the prediction is borne out.[76]

Furthermore, the reader may wonder why forwards anaphora is available in sentences like (67) that are obtained by reversing the order of the clauses in the *poka*-sentences (65).

(67) Ivan$_i$ el jabloko, poka on$_i$ smotrel novosti.
Ivan ate.Imp apple while he.Nom watched.Imp news
Ivan was eating an apple while he was watching the news.

The reason is that at the point when the pronoun is encountered and searches for a file card to link to, it can be directly linked to an already existing file card created by the NP *Ivan*. The evaluated pronoun is then used to anchor the embedded clause to discourse and to evaluate the main event. Thus, the procedure is the same as in (65), except that there an attempt to assign a discourse referent to the pronoun (other than a new file card with an unspecified reference) fails because the file card stack is empty.

Finally, it is important to ensure that the *poka*-constraint that we propose is learnable. Recall that the reason that makes the *poka*-constraint possible in Russian, but not in English or Dutch, is the lack of a syntactically coded dependency between the main and embedded tenses. In order to note that there main and the embedded tenses do not enter in a syntactic dependency the child needs to notice that T_{emb} cannot establish a dependency with T_{matr} in broad syntax. This is possible based on positive evidence, e.g. by encountering sentences such as (68).

(68) Ivan$_i$ skazal, chto Masha est jabloko.
Ivan said.Perf that he eats.Imp apple.
Meaning: Ivan said that Masha was eating an apple.

In (68) the main predicate is past and the embedded predicate is present, although the sentence is true only in a situation where Masha was eating the apple at the time of Ivan's utterance, i.e. in the past. Thus, although the interpretation of the embedded clause tense is clearly affected by the tense of the matrix clause, this is

[76] The availability of coreference between *he* and *Michael* is surprising for Avrutin & Reuland's account: if the relative clause in (66) needs to be linked to discourse like any other clause, it must do so through its subject *on* 'he', as the clause is imperfective, and therefore the pronoun should get a file card with some unspecified referent. If so, coreference of the pronoun with a following NP should lead to an Accessibility condition violation.

not encoded using morpho-syntactic means, i.e. there is no formal dependency between T_{emb} and T_{matr}.[77]

The child who knows this fact and understands that in sentences with temporal embedded clauses the main event must be evaluated against the interval defined by the embedded clause has no other way but to appeal to form such a relation via discourse.[78] Once in discourse, the child will use the general principles of introducing file cards and resolving reference. The key point – why coreference is unavailable in the *poka*-sentences like (65) is also derivable from general principles (either that there is no circuicity in evaluation, or that a dependency that failed in syntax cannot be re-established in discourse).

4.1.3.3 Constraints on Coreference in Time

We have argued that Russian has (at least) two independent constraints that regulate long-distance dependencies between a pronoun and a following noun: Principle C and the *poka*-constraint. We have suggested that unlike Principle C, which is a structural constraint, the *poka*-constraint should be considered as operating at the discourse level. Everything that we have said so far was based on adult speakers' judgments of whole sentences. If the time dimension is added to the picture, two types of questions arise.

One question is when these constraints on coreference emerge in the process of language development. The developmental profile of Principle C and the *poka*-constraint can provide evidence for or against theoretical accounts of the constraints. For example, if there is a single underlying reason for the unavailability of backwards anaphora in sentences like (7) & (8), the most natural

---

[77] Norbert Hornstein (p.c.) points out that there is an extra detail that is important for the explanation how the *poka*-constrant is learnt. The matter is that the types of sentences which unambiguously demonstrate the lack of the tense concord in Russian by lacking a morpho-syntactic dependency between the main and embedded tenses all involve complement clauses. On the other hand, *poka*-sentences involve adjunct clauses and on the surface do not show the same strong lack of a temporal relation between the main and the embedded clauses, e.g., the embedded predicate cannot be in the present tense and be 'shifted' into the past based on the past tense of the main predicate. [compare (68) with (i) below]:

(i) *Ivan pel poka Masha est jabloko.
Ivan sang.Imp while Masha eats.Imp apple
Intended meaning: Ivan sang while Masha was eating an apple.

Hence, the Russian-speaking child must notice tense concord based on clausal complementation and then transfer this knowledge to adjunct clauses in which the same phenomenon is less apparent. The reason why temporal adjunct clauses more 'faithfully' describe the event time (i.e. must use a past tense whenever the embedded event is in the past) could be due to the fact that the temporal interval which they define must be established independently in order to evaluate the main event (section 4.1.3.2.1). In contrast, complement clauses are not used for evaluation of the main event.

[78] See Kazanina & Phillips (submitted) for evidence that Russian-speaking children are aware of the fact that in sentences with temporal embedded clauses the main event must be evaluated against the interval defined by the embedded clause already at the age of 3.

prediction is that both constraints become operational in the child's language relatively close in time. On the other hand, if we are right in arguing that the two constraints are of a different nature, their simultaneous emergence in the child's grammar could only result from a pure (and presumably rather unlikely) coincidence.

The other question concerns the representation of these constraints by the parser. Can consequences of both constraints be calculated online? This is an extension of the question we asked of Principle C in English in Chapter 3, i.e. if and how it is that illicit representations are ruled out by the parser during real-time processing. We have argued that Principle C applies immediately during sentence processing which makes it a candidate for a constraint on structure generation. However, so far we have only had an opportunity to compare sentences that are subject to Principle C with controls in which backwards anaphora are not subject to any constraints. Russian provides us with an opportunity to make a closer and perhaps a more informative comparison, i.e. how different types of illicit antecedents (i.e. antecedents that are subject to Principle C to those that are subject to the *poka*-constraint) are ruled out by the parser.

In what follows I will discuss the issue of how Principle C and the *poka*-constraint emerge in children (section 4.2), and how it applies during online processing (section 4.3). I will show that the *poka*-constraint is 'delayed' on both time-scales relative to Principle C: it emerges later in the development and it applies later in the course of real-time sentence processing.

4.2 DEVELOPMENT OF COREFERENCE CONSTRAINTS: PRINCIPLE C AND THE POKA-CONSTRAINT IN CHILD RUSSIAN

Let us start by once again listing in (69)-(72) all possible permutations of the name-pronoun order and the main-subordinate clause order for bi-clausal sentences with the embedded clause introduced by the subordinator *while* (repeated from (1)-(4) above). Both forwards anaphora sentences (69) & (70) allow the coreference interpretation, and so does the backwards anaphora sentence (72) in which the embedded clause precedes the main clause.[79] The only member of the paradigm that in which coreference is unacceptable is (71) that which is subject to Principle C.

(69) Pooh$_i$ ate the apple while he$_i$ was reading a book.
(70) While Pooh$_i$ was reading a book, he$_i$ ate the apple.
(71) *He$_i$ ate the apple while Pooh$_i$ was reading a book.
(72) While he$_i$ was reading a book, Pooh$_i$ ate the apple.

[79] We will continue to use the terms 'forwards anaphora' and 'backwards anaphora' primarily to refer to the linear order of a pronoun and an R-expression. The use of these terms does not necessarily entail that a referential dependency between the pronoun and the R-expression is licit.

The Russian counterparts of the English (69)-(72) are given in (73)-(76). Similar to English, coreference is licit in forwards anaphora sentences (73) & (74) and impossible in (75) due to Principle C. Unlike English, coreference is also excluded in (76), due to the Russian-specific *poka*-constraint.

(73) Pooh$_i$ s"el jabloko, poka on$_i$ čital knigu.
Pooh ate.perf the apple while he was reading.imp the book
(74) Poka Pooh$_i$ čital knigu, on$_i$ s"el jabloko.
while Pooh was reading.imp the book he ate.perf the apple.
(75) *On$_i$ s"el jabloko, poka Pooh$_i$ čital knigu.
he ate.perf the apple while Pooh was reading.imp the book.
(76) *Poka on$_i$ čital knigu, Pooh$_i$ s"el jabloko.
while he was reading.imp the book Pooh ate.perf the apple

We argued based on theoretical considerations that despite the surface similarity of sentences (75) & (76), the underlying constraints clearly differ in their source. If this is correct, we expect to see traces of this difference in how the child acquires the constraints. To this end, we will investigate young Russian-speaking children's comprehension of referential relations in sentences like (73)-(76). The age of our participants ranged from under 3 years of age to 6+ years of age and was chosen so as to make it possible to track the status of each constraint at different points of language development.

We will start by reviewing the existing literature on the status of (constraints on) backwards anaphora in language acquisition in section 4.2.1. Most of this literature converges on the claim that Principle C is respected by children already at the earliest testable age. We will then show in section 4.2.2 that this claim is supported by the data from a new language, Russian, in which young children respect Principle C from at least age 3. We will also show that the Russian-specific *poka*-constraint appears in the child's grammar much later, not until the age of 5-6, which we take as another piece of evidence for differences in the nature of the two constraints.

4.2.1 Previous research on backwards anaphora and Principle C in children

This section reviews previous research on the status of backwards anaphora and constraints on it in the child's grammar. The interest in backwards anaphora was inspired by its relevance to the argument regarding the child's initial hypothesis about language, namely, whether the child immediately takes language representations to be organized in hierarchical, rather than linear, terms. In the area of backwards anaphora, for example, children may start by having a purely linear prohibition against coreference between a pronoun and a linearly following

R-expression or they may start by having adult-like structure-dependent constraints on backwards anaphora.

Consequently, two main issues have been investigated. The first issue is whether children do or do not have an across-the-board prohibition against backwards anaphora. If such a prohibition existed, if could be taken as evidence for a linear constraint in the child's grammar. On the other hand, if the child does not have a prohibition on backwards anaphora, a further question is whether the child abides by a universally valid structure-dependent constraint on coreference, Principle C. Most of the studies that we discuss below explore these two questions in parallel.

In her pioneering study, C. Chomsky (1969) investigated children's comprehension of sentences such as (77)-(79) with licit or illicit anaphora.

(77) *He$_i$ found out that Mickey$_i$ won the race.
[backwards anaphora, Principle C]

(78) If he$_i$ wins, Pluto$_i$ will be happy.
[backwards anaphora, unrestricted]

(79) Mickey$_i$ said he$_i$ was hungry.
[forwards anaphora, unrestricted]

The experiment was run in the form of an interview where the child answers a question asked by an experimenter. The experimenter introduced the child to two toys – Pluto and Mickey Mouse – that were placed on the table in front of the child. Then the experimenter made a statement about one or both of the toys, and asked the child a question about that statement.

Experimenter: Mickey wants to go to the movies. Who wants to go to the movies?

Or, using (77) for an example:

Experimenter: He found out that Mickey won the race. Who found out? And who won the race?

To respond the question the child could either say the character's name or point to one of the toys. Chomsky tested forty-two 5-10 year old children. Each child saw 5 trials from each of the three conditions.

Chomsky found that all children overwhelmingly preferred the coreference reading in the forwards anaphora condition (79) in 84% (177/210) of trials. Children also showed essentially adultlike behavior in the Principle C condition by giving a disjoint response to (77) in 92% (194/210) of trials. More precisely, 33/42 children correctly treated the pronoun and the R-expression in (77) as disjoint in all of their trials (165/165). All of the errors were due to 9 children who accepted coreference in (77) in 36% (16/45) of their trials. Seven of the Principle C violators were among the youngest children in the study (between 5;0 and 5;3 years old), and the remaining two children only made one error each. Finally, in sentences like (78) in which backwards anaphora is licit children chose the

coreference interpretation in 36% (75/210) of all trials (36% for Principle C respecters, 36% for Principle C violators).
Chomsky concludes that those children who always rejected Principle C sentences on coreference (i.e. Principles C respecters) did so due to their knowledge of the specific constraint rather than due to a general prohibition against backwards anaphora, since the same children did accept coreference in (78) in 36% of trials. On the other hand, the remaining 9 children (Principle C violators) effectively did not make a difference between sentences with illicit (77) and licit (78) backwards anaphora and accepted either condition in 36% of trials. Since these 9 children were by-and-large the youngest children in the experiment, Chomsky suggested that they have not yet learnt Principle C (she proposed 5;6 to be a mean age for acquisition of Principle C).

A different conclusion was reached by Lust, Loveland & Kornet (1980)[80]. They tested 82 children between 3;6 and 7;6 distributed across 8 age groups in a production (elicited imitation) task and a comprehension (act-out) task using sentences like (80)-(81):

(80) $He_{*i/j}$ turned around when $Snuffles_i$ found the penny. [Principle C]
(81) When $he_{i/j}$ closed the box, $Cookie\ Monster_i$ lay down. [optional]

The authors found that children generally gave few coreference interpretations for either sentence. However, they allowed coreference in sentences like (80) significantly less often than in (81) (14.5% vs. 24%). Comparisons inside individual age groups showed that this pattern held in 5/8 groups. In 2/8 groups the pattern went numerically in the opposite direction (age groups 1 and 8) and in group 7 the performance was almost identical in the two conditions.

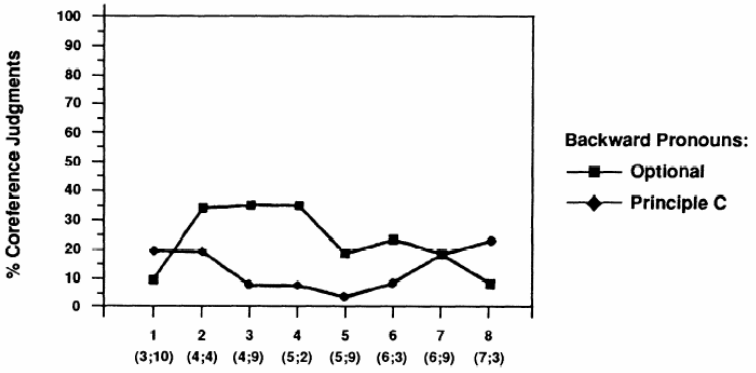


FIGURE 2. Coreference judgments: Act-out task without pragmatic lead.

Figure 12. Coreference judgments in the act-out task from Lust et al. (1980). Adopted from Lust et al. (1992).

[80] Reported here according to the review in Lust et al. (1992).

Tavakolian (1977) (presented here based on the description in Crain & McKee 1986 and Gordon 1996) conducted an act-out task with 3-5 year old children. In acting out sentences like (82) children frequently selected as a referent for the pronoun an animal that was not mentioned in the sentence but was present in the experimental workspace.

(82) For him to kiss the lion would make the duck happy.

Based on the results of Tavakolian's (1977) and Lust et al.'s (1980) studies various researchers (e.g., Solan 1983, Tavakolian 1977) claimed that children initially have a general prohibition against backwards anaphora. Crain & McKee (1986) argued that the results at most suggest that children have a strong *preference* to avoid backwards anaphora. They point out that some dispreferred representations may nevertheless be grammatical and compare the child's preference to avoid backwards anaphora with the adult's preference to avoid triply center-embedded sentences.

To test their interpretation of the previous results Crain & McKee (1986) investigated whether young English-speaking children allow backwards anaphora in contexts such as (83) using a different technique, a Truth Value Judgment task, which provided a better opportunity to reveal ungrammaticality (as opposed to a dispreferred status) of a representation. As made apparent by the name of the task, the child judges the truth-value of a sentence, i.e. whether a certain sentence is true or false in a given context. The implication is that all interpretations that are grammatical in a given situation, including those that are dispreferred, will be judged as true provided that the child has the same grammar as the adult.

The crucial conditions from Crain & McKee's experiment are given in (83)-(84). In (83) backwards anaphora is licit, whereas in (84) it is ruled out by Principle C.

(83) a. When $he_{i/j}$ stole the chickens, the $lion_i$ was in the box.
b. When $she_{i/j}$ was outside playing, Strawberry $Shortcake_i$ ate an ice cream cone.

(84) a. *He_i washed Luke $Skywalker_i$.
b. *He_i ate the hamburger when the $Smurf_i$ was in the fence.

The sentence (83) is ambiguous between a coreference and disjoint interpretation. Consequently, for an adult speaker of English (83) is true in a situation that supports at least one of the interpretations. Children's representation of (83) was tested in two scenarios. The first scenario made the coreference reading true (e.g. the lion stole some chickens from inside a box), whereas the second scenario made the disjoint reading true (a man stole some chickens while the lion was inside the box). Principle C sentences (84) were only tested in the scenario that supported the coreference reading.

Sixty-two children (mean age = 4;2, 4 scenarios per child, 2 of each type) were tested. The children appropriately accepted (83) in either the disjoint (81%

acceptance) or the coreference (73% acceptance) scenario. Crucially, the same children rejected (84) in the coreference scenario 88% of the time, thus showing a clear contrast in their treatment of (83) vs. (84) and knowledge of Principle C. The authors separated out the results from the 7 youngest children in the study (mean age = 3;1) to show that the same pattern of results obtains even with the youngest children: the youngest group overwhelmingly accepted (83) in either coreference (86% 'yes') or disjoint (71%' yes') scenario, whereas they rejected (84) in the coreference scenario (79% 'no').

Crain & McKee (1986) conclude that young English-speaking children have an essentially adultlike representation of backwards anaphora: they do not have a general linear prohibition on backwards anaphora (demonstrated by their acceptance of (83) on coreference), but they reject backwards anaphora whenever it is subject to Principle C (84). The authors claim that 3-year-old children's adherence to Principle C provides strong support for the claim that this structural principle is innate.

A similar conclusion is made by Guasti & Chierchia (1999/2000) who tested the representation of backwards anaphora in Italian children aged 3-6. They used sentences such as (85)-(86) that are structurally similar to the English sentences (83)-(84). (85) allows both the coreference and disjoint readings, whereas in (86) the coreference reading is ruled out by Principle C.

(85) Mentre ballava, un pagliaccio suonava la chitarra.
[$pro_{i/k}$ mentre ballava, un pagliaccio$_i$ suonava la chitarra.]
While (he) was dancing, a clown was playing the guitar.

(86) Andava sul cavallo a dondolo, mentre un musicista suonava la chitarra.
[$pro_{*i/k}$ andava sul cavallo a dondolo, mentre un musicista$_i$ suonava la chitarra.]
(He) was riding a rocking horse, while a musician was playing a guitar.

In Experiment 1 Guasti & Chierchia investigated children's production of sentences like (85)-(86) in an elicited imitation task. The child watched a story acted out with small props with a frog puppet who was learning Italian. To practice her Italian, the puppet would describe what happened in the story using sentences like (85) or (86). The child had to then report what she said to another puppet who was unable to see the act-out. Given the nature of the task, the child could either provide a verbatim report on what the frog had said or report what the frog meant to say. The authors reasoned that if the child understood the task in the first way, there would be either no change in the child's repetition of the target sentence or an equal amount of change for both conditions. Alternatively, if the child reported what the frog meant, he could correct the target sentence such that it reflected what happened in the story. The stories were such that they always supported the coreference reading, so if children have an adultlike representation of (85) & (86), they were expected to produce more structural changes in (86), which is ungrammatical on the coreference reading, than in (85) . This is exactly

what the authors found: whereas the children made structural changes to (85) in 54% of trials, they modified sentences like (86) significantly more often, in 84% of all trials.[81] The authors concluded that children did not treat the two types of sentences as equivalent and that they adjusted (86) more often because it was inappropriate in the coreference context. Since the reason for the inappropriateness of (86) is Principle C, one must concede that the children had access to this principle in order to explain the findings.

Experiment 2 assessed children's knowledge of Principle C in the same sentences using a Truth Value Judgment task. In this experiment in addition to testing the children's knowledge of the ungrammaticality of (86) on the coreference reading (and using (85) as a control for that), the authors also tested children's acceptance of the disjoint reading in (85), in which the antecedent for the null pronoun is extra-sentential. The procedure was similar to Crain & McKee (1986): both sentences (85) & (86) were tested in a context that made the coreference reading true and (85) was additionally tested in a context which made the disjoint reading true. If the child has an adultlike grammar including Principle C, he should accept (85) in both types of situation, but reject (86) on coreference. The results from 20 children between 3;2 and 57 years of age (mean age = 4;5) show that children correctly reject (86) in 80% cases while accepting (85) in 92% of the times in either type of situation.[82] Thus, Italian-speaking children rule out coreference in Principle C contexts at least from age 3;10, although they accept other, licit instances of backwards anaphora.

Summarizing the existing literature, we conclude that although children may disprefer structures with backwards anaphora, they do not have an across-the-board prohibition on backwards anaphora. However, they do reject backwards anaphora in sentences that are subject to Principle C from a very early age.[83]
We will continue the discussion of the status of backwards anaphora in child language in the next section, which presents an experiment that tested structures with backwards anaphora similar to English (83)-(84) or Italian (85) & (86) in Russian.

[81] The results are based on 12 children (mean age = 4;7 ranging from 4;0 to 5;5). Each child was tested on 6 stories distributed equally between two conditions.
[82] The analysis of individual subject data showed that there were two children who incorrectly accepted (86) in 4/4 cases; additional investigation of these two children revealed that this was a result of a response strategy that was overcome using a variation in the experimental technique rather than a result of any grammar limitations.
[83] C. Chomsky's (1969) study is an exception to the general pattern in the respect that Condition C effects seemed to come at a later age in that study.

4.2.2 Acquisition of Constraints on Backwards Anaphora in Russian

Participants

Participants were 56 children aged 2;8 to 4;11 from two day-care centers in Moscow.[84] Of them 12 children were excluded from the analyses because they gave incorrect answers on one or more filler trials. A further 6 older children (age 5-6) were tested in a within-subject design. The parents of all children gave informed consent for their participation.

Methods & Design

We used a Truth Value Judgment Task (TVJT) (Crain & McKee 1985, Crain & Thornton 1998, Gordon 1998) to test Russian children's knowledge of coreference possibilities in sentences like (69)-(72). We tested coreference judgments in a 2x2 between subjects design, in which we independently varied the name-pronoun order (forwards vs. backwards anaphora) and the main-subordinate clause order, thus testing the full paradigm in (69)-(72). The use of a between-subject design allowed us to test more than one token of each condition per child. This proved useful in assessing whether individual children were consistent in their judgments of similar structures.

The experiment was run in one or two sessions lasting for 15-25 minutes. At the first visit the child was assigned randomly to one of the 4 conditions. Each child saw the same condition across four experimental stories, interspersed with filler stories. Fillers were assigned after each test item in such a way that the child should give the opposite response to the preceding test item. The filler stories were used to prevent the children from giving identical responses across all trials, and as an independent measure of whether each child understood the task.

The stories were acted out by one experimenter using toy figures, while a second experimenter manipulated a puppet. Following each story, the puppet described one thing that he thought happened in the story. The child's task was to reward the puppet with a strawberry or a candy if he was correct, or to feed him a candy wrapper if he was incorrect. When the child said that the puppet was wrong, we asked the child to explain why, in order to assess the child's understanding of the story. For the same reasons, we asked the child to say and explain again what the puppet said, even if the child agreed with the puppet's statement.

Crucially, each of the four stories was used in exactly the same form in all four experimental conditions. The only difference across conditions was the puppet's final statement that the child judged. Below we give a sample story, which was

[84] A number of additional children, predominantly 3-year olds, were excluded from the experiment because they were unable to perform the task.

one of the two stories in which the test sentence contained the conjunction *poka* 'while'; the two remaining stories used the conjunction *kogda* 'when'.[85] The remaining stories are listed in Appendix 4-A.

EXAMPLE STORY (translated from Russian)

SETTING: *A room. There is a table in the corner of the room, and there is an apple and two bananas on the table. Eeyore is in the other corner of the room reading a book.*
POOH: "Hi, Eeyore. You are reading a book, I see. I wonder what I should do? – *Pooh walks around the room, notices the apple.* - Oh, what a nice apple! I shall eat it right now."
EEYORE: "No, Pooh, you can't eat it: that's my apple." *Eeyore continues to read the book.*
POOH: "OK, I can't eat the apple, because it's Eeyore's apple. Then I shall have to eat a banana instead." *Pooh eats a banana.*
EEYORE: "OK, Pooh, I've finished reading the book, so you can read the book now." *Pooh starts reading the book.*
EEYORE: *Eeyore walks across the room to the table with the apple*. "Here is my apple. I think I shall eat it right now." *Eeyore takes the apple to his mouth to eat it, but just before biting into it he stops and says*: "I shouldn't be such a greedy donkey! Pooh wants the apple and so I think I should give it to him. As for me, I can have a banana instead." *Eeyore drags the table with the apple to Pooh who is reading the book.* "Pooh, here is the apple and you can have it."
POOH: "Oh, I'm such a happy bear! I have a book to read and an apple to eat! I am going to read the book and eat the apple!"
FINAL SETTING: *Pooh continues reading the book. At the end of the story, there is an apple leaf next to Pooh and the book, to remind the child that it was Pooh who ate the apple, while reading the book.*

PUPPET: "That was a story about Eeyore and Pooh. First Eeyore was reading the book and then Pooh was reading the book. I know one thing that happened …

	While Pooh was reading the book, he ate the apple.	Forwards Anaphora Embedded clause first
or	Pooh ate the apple, while he was reading the book.	Forwards anaphora Matrix clause first
or	He ate the apple, while Pooh was reading the book.	Backwards Anaphora Matrix clause first (Principle C)
or	While he was reading the book, Pooh ate the apple.	Backwards Anaphora Embedded clause first (*poka*-constraint)

[85] Although Russian adults disallowed coreference 100% of the time in our test sentences, in general *kogda* seems to be less restrictive than *poka*.

In order to use exactly the same story across all four experimental conditions, while satisfying the requirements of the TVJT (Crain & Thornton 1998), it was necessary to satisfy the Condition of Plausible Denial *twice*. Pooh nearly ate the apple while Eeyore was reading the book, and Eeyore nearly ate the apple while Pooh was reading the book, but in both instances the apple-eater chose a banana instead at the last minute. These circumstances made it felicitous for children to reject any of the test sentences under a disjoint reference reading. In all other respects, the stories were designed to bias children towards a coreferential interpretation of the pronoun. In this way, children's 'no' responses could be interpreted as clear evidence for a constraint blocking coreference.
In Russian, as in many other languages, the likelihood of coreference in sentences as above greatly depends on the intonational contour, specifically on whether the pronoun is stressed. In order to attenuate the effects of this factor the puppet spoke in a somewhat 'automatic' voice where every word was stressed (a type of voice that is extensively used in children's cartoons).

Results

Figure 13 shows overall results for all conditions.

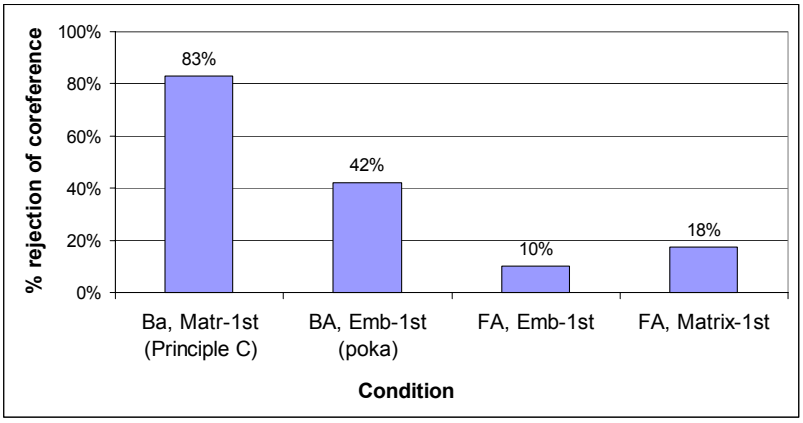


Figure 13. Rejection of the coreference reading by condition in 44 Russian children, aged 2;8-4;11 (mean age = 4;0).

Figure 13 shows that children accepted coreference readings (by responding 'yes') on more than 80% of the control trials involving forwards anaphora. They rejected coreference readings on the Principle C trials at a rate of 83%, providing further cross-linguistic evidence for the claim that even very young children respect Principle C. Russian children reject coreference in the critical *poka*-conditions at a rate of 42% (recall that these in fact include examples with both *poka* 'while' and *kogda* 'when'). On the one hand, this is a significantly higher rejection score than in the two forwards anaphora conditions; on the other hand, this is a much higher rate of acceptance than is found in the Principle C trials. Pairwise comparisons of the condition-scores show a significant effect of clause order within backwards anaphora conditions (2-tailed 2-sample equal-variance t-

test, $t(1,11) = 4.2$, $p = .001$), but not the two forward anaphora control conditions, which were not reliably different, as expected (t (1, 9) = 0.90, p >.1). The mean age differences within each pair of conditions were not significant (all t's<1, p's>.1)

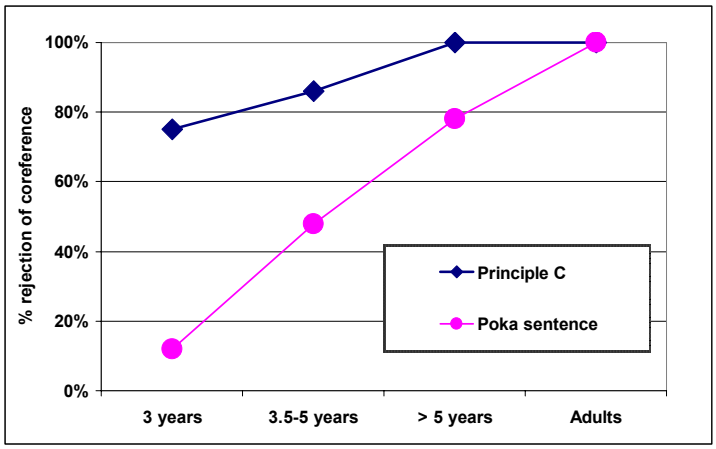


Figure 14. Rejection of coreference in backwards anaphora condistions, ages 3 years old to adults.

Figure 14 shows a breakdown by age-group of the rejection of Principle C and *poka*-constraint conditions in Russian children and adults. The figure shows that whereas rejection of Principle C violations is consistent at all ages tested, rejection of *poka*-constraint violations shows a clear change over time. Around age 3, children show a high level of acceptance of coreference in *poka*-constraint contexts; 3-4 year olds show 48% rejection of coreference. It is not until after age 5 that children consistently disallow coreference in *poka*-constraint contexts. Our group of 5-year olds rejected coreference in both Principle C (100%) and *poka*-constraint contexts (78%).

It can also be worthwhile looking at the individual subjects' data to assess how consistent the child's performance was on multiple trials for a condition. Recall that we used a between-subject design, whereby each child gave 4 data points for the condition that he was randomly assigned to. Thus a child could either accept the puppet's sentence in all trials (Yes/No score = 4/0), or consistently reject it (Yes/No = 0/4), or accept it on some trials but not on others (Yes/No scores 1/3, 2/2 or 3/1). In Table 6 we present the distribution of such Yes/No scores for each condition.

Table 6 reveals that the behavior of the majority of children in every condition was above the chance level which was 50%. Moreover, there were no children with an above-chance behavior who either accepted Principle C sentences or rejected the forwards anaphora conditions. The *poka*-condition represents a special case in that some children accepted this condition on the coreference reading in most trials, whereas other children rejected it. These results suggest

that whereas Russian-speaking children possess an adultlike representation for the structures in (69)-(71) by the age of 3, they acquire the *poka*-constraint later, between the ages 3 and 6.

Table 6. Individual children's performance in by condition. Each cell represents the number of children who had a corresponding Yes/No score in the corresponding condition.

Yes/No	Principle C *n=12*	*Poka*-sent *N=12*	While Pux .. he *n=10*	Pux … while he *n=10*
0/4	7	1	0	0
1/3	2	3	0	0
2/2	3	1	1	2
3/1	0	5	2	3
4/0	0	2	7	5

Before discussing the implications of our results, let us note that whenever children rejected the Principle C or *poka*-sentences they apparently did so for an appropriate reason. If the child rejected the puppet's statement, he/she was asked to explain the reason behind the rejection. The most typical explanation in the Principle C condition involved the child's naming or pointing to the extra-sentential referent; in the *poka*-condition they usually corrected the proposition of the main clause as in (87).

(87) Children's explanations of their rejection of the puppet's statement

Principle C condition
Puppet: … He ate the apple while Pooh was reading the book.
Child: No, him (points to Eeyore)

Poka-condition
Puppet: … While he was reading the book, Pooh ate the apple.
Child: No, a banana!

Discussion

The results of the experiment show a clear developmental dissociation between Principle C and the *poka*-constraint in Russian. Whereas even the youngest Russian-speaking children respect Principle C, they violate the *poka*-constraint at least until the age of 5. This dissociation is particularly striking, because the sentences in which they come into play are minimally different ((71) vs. (72)) and because Russian adults equally strongly disallow coreference in both types of sentences (additional data from adults will be presented in section 4.3.1). The delay lends further support to the notion that the two constraints are independent,

but leaves open the question of what the *poka*-constraint actually is, and how Russian children succeed in learning it by around age 5.

Our study with Russian-speaking children can be added to the list of previous studies that have shown Principle C to be operational early in development in different languages, including English (Crain & McKee 1986) and Italian (Guasti & Chierchia 1999/2000). The growing list of cases showing early presence of Principle C in the child's grammar can be taken as support for the innateness of this principle, as proposed by nativist approaches to language (Pinker 1984, Crain 1991). Considering Principle C as part of the child's innate endowment is justified by its cross-linguistic validity.

The developmental curve for the *poka*-constraint in Russian looks dramatically different (Figure 14). Russian 3-year olds overwhelmingly accept *poka*-sentences on coreference. The 4-year olds' reject them more frequently, but are still far from an adult-like level. Finally, Russian 5-year olds are rather similar to adults in their overwhelming rejection of *poka*-sentences on the coreference reading. Hence, the *poka*-constraint unambiguously arises in the process of and as a result of linguistic development of a Russian-speaking child.

The judgments of the youngest group of Russian children look strikingly like those of English children, as suggested by a comparison of our findings with the findings from Crain & McKee (1986). Figure 15a depicts the results from the youngest group of children in both studies, the 3-year olds, who reject coreference in the Principle C condition, but allow it in the *poka/while*-condition. Although the size of the youngest group of Russian children is small (n=5), the difference between conditions is significant (Mann-Whitney U-test, $p<0.01$). Furthermore, in a later experiment we tested an additional four children aged between 2;11 and 3;4 on the Principle C and *poka*-conditions (between-subject design, two additional children per condition). The data from these children confirms the original pattern. The cumulative data based on the results of all 9 children is given in Figure 15b (compared to the same English group from Crain & McKee 1986).

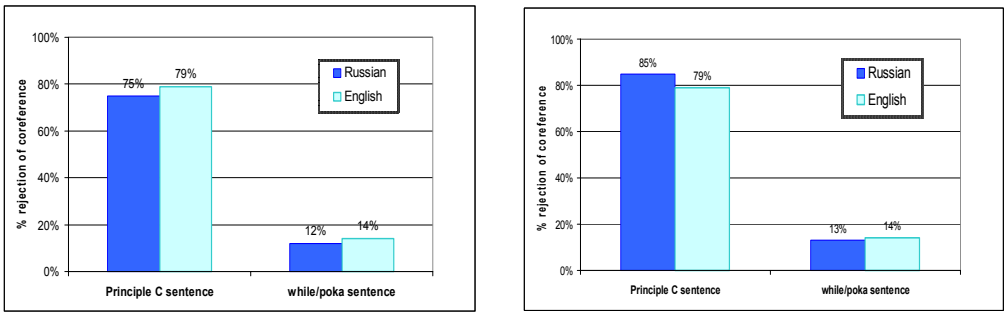


Figure 15: Backwards anaphora judgments of 3-year olds. (a) Russian (n=5, mean 2;9), English (n=7, mean 3;1, Crain & McKee, 1985); (b) expanded Russian group (n=9, mean 3;2), English (n=7, mean 3;1, Crain & McKee 1986)

The older groups of Russian and English-speaking children, aged 3-5 years, continue to give very similar judgments in Principle C sentences. However, Russian children begin to reject coreference in *poka*-sentences more often than their English-speaking peers do in *while*-sentences.[86] With age, the split in the two language groups on the *poka/while*-sentences will become even bigger : English-speaking adults still overwhelmingly accept them on coreference, whereas Russian adults will overwhelmingly reject them.

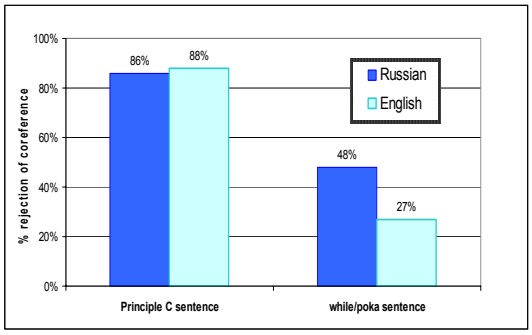


Figure 16: Backwards anaphora judgments of 4-year olds. Russian (n=39, mean 4;2), English (n=62, mean 4;2, Crain & McKee, 1986)

The dissociation between the two constraints in Russian and the fact that Russian-speaking children initially resemble their English-speaking peers rather than the adult speakers of their own language may be viewed as support for the nativist view on language acquisition. In such approach all children start with the same set of general principles (UG) and only later acquire language-specific idiosyncrasies, such as the *poka*-constraint. Thus, young children are expected to differ from their corresponding adults in phenomena where the adult language deviates from the universally-attested principles.

The data from Russian children is also interesting in light of the suggestion that young children have a linear constraint that disallows backwards anaphora across-the-board (Tavakolian 1977, Solan 1983). Previously, Crain & McKee (1986) and Guasti & Chierchia (1999/2000) challenged this claim by showing that English and Italian speaking children have the same representation of backwards anaphora as adult speakers of the respective languages. The Russian findings provide even stronger evidence against the claim that children have an across-the-board linear prohibition against backwards anaphora: strikingly, young Russian-speaking children allow backwards anaphora more liberally than older children or adults in *poka*-sentences. Thus we find no support for the notion that the youngest children

[86] Note that within this older group of Russian children it is not the case that children give a mix of 'yes' and 'no' responses: most of them either consistently accept or consistently reject coreference in *poka*-sentences as shown in Table 6.

have a *linear* constraint against backwards anaphora.[87] However, children do show sensitivity to the order of the pronoun and noun in *hierarchical* terms. Among all cases where the pronoun precedes the noun, they rule out coreference only if the pronoun also c-commands the noun.[88]

We advocate an interpretation of the results whereby differences in the developmental profile of Principle C and the *poka*-constraint are due to differences in the source of the constraints. The child may be pre-disposed towards the structural and universal Principle C constraint, whereas the Russian-specific *poka*-constraint that operates at the level of discourse has to be 'figured out' by the child. However, there may be an alternative interpretation that is compatible with the view that all constraints on coreference have the same pragmatic underlying reasons (van Hoek 1997). In this view early rejection of Principle C sentences on coreference can be explained by the fact that these sentences are 'prototypically bad': in these sentences various cognitive factors that affect the availability of coreference all point in the direction of suggesting that coreference is unlikely, e.g. the pronoun both linearly precedes the name (linear order factor) and is foregrounded as part of the main clause (prominence factor). In *poka*-sentences, on the other hand, these factors pull in opposite directions: whereas the pronoun still precedes the name, it is now backgrounded as part of the embedded clause. Hence, the child may need a longer period of time to start rejecting *poka*-sentences on coreference, since this requires figuring out that the linear order factor may outweigh the prominence factor. Principle C sentences, on the other hand, are illicit on coreference irrespective of the weighting of the two factors and hence can be rejected on coreference without reaching a conclusion on the weights of the two factors. Obviously, what this approach owes is an explanation of how a child might infer the correct 'weightings'.

Finally, we would like to consider whether the 'delay of the *poka*-constraint' effect that we have observed could have the same cause as the well-studied 'delay of Principle B' effect. A number of studies of children's coreference judgments have shown that children incorrectly allow Condition B violations involving referential antecedents (88)a, but correctly reject Condition B violations involving quantificational antecedents (88)b (e.g., Chien & Wexler 1990; Philip & Coopmans 1996, Crain & Thornton 1998, Thornton & Wexler 1999, Avrutin &

[87] Note that if the child indeed started with a prohibition against all instances of backwards anaphora and then relaxed it as needed, it would be difficult to maintain 'negatively' formulated constraints such as Principle C ('coreference is available unless the pronoun c-commands the R-expression'), since what is discovered by the child is when coreference is available.

[88] Proponents of cognitive grammars note that c-command is only one possible way of representing structure. In the Cognitive Grammar (CG) style approaches the traditional notion of c-command may be derived from other parameters, such as figure-ground relations, prominence, etc, which are claimed to be an alternative for c-command. The reader is referred to section 2.4.2 for the discussion why CG approaches to anaphora may not be adequate.

Wexler 1999; but see Elbourne2005). A family of accounts of this effect run as follows. Adults reject both (88)a and (88)b because they correctly analyze both as cases of variable binding, which is regulated by Condition B. Young children incorrectly analyze (88)a as involving coreference, which is not regulated by Condition B, and hence escape Condition B in (88)a. Since the pronoun in (88)b can only be linked to the quantificational subject by means of variable binding, Condition B cannot be escaped. In other words, children can use covaluation (as opposed to variable binding) in configurations where adults cannot (either because they do not know the constraint, or because they are unable to compute its consequences).

(88) a. Mama Bear$_i$ scratched her$_i$.
b. Every bear$_i$ scratched her$_i$.

We can ask whether Russian children's over-acceptance of coreference in *poka*-sentences also results from the use of coreference relations where variable binding is required. We consider this possibility unlikely. First, if this strategy were available to Russian children in *poka*-sentences, it should also be available to them in Principle C contexts (in which the pronoun-name order is identical to *poka*-sentences). But this clearly is not the case, since even the youngest children respect Principle C. Second, the issue of whether a bound variable representation is available is irrelevant in the critical *poka*-sentences, since the main clause subject neither c-commands nor is c-commanded by the subject of the *poka*-clause; the bound variable reading in such structures is unavailable for Russian-speaking adults (see also Avrutin & Reuland 2003). Therefore, we consider it unlikely that the 'delay of the *poka*-constraint' effect and the 'delay of Condition B' effect have the same source.

4.2.3 What can long-distance dependencies tell us about learning?

This section is a digression from the previous discussion in that it does not draw direct inferences from the results of our experiments. Rather, we would like to step back and show how our data may pertain to the discussion of more general issues in linguistic theory and language acquisition.

In previous sections we have been assuming that children may have some principles of language organization as part of their innate endowment. The necessity of innate knowledge has traditionally been linked to the problem of the poverty of the stimulus, i.e. the assumption that the information in the input by itself is not sufficiently rich to allow the child to attain adult competence.[89] Strong

[89] The poverty of the stimulus argument can be illustrated using the well-known example of auxiliary fronting in English interrogatives such as (1) (Chomsky 1965).

(1) The dog that is in the corner **is** hungry. → Is the dog that is in the corner hungry?

connectionist and statistical approaches to language challenge the poverty of the stimulus argument and claim that all facts about language are derivable from the input. They challenge the notion that there are language specific rules and/or constraints, apart from general cognitive constraints that may be applicable in the area of language and language learning.

In what follows I would like to argue against one particular approach which claims the child learns the language by 'blindly' collecting statistics on co-occurences of words (N-grams) in the input; these statistics are assumed to be sufficient to derive all facts about the language. Followers of such an approach may agree with the poverty-of-stimulus argument in that the input to the child is sparse and that certain grammatical constructions may never occur in child-directed speech. However, they challenge the poverty-of-stimulus argument by claiming that linguistic regularities that are not directly observable from the input can nevertheless be discovered even in the absence of pre-existing hypotheses about language representations. The means of deducing the 'missing' regularities is through collecting statistical information on distributions in language. A big part of this statistics deals with co-occurrences of words in the input. Various regularities about the language, including local and non-local dependencies, are considered to be reducible to transitional probabilities between adjacent elements in the input (N-grams).[90] The basic idea is that the combined probability of the N-grams in an ungrammatical sentence is reliably lower than in a corresponding grammatical sentence.

Reali & Christiansen (2003) is a concrete example of such an approach. It is a case study of auxiliary fronting in English (see fn. 89) that shows that the transitional probability of a sentence with correct interrogative fronting (e.g. *Is the man who is hungry _ ordering dinner?*) is higher than that of a sentence with an incorrect interrogative fronting (e.g. *Is the man who _ hungry is ordering dinner?*) using bigram and a trigram models[91]. Given that infants have shown an ability to

---

(2) a. The dog is hungry. → Is the dog hungry?
b. Pluto is the dog that is brushed by Mary. → Is Pluto the dog that is brushed by Mary?
Chomsky claims that children front an appropriate auxiliary (i.e. 2nd auxiliary) on their first production of the construction like (1), although it is unlikely they have been previously exposed to any examples like (1). If so, the child's consistently grammatical behavior on structures like (1) is surprising under the view that all learning is observational. Furthermore, Chomsky argued that if the child treated sentences purely in linear terms, he would never succeed in learning the grammatical way of forming questions for sentences like (1), in light of existence of structures like (2) which suggest that the first auxiliary must be fronted. He concluded that the child must have the notion of a hierarchical structure in order to derive the rule underlying the auxiliary fronting in English.

[90] An N-gram probability is defined as the probability of the n^{th} word given a preceding sequence of n-1 words, i.e. $P(w_n|w_1,w_2,w_3...w_{n-1})$.

[91] The probability P(s) of a sentence s was defined as the product of the probabilities of the words in the sentence, with each word probability conditional to the preceding word (bigram model) or to two preceding words (trigram model). For example, the probability of the sentence s='Is the man who is hungry _ ordering dinner' equals the product of the probabilities of the following

extract certain transitional probabilities from the input (Saffran, Aslin & Newport 1996), this mechanism might explain how a child learns the correct procedure for auxiliary fronting in English without any knowledge of hierarchical structure and even in the absence of sentences like (1) from the child's input. A similar account is sketched by Pullum & Scholz (2002) for the *that*-trace effect in English.

One may challenge the robustness of this result and whether it carries over to other languages. For example, Kam, Stoyneshka, Tornyova, Sakas & Fodor (2005) suggest that the success of the N-gram account of Reali & Christiansen (2003) is an artifact of homophony between the relativizer *that* and the demonstrative pronoun *that* in English. I would like to take a different avenue here and show that this approach cannot serve as a global learning strategy since it fails to capture the notion of ungrammaticality for interpretative dependencies, e.g. referential dependencies.

The reason why the mechanism proposed by Reali & Christiansen (2003) cannot be helpful in the case of referential dependencies is because it is crucially based on the notion that a licit construction will have a less probable form than its illicit counterpart. However, the probability of the surface sentence form by itself will not be sufficient for the child to make a decision about whether a certain pronominal dependency is licit, because the illicitness of a dependency concerns the interpretation of the sentence, rather than the choice or order of words in the sentence. For example, the illicit coreference dependencies that we looked at, the Principle C and *poka*-sentences, are perfectly well-formed and are all possible sentences. The ill-formedness of the Principle C and *poka*-sentences appears only as a result of form-meaning pairing, specifically, when they are evaluated against a coreference scenario. The probability of the ungrammatical sentence *He$_i$ read a book while John$_i$ ate an apple* will always equal the probability of the grammatical *He$_i$ read a book while John$_j$ ate an apple* on any N-gram model, since the two sentences involve an identical set of word forms in an identical order, and therefore the child will never be able to mark one of the interpretations as ungrammatical.[92] No matter how large N is, N-grams that solely track forms cannot be useful in cases of coreference dependencies that are 'homophonous' between a grammatical and an ungrammatical reading. We conclude that statistical models based on calculation of N-grams of surface forms, or any other type of 'form'-based statistics, fall short of explaining acquisition of pronominal dependencies.[93]

---

bigrams: [is-the], [the-man], [man-who], [who-is], [is-hungry], [hungry-ordering], [ordering-dinner].

[92] Needless to say, the opposite scenario can also be true: many sentences that are unacceptable on the disjoint reading (e.g. **After PRO$_i$ getting up, John$_j$ jogged in the park*), are perfectly acceptable on the coreference interpretation.

[93] Statistical models of acquisition that do address the form-meaning issue exist, but they operate with a qualitatively different type of statistical assumptions, i.e. Bayesian inferences (e.g., Tenenbaum & Griffiths 2001). The key insight is that a Bayesian learner must entertain various

We hereby end the discussion on how Principle C and the *poka*-constraint develop in the grammar of the Russian speaker. In the next section we address the second issue on our agenda that concerns examining the effect of these constraints during real-time sentence processing.

4.3 PRINCIPLE C AND THE POKA-CONSTRAINT IN ONLINE PROCESSING

We now turn to an investigation of the second timeline that we alluded to in section 4.1.3.3 which puts the constraints on backwards anaphora into a real-time perspective. Knowing how Russian speakers process sentences like (89)-(90) (repeated from above) online is interesting in several respects.

(89) *On$_i$ čital knigu, poka Ivan$_i$ el jabloko. [Principle C]
He was reading a book while Ivan was eating an apple.

(90) *Poka on$_i$ čital knigu, Ivan$_i$ el jabloko. [*poka*]
While he was reading a book, Ivan was eating an apple.

First, by looking at Russian speakers' processing of sentences like (89) we will extend our investigation of Principle C to a cross-linguistic dimension. Knowing whether candidate antecedents that are subject to Principle C are ruled out in the same way by parsers in different languages could distinguish between competing views of Principle C. If Principle C is a universal configurational constraint on coreference, as we claim it to be, we should find no differences in how the parser processes sentences that are subject to Principle C in English vs. in Russian, as long as the structures that underlie the tested sentences are identical between the two languages. Under approaches that are based on Cognitive Grammar (e.g., van Hoek 1997 discussed in Chapter 2), on the other hand, differences in the processing of sentences that are subject to Principle C in the two languages are possible (although not necessary). According to CG, languages may differ in preferences and weights that they assign to various factors; although eventually the combination of these factors rules out coreference in both English and Russian, the exact amount of effort need not be the same in both languages. This possibility is especially plausible in light of the fact that differences in the preferences chosen by the two languages are strong enough to yield opposite decisions on whether backwards anaphora is licit in structures like (90).

More importantly perhaps, Russian provides a rare opportunity to compare the parser's treatment of antecedents that are illicit due to Principle C and those that are illicit due to another constraint (i.e. the *poka*-constraint) using closely matched stimuli. Sentences (89) & (90) contain identical lexical material, and differ only in

---

hypotheses about language representations that he it trying to understand. This makes Bayesian models of learning qualitatively different from 'blind' and 'tabula rasa' approaches to learning, e.g. frequentist statistical approaches such as the N-gram approach described above, connectionist self-training networks, etc.

the position of the subordinator *while*. In both types of sentences the pronoun linearly precedes an R-expression which eventually is judged as an illicit antecedent for the pronoun in either case. Comparison of when and how different types of inaccessible antecedents are ruled out by the parser could help us understand whether Principle C is special in being immediately respected by the parser, or whether this is true of any constraint on backwards anaphora.

Apart from conditions (89) & (90), the experiment also included a third condition, the *no-constraint* condition (91) that was similar to the *poka*-condition in that the embedded clause preceded the matrix clause, but was introduced by a different subordinator, *do togo kak* 'before' or *posle togo kak* 'after'. As discussed in section 4.1, (91) does not trigger any constraints on coreference and can be used as a control demonstrating how a pronominal dependency is established in a constraint-free environment in Russian.

(91) Do togo kak on$_i$ pročital knigu, Ivan$_i$ s"el jabloko. [no-constraint]
Before he read a book, Ivan ate an apple.

We have argued based on the results of English Experiments 1-3 that backwards anaphora resolution involves an active mechanism that aims at finding an antecedent as soon as possible and yet respects grammatical constraints on coreference. In Experiment 4 below we will test this claim using similar configurations in Russian (i.e. the Principle C and no-constraint conditions). We will broaden our inquiry by adding another condition pair which involves a non-configurational constraint on coreference (the *poka*-condition).

4.3.1 Experiment 4a

The goal of the experiment was to confirm that offline Russian speakers accept coreference between the cataphoric pronoun and the 2nd subject in the no-constraint condition, but reject it in the Principle C or *poka*-conditions.

Participants
40 native speakers of Russian from Moscow were recruited for the experiment.

Materials & Design

The experiment was administered in the form of a pen-and-paper questionnaire.

The questionnaire contained 12 target sentences, equally distributed among three conditions (89)-(91). For the no-constraint condition (91), two sentences contained the subordinator *before* and the other two had the subordinator *after*.

Procedure

We used the same procedure as in the English Experiments 1a and 2a. Participants were asked to judge whether each sentence is acceptable on the coreference reading between a pronoun and a name that were highlighted in bold (*Can the pronoun in bold and the noun in bold refer to the same person?*). The subjects' rated sentences using a 1-to-5 scale, where 1=impossible, 5=absolutely natural.

Results

The results of Experiment 4a are summarized in Figure 17. Speakers rejected coreference in both the Principle C and the *poka*-conditions (average rating score 1.1 and 1.3 respectively). Both scores were rather close to the minimum score of 1, although the difference was statistically significant (2-tailed t-test, $p < .01$). Coreference acceptance in the no-constraint condition (3.6 on average) was significantly higher than in either the Principle C or *poka*-condition ($p<.001$), although not as high as in the forwards anaphora condition (4.3 on average).

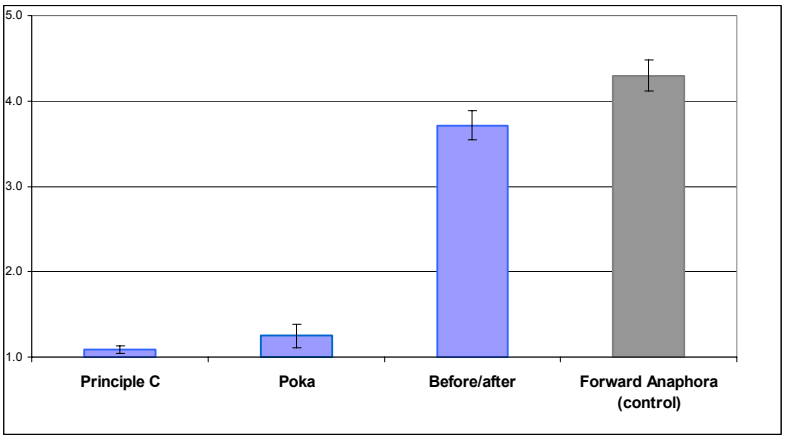


Figure 17. Results from an offline Experiment 4a.

Discussion

The results of Experiment 4a confirm that Russian speakers strongly disallow coreference in sentences like (89) that are subject to Principle C and in *poka*-sentences like (90). Importantly, the same speakers who rejected the coreference reading in the Principle C and *poka*-conditions accepted it in the no-constraint condition.

4.3.2 Experiment 4b: Self-paced reading

Having confirmed that offline Russian speakers reject backwards anaphora in the Principle C and *poka*-sentences, and accept it in sentences introduced by subordinators *before* or *after*, we turn to the question of how decisions regarding availability of backwards anaphora are made during online processing.

Participants

42 speakers of Russians (age 18-28, mean age 23.6 years, 22 male/20 female) were run in Moscow, Russia.[94] All participants were right-handed with normal or corrected-to-normal vision and no history of language disorders. They gave informed consent and were paid $10 for participation in the experiment.

Materials & Design

Twenty-four sets of items were constructed using a 2x2 design with the factors *constraint type* (Principle C vs. *poka*-constraint) and *congruency* (gender-match vs. gender-mismatch) between the pronoun and the 2nd clause subject ((92)a-d). In addition, we constructed 12 sets of the no-constraint condition, in which the sentence-initial embedded clause was introduced by the conjunction *do togo kak* 'before' or *posle togo kak* 'after' ((92)e,f).

(92) Simplified stimuli from Experiment 4b.

a. Principle C, gender-match
Ona učila bilety po fizike, poka **Vera** slušala prjamoj reportaž s "Evrovidenija".
b. Principle C, gender-mismatch
Ona učila bilety po fizike, poka **Ivan** slušal prjamoj reportaž s "Evrovidenija".
'She was studying for her physics exam while Vera/Ivan was listening to a live broadcast from "Eurovision".'

[94] This was a different group of participants from Experiment 4a.

(93) *c. Poka, gender-match*
Poka **ona** učila bilety po fizike, **Vera** slušala pryamoj reportaž s "Evrovidenija".
d. Poka, gender-mismatch
Poka **ona** učila bilety po fizike, **Ivan** slušal pryamoj reportaž s "Evrovidenija".
'While she was studying for her physics exam Vera/Ivan was listening to a life broadcast from "Eurovision".'

e. No-constraint, gender-match
Do togo kak **ona** vošla v sostav sbornoj, **Nataša** byla soveršenno neizvestna specialistam.
f. No-constraint, gender-mismatch
Do togo kak **ona** vošla v sostav sbornoj, **Mixail** byl soveršenno neizvesten specialistam.
'Before she entered the national team, Natasha/Michael was completely unknown to specialists.'

The gender of the pronoun in the first clause was balanced across all items: half of the sets were built around the masculine pronoun *on* 'he' and the other half were built around the feminine pronoun *ona* 'she'. Therefore, the only difference between the gender-match and gender-mismatch variants in each pair was the gender of the subject of the second clause (and the following predicate, due to the subject-verb agreement in Russian). The 2nd subject was always a gender-unambiguous proper name, controlled for the number of letters and syllables in the gender-matching and gender-mismatching counterparts.

In creating the stimuli for the Principle C and *poka*-conditions we chose the main and embedded predicates such that they could plausibly be performed by either different agents or the same agent. We also did our best to choose pairs of events such that the plausibility of the disjoint or coreference interpretation did not differ between the Principle C and the *poka*-conditions. This was to ensure that, if found, a difference in the reading profiles between the conditions could safely be taken to reflect differences in the online application of the two constraints, rather than a bias in the stimuli. This point will be explained in greater detail in Experiment 4c, which was designed to control for this issue.

As in all of the previous on-line experiments in this project, we provided an inter-sentential grammatical antecedent for the cataphoric pronoun. We used the same method as in the English Experiment 1: each of the sentences in (92) was embedded into a bigger sentence introduced by the conjunction *xotja* 'although' or *poskol'ku* 'since', as illustrated in (94). In the Principle C and *poka*-conditions the subject of the 3rd clause (the main clause of the resulting sentence) always matched in gender with the pronoun and served as a licit antecedent for that pronoun. In the no-constraint condition the gender of the 3rd clause subject was

chosen such that the sentence had a unique antecedent for the pronoun. Hence, whenever the pronoun matched the 2nd subject in gender, it mismatched with the 3rd subject and vice versa (see (94)e,f).

(94) A full sample stimulus set from Experiment 4b.

a/b. Principle C conditions, gender-match/gender-mismatch
Xotja v voskresenje ona učila bilety po fizike, poka **Vera**/**Ivan** slušala/slušal prjamoj reportaž s "Evrovidenija", Rita umudrjalas' ne obraš'at' vnimanija na proisxodjaščee.
'Although on Sunday she was studying for her physics exam while Vera/Ivan was listening to a live broadcast from "Eurovision", Rita managed not to pay attention to what was going on.'

c/d. Poka-conditions, gender-match/gender-mismatch
Xotja v voskresenje, poka ona učila bilety po fizike, **Vera/Ivan** slušala/slušal pryamoj reportaž s "Evrovidenija", Rita umudrjalas' ne obraš'at' vnimanija na proisxodjaščee.
'Although on Sunday while she was studying for her physics exam Vera/Ivan was listening to a life broadcast from "Eurovision", Rita managed not to pay attention to what was going on.'

e/f. No constraint-conditions, gender-match/gender-mismatch
Xotja do togo kak ona vošla v sostav sbornoj, **Nataša/Mixail** byla/byl sovershenno neizvestna/neizvesten specialistam, Mixail/Natasha nikogda ne somnevalsja/somnevalas' v ee/ego vydajuščixsja trenerskix sposobnostjax.
'Although before she entered the national team, Natasha/Michael was completely unknown to specialists, Michael/Natasha had never doubted her/his talent of a coach.'

In addition to the target sentences, the experiment also contained 84 filler sentences of various length and complexity. Some fillers contained proper names and some started with a subordinator to create items that were superficially similar to the targets. Every sentence was followed by a Yes/No comprehension question. Exactly half of the sentences were followed by a comprehension question to which 'yes' was the correct response, and the remaining half had 'no' as a correct response. The full list of experimental materials is given in Appendix 4-B.

Target items were distributed among four presentation lists in a Latin Square design. Each list contained 36 experimental sentences (12 items per condition) and 84 filler sentences. Filler items were identical across all four lists. Participants were assigned to one of the lists according to their number; the order of the stimuli within the presentation list was randomized for each participant.

Procedure

Participants were tested using a laptop running the Windows-based version of the Linger software (Doug Rohde, MIT). Sentences were presented in a standard non-cumulative word-by-word moving window paradigm with the font Courier New Cyrillic 20. Each trial started with a blank screen. Upon pressing the space-bar, a sentence masked by dashes appeared on the screen. The masking extended to all letters and punctuation signs, but left spaces between words visible. As the participants pressed on the spacebar, a new word appeared on the screen, whereas the previous word was re-masked by dashes.

The comprehension question appeared all at once and unmasked. To answer the question the subject pressed the *f* key for 'yes' and the *j* key for 'no'. If the question was answered incorrectly the word HEBEPHO 'incorrect' briefly appeared in the center of the screen.

Participants were instructed to read sentences at a natural pace and to respond to the comprehension questions as accurately as possible. The testing session lasted 45 minutes on average. Following the online experiment, each participant filled out a pen-and-pencil questionnaire aimed at confirming the plausibility of the target sentences (Experiment 4c).

Analysis

Two subjects were excluded from the final analysis. One subject was excluded due to low accuracy on the comprehension questions (< 50%) and the other due to a reading rate that exceeded the mean rate by more than 2.5 standard deviations. The analysis was performed on the remaining 40 subjects, who were equally distributed between the four experimental lists. Sentences for which the comprehension question was answered incorrectly were excluded from the analysis. Furthermore, reading times that exceeded the mean value for a given condition and region by more than 2.5 standard deviations were winsorized (i.e. replaced by that threshold). This procedure affected 2.1% of all data points (1.9-2.5% for individual conditions).

Reading times were analyzed in regions that corresponded to a single word (except for the last region in each clause that represents an average of reading times for all remaining words in that clause). Reading times from the *poka*- and Principle C conditions that were closely matched on lexical material were entered into a 2x2 ANOVA with the factors *constraint* (Principle C vs. *poka*-constraint) and *congruency* (gender-match vs. gender-mismatch). Reading times from the no-constraint condition were analyzed in a separate 1-way ANOVA with two levels of the factor *congruency* (gender-match vs. gender-mismatch). In all cases two ANOVA's were calculated, with participants (F1) or items (F2) as a random factor.

Results

Comprehension question accuracy

The mean comprehension question response accuracy for filler items was 93.1%. For target items the average comprehension question response accuracy was 88.5% and did not differ significantly among the conditions (84.2-92.5% for individual conditions, all *F*s < 1.5).

Self-paced reading

Principle C and Poka conditions

Results from the Principle C and *poka*-conditions are diagrammed in Figure 18.

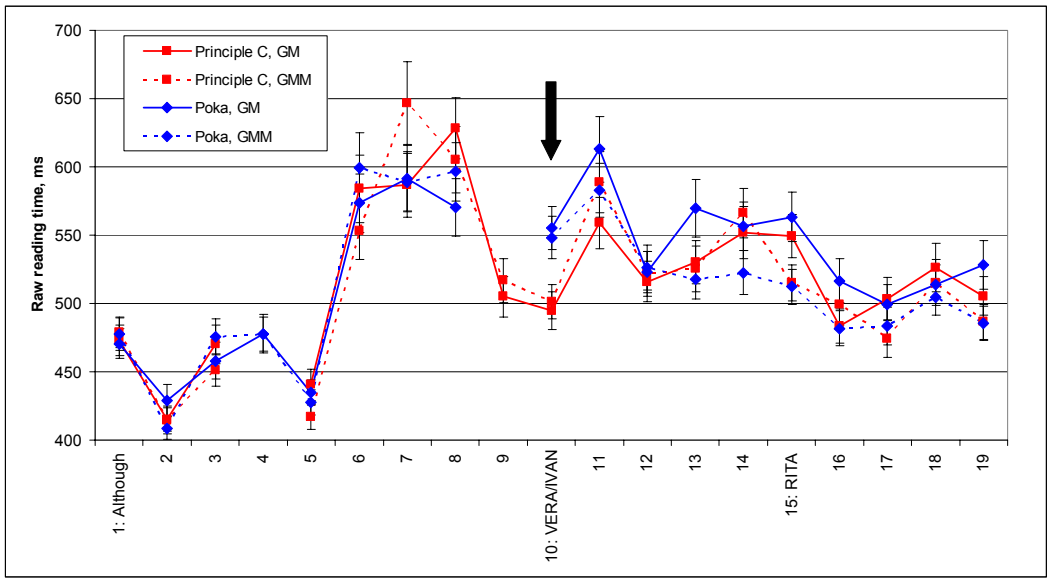


Figure 18. Experiment 4b. Raw reading times from the Principle C and *poka*-conditions. The arrow marks the position of the critical 2nd subject. Full information on regions is given in Figure 19 & Figure 20.

In the main 2x2 ANOVA a main effect of *congruency* was significant in the participants analysis and marginally significant in the items analysis ($F1(1,39)=4.3$, $p < .05$, $F2(1,23)=3.2$, $p = .09$) at the pronoun subject (region 5). This effect was driven by the Principle C conditions, which showed an unexpected effect of *congruency* in the same region that was significant in the participants analysis only ($F1(1,39) = 5.42$, $p < .05$, $F2(1,23) = 2.78$, $p = .10$).

There was a main effect of *constraint* ($F1(1,39)=19.1$, $p < .001$, $F2(1,23)=15.9$, $p = 0.001$) at the critical 2nd subject (region 10) due to differences in the material in the immediately preceding region, i.e. the subordinator *while* in the Principle C conditions vs. the end of the first clause in the *poka*-conditions. There was a marginally significant effect of *congruency* in region 13 ($F1(1,39)=3.1$, $p = .09$, $F2(1,23)=3.8$, $p = 0.07$). Planned pairwise comparisons within each level of the

factor *constraint* revealed a significant effect of *congruency* in region 13, three words downstream from the critical 2nd subject (F1(1,39) = 5.4, p<.05, F2(1,23) = 4.7, p<.05) in the *poka*-conditions (Figure 20). The effect was due to longer average reading times when the 2nd subject matched in gender with the preceding pronoun (570 vs. 517 ms), i.e. a GME. There was no effect of *congruency* in the Principle C condition in region 13, or anywhere else in the 2nd clause (all *F*'s<1.5 , all *p*'s>.2) (Figure 19).

A main effect of *congruency* was found in the main ANOVAs at the subject of the 3rd clause (region 15): F1(1,39)=10.2, p < .01, F2(1,23)=9.2, p < 0.01). Pairwise comparisons within each level of the factor *constraint* revealed an effect of *congruency* in the Principle C condition that was significant in the participants analysis and marginally significant in the items analysis (F1(1,39) = 4.9, p<.05, F2(1,23) = 2.8, p = .1). The effect of *congruency* was also significant in the *poka*-conditions (F1(1,39) = 7.3, p<.01, F2 (1,23) = 5.2, p <.05). In both cases, the 3rd subject was read more slowly if the 2nd subject matched the gender of the preceding pronoun. Moreover, the *constraint* x *congruency* interaction in region 16 was significant in the participants analysis and marginally significant in the items analysis (F1(1,39)=5.2, p < .05, F2(1,23)=3.3, p = 0.07). Resolving this interaction revealed that it emerged mostly due to the *poka*-conditions, in which the effect of *congruency* was significant in the participants analysis (F1(1,39) = 5.1, p < .05, F2 (1, 23) = 2.5, p = 0.12).

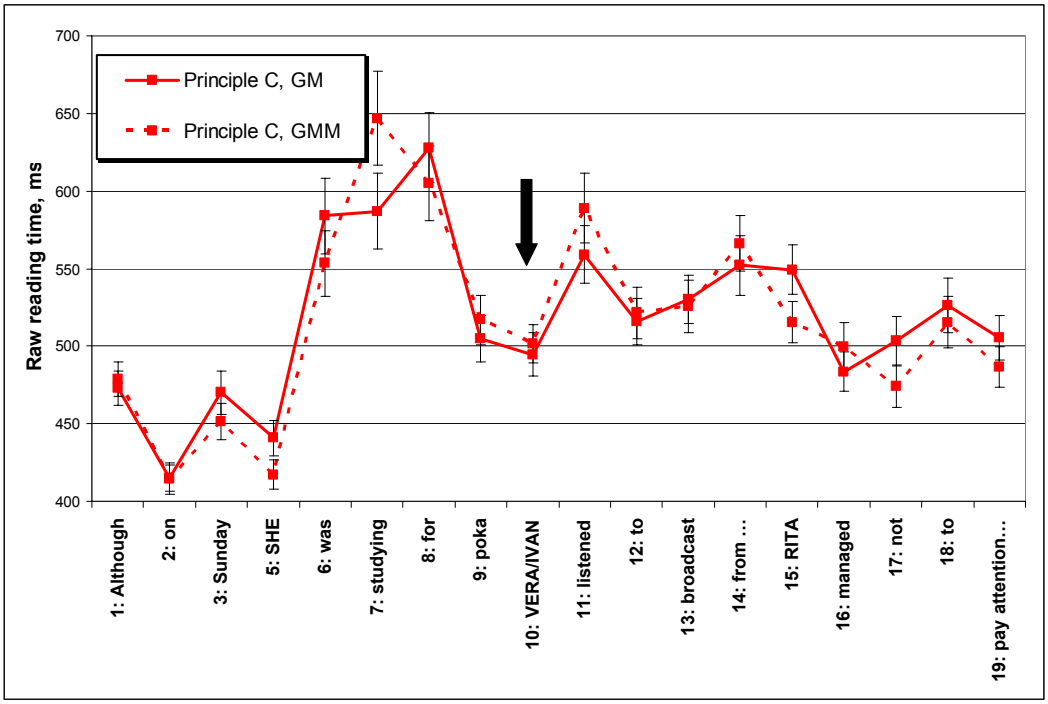


Figure 19. Raw reading times from the Principle C conditions, Experiment 4b. The arrow marks the position of the critical 2nd subject.

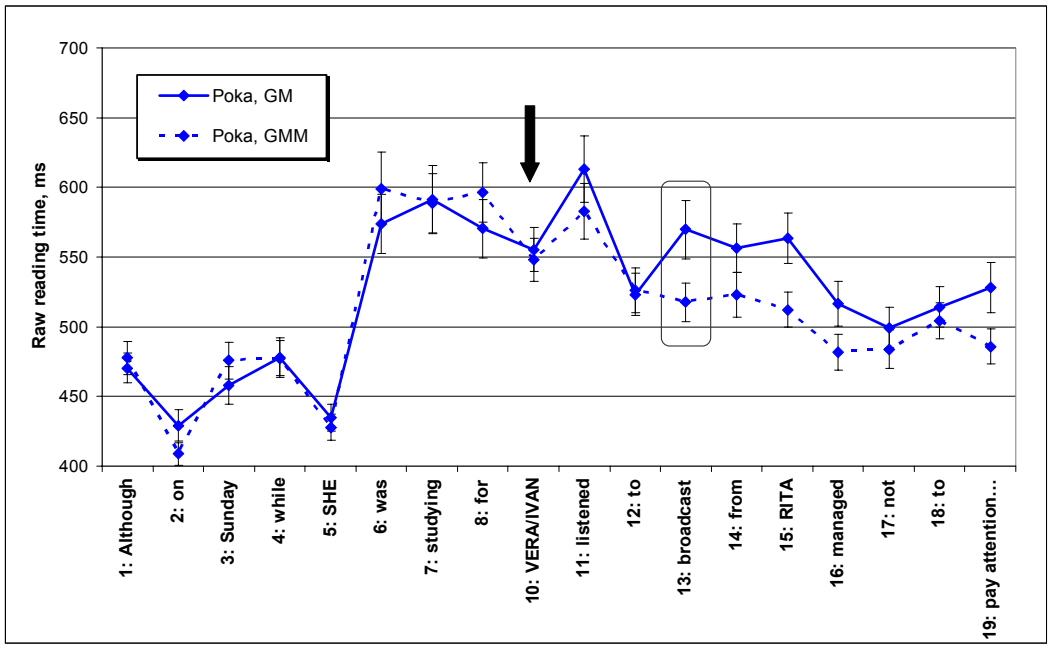


Figure 20. Raw reading times from the *poka*-conditions, Experiment 4b. The arrow marks the position of the critical 2nd subject. A GME was found three words down from the critical subject (marked by a box).

No-constraint conditions

The results from the no-constraint conditions are diagrammed in Figure 21.

There were no significant effects in the first clause (all *F*s>2.6, *p*s>.1). The effect of *congruency* was not significant at the critical 2nd subject (region 10) or in the following region (all *F*'s<1.7, *p*'s>.2). However, in region 12, two words downstream from the critical 2nd subject, we found a GMME that was significant in the participants analysis and marginally significant in the items analysis ($F1(1,39) = 4.2$, $p < .05$, $F2(1,11) = 3.4$, $p = .07$). Mean reading times in region 12 were longer in the gender-mismatching sentences than in the gender-matching sentences (608 vs. 553 ms). There were no other significant effects in the remainder of the sentence (all *F*s < 2.2, *p*s > .1).

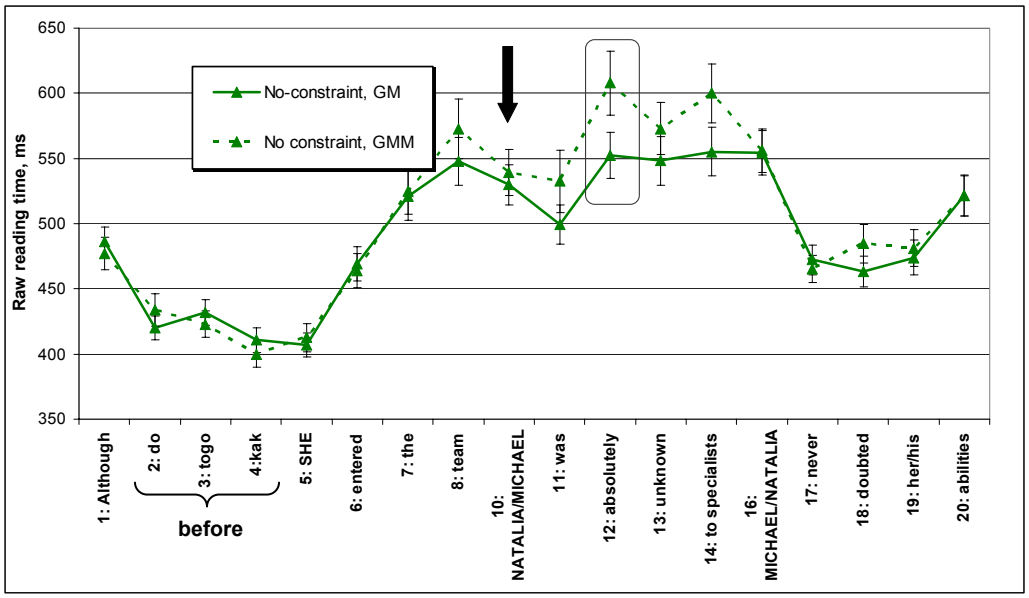


Figure 21. Raw reading times from the no-constraint conditions, Experiment 4b. A GMME was found two words downstream from the critical subject (marked by a box). Region marking:Xotja$_1$ do$_2$ togo$_3$ kak$_4$ ona$_5$ vošla$_6$ v$_7$ sostav$_8$ (sbornoj,)$_9$ Nataša/Mixail$_{10}$ byla/byl$_{11}$ soveršenno$_{12}$ neizvestna/neizvesten$_{13}$ specialistam$_{14}$ (…)$_{15}$, Mixail/Nataša$_{16}$ nikogda$_{17}$ ne$_{18}$ somnevalsja/somnevalas'$_{19}$ (v ee/ego vydajuščixsja trenerskix sposobnostjax.)$_{20}$
'Although before she entered the national team, Natasha/Michael was completely unknown to specialists, Michael/Natasha had never doubted her/his talent of a coach.'

Discussion

In Experiment 4b we found that a gender manipulation in the 2nd subject caused a GMME in the no-constraint conditions and a GME in the *poka*-conditions. There was no effect of the gender congruency of the 2nd subject in the Principle C condition.

One of the most interesting findings in the experiment is that the Principle C and *poka*-conditions yielded different parsing profiles: the 2nd subject that is eventually judged to be an inaccessible antecedent in either condition was treated differently as the parser incrementally stepped through the sentence. Before going into interpreting these results, it is important to be sure that differences in processing of the Principle C and *poka*-conditions were indeed due to the constraints, rather than to artifacts of the stimuli that we used. Let us explain which potential artifacts we have in mind.
The stimuli in the Principle C and *poka*-conditions differed in the position of the subordinator *while*. Ideally, this manipulation should have had no other effect but switching the sentence structure so that the sentence either invokes Principle C or the *poka*-constraint. However, this structural change is accompanied by a shift in the figure/ground relation of the events in the sentence, as demonstrated in (95).

(95) Principle C: *E1, while E2* [E1 = figure, E2 = ground]
poka-condition: *While E1, E2* [E1 = ground, E2 = figure]

The subordinator *while* establishes a relation in which the embedded event serves as a ground for the main event, which is a figure (also known as foregrounding/ backgrounding, e.g. Matthiessen & Thompson 1988). If so, in addition to the structural differences, the conditions in (95) also differ in which of the events E1 and E2 is figure or ground. Below we show why the figure/ground relation of the events is important for us.

It is obvious that only some pairs of events meet a criterion whereby either event in the pair can serve both as figure (or as ground) for the other event in a *while*-sentence. (96) is a pair of events that meet this criterion, whereas (97) does not.

(96) E1= reading a letter, E2= eating an apple
a. Jane read a letter while Bill ate an apple. [E1 figure, E2 ground]
b. While Jane read a letter, Bill ate an apple. [E1 ground, E2 figure]
(97) E1=breaking a glass, E2= reading a letter
a. Jane broke a glass while Bill ate an apple. [E1 figure, E2 ground]
b. #While Jane broke a glass, Bill ate an apple.[E1 ground, E2 figure]

All stimuli that were used in Experiment 4b were built using 'symmetric' event pairs such as (96). However, events that make up a symmetric figure/ground pair if performed by two different agents might not be such if they are performed by the same agent. An example is given in (98) & (99) (we use forwards anaphora to convey coreference).

(98) E1=feeling dizzy, E2=cleaning the floor
a. Jane was feeling dizzy while Bill was cleaning the floor. [E1 figure, E2 ground]
b. While Jane was feeling dizzy, Bill was cleaning the floor. [E1 ground, E2 figure]

(99) E1= feeling dizzy, E2=cleaning the floor
a. Jane$_i$ was feeling dizzy while she$_i$ was cleaning the floor. [E1 figure, E2 ground]
b. ?While Jane$_i$ was feeling dizzy, she$_i$ was cleaning the floor.[E1 ground, E2 figure]

(98) shows that a chosen pair of events E1 & E2 is 'symmetric'. Events E1 and E2 are such that either of them can be a plausible figure (or ground) for the other one, as suggested by (98)a and (98)b being equally plausible. Note that in (98) E1 and E2 are performed by different people. In (99) we use the same events, but now they are performed by the same person. (99)a is quite natural on the coreference reading and suggests that E1 and E2 can in principle be performed

simultaneously by the same agent. However, the coreference reading in (99)b is less plausible, for the reason that *feeling dizzy* is a not an ideal ground for the event *cleaning the floor* in a frame where these events describe a simultaneous activity by the same person.[95]

The existence of asymmetries as in (98)-(99) led us to check that our stimuli were not biased in a similar way. Specifically, we wanted to ensure that the pairs of events E1 & E2 in each set were such that switching their figure/ground relation was not accompanied by introduction of a bias towards a coreference or disjoint interpretation, as in (99)a vs. (99)b. If both figure/ground combinations of the events were equally plausible on coreference interpretation, we could confidently attribute differences in the processing of the Principle C vs. the *poka*-conditions specifically to the fact that they invoke different constraints on coreference. We address this issue in Experiment 4c.

4.3.3 Experiment 4c

Experiment 4c was a pencil-and-paper questionnaire. The aim of the experiment was to determine whether the pairs of events that we used in Experiment 4b were balanced so that switching their figure/ground relation did not have a bias on the sentence interpretation in terms of coreference resolution. Participants rated the plausibility of coreference in a subset of sentences from Experiment 4b, in which the order of the pronoun and the name was swapped to obtain forwards anaphora. Thus, stimuli from the Principle C condition were transformed into forward anaphora sentences in which the main clause preceded the embedded clause (i.e. FA/matrix-first), whereas the *poka*-conditions were transformed into sentences with forwards anaphora in which the embedded clause preceded the main clause (FA/embedded-first). Reversing the order of the pronoun and the name (i.e. switching from backwards anaphora to forwards anaphora) renders Principle C and the *poka*-constraint irrelevant and thus makes it possible to identify any potential effect of the figure/ground relation of the events on sentence interpretation.

Participants judged the plausibility of each sentence on the coreference reading on a 1-to-5 scale (in each sentence two NP's, a name and a pronoun, that were meant to be taken as coreferent were highlighted in bold as shown in (100)). Besides the conditions mentioned above (FA/matrix-first & FA/embedded-first) there were filler sentences that were either plausible (*filler-plausible*) or implausible (*filler-implausible*) on the coreference reading. In addition to serving as distractors, the fillers also served as a control that the participant understood the task correctly.

[95] The example in (99) involves an experiencer predicate in the first clause. This is only done for the purposes of illustration. All predicates in the 1st and 2nd clause in Experiment 4 were agentive, which made the *poka*-conditions subject to the *poka*-constraint.

(100) Stimuli from offline Experiment 4c.

a. FA/matrix-first
Vera učila bilety po fizike, poka **ona** slušala/slušal prjamoj reportaž s "Evrovidenija".
Vera *was studying for her physics exam while* ***she*** *was listening to a live broadcast from "Eurovision".*

b. FA/embedded-first
Poka Vera učila bilety po fizike, **ona** slušala/slušal prjamoj reportaž s "Evrovidenija
While ***Vera*** *was studying for her physics exam* ***she*** *was listening to a live broadcast from "Eurovision".*

c. Filler-plausible
Tak kak za poslednie tri goda **ona** ni razu ne brala otpuska, **Olesja** tverdo rešila, čto v etom godu uedet otdyxat' na more ne menee chem na poltora mesjaca.
Since in the last three years ***she*** *never took time off,* ***Olesja*** *firmly decided to go to a sea resort for at least a month-and-a-half this year.*

d. Filler-implausible
Poka **Inna** naxodilas' po bedro v gipse, **ona** bez truda begala po lestnicam.
While ***Inna's*** *leg was in cast,* ***she*** *could easily run stairs.*

Experiment 4c was administered as a post-test to Experiment 4b. The results are based on the same 40 participants who were included in the final analysis in Experiment 4b.

Results
The results of Experiment 4c are diagrammed in Figure 22.

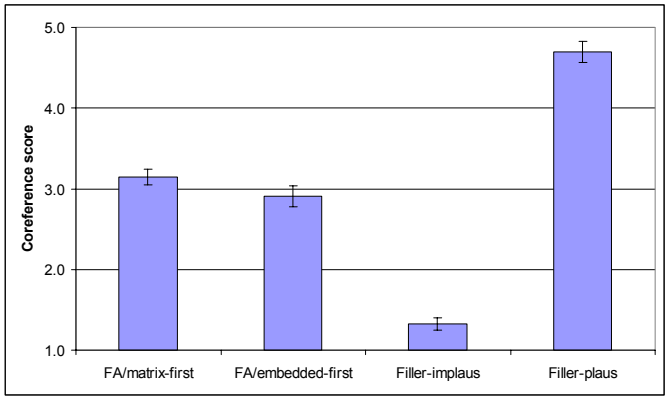


Figure 22. Results of Experiment 4c by condition.

As expected, the implausible fillers received the lowest rating (average rating score = 1.3), whereas the plausible fillers received the highest rating (average rating score = 4.7). The critical conditions FA/matrix-first and FA/embedded-first received intermediate scores in the middle of the range (average rating score 3.1 & 2.9 respectively). Importantly, the difference in the plausibility scores was not significant (2-tailed paired t-test, $p > .1$).

Discussion

Experiment 4c confirmed that the stimuli used in the Principle C and *poka*-conditions were balanced in terms of the figure/ground relation between the events. There was no bias towards higher/lower plausibility of the coreference reading as a result of a reversal of the figure/ground relation of the events in the Principle C vs. the *poka*-conditions. In light of this finding we conclude that the differential treatment of the 2nd subject in the two conditions in the online experiment (Experiment 4b) is due to the differences in the constraints that apply in each condition.

4.3.4 General Discussion

The main findings of Experiment 4 concern the parser's behavior following the 2nd subject position in which the critical gender manipulation occurred. We found a GMME following the critical 2nd subject position in the no-constraint condition in Russian, thus replicating our earlier finding with similar sentences in English. Also similar to English, there was no effect on the gender congruency of the 2nd subject in the Principle C condition. Finally, in the Russian-specific *poka*-condition manipulation of gender congruency in the 2nd subject yielded a GME, i.e. reading times were longer when the 2nd subject matched in gender with the preceding pronoun.

In other words, the gender manipulation in the 2nd subject had an effect on processing in two of the three conditions and the direction of the effect was in opposite directions: an incongruous 2nd subject led to longer reading times in the no-constraint conditions and to shorter reading times in the *poka*-conditions. Both of these pairs of conditions contrast with the Principle C conditions, in which no difference was found anywhere in the 2nd clause. We will next discuss the implications of these results by comparing conditions to each other in pairwise fashion.

Principle C conditions (null-effect) vs. Poka-conditions (GME)

Differences in the parsing profiles for the Principle C and the *poka*-conditions suggest that the parser took different routes to arrive at its final decision on ruling out coreference between the pronoun and the 2nd subject in either condition. We

take the gender-match effect |in the *poka*-condition to suggest that the parser entertains a relation between the 2nd subject and the pronoun (hence, there is interest in the content of that position). The lack of an effect from the 2nd subject in the Principle C conditions then suggests that the parser never considers that position as a candidate antecedent position for the preceding pronoun.

In Chapter 3 we claimed that Principle C applies immediately during processing based on a comparison of the parser's treatment of antecedents that are subject to Principle C with those that are not subject to any constraints. The findings from the Russian Experiment 4 present more compelling evidence for this claim. These findings suggest that some types of antecedent that are eventually inaccessible are nevertheless temporarily considered by the parser in the course of processing, whereas antecedents that are subject to Principle C are immediately disregarded from the earliest stage of processing.

Let us note that one can be confident that the GME in the *poka*-condition was not merely a general interference caused by the co-occurrence of a pronoun and a referent of the same gender. If that was the case, we should have obtained the same effect in the Principle C condition. Thus, the GME in the *poka*-condition must be attributed to the parser's consideration of a relation between the 2nd subject and the pronoun. We will return to what exactly is involved in this process later.

No-constraint conditions (GMME) vs. Poka-conditions (GME)

In both pairs of conditions the gender manipulation in the critical 2nd subject position had an effect on processing. However, as we said earlier, the effects went in the opposite direction: a gender-incongruent 2nd subject required more processing effort than its congruent counterpart when it was an accessible antecedent (no-constraint conditions), but less effort when it was subject to the *poka*-constraint.

Comparison of these conditions can lead us to make inferences about the time-course and manner of application of the *poka*-constraint. At first, the GME in the *poka*-condition may be taken to suggest that the 2nd subject is initially treated by the parser as a viable antecedent candidate that is ruled out at a later stage.[96] However, this hypothesis falls short of explaining our results. If the parser temporarily considered the 2nd subject in the *poka*-condition as a perfectly licit antecedent (i.e. an antecedent that is not subject to any constraints on coreference), this subject should have caused the same effect as its counterpart in the no-constraint condition (i.e. GMME). Yet this was not the case. Hence, the parser should have detected that the sentence could be subject to the *poka*-constraint

[96] This was how Badecker & Straub (2002) explained a GME obtained as a result of gender manipulation in the inaccessible antecedent in their study.

early, presumably, after encountering the subordinator *poka* 'while' followed by an agent subject pronoun in the 1st clause.

One more point is needed here concerning NP positions that are predictable when a cataphoric pronoun is encountered. It would not be implausible to suggest that the parser always considers such positions as a candidate antecedent regardless of other factors. This claim cannot be assessed based on our English experiments, since there the existence of a prediction at the time of the pronoun always co-varied with the accessibility of that position as an antecedent for the pronoun. In Experiment 1 (which was the only experiment in which such an early prediction for an antecedent was available) the 2nd subject could be anticipated at the time when the pronoun was encountered only in the no-constraint conditions in which this position was an accessible antecedent position; in the Principle C conditions, on the other hand, the 2nd subject, an inaccessible antecedent, could not be anticipated at the time when the pronoun was processed. In the Russian experiment the two factors were separated: a position that was predictable at the pronoun could contain either a licit (the no-constraint conditions) or an illicit (the *poka*-conditions) antecedent for that pronoun. The results suggest that the existence of a prediction for an NP position could not have been the sole factor in determining its status as a potential antecedent: opposite effects (GMME vs. GME) were obtained in the no-constraint vs. *poka*-conditions, despite the fact that they make an identical prediction for an upcoming main clause. Hence, other factors, such as the presence of a constraint, must have been taken into consideration by the parser.

In the discussion above I assumed that the parser makes a top-down prediction for an antecedent in the 2nd subject position in the no-constraint and the *poka*-conditions. An alternative to this view is that the parser makes no anticipatory predictions and rather checks whether an NP could serve as an antecedent for the cataphoric pronoun as that NP becomes available from the input. We cannot completely rule out this possibility, especially in light of the fact that unlike in English, the GMME to an accessible antecedent in Russian was somewhat delayed and did not appear until two words down from the critical region. Crucially, regardless of whether the parser makes a top-down prediction, it must be conceded that the parser initially considers the accessibility of the position of an NP before checking whether that NP and the pronoun match in gender. Otherwise, as we said repeatedly above, we would expect, contrary to fact, the same reaction by the parser to any morphologically incongruent NP independent of whether it is in a binding-accessible or inaccessible position.

Why GME in the poka-condition? What is computed first: gender or position?

An objection may be raised here concerning an apparent inconsisteny in our interpretation of the results in the *poka*-condition: if the parser is aware of the *poka*-constraint early on, why is there a GME following the critical subject, rather than a null-effect as in the Principle C condition?

The GME in the *poka*-condition indicates an increased processing effort when the gender of the 2nd subject matches the gender of the pronoun. We suggest that although the parser becomes aware of the *poka*-constraint early, as soon as it encounters a fronted *poka*-clause with a pronoun subject, it nevertheless does not completely discard the 2nd subject from the list of candidate antecedents until later. In what follows I will discuss various reasons that could explain this 'conservative' behavior by the parser.

In line with the fact that constraints on coreference can vary in their nature, they may also vary in their 'strength'. Principle C is a configurational constraint that is valid cross-linguistically. The *poka*-constraint operates at the level of discourse and is highly language-specific. We saw earlier that the constraints have very different developmental profiles: whereas Principle C is operational from a very early age and is possibly innate, the *poka*-constraint needs to be learned by the child. The same factors could affect the real-time application of the constraint, e.g. if the parser relies more willingly on constraints that are configurational and universal.

Alternatively, there may be no qualitative differences in the application of Principle C and the *poka*-constraint, each constraint taking effect as soon as all prerequisite conditions are met. In order to finalize a coreference relation, at least two requirements must be met: the antecedent should be accessible, i.e. not invoke any (syntactic or discourse) constraints on coreference, and its morphological features should match those of the pronoun. An efficient parser can potentially rule out an antecedent based on either of the factors, whichever may come first. As we claimed in Chapter 3, in the case of Principle C an inaccessible antecedent can be ruled out on the basis of its structural position, information about which is available using top-down prediction and prior to when the gender of that antecedent becomes known. In contrast, in the case of *poka*-sentences, the information on the gender of the critical 2nd subject becomes available earlier than the 'accessibility' information, which in this case requires knowing the thematic role of that NP (section 4.1). The thematic role of the 2nd NP can only be determined later upon encountering the predicate that follows that NP. Thus, the parser excludes coreference based on gender whenever possible, i.e. when the gender of the cataphoric pronoun mismatches that of the 2nd subject. However, when the two NPs do match in morphological features, the parser needs to wait longer until it gets the predicate in order to check the eligibility of the 2nd subject. So a GME in the *poka*-condition indicates an increased effort in establishing whether the 2nd subject is a licit antecedent when it matches in gender with the pronoun.

4.4 CONCLUSION

We have shown that in addition to Principle C Russian has another constraint on backwards anaphora, the *poka*-constraint, that operates in complex sentences introduced by the subordinator *poka* 'while'. We argued that unlike Principle C, the *poka*-constraint operates at the level of discourse. We then showed that the *poka*-constraint is delayed compared to Principle C in language development and that it is also delayed in terms of when it applies during sentence processing.

Appendix 4-A. List of stories from the acquisition experiment with Russian speaking children.

1. Winnie-the-Pooh and Eeyore

Setting: A room. There is a table in the corner of the room, and there is an apple and bananas on it. Eeyore is in the other corner of the room reading a book.
Winnie-the-Pooh: Hi, Eeyore. You are reading the book, I can see. What shall I do? – *Pooh walks around the room. He notices the apple.* - Oh, what a nice apple! I am going to eat it right now.
Eeyore: No, Pooh, you can't eat it, it's my apple.
Pooh: OK, I can't eat the apple, it's Eeyore's apple, I have to eat a banana then…
Pooh eats a banana.
(Some time later) *Eeyore*: OK, Pooh, I've read the book up, you can read the book now."
Winnie-the-Pooh starts reading the book. Eeyore walks across the room and comes up to the apple.
Eeyore: Here is my apple. I'm gonna eat it right now. *Eeyore takes the apple up to his mouth to eat it, but at this moment he stops.* I don't have to be such a greedy donkey! Pooh wants the apple and I'd rather give it to him. Myself, I can have a banana instead.
Eeyore pushes the table with the apple to Pooh who is reading the book.
Eeyore: Pooh, here is the apple and you can have it
Pooh: Oh, I'm such a happy bear! I have a book to read and an apple to eat! I am gonna read the book and eat the apple!
Final scene: Pooh keeps reading the book. There is the table with an apple leave next to him to remind the kids that it was Winnie-the Pooh who ate the apple.
Puppet: OK, it was a story about Eeyore and Winnie-the-Pooh. First Eeyore was reading the book and then Winnie-the-Pooh was reading the book. I know one thing that happened…

Poka	Poka on čital knigu, Vinni-Pux s'el jabloko. *While he was reading the book, Winnie-the-Pooh ate the apple.*
Principle C	On s'el jabloko, poka Vinni-Pux čital knigu. *He ate the apple while Winnie-the-Pooh was reading the book.*
Forwards anaphora embedded-first	Poka Vinni-Pux čital knigu, on s'el jabloko. *While Winnie-the-Pooh was reading the book, he ate the apple.*
Forwards anaphora main-first	Vinni-Pux s'el jabloko, poka on čital knigu. *Winnie-the-Pooh ate the apple while he was reading the book.*

2. Mickey Mouse (MM) and Donald Duck (DD)

Setting*: A room with a big red egg ('Egg of Luck') and a pot with a blossoming plant in it.*
MM: I know what this egg is – it's the Egg of Luck. Whoever is inside the egg can dream of anything he wants and the dream will come true.
DD: Really? Let me try then! DD gets inside the egg and starts dreaming aloud.
MM: In the meantime I really want to pick this flower for my mom – it's her birthday today and I would really like to give her this flower. *MM tries to pluck the flower.* Wow, the stem is quite thin, it turns out. I thought I could do it, but I guess I am not strong enough for it.

DD: Well, I think I had enough of it. MM, it's your turn to sit inside the Egg of Luck now.
MM gets inside the Egg, DD approaches the flower.
DD: What a beautiful flower, I can pluck it for my sister Minnie. I am sure I would be successful if I use my beak for it. – *DD reaches the flower to pick it, then pauses.* On the other hand, Minnie is not around this week – well, I'd leave it here for someone else.
MM: I am glad to be in here and to dream a little bit. Maybe I can get stronger so that I can pick that flower for my mom. Oh, I start feeling stronger, indeed – let me reach that flower. *MM reaches the flower from inside the egg and plucks it.* Wow, that's a miracle, I got it, and all due to the Egg of Luck!

Poka	Poka on sidel v Jajce Udači, Miki Maus sorval svetok.
	While he was inside the Egg of Luck, MM plucked the flower.
Principle C	On sorval svetok, poka Miki Maus sidel v Jajce Udači.
	He plucked the flower while MM was inside the Egg of Luck.
Forwards anaphora	Poka Miki Maus sidel v Jajce Udači, on sorval svetok.
embedded-first	*While MM was inside the Egg of Luck, he plucked the flower.*
Forwards anaphora	Miki Maus sorval svetok, poka on sidel v Jajce Udači.
main-first	*MM plucked the flower while he was inside the Egg of Luck.*

3. Wolf and Rabbit

Setting*: Rabbit is swinging on the swing. There is a pillar not far from the swing with a balloon tied to its top.*
Wolf: Aha, Rabbit is having fun there on the swing. I want to swing too. Rabbit, can I swing now?
Rabbit: Could you please wait a little bit, I would like to swing some more.
Wolf: That's ok, I am a nice wolf. What do I do then? Wow, look at that balloon up there – I want to pierce it. I'll need to jump high for it.
Wolf jumps as high as he can, but he cannot reach the balloon.
Rabbit: I am done now, Wolf, you can do some swinging now, and I am going to jump the rope now.
Wolf embarks on the swing and starts swinging.
Rabbit: now that Wolf if swinging, I can go ahead and pop up the balloon – I'd like to hear a loud noise from the burst! And I can jump high, so I'm sure I'll succeed. *The Rabbit jumps up, almost enough to pop up the balloon, but not quite enough.* OK, I quit my attempts here, I'd rather go play instead.
Wolf: Look, Rabbit. Now that I am swinging so high, I can reach the balloon. One, two, three – *the Wolf reaches the balloon and pops it up with his finger.*
***Puppet*:** It was a story about Rabbit and Wolf. First Rabbit was swinging on the swing and then Wolf did so. I know one thing that happened:

Poka	Poka on kačalsja na kačeljax, Volk prokolol šarik.
	While he was swinging on the swing, Wolf popped the balloon.
Principle C	On prokolol šarik, poka Volk kačalsja na kačeljax.
	He popped the balloon, while Wolf was swinging on the swing.
Forwards anaphora,	Poka Volk kačalsja na kačeljax, on prokolol šarik.
embedded first	*While Wolf was swinging on the swing, he popped a balloon.*
Forwards anaphora,	Volk prokolol šarik, poka on kačalsja na kačeljax.

main first	*Wolf popped the balloon, while he was swinging on the swing.*

4. Little Red Riding Hood and Princess

Setting: *Princess is riding a horse and meets Little Red Riding Hood. There is a big gift box with a bow on the ground.*
Little Red Riding Hood: Good morning, my Princess. What are you doing?
Princess: I'm riding my favorite horse today.
Little Red Riding Hood: (*notices the box on the sidewalk*) Oh, what a beautiful box! I am so curious to see what's inside! I'm gonna open it and see what's inside.
Princess: No, my dear, you can't touch it. It's my royal box and I prefer to open it myself later.
Little Red Riding Hood: OK, then I'd rather just put it on the stand over here so that it doesn't get dirty on the ground.
Little Red Riding Hood puts the box on the stand.
(some time later) Princess: OK, Little Red Riding Hood, you can ride my horse for a while if you want to.
Little Red Riding Hood: Sure, I do!
Princess gets off the horse and Little Red Riding Hood starts riding the horse.
Princess comes up to the box: OK, I'm gonna open the box now and see what's inside. First I have to detach the bow. *(takes off the bow)* Oh, but I'm a generous princess, amn't I? Little Red Riding Hood also wants to open it and be the first to see what's inside - I'd better let her do it."
Princess brings the box to Little Red Riding Hood and gives it to her.
Princess: Take the box, Little Red Riding Hood! You can open it and see what's inside.
Little Red Riding Hood: Oh, my Princess, you are so kind! Thank you for letting me open the box…. *(Little Red Riding Hood opens the box while riding the horse)* Oh, there is such a nice candy in here!
Final scene: Little Red Riding Hood is on the horse and holds the opened box in her hands.
Puppet: So, it was a story about Princess and Little Red Riding Hood. First Princess was riding the horse, and then Little Red Riding Hood was riding the horse. I know one thing that happened.

Poka	Poka ona katalas' na kone, Krasnaja Šapočka raskryla korobočku. *While she was riding the horse Little Red Riding Hood opened the box.*
Principle C	Ona raskryla korobočku, poka Krasnaja Šapočka katalas' na kone. *She opened the box while Little Red Riding Hood was riding the horse.*
Forwards anaphora embedded-first	Poka Krasnaja Šapočka katalas' na kone, ona raskryla korobočku. *While Little Red Riding Hood was riding the horse she opened the box.*
Forwards anaphora main-first	Krasnaja Šapočka raskryla korobočku, poka ona katalas' na kone. *Little Red Riding Hood opened the box while she was riding the horse.*

Appendix 4-B. Full list of experimental stimuli for Experiment 4b.

The experiment contained 24 'poka' sets of four conditions (the Principle C and poka-consditions) and 12 'before' sets of two conditions (no-constraint conditions). Conditions were as follows:

24 'poka'-sets:
Condition a: Principle C, gender-matched
Condition b: Principle C, gender-mismatched
Condition c: poka-constraint, gender-matched
Condition d: poka-constraint, gender-mismatched

12 'before' sets
Condition a: no-constraint, gender-matched
Condition b: no-constraint, gender-mismatched

1a. Поскольку перед Рождеством он сбывал с рук драгоценности, пока Марат пытался получить визы в Аргентину, Стас опасался вызвать подозрение у окружающих.
? Пытался ли Марат получить визы в Бразилию? N
1b. Поскольку перед Рождеством он сбывал с рук драгоценности, пока Алина пыталась получить визы в Аргентину, Стас опасался вызвать подозрение у окружающих.
? Пыталась ли Алина получить визы в Бразилию? N
1c. Поскольку перед Рождеством, пока он сбывал с рук драгоценности, Марат пытался получить визы в Аргентину, Стас опасался вызвать подозрение у окружающих.
? Пытался ли Марат получить визы в Бразилию? N
1d. Поскольку перед Рождеством, пока он сбывал с рук драгоценности, Алина пыталась получить визы в Аргентину, Стас опасался вызвать подозрение у окружающих.
? Пыталась ли Алина получить визы в Бразилию? N

2a. Поскольку уже в полдень он скрупулезно изучал внутреннюю структуру купола, пока Михаил делал карандашные зарисовки постройки, архитектор рассчитывал иметь всю необходимую информацию о храме к концу дня.
? Надеялся ли архитектор закончить сбор необходимой информации до конца дня? Y
2b. Поскольку уже в полдень он скрупулезно изучал внутреннюю структуру купола, пока Тамара делала карандашные зарисовки постройки, архитектор рассчитывал иметь всю необходимую информацию о храме к концу дня.
? Надеялся ли архитектор закончить сбор необходимой информации до конца дня? Y
2c. Поскольку уже в полдень, пока он скрупулезно изучал внутреннюю структуру купола, Михаил делал карандашные зарисовки постройки, архитектор рассчитывал иметь всю необходимую информацию о храме к концу дня.
? Надеялся ли архитектор закончить сбор необходимой информации до конца дня? Y
2d. Поскольку уже в полдень, пока он скрупулезно изучал внутреннюю структуру купола, Тамара делала карандашные зарисовки постройки, архитектор рассчитывал иметь всю необходимую информацию о храме к концу дня.
? Надеялся ли архитектор закончить сбор необходимой информации до конца дня? Y

3a. Поскольку перед началом спектакля она разговаривала по рации с охраной, пока Раиса рассаживала по местам высокопоставленных гостей, Инна не видела приключившегося с генералом казуса.
? Состоялся ли разговор с охраной после спектакля? N
3b. Поскольку перед началом спектакля она разговаривала по рации с охраной, пока Сергей рассаживал по местам высокопоставленных гостей, Инна не видела приключившегося с генералом казуса.
? Состоялся ли разговор с охраной после спектакля? N
3c. Поскольку перед началом спектакля, пока она разговаривала по рации с охраной, Раиса рассаживала по местам высокопоставленных гостей, Инна была очень благодарна ей за помощь.
? Состоялся ли разговор с охраной после спектакля? N
3d. Поскольку перед началом спектакля, пока она разговаривала по рации с охраной, Сергей рассаживал по местам высокопоставленных гостей, Инна была очень благодарна ему за помощь.
? Состоялся ли разговор с охраной после спектакля? N

4a. Поскольку перед эфиром она просматривала тексты сообщений, пока Марина гримировалась к началу сьемок, Зоя первой узнала сенсационную новость.
? Узнала ли Зоя сенсационную новость до начала эфира? Y
4b. Поскольку перед эфиром она просматривала тексты сообщений, пока Даниил гримировался к началу сьемок, Зоя первой узнала сенсационную новость.
? Узнала ли Зоя сенсационную новость до начала эфира? Y

4c. Поскольку перед эфиром, пока она просматривала тексты сообщений, Марина гримировалась к началу сьемок, Зоя сама определила порядок репортажей в выпуске.
? Был ли определен порядок новостей Зоей? Y
4d. Поскольку перед эфиром, пока она просматривала тексты сообщений, Даниил гримировался к началу сьемок, Зоя сама определила порядок репортажей в выпуске.
? Был ли определен порядок новостей Зоей? Y

5a. Поскольку вчера днем он подключал измерители давления, пока Артур настраивал специальную вакуумную камеру, Виталий был уверен в полной синхронизации установки.
? Сомневался ли Виталий в синхронизации установки? N
5b. Поскольку вчера днем он подключал измерители давления, пока Клара настраивала специальную вакуумную камеру, Виталий был уверен в полной синхронизации установки.
? Сомневался ли Виталий в синхронизации установки? N
5c. Поскольку вчера днем, пока он подключал измерители давления, Артур настраивал специальную вакуумную камеру, Валерий надеялся наладить оборудование до наступления сумерек.
? Потерял ли Валерий надежду настроить оборудование до темноты? N
5d. Поскольку вчера днем, пока он подключал измерители давления, Клара настраивала специальную вакуумную камеру, Валерий надеялся наладить оборудование до наступления сумерек.
? Потерял ли Валерий надежду настроить оборудование до темноты? N

6a. Поскольку всю неделю перед парным финалом он беспечно заигрывал с многочисленными поклонницами, пока Всеволод отрабатывал на корте ключевые удары, Дмитрий отыграл в финальном матче значительно хуже своего партнера.
? Играл ли Дмитрий в индивидуальном финале? N
6b. Поскольку всю неделю перед парным финалом он беспечно заигрывал с многочисленными поклонницами, пока Кристина отрабатывала на корте ключевые удары, Дмитрий отыграл в финальном матче значительно хуже своей партнерши.
? Играл ли Дмитрий в индивидуальном финале? N
6c. Поскольку всю неделю перед парным финалом, пока он беспечно заигрывал с многочисленными поклонницами, Всеволод отрабатывал на корте ключевые удары, Дмитрий не мог не признать вклад своего партнера решающим в их победе.
? Играл ли Дмитрий в индивидуальном финале? N
6d. Поскольку всю неделю перед парным финалом, пока он беспечно заигрывал с многочисленными поклонницами, Кристина отрабатывала на корте ключевые удары, Дмитрий не мог не признать вклад своей партнерши решающим в их победе.
? Играл ли Дмитрий в индивидуальном финале? N

7a. Поскольку после контрольной она просматривала ученические тетради, пока Ольга вводила в компьютер статистику по ошибкам, учительница могла сравнивать результаты между классами.
? Вводила ли Ольга статистику по ошибкам в компьютер? Y
7b. Поскольку после контрольной она просматривала ученические тетради, пока Борис вводил в компьютер статистику по ошибкам, учительница могла сравнивать результаты между классами.
? Вводил ли Борис статистику по ошибкам в компьютер? Y
7c. Поскольку после контрольной, пока она просматривала ученические тетради, Ольга вводила в компьютер статистику по ошибкам, учительница знала подробные результаты сразу по окончании проверки.
? Вводила ли Ольга статистику по ошибкам в компьютер? Y
7d. Поскольку после контрольной, пока она просматривала ученические тетради, Борис вводил в компьютер статистику по ошибкам, учительница знала подробные результаты сразу по окончании проверки.
? Вводил ли Борис статистику по ошибкам в компьютер? Y

8a. Хотя в прошлом году он впервые тестировал новый вид диеты, пока Ярослав готовился к ответственному этапу чемпионата мира, Глеб без опасений внедрил свои новшества в схему питания спортсмена.
? Испытывал ли Глеб опасения относительно новой диеты? N
8b. Хотя в прошлом году он впервые тестировал новый вид диеты, пока Варвара готовилась к ответственному этапу чемпионата мира, Глеб без опасений внедрил свои новшества в схему питания спортсменки.
? Испытывал ли Глеб опасения относительно новой диеты? N
8c. Хотя в прошлом году, пока он впервые тестировал новый вид диеты, Ярослав готовился к ответственному этапу чемпионата мира, Глеб без опасений внедрил свои новшества в схему питания спортсмена.
? Побоялся ли Глеб испытать новый вид диеты перед чемпионатом мира? N

8d. Хотя в прошлом году, пока он впервые тестировал новый вид диеты, Варвара готовилась к ответственному этапу чемпионата мира, Глеб без опасений внедрил свои новшества в схему питания спортсменки.
? Побоялся ли Глеб испытать новый вид диеты перед чемпионатом мира? N

9a. Хотя в январе она готовилась к квалификационным экзаменам, пока Алла отдыхала на Канарских островах, Вика не сетовала на судьбу.
? Отдыхала ли Алла на Канарах в августе? N
9b. Хотя в январе она готовилась к квалификационным экзаменам, пока Иван отдыхал на Канарских островах, Вика не сетовала на судьбу.
? Отдыхал ли Иван на Канарах в августе? N
9c. Хотя в январе, пока она готовилась к квалификационным экзаменам, Алла отдыхала на Канарских островах, Вика не сетовала на судьбу.
? Отдыхала ли Алла на Канарах в августе? N
9d. Хотя в январе, пока она готовилась к квалификационным экзаменам, Иван отдыхал на Канарских островах, Вика не сетовала на судьбу.
? Отдыхал ли Иван на Канарах в августе? N

10a. Хотя в воскресенье она учила билеты по физике, пока Варя слушала прямой радиорепортаж с `Евровидения`, Рита умудрялась не обращать внимание на происходящее.
? Транслировали ли `Евровидение` по радио в прямом эфире? Y
10b. Хотя в воскресенье она учила билеты по физике, пока Вова слушал прямой радиорепортаж с `Евровидения`, Рита умудрялась не обращать внимание на происходящее.
? Транслировали ли `Евровидение` по радио в прямом эфире? Y
10c. Хотя в воскресенье, пока она учила билеты по физике, Варя слушала прямой радиорепортаж с `Евровидения`, Рита умудрялась не обращать внимание на происходящее.
? Транслировали ли `Евровидение` по радио в прямом эфире? Y
10d. Хотя в воскресенье, пока она учила билеты по физике, Вова слушал прямой радиорепортаж с `Евровидения`, Рита умудрялась не обращать внимание на происходящее.
? Транслировали ли `Евровидение` по радио в прямом эфире? Y

11a. Хотя почти все утро он искал представителей редкого вида гусеницы, пока Егор снимал на камеру наиболее красивые уголки парка, Аркадий так и не нашел нужный ему экземпляр.
? Правда ли, что поиски гусеницы велись вечером? N
11b. Хотя почти все утро он искал представителей редкого вида гусеницы, пока Маша снимала на камеру наиболее красивые уголки парка, Аркадий так и не нашел нужный ему экземпляр.
? Правда ли, что поиски гусеницы велись вечером? N
11c. Хотя почти все утро, пока он искал экземпляр редкого вида гусеницы, Егор снимал на камеру наиболее красивые уголки парка, Аркадий переживал, что мальчик скучает.
? Правда ли, что поиски гусеницы велись вечером? N
11d. Хотя почти все утро, пока он искал экземпляр редкого вида гусеницы, Маша снимала на камеру наиболее красивые уголки парка, Аркадий переживал, что девушка скучает.
? Правда ли, что поиски гусеницы велись вечером? N

12a. Хотя каждое воскресенье она пила утренний кофе, пока Оля увлеченно решала какой-нибудь кроссворд, Марина никогда не предлагала помочь.
? Правда ли, что Марина никогда не предлагала своей помощи в решении кроссвордов? Y
12b. Хотя каждое воскресенье она пила утренний кофе, пока Юра увлеченно решал какой-нибудь кроссворд, Марина никогда не предлагала помочь.
? Правда ли, что Марина никогда не предлагала своей помощи в решении кроссвордов? Y
12c. Хотя каждое воскресенье, пока она пила утренний кофе, Оля увлеченно решала какой-нибудь кроссворд, Марина никогда не принимала участия в разгадке вопросов.
? Правда ли, что Марина никогда не принимала участия в разгадке кроссвордов? Y
12d. Хотя каждое воскресенье, пока она пила утренний кофе, Юра увлеченно решал какой-нибудь кроссворд, Марина никогда не принимала участия в разгадке вопросов.
? Правда ли, что Марина никогда не принимала участия в разгадке кроссвордов? Y

13a. Поскольку в магазине он болтал с симпатичными кондитершами, пока Максим пробовал разные виды пирожных, Дима получил огромное удовольствие от похода за покупками.
? Правда ли, что в магазине работали симпатичные кондитерши? Y

13b. Поскольку в магазине он болтал с симпатичными кондитершами, пока Наташа пробовала разные виды пирожных, Дима получил огромное удовольствие от похода за покупками.
? Правда ли, что в магазине работали симпатичные кондитерши? Y
13c. Поскольку в магазине, пока он болтал с симпатичными кондитершами, Максим пробовал разные виды пирожных, Дима захотел тоже попробовать что-нибудь из сладкого.
? Правда ли, что в магазине можно было попробовать выпечку? Y
13d. Поскольку в магазине, пока он болтал с симпатичными кондитершами, Наташа пробовала разные виды пирожных, Дима захотел тоже попробовать что-нибудь из сладкого.
? Правда ли, что в магазине можно было попробовать выпечку? Y

14a. Хотя уже в полшестого она украшала зеленью салаты, пока Галина спешно накрывала на стол, хозяйка так и не успела закончить приготовления до прихода гостей.
? Украшались ли салаты лимоном? N
14b. Хотя уже в полшестого она украшала зеленью салаты, пока Сережа спешно накрывал на стол, хозяйка так и не успела закончить приготовления до прихода гостей.
? Украшались ли салаты лимоном? N
14c. Хотя уже в полшестого, пока она украшала зеленью салаты, Галина спешно накрывала на стол, хозяйка не надеялась закончить приготовления до прихода гостей.
? Украшались ли салаты лимоном? N
14d. Хотя уже в полшестого, пока она украшала зеленью салаты, Сережа спешно накрывал на стол, хозяйка не надеялась закончить приготовления до прихода гостей.
? Украшались ли салаты лимоном? N

15a. Хотя всю субботу она полола разбушевавшиеся сорняки, пока Света срезала сухие ветки с кустов, Валентина, в отличие от подруги, совершенно не чувствовала усталости.
? Утомила ли Валентину работа в саду? N
15b. Хотя всю субботу она полола разбушевавшиеся сорняки, пока Вадим срезал сухие ветки с кустов, Валентина, в отличие от мужа, совершенно не чувствовала усталости.
? Утомила ли Валентину работа в саду? N
15c. Хотя всю субботу, пока она полола разбушевавшиеся сорняки, Света срезала сухие ветки с кустов, Валентина считала уборку сада полностью своей заслугой.
? Была ли Валентина объективна в своей оценке? N
15d. Хотя всю субботу, пока она полола разбушевавшиеся сорняки, Вадим срезал сухие ветки с кустов, Валентина считала уборку сада полностью своей заслугой.
? Была ли Валентина объективна в своей оценке? N

16a. Поскольку весь вечер она составляла квартальный отчет, пока Лида смотрела повтор `Рождественских встреч`, Марина считала себя вправе не готовить ужин.
? Смотрела ли Лида повтор `Рождественских встреч`? Y
16b. Поскольку весь вечер она составляла квартальный отчет, пока Вася смотрел повтор `Рождественских встреч`, Марина считала себя вправе не готовить ужин.
? Смотрел ли Вася повтор `Рождественских встреч`? Y
16c. Поскольку весь вечер, пока она составляла квартальный отчет, Лида смотрела повтор `Рождественских встреч`, Марина не могла полностью сосредоточиться на работе.
? Смотрела ли Лида повтор `Рождественских встреч`? Y
16d. Поскольку весь вечер, пока она составляла квартальный отчет, Вася смотрел повтор `Рождественских встреч`, Марина не могла полностью сосредоточиться на работе.
? Смотрел ли Вася повтор `Рождественских встреч`? Y

17a. Хотя в ресторане она курила на террассе, пока Алина обсуждала с официантом прелести японской кухни, Галя слышала все детали их разговора.
? Обсуждала ли Алина с официантом итальянскую кухню? N
17b. Хотя в ресторане она курила на террассе, пока Борис обсуждал с официантом прелести японской кухни, Галя слышала все детали их разговора.
? Обсуждал ли Борис с официантом итальянскую кухню? N
17c. Хотя в ресторане, пока она курила на террассе, Алина обсуждала с официантом прелести японской кухни, Галя настояла на заказе исключительно европейских блюд.
? Обсуждала ли Алина с официантом итальянскую кухню? N
17d. Хотя в ресторане, пока она курила на террассе, Борис обсуждал с официантом прелести японской кухни, Галя настояла на заказе исключительно европейских блюд.

? Обсуждал ли Борис с официантом итальянскую кухню? N

18a. Поскольку в самолете она расшивала блестками костюм, пока Алена заучивала наизусть слова финальной песни, Жанна старалась не отвлекать дочь своими просьбами и расспросами.
? Правда ли, что Жанна старалась не беспокоить дочь? Y
18b. Поскольку в самолете она расшивала блестками костюм, пока Гриша заучивал наизусть слова финальной песни, Жанна старалась не отвлекать сына своими просьбами и расспросами.
? Правда ли, что Жанна старалась не беспокоить сына? Y
18c. Поскольку в самолете, пока она расшивала блестками костюм, Алена заучивала наизусть слова финальной песни, Жанна невольно выучила текст вместе с дочерью.
? Выучила ли Жанна текст финальной песни? Y
18d. Поскольку в самолете, пока она расшивала блестками костюм, Гриша заучивал наизусть слова финальной песни, Жанна невольно выучила текст вместе с сыном.
? Выучила ли Жанна текст финальной песни? Y

19 a
Поскольку все утро он пылесосил ковры в лоджии, пока Денис обзванивал друзей относительно предстоящей встречи, Семен ни капли не сожалел о решении встретиться у них дома.
? Были ли подвергнуты чистке ковры в лоджии? Y
19b. Поскольку все утро он пылесосил ковры в лоджии, пока Ксюша обзванивала друзей относительно предстоящей встречи, Семен ни капли не сожалел о решении встретиться у них дома.
? Были ли подвергнуты чистке ковры в лоджии? Y
19c. Поскольку все утро, пока он пылесосил ковры в лоджии, Денис обзванивал друзей относительно предстоящей встречи, Семен плотно закрыл все двери и старался не шуметь.
? Были ли подвергнуты чистке ковры в лоджии? Y
19d. Поскольку все утро, пока он пылесосил ковры в лоджии, Ксюша обзванивала друзей относительно предстоящей встречи, Семен плотно закрыл все двери и старался не шуметь.
? Были ли подвергнуты чистке ковры в лоджии? Y

20a. Хотя весь день он выписывал основные цитаты, пока Илья писал назначенный на четверг доклад, Кирилл не хотел вникать в суть его работы.
? Вникал ли Кирилл в суть доклада? N
20b. Хотя весь день он выписывал основные цитаты, пока Лина писала назначенный на четверг доклад, Кирилл не хотел вникать в суть её работы.
? Вникал ли Кирилл в суть доклада? N
20c. Хотя весь день, пока он выписывал основные цитаты, Илья писал назначенный на четверг доклад, Кирилл не хотел вникать в суть его работы.
? Вникал ли Кирилл в суть доклада? N
20d. Хотя весь день, пока он выписывал основные цитаты, Лина писала назначенный на четверг доклад, Кирилл не хотел вникать в суть её работы.
? Вникал ли Кирилл в суть доклада? N

21a. Хотя два дня назад он обрабатывал собранные ранее данные, пока Игорь проводил повторный контрольный эксперимент, Геннадий не захотел делиться с ним результатами.
? Правда ли, что контрольный эксперимент проводился повторно? Y
21b. Хотя два дня назад он обрабатывал собранные ранее данные, пока Света проводила повторный контрольный эксперимент, Геннадий не захотел делиться с ней результатами.
? Правда ли, что контрольный эксперимент проводился повторно? Y
21c. Хотя два дня назад, пока он обрабатывал собранные ранее данные, Игорь проводил повторный контрольный эксперимент, Геннадий не захотел делиться с ним результатами.
? Правда ли, что контрольный эксперимент проводился повторно? Y
21d. Хотя два дня назад, пока он обрабатывал собранные ранее данные, Света проводила повторный контрольный эксперимент, Геннадий не захотел делиться с ней результатами.
? Правда ли, что контрольный эксперимент проводился повторно? Y

22a. Поскольку в прошлую пятницу он проветривал рабочее помещение, пока Яков распечатывал копии праздничных рекламных листовок, Костя винил себя в простуде коллеги.
? Правда ли, что Яков распечатывал копии бухгалтерских форм? N
22b. Поскольку в прошлую пятницу он проветривал рабочее помещение, пока Элла распечатывала копии праздничных рекламных листовок, Костя винил себя в простуде коллеги.

? Правда ли, что Элла распечатывала копии бухгалтерских форм? N
22c. Поскольку в прошлую пятницу, пока он проветривал рабочее помещение, Яков распечатывал копии праздничных рекламных листовок, Костя наконец познакомился с новым сотрудником.
? Правда ли, что Элла распечатывала копии бухгалтерских форм? N
22d. Поскольку в прошлую пятницу, пока он проветривал рабочее помещение, Элла распечатывала копии праздничных рекламных листовок, Костя наконец познакомился с новой сотрудницей.
? Правда ли, что Элла распечатывала копии бухгалтерских форм? N

23a. Хотя с утра он делал уборку в квартире, пока Коля слушал новый альбом `Алисы`, Иван не стал упрекать его ни в чем.
? Правда ли, что Иван ни в чем не упрекнул Колю? Y
23b. Хотя с утра он делал уборку в квартире, пока Зина слушала новый альбом `Алисы`, Иван не стал упрекать её ни в чем.
? Правда ли, что Иван ни в чем не упрекнул Зину? Y
23c. Хотя с утра, пока он делал уборку в квартире, Коля слушал новый альбом `Алисы`, Иван не стал упрекать его ни в чем.
? Правда ли, что Иван ни в чем не упрекнул Колю? Y
23d. Хотя с утра, пока он делал уборку в квартире, Зина слушала новый альбом `Алисы`, Иван не стал упрекать её ни в чем.
? Правда ли, что Иван ни в чем не упрекнул Зину? Y

24a. Хотя в парке она уныло грызла тыквенные семечки, пока Вера читала объявления на стенде, Юля не хотела торопить подругу.
? Правда ли, что Вера читала объявления в журнале? N
24b. Хотя в парке она уныло грызла тыквенные семечки, пока Витя читал объявления на стенде, Юля не хотела торопить друга.
? Правда ли, что Витя читал объявления в журнале? N
24c. Хотя в парке, пока она уныло грызла тыквенные семечки, Вера читала объявления на стенде, Юля не захотела присоединиться к ней.
? Правда ли, что Вера читала объявления в журнале? N
24d. Хотя в парке, пока она уныло грызла тыквенные семечки, Витя читал объявления на стенде, Юля не захотела присоединиться к нему.
? Правда ли, что Витя читал объявления в журнале? N

1a. Хотя до того как она открыла собственную фирму, Анна хорошо зарабатывала переводами в Лукойле, Дима не сомневался в правильности её решения.
? Была ли Анна недовольна зарплатой в Лукойле? N
1b. Хотя до того как она открыла собственную фирму, Дима хорошо зарабатывал переводами в Лукойле, Анна не сомневалась в необходимости перехода мужа в семейный бизнес.
? Был ли Дима недоволен зарплатой в Лукойле? N

2a. Хотя после того как он закончил генеральную репетицию, Владимир был полностью доволен спектаклем, Кристина не разделял его энтузиазма.
? Различались ли мнения Кристины и Владимира в отношении спектакля? Y
2b. Хотя после того как он закончил генеральную репетицию, Кристина была полностью довольна спектаклем, Владимир не разделял её энтузиазма.
? Различались ли мнения Кристины и Владимира в отношении спектакля? Y

3a. Хотя после того как она написала заказанную статью, Валентина несколько раз правила текст, Александр все-таки был недоволен ее подачей материала.
? Правила ли Валентина первоначальный текст статьи? Y
3b. Хотя после того как она написала заказанную статью, Александр несколько раз правил текст, Валентина больше всего гордилась своим первоначальным вариантом.
? Правил ли Александр первоначальный текст статьи? Y

4a. Хотя до того как он опубликовал обзор в Литературной Газете, Гавриил был неизвестен в писательских кругах, Надежда не сомневалась в появлении интереса к нему после выхода обзора.
? Был ли Гавриил широко известен писателям до публикации в Литературной Газете? N
4b. Хотя до того как он опубликовал обзор в Литературной Газете, Надежда была неизвестна в писательских кругах, Гавриил не сомневался в появлении интереса к ней после выхода его обзора.
? Была ли Надежда широко известна писателям до публикации в Литературной Газете? N

5a. Хотя после того как он выиграл зимние соревнования, Павел недвусмысленно намекнул на участие в Олимпиаде, Настя отказалась от любых комментариев относительно своего подопечного.
? Намекнул ли Павел на участие в Олимпиаде? Y
5b. Хотя после того как он выиграл зимние соревнования, Настя недвусмысленно намекнула на участие в Олимпиаде, Павел отказался от любых комментариев относительно её высказывания.
? Намекнула ли Настя на участие в Олимпиаде? Y

6a. Хотя до того как она вошла в состав сборной, Наташа была совершенно неизвестна специалистам, Михаил хорошо понимал значимость ее успеха.
? Была ли Наташа знаменитой? N
6b. Хотя до того как она вошла в состав сборной, Михаил был совершенно неизвестен специалистам, Наташа никогда не сомневалась в его выдающихся тренерских способностях.
? Был ли Михаил знаменитым? N

7a. Поскольку после того как он переехал работать на север, Юрий стал получать анонимные угрозы, Катя всеми силами пыталась уговорить его вернуться в деревню.
? Пыталась ли Катя уговорить Юрия уехать с севера? Y
7b. Поскольку после того как он переехал работать на север, Катя стала получать анонимные угрозы, Юрий решил забрать сестру к себе в Норильск.
? Планировал ли Юрий забрать к себе сестру? Y

8a. Поскольку до того как он не получил печать, Дима не подписывал никаких бумаг, Нина была вынуждена отложить заключение контрактов.
? Правда ли, что все контракты были подписаны без отлагательств? N
8b. Поскольку до того как он не получил печать, Нина не подписывала никаких бумаг, Дима был вынужден отложить переговоры о заключении контракта.
? Правда ли, что переговоры были проведены в изначально планируемое время? N

9a. Поскольку до того как она не разобралась во всех документах, Оксана не проводила переговоров с клиентами, Никита не мог сразу оценить ее деловые качества.
? Требовалось ли Оксане время на ознакомление с документами? Y
9b. Поскольку до того как она не разобралась во всех документах, Никита не проводил переговоров с клиентами, Оксана не могла рассчитывать на премию.
? Верно ли, что Оксана не надеялась на премию? Y

10a. Поскольку до того как он устроился в охранное агенство, Николай постоянно пересказывал прочитанные детективы, Лариса считала, что новая работа придется ему по душе.
? Правда ли, что Лариса устроилась на работу в охранное агенство? N
10b. Поскольку до того как он устроился в охранное агенство, Лариса постоянно пересказывала прочитанные детективы, Николай надеялся сразить подругу захватывающими случаями из своей практики.
? Правда ли, что Лариса устроилась на работу в охранное агенство? N

11a. Поскольку после того как она перешла на другую работу, Даша стала забирать детей из садика, Олег был очень рад появившемуся у него свободному времени.
? Забирала ли Даша детей из сада до того как она перешла на новую работу? N
11b. Поскольку после того как она перешла на другую работу, Олег стал забирать детей из садика, Даша была очень рада появившемуся у неё свободному времени.
? Была ли Даша огорчена появившемуся свободному времени? N

12a. Поскольку до того как она начала заниматься латиноамериканскими танцами, Ирина профессионально занималась классическим балетом, Антон сразу же по достоинству оценил будущую партнершу.
? Правда ли, что Ирина занималась балетом? Y
12b. Поскольку до того как она начала заниматься латиноамериканскими танцами, Антон профессионально занимался классическим балетом, Ирина с большим трудом уговорила его стать ее партнером.
? Правда ли, что Антон занимался балетом? Y

5 CHAPTER 5. PARSING OF BINDING & CONTROL STRUCTURES

In Chapters 3 and 4 I argued based on the results of a series of sentence reading experiments that upon encountering a referentially dependent element, e.g. a pronoun, the parser starts an active search for its antecedent. I also argued that the active search by the parser avoids false alarms: it skips positions that would violate the Binding Principles. Support for this claim came from the results of sentence-processing experiments in which we found a GMME only at antecedents that were in a binding-accessible position for the pronoun, such as the 2nd subject position in (1), but not in positions that were binding-inaccessible due to Principle C, such as the 2nd subject position in (2).

Developing a parsing algorithm that would be capable of reproducing these findings is one of the goals of this chapter.

(1) After he moved to Holland, Jose/Leticia quit smoking.
[no-constraint, GMME]

(2) He quit smoking, after Jose/Leticia moved to Holland.
[Principle C, null-effect]

The second goal, which can be viewed as a development of the one above, is to extend the algorithm to encompass another type of dependency, known as the *control* dependency, exemplified in (3).

(3) Leticia persuaded Jose to quit smoking.

In (3) the unpronounced subject of the embedded clause (the quitter) must be understood as *Jose*. In other words in (3) there is a referential dependency between a silent embedded subject (dubbed as PRO since Chomsky (1981)) and the object of the main clause. In general terms, the relevance in control is due to it being another type of referential dependency. More specifically, some control structures, e.g. (4), bear a striking resemblance to structures with binding dependencies such as (1), and were shown to be processed similarly by the human parser (Sturt, Lombardo & Betancort 2004).

(4) After moving to Holland, Jose hopes to quit smoking.

It is capturing this parsing similarity, as well as handling other control structures, e.g. (3), that I will concentrate on. The parsing algorithm that I develop is based on a unified movement-based account of binding and control by Hornstein (1999, 2001) and covers the parser's processing of reflexives, bound pronouns and PROs.

The chapter is organized as follows. Section 5.1 presents the background on a minimalist parser SPARSE (Schneider 1999) that I will adopt as a model of incremental sentence structure building. I will review the process of incremental structure building for simple sentences and extend the model to handle adjunct

clauses, which will lay the groundwork for subsequent investigation of sentences that involve referential dependencies. Section 5.2 'teaches' SPARSE to deal with binding dependencies, and section 5.3 extends the parser's coverage to cases of control dependencies. The algorithm that I develop for handling these dependencies borrows its main insights from Hornstein's (1999, 2001) theory of binding and control, e.g. reflexives and PROs are considered to be residues of movement from their position to a position of their antecedent/controller. The search procedure that is used to derive which positions are licit targets for such movement (in other words, which nominals are licit antecedents for the dependent element) is, at least partially, independently motivated by phenomena that are unrelated to binding and control (e.g. strictly incremental incorporation of incoming material into the structure, ambiguity resolution considerations, etc). Section 5.4 concludes.

5.1 SPARSE (SCHNEIDER 1999)

5.1.1 SPARSE Basics

To generate syntactic structures incrementally I will use the feature-based parser SPARSE proposed in Schneider (1999). SPARSE builds bare-phrase structure (Chomsky 1995) and hence is similar to Stabler's Minimalist parser (Stabler 1997). SPARSE builds sentence structures in a bottom-up manner, by expanding the tree as each word is encountered in the input. SPARSE maintains only the information in a single tree and does not keep information about alternatives and choice points (as in backtracking parsers and the variant thereof in Inoue & Fodor 1995) or about dispreferred parses (Gibson 1991). When incoming words cannot be directly added to the existing tree the structure is revised through a process of repair.

One of the key properties of SPARSE is that it is not forced to make a prediction for a specific upcoming syntactic category. This is done to minimize the amount of reanalysis needed in head-final languages, in which it is risky to predict the following syntactic category, e.g. a case-marked NP may be licensed by either verb or a postposition. Thus, the minimal building block accessible to the parser is the syntactic feature, e.g. [Case: Acc].

In Schneider (1999) the parser is allowed to predict structure only if the parser cannot build a fully connected tree otherwise. Predicted features are therefore those which are required to license the current word into the structure. However, we have argued that when the parser encounters a dependent element that requires an antecedent, it actually does make a forward prediction even when such prediction is not mandated in order to incorporate the current word in the tree. The aim of the following sections is to extend the SPARSE model to extend to

such cases of 'unforced prediction' for an antecedent in sentences with binding and control dependencies. In modeling these cases I will rely on the experimental results from chapters 3 and 4.

I will first demonstrate the main features of SPARSE using the example in (5). The presentation here closely follows the discussion in chapter 2.5.1 in Schneider (1999).[97]

(5) Dorothy will see munchkins. (from Schneider 1999)

Upon encountering the first word in the sentence, *Dorothy*, the parser does not need to do anything beyond activating the representation of that word in the lexicon.

(6) *Dorothy*:
[Cat: Noun], [Case: Nom, Acc], [Num: Sg], [Person: 3]

The next word, *will,* is then encountered. Its lexical entry contains at least the following contents:

(7) *will*:

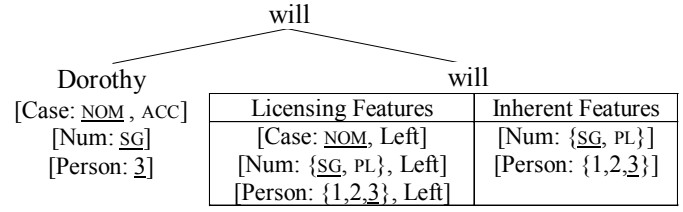
Inherent Features	Licensing Features
[Cat: T]	[Case: Nom, Left] [Num: {Sg, Pl}, Left] [Person: {1,2,3}, Left] [Cat: V, Right] [VForm: Infin, Right]

(7) illustrates two types of features that can be stored in the lexical entry for a word. The inherent features for an item include a type of the feature (e.g., Case, Cat(egory), etc.) and the value of the feature (e.g., Nom or Acc for the feature Case). Inherent features specify the grammatical category of a word, its morphological features (number, person, case, gender), etc. Licensing features define which type of elements a given word may find to the right or left of itself in the structure and thus allow a head to combine with another head. Licensing features represent relations like case assignment, complement selection or theta-role assignment, e.g. [Case: Nom, Left] on a certain item requires it to have a Nominative noun to its left.

[97] I will limit the discussion only to the most relevant points. In the area of NP/DP I will assume Abney's DP-hypothesis, i.e. [$_{DP}$ D [$_{NP}$ ..N..]]. I will assume that every time the parser encounters a noun that is not preceded by a determiner, it projects a null-D head. As a shorthand, I will represent bare nouns as D's. See also fn. 105 for a similar point.

When the parser encounters a new word, its licensing features are inspected to determine whether they can license the previous word, *Dorothy*. *Dorothy* is attached to *will* as a result of licensing the [Case: Nom] feature of *Dorothy* by the [Case: Nom, Left] feature of *will*. The label for the new projection is inherited from the licensor.[98]

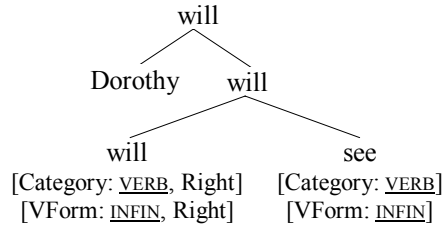
(8)



When a head that has a certain licensing feature is combined with a head that has an inherent feature of the same type, there must be a non-null intersection of the values in the two sets. This is how head-complement or head-specifier agreement is reached. For example, in (8) the value '3rd person' is chosen from the set of person features for *will* as a result of an unambiguous person feature on the noun *Dorothy*. (8) lists the features of *Dorothy* and *will* that are evaluated at the current step. Underlined values in the feature sets are those that have been checked and specified by intersecting the feature sets of the combined heads.

Upon encountering the next word, *see*, the parser examines its features to determine whether it can be attached to the previous word, *will*. *See* contains the inherent features [Cat: verb] and [VForm: Infin] and therefore can be licensed as the complement of *will*, as shown in (9):

(9)

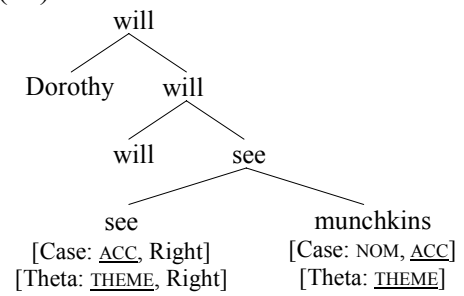


The final word, *munchkins*, has the same sets of features as *Dorothy* in (6), with the exception of the number feature which is plural. Its inherent feature [Case: Nom, Acc] is intersected with the licensing feature [Case: Acc, Right] on the preceding verb *see* and *munchkins* is linked as the complement of *see*.[99]

[98] I will also occasionally use the category of the licensor, rather than the word itself, as a label for convenience.

[99] In what follows I will abstract away from the details of how each thematic role of a verb is checked (Schneider (1999) does not discuss it either). The following example demonstrates how the Agent role can be checked against the verb in *Dorothy will see munchkins*. The verb *see* has

(10)



In the following few sections I will review how adjuncts are handled in SPARSE. In section 5.1.2 I review the basic idea for how adjuncts are incorporated into the structure. I then expand the current coverage of SPARSE that is limited to single post-posed adverbs to cases of preposed adverbs (sections 5.1.3) and preposed adjunct clauses (section 5.1.4). The motivation for looking into the details of how preposed adjuncts are handled is because it is a necessary step towards one of our final goals – modeling how sentences with dependent elements inside an adjunct clause, e.g. (1), are processed.

5.1.2 Adjuncts in SPARSE

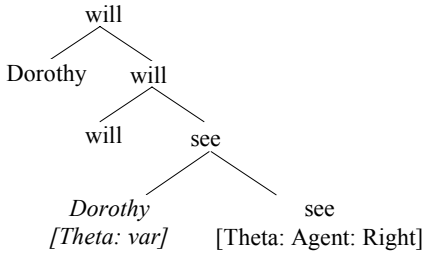
Schneider (1999) adopts the assumption from Categorial Grammar (Steedman 1996, 2000) according to which an adjunct selects its host. This information can be represented most easily for words in the lexicon that are always adjuncts, e.g. adverbs.

(11) *tomorrow*: *[Cat: V, Left]

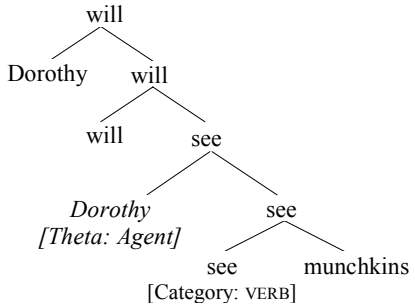
---

the licensing features [Theta: Agent, Left] and [Theta: Theme, Right] in its repository. The theme feature of the verb *see* is checked off via an intersection with the variable that is the value of the theta feature for *munchkins,* which is assigned the value *theme* and marked as checked. Theta-checking of the agent role of *see* is not explicitly discussed. Since features are checked off strictly locally, I suggest that it requires copying *Dorothy* into a lower position after the verb is encountered, as in (i). The final structure is given in (ii), with lower (a.k.a. unpronounced) copies italicized.

(i)



(ii)



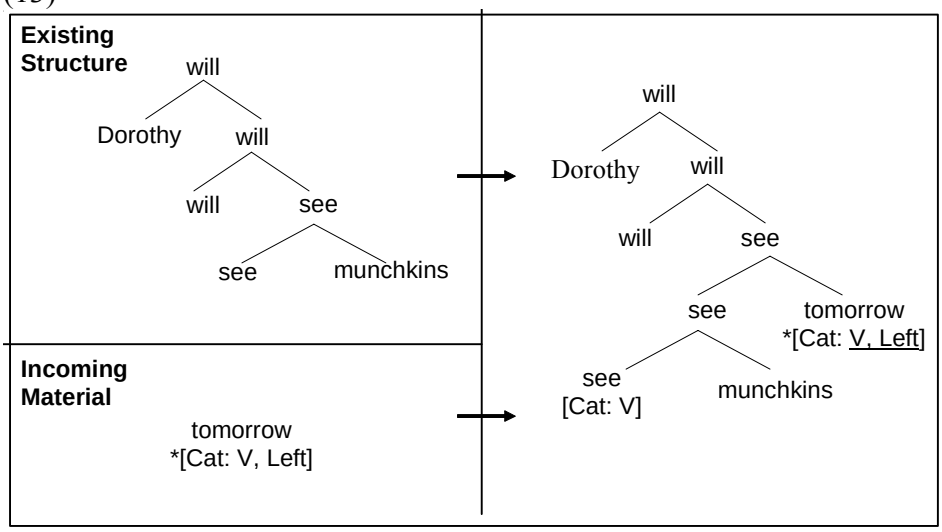
For example, the lexical entry for the adverb *tomorrow* contains a licensing feature that selects a category V to its left (as in (11)). The asterisk on the licensing feature [Cat: V, Left] indicates that the lexical item does not project, i.e. a projection resulting from checking off the feature *[Cat: V, Left] on the adverb must be headed by the category that *tomorrow* adjoins to, i.e. V.

Let us illustrate how adverbs are parsed using (12).

(12) Dorothy will see munchkins tomorrow.

(12) is different from (5) above in that it contains the sentence final adverb *tomorrow* that modifies the verb phrase. By the time when *tomorrow* is encountered the parser has built the structure in (10) for the preceding words. Unlike all cases we have seen so far, *tomorrow* cannot be attached to the most recently processed word. When a new word cannot be added to the structure based on its licensing features, the parser searches the tree upwards along the right spine of the tree for a possible licensor (details of how the search proceeds are given in section 5.2). The next constituent that is encountered is headed by *see*, so the parser attempts to compare the features of *tomorrow* against the features on *see.* As a result, the licensing feature *[Cat: Verb, left] of *tomorrow* enables it to combine with *see*. *See* projects up to head the new constituent, as shown in (13).

(13)



5.1.3 Preposed Adverbs[100]

Schneider (1999) does not discuss other cases of adjunct attachment beyond post-posed adverbs. In what follows I will attempt to see which additional assumptions are needed to handle fronted adjuncts and clausal adjuncts. I will start by considering a simple modification of the case in (12), shown in (14). In going through this case I will also discuss some additional features that were not part of Schneider (1999).

(14) Tomorrow Carmen will dance flamenco.

(14) is almost identical to (12) in all relevant respects, with the difference that the adverb *tomorrow* is fronted. This change leads to another difference: unlike (12) in which each new word could be directly added to the previously built structure, incremental processing of (14) requires projecting pieces of structure beyond what is directly supplied by the input.[101]

Let us show word-by-word structure building for (14).

→ *tomorrow*

The licensing feature for *tomorrow* in Schneider (1999) (repeated in (15)a) will not suffice to handle (14), since no verb can occur to the left of *tomorrow*. (15)a then needs to be expanded to (15)b, whereby *tomorrow* can be licensed either by a verb to its left or by a Tense-head or verb to its right.

(15) *tomorrow*:
a. *[Cat: V, Left]
b. *[Cat: {[V, Left], [T, Right]}]

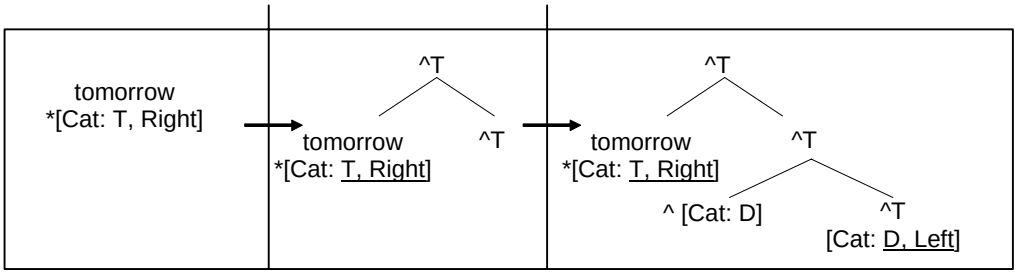
Since *tomorrow* is the first word in the input, the licensing feature [V, Left] is not applicable and is discarded. A T-head is then projected based on the remaining licensing feature *[Cat: T, Right]. Moreover, in English TP's must have an overtly realized specifier (Extended Projection Principle (EPP), Chomsky 1981). I will assume that this is the basis that enables a real or predicted T-head to predict

[100] I will use terms 'preposed' and 'post-posed' with respect to adjuncts in a non technical sense, simply to indicate their surface position in the sentence. I will not discuss adjunct reconstruction into a post-verbal position here for reasons of time and space. Should one attempt to incorporate reconstruction into the model, the biggest part of the task would be the formulation of conditions on obligatory reconstruction of adjuncts, e.g. whether it interacts with binding considerations, avoidance of Principle C violations, etc.

[101] The example below shows that the sentence initial adverb is indeed immediately connected into the main tree, since it can form a constituent with the following NP, as in *Tomorrow Carmen and on Tuesday Kitri will dance flamenco.*

a subject by projecting its distinguished licensing feature [Cat: D, Left] early on.[102]

This step is different from the original version of SPARSE in two respects. First, Schneider (1999) used [Case: Nom, Left] as a distinguished feature for T-heads. I will assume that this feature is [Cat: D, Left]. Indeed, not every T-head is capable of projecting nominative case, e.g. nominative case is not licensed in non-finite clauses by *–ing* and *to*. Changing the distinguished licensing feature for T-heads to [Cat: D, Left] enables uniformity in projecting the subject position in finite and non-finite clauses and represents a direct parallel of the EPP feature in English postulated in many theories.[103] Moreover, the subject position is projected as soon as there is a real or a predicted T-head added to the structure. In SPARSE, on the other hand, no heads were predicted unless they were necessary to incorporate a current word in to the tree. This modification will be critical for explaining the results of the self-paced reading experiments presented in Chapters 3 & 4.



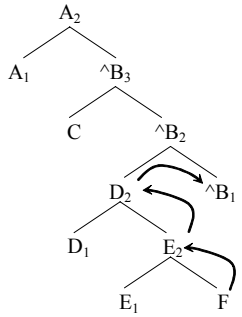
→ *Carmen*

Carmen cannot be attached directly to the preceding word, *tomorrow*. The parser therefore searches the tree for a position for *Carmen*, and finds no available position.[104] The parser then conducts a search in the lexicon for a predicted head

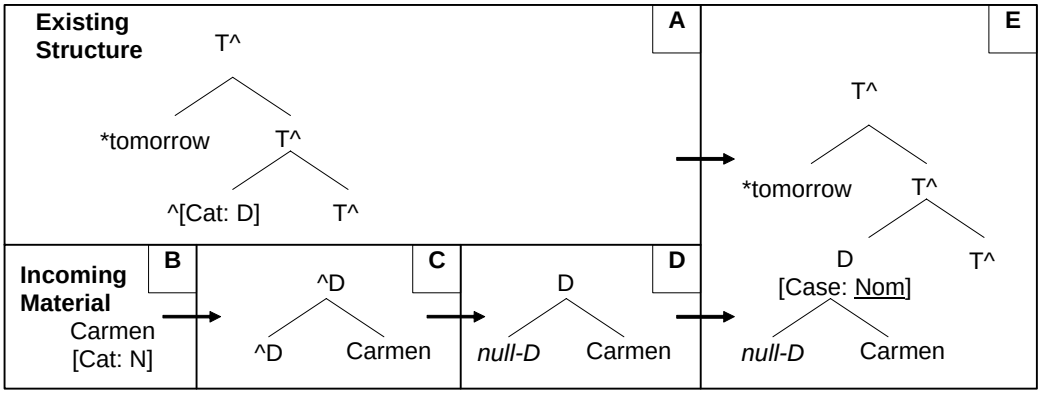
[102] Predicted heads must eventually be instantiated with by a real head (either a head from the input or a null-head from the lexicon). A predicted head is instantiated by a real head that is compatible with all features of the predicted head.

[103] The Extended Projection Principle (EPP) was first proposed in Chomsky (1981) under the name 'Principle p' as a "structural requirement that certain configurations … must have subjects" [Chomsky 1981: 27].

[104] The search for a position for the current word begins only if it cannot be attached directly to the previous word and proceeds along the right edge of the tree, as shown in the diagram below.



that can license *Carmen*, which returns a D-head (panel C). Moreover, the D-head can be instantiated by a null determiner which is assumed to be among the lexical entries in the lexicon (panel D). [I am in no way committed to the order of steps in panels C and D]. The resulting D-head can straightforwardly substitute into ^[Cat: D]; its Case feature is set to Nominative.[105]

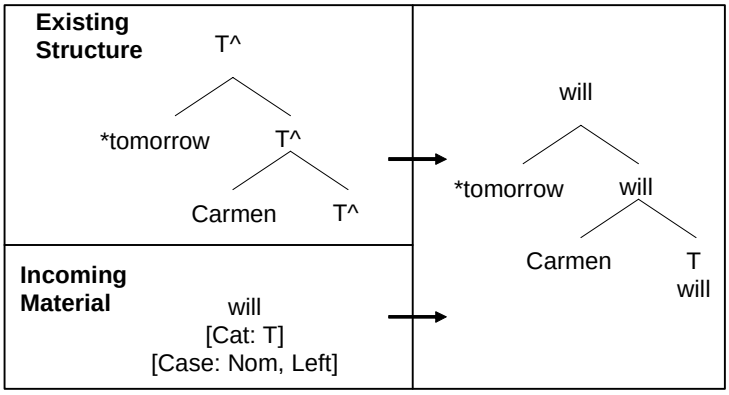


→ *will*

Will has an inherent feature [Cat: T]. It therefore can take a spot to the right of the DP *null-D Carmen* in place of the predicted head ^T. *Will* checks off nominative case on the DP *null-D Carmen*.

"The search of the tree stops when any of the following types of heads is reached: the root of the tree, a predicted head, or a head/constituent still missing a required element on its right (e.g. an obligatorily transitive verb that has not yet received its complement, which would be represented by an unchecked "right"-pointing feature). If any such head is reached before an attachment site is found, the search fails. The search must stop at the root for obvious reasons—there is nothing else to search. It must stop upon reaching a predicted head because the predicted head must be instantiated with a real head. If the incoming word is attached to the right of the predicted head, there is no way for another incoming word to instantiate the predicted head without violating the linear order of the sentence. Likewise, if an obligatory complement is passed up, there is no way to attach a complement later in the derivation without violating the linear order of the input sentence." [Schneider 1999: 54]. We will return to the details of the search procedure in section 5.2.

[105] In what follows for reasons of time I will skip the details of how DP's are formed based on a bare noun and will take the liberty of straightforwardly replacing a predicted head ^[Cat: D] with either a determiner or a bare noun, which I will understand as a shorthand for the procedure spelled out in this example. I will assume that the parser adds the null D to the left of the name as soon as it realizes that the noun can not be directly attached to the preceding word.



The rest of the derivation will proceed as in (5).

To recap, the crucial fact about adjuncts is that they must be marked as such in the lexicon. This can be done most straightforwardly when an adjunct is stored in the lexicon as a whole, as in cases of single adverbs above. In the next sections we will implement how adjunct PPs are assembled, e.g. *in the garden*, *at five p.m* or *before Dan answered the question*.[106] Such PPs are not stored pre-coined in the lexicon, and therefore the whole phrase cannot be marked by the adjunct asterisk. A common sense solution is that a preposition that selects a DP or TP complement also contains an adjunct marking in their lexical entry, which enables attachment of the PP to the main tree. I will go through the details of this solution in the next section.

5.1.4 Preposed Adjunct Clauses

In this section I examine sentences with preposed adjunct clauses, such as (16) & (17). The motivation for spelling out the details of the procedure for building adjunct clauses is a necessary step on the way to our ultimate goal, i.e. modeling referential dependencies in which one of the members of the dependency is inside an adjunct.

(16) After the breakfast John kissed Mary.
(17) After the breakfast ended John kissed Mary.

As mentioned above, I will assume that the adjunct clause in (16) & (17) is a PP. Lexical entries for a few prepositions that can take a DP or a TP as their complement are given in (18) & (19). *While* is different from *after/before* in that it cannot take a DP complement.[107]

[106] I will assume that adjuncts like *before Dan answered the question* are PPs, although nothing specific is tied to this choice.
[107] In what follows I will assume Abney's DP hypothesis, i.e. [$_{DP}$ D [$_{NP}$ …N…]] and a lexicon that in addition to overt D's, e.g. *a* or *the*, also contains a null D-head.

(18) *after/before*[108]

Inherent Features	Licensing Features
[Cat: P]	[Cat: {[T, Right], [D, Right]}] *(embedded clause)* *[Cat: {[V, Left], [T, Right]}] *(main clause)*

(19) *while*

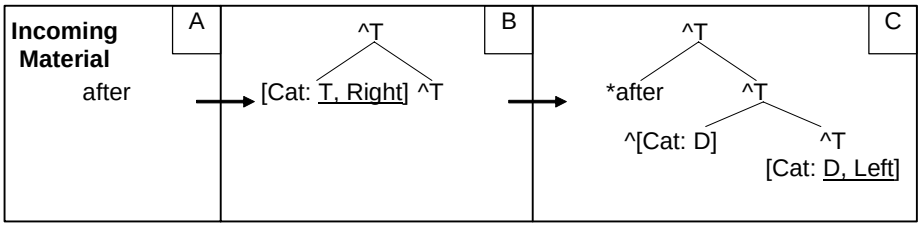
Inherent Features	Licensing Features
[Cat: P]	[Cat: T, Right] *(embedded clause)* *[Cat: {V, Left}, {T, Right}] *(main clause)*

Let us start by going through the derivation of (16).

(16) After the breakfast John kissed Mary.

→ *after*: *[Cat: V, Left] is excluded, projects *[Cat: T, Right], which in turn projects ^[Cat: D] to its left.

There are two things going on once *after* is extracted from the lexicon. First it projects *[Cat: T, Right], i.e. the T-head of the main clause. This is possible because this is the only licensing feature left in the relevant set. Second, ^T-head right away projects a subject position based on its licensing feature [Cat: D, Left] (the EPP feature).



Projecting heads at this point is a divergence from the original formulation of SPARSE model (Schneider 1999) in which the parser never made any predictions unless they were absolutely mandatory to keep incremental processing of the sentence. Here the predictions are made more freely, but only to the extent that they are unambiguous based on the lexical information from the element. Thus, it retains the spirit of SPARSE in the sense that the model avoids making unwarranted predictions.

To summarize this step, *after* projects *[Cat: T, Right], which in turn projects ^[Cat: D] to its left. Hence, the main subject position is projected immediately after *after* is processed and long before it appears bottom-up. This step will be

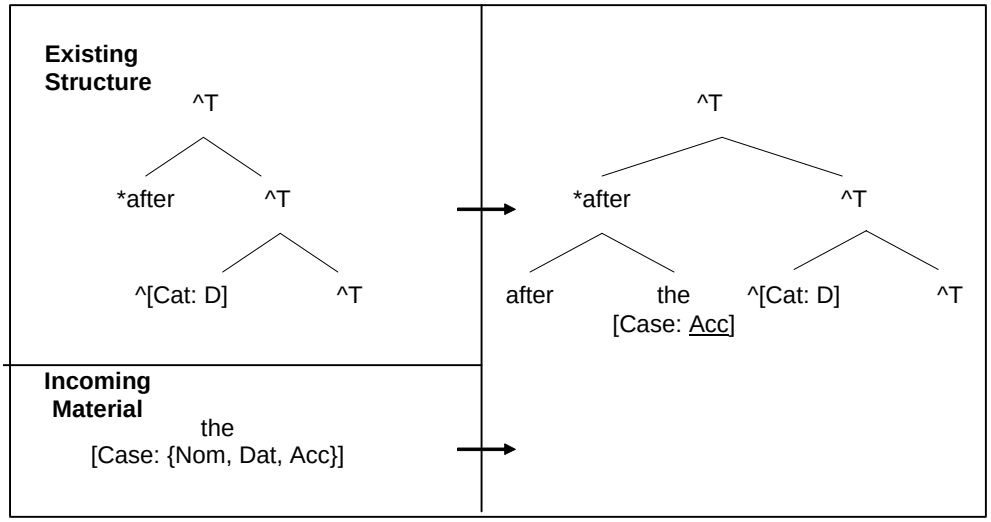
[108] Curly brackets are used to denote sets of alternative mutually-exclusive features, e.g. [Cat: {[T, Right], [D, Right]}] can either license a T on the right, or a D on the Right.

critical later in explaining the results of the self-paced reading experiments that we presented in Chapters 3 & 4, e.g. why the parser anticipates an antecedent for the cataphoric pronoun in sentences like (1) in the main subject position.

→ *the*

Inherent Features	Licensing Features
[Cat: D] [Case: {Nom, Dat, Acc}]	[Cat: N, Right] (noun to the Right)

The is added to the structure as the complement of *after*, i.e. the [Cat: D, Right] feature of *after* is checked off.[109]

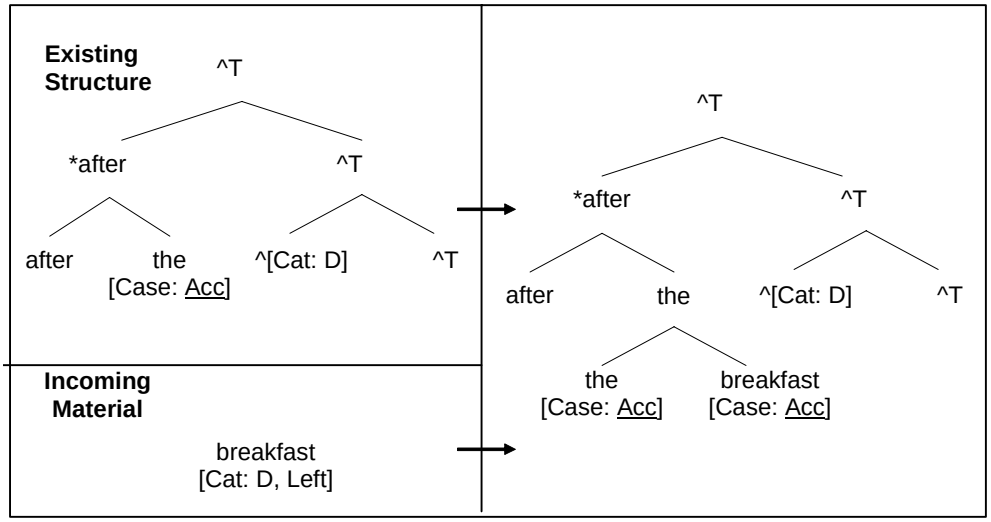


→ *breakfast*

Breakfast can be added to the previous word *the* based on the licensing feature [Cat: N, Right] of the determiner *the*.

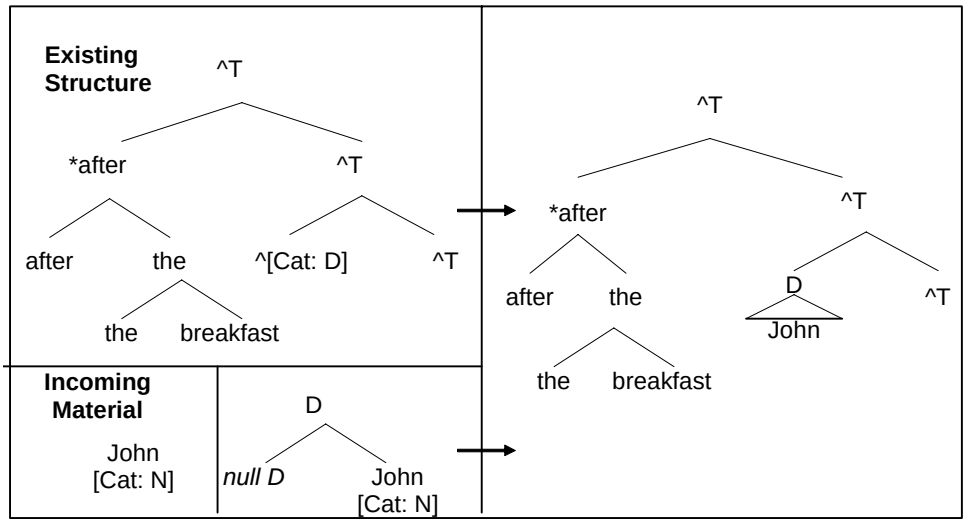
Let us note here that whenever two heads enter into a licensing relation, each of the inherent features that they have in common must be matched on its value. At this step, for example, the Case feature of *breakfast* is set to the value [Case: Acc] in accordance with the case on the determiner.

[109] According to the SPARSE algorithm (see Appendix 5-A) the parser initially always attempts to attach the new word to the previously processed word, and only then starts searching the tree for other possibilities. This is the reason why *the* attaches to *after*, rather than replaces the predicted ^[Cat: D] head in the main subject position.



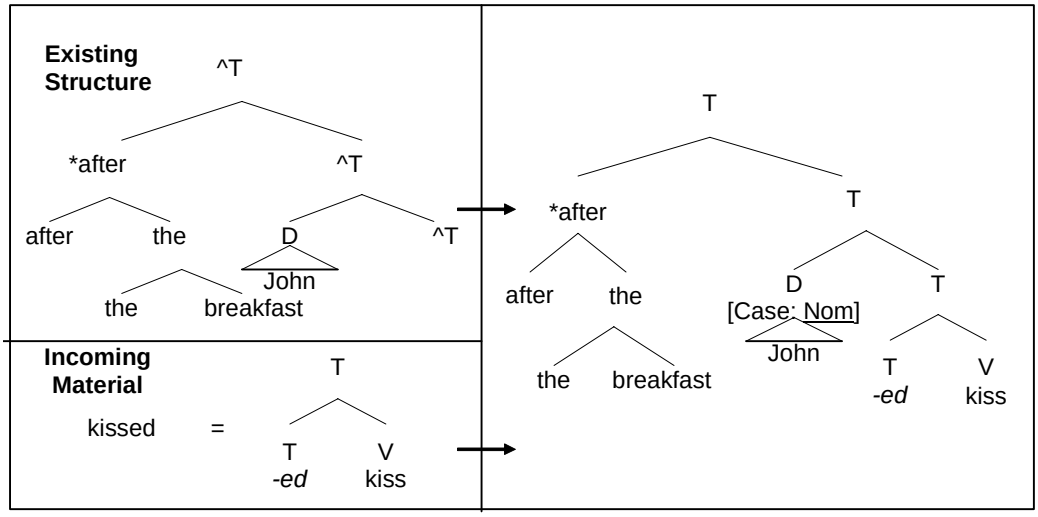
→ *John*

John cannot be attached to the preceding word, *breakfast*. The null D-head is then added to the left of the noun (see fn. 105 on page 199). A D-complex can be substituted into the predicted head ^[Cat: D].



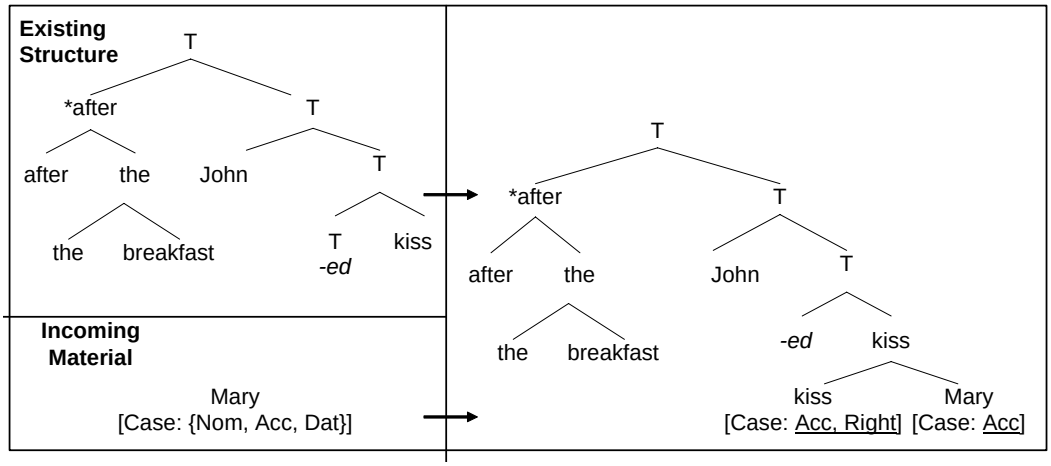
→ *kissed*

Schneider (1999) assumes that inflected verbs are morphologically complex and are represented in the lexicon as a V-T complex, as shown below. The upper T-head of this complex can therefore attach to the right of the previous word, *John*, and substitute the predicted ^T-head. The DP in the subject position receives nominative case from the *-ed* T-head.



→ *Mary*: [Case: {Nom, Dat, Acc}]

Mary is attached as a complement of *kiss*. The accusative case and the patient theta-role of *kiss* is checked off.



Next, let's consider the derivation for (17), repeated below for convenience.

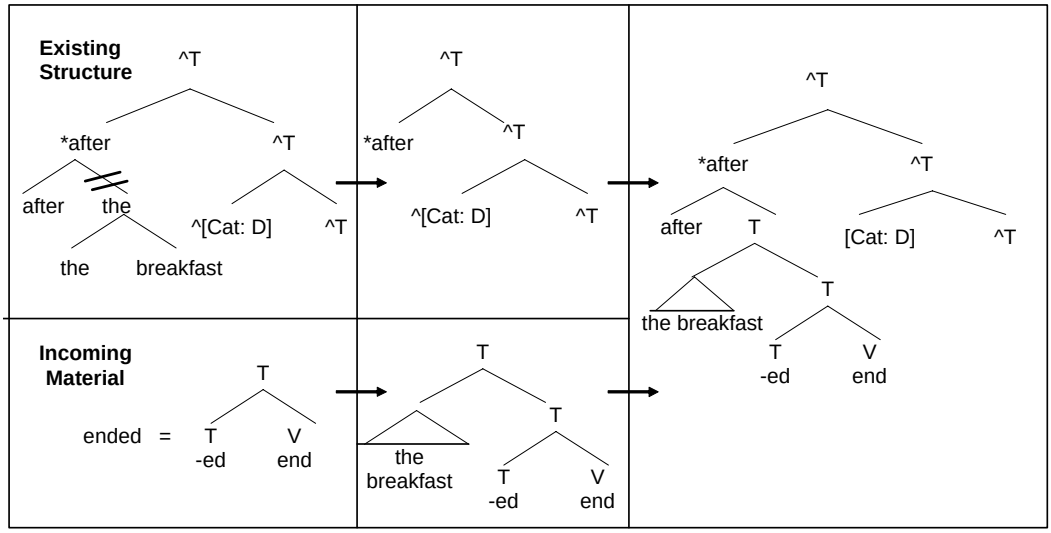
(17) After the breakfast ended, John kissed Mary.

→ *After* + *the* + *breakfast*

These first three words will be processed identically to (16) above.

→ *ended*

The verb *ended* is again a V-T complex. This complex cannot be attached directly to the DP *the breakfast* in the existing tree[110], since the DP has received accusative case from the preposition *after*. The parser therefore searches the right edge of the tree for another attachment site. Note that the T of the V-T complex cannot substitute into the ^T head in the main tree, since then the [Cat: D] feature of T would remain unchecked (actually, accorsing to the SPARSE algorithm, the matrix ^T-head will not be on the search path of the parser, since it is to the right of another predicted head).[111]



Since there is no licit attachment site for the verb, the parser is required to reanalyze the previously built structure. The V-T complex has a licensing feature [Cat: D, Left] which could be satisfied by the DP *the breakfast*. The DP is therefore disconnected from the tree and attached as the subject of *ended*.

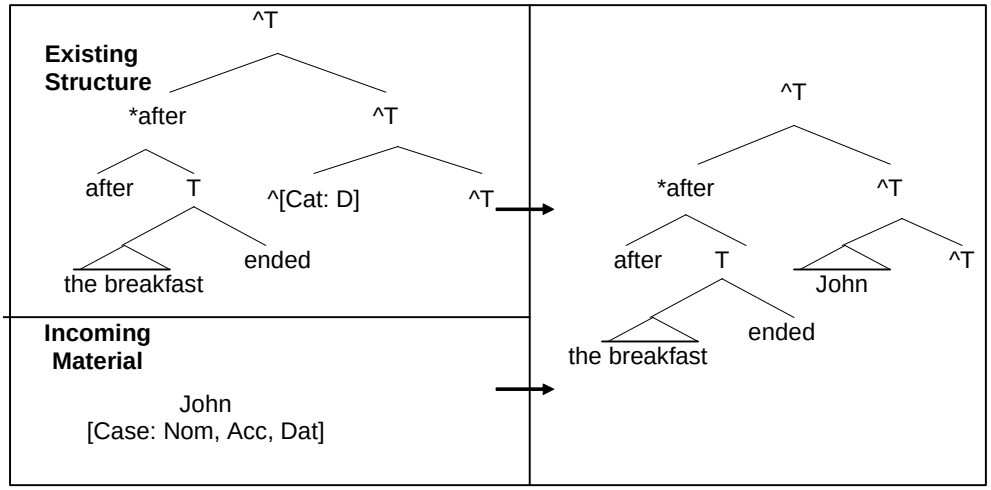
→ *John*

John cannot attach to *ended* and has to search along the right edge of the tree for a position.[112] The predicted head ^[Cat: D] is found and replaced by *John*.

[110] This step assumes that once D and N are put together, the DP is closed off and treated as a single element by the parser. The consequences of this assumption need to be explored, e.g. in cases of relative clauses.

[111] Reanalysis may not be needed if semantics disambiguates beforehand, e.g. in *After John…* the NP *John* is more likely to be the subject of a TP rather than a DP complement of *after*.

[112] I assume here that the intransitive version of the verb *ended* was chosen on the basis of the semantic properties of the subject *the breakfast* before *John* was encountered in the input.



The rest of the derivation proceeds straightforwardly as described in previous examples.

5.2 BINDING IN SPARSE

Having developed the details of how 'simple' adjunct clauses can be handled we are ready to turn to more challenging cases that are of direct interest to us, e.g. sentences like (1) & (2) that involve adjunct clauses and also contain a referentially dependent element. We would like to show which steps during the structure building for (1) & (2) are crucial in explaining the parser's reaction to the 2nd subject in either case (i.e. GMME in (1) and a null-effect in (2)).

In order to deal with sentences like (1) & (2) that contain referential dependencies, we need to 'explain' the relevant principles to the parser, i.e. to encode the mechanism that is used in order to establish a referential dependency and constraints on these dependencies. To do so, I will borrow insights from Hornstein's (1998, 2001) theory, in which binding dependencies are considered to be formed via movement. After reviewing the core ideas of Hornstein's theory, I discuss how his insights may be re-cast in terms of procedures that are independently needed in parsing (e.g. search procedures).

5.2.1 Movement Approach to Binding

Hornstein (1999, 2001) presents a movement-based approach to binding in the Minimalist framework.[113] Binding phenomena are explained via a basic operation of movement (MOVE). This claim is inherently intertwined with the idea that thematic roles are features that are checked off by NPs and that a single NP can check more than one thematic role. Reflexives are NP-traces that are related to their antecedents via A-movement. For example, in (20) a complex reflexive *Alice-self* first checks off the patient role of the verb, and then its part (*Alice*) moves into the subject position to check off the agent role of the verb.

(20) Alice_i pinched herself_i (to make sure she was not dreaming).
Derivation: [$_{TP}$ Alice T [$_{VP}$ [pinched [[Alice]self]]]]

The biggest appeal of Hornstein's analysis is its parsimony and reduction in the number of different types of nominals. Various nominal categories that are traditionally considered to be regulated by separate modules – reflexives, bound pronouns, obligatory and non-obligatory control PRO, null pronoun *pro* and NP-traces – are all united by the notions of movement and economy. All these categories are viewed as grammatical formatives (rather than lexical elements) which are defined by the structural position that they appear in (in terms of features that can be checked off in a given position); those that occur in similar positions are in competition with one another. Traces are considered to be the default type of nominals and inserted into derivations as first choice. Categories that are not traces of movement, e.g. bound pronouns, are inserted as last resort when their less costly competitors that are residues of movement, e.g. reflexives, are illicit.

Viewing binding as movement allows Hornstein to dispense with Principles A and B of the Binding Theory. These principles now directly follow from the definition of a reflexive or a bound pronoun. Principle A follows from the facts that (i) reflexives are traces of A-movement that obeys locality and (ii) reflexives cannot occur in the subject position (see (21)i). Pronouns obey Principle B (i.e. they are locally free) because they are inserted as last-resort whenever a reflexive is illicit. In other words, an antecedent for a pronoun cannot occur in a position that is attainable via movement, or else a reflexive would have been used in place of the pronoun.

[113] In this section the review of Hornstein's account will be rather minimal and will only cover the part that directly concerns binding. Further discussion of the movement approach and how it applies to other referential dependencies, e.g. control, is given in section 5.3.1.

5.2.2 Algorithm for Building Binding Dependencies[114]

The algorithm that we will develop for the parser to build binding dependencies relies on the principles listed in (21) which largely follow Hornstein (1999, 2001):

(21) The parser's knowledge base:
- a Traces of A-movement can check off thematic roles;
- b A nominal position is [+Theta] or [+Case];
- c In English [Case] must be checked by an overt NP;
- d [Spec TP] must be filled;
- e A-movement is local, i.e. into the closest A-position (and is not limited to c-commanding positions);
- f A trace is the default nominal category (whenever applicable);
- g Reflexives are NP-traces;
- h Bound pronouns are inserted as a last resort into positions from which movement is illicit;
- i Reflexives cannot occur in positions construed with agreement (Woolford 1999);
- j Principle C: a dependent element must not c-command its antecedent.

The statements in (21)a-h were all mentioned briefly in one or another way in the previous section. (21)i requires some explanation. Based on an extensive cross-language survey Woolford (1999) reports that anaphora do not occur in positions where there is agreement, thus supporting an earlier observation by Rizzi (1990). In English a position that is construed with agreement is the subject (due to subject-verb agreement). Hornstein (2001) derives the generalization from the requirement on feature checking by the head of the morphologically complex reflexive, e.g. *him+self*. In this work, however, I will consider reflexives as morphologically simplex and will use this fact as given in order to explain why reflexives cannot occur in the subject position.

One of the important parts of instantiating (21) is the notion of movement and the fact that it must be local. This is achieved in the algorithm by implementing movement as parasitic on the parser's search of the tree, which in turn is local. In doing so I will make use of the search procedure that is independently motivated by other phenomena that are unrelated to binding. This procedure was developed by Schneider (1999) for dealing with phenomena such as incremental incorporation of a new word into the structure, low vs. high attachment of adjuncts, etc. Below I review Schneider's search procedure that was designed to

[114] A more appropriate name for this section is 'A tentative procedure for building binding dependencies'. Developing a formal algorithm based on the proposal here is postponed for another occasion.

accommodate a new head into the tree and complement it with another part that does not aim at accommodating any new heads.

According to Schneider's (1999) algorithm (see also fn. 104 on page 198) the search starts at the most recently processed word (*F* in Figure 23) and proceeds up along the right edge of the tree, as shown by arrows in Figure 23. The search abides locality, i.e. skipping heads along the right edge of the tree is not allowed.

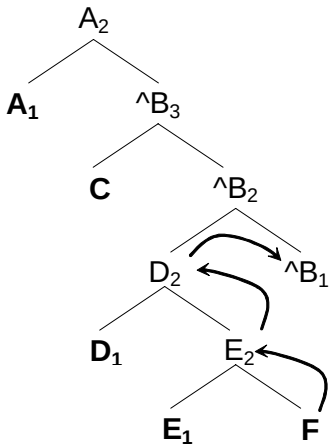


Figure 23. Search along the right edge of a (hypothetical) tree. Terminal nodes that are filled by lexical material are in bold; F is the current word that was added to the structure at the latest step. The search stops at the first predicted head B_1.

Schneider's 'right edge' algorithm was designed to look for a position for a new word. Whether the new word is attached by adding a new node or by filling out a yet unfilled predicted head, it had to always be added to the right of the most recently incorporated word due to considerations of linear word order. In other words, the parser deliberately avoids searching nodes that already exist in the tree and are to the right of the most recently incorporated word.

Quite on the contrary, the existing terminal nodes must be on the primary search path when the parser looks for an antecedent for a dependent element, since the antecedent may either linearly precede or follow the dependent element. Thus, the 'right edge' search procedure must be complemented by a 'left edge' search that goes through the positions in the tree that precede the current element. The search from the current word upwards along the left edge of the tree (i.e. searching heads that are to the left of the main spine of the tree and attached to the spine) will give the desired outcome, as shown in Figure 24a.[115] The fact that only nodes that are attached to the main spine of the tree, but not nodes below them, are visible to the parser has the effect that among all preceding nominal position only those that c-command the current element are on the search path. This is additionally illustrated in Figure 24b.

[115] The main spine is constituted by a sequence of T and V nodes down from the main root node.

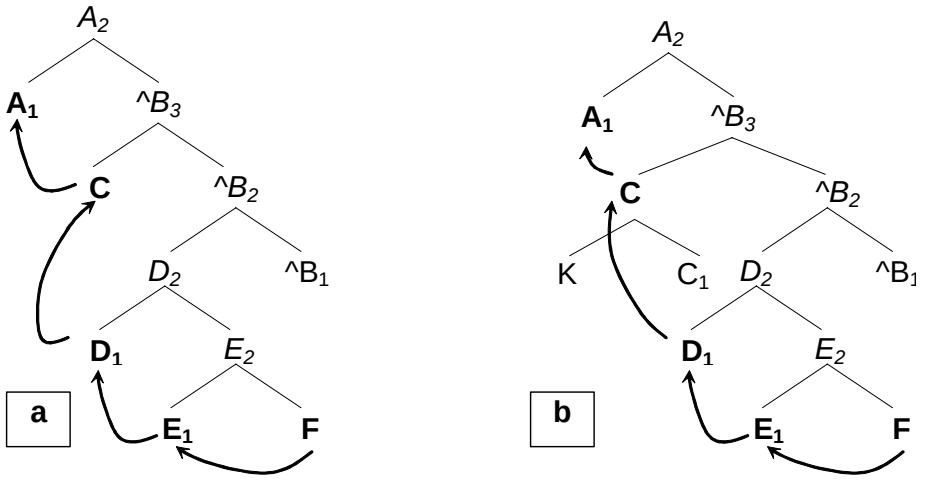


Figure 24. Search along the left edge of a (hypothetical) tree. The nodes constituting the main spine of the tree are italicized.

Having outlined the basic search procedure, we can now discuss it in relation to binding. As mentioned above, the operation of movement is defined for the parser as a procedure of searching for the antecedent of a trace, which is held in accordance with the search procedure outlined above. In order to account for the preference for forwards vs. backwards anaphora we add an additional condition whereby the parser searches the left edge of the tree before the right edge.

Thus, whenever the parser categorizes a dependent element as a trace of movement (e.g., a reflexive), it initiates a search for its antecedent along the left edge of the tree until the first potential antecedent is found. This nominal head is appointed as the antecedent for the trace. If an antecedent was not found along the left edge, e.g. when the dependent element precedes all other nominals in the sentence, the parser continues its search along the right edge of the tree.

The search for an antecedent for a dependent element that is not a trace, e.g. a bound pronoun, is different in two respects. First, during the left and right edge search the parser excludes the position that is most local, i.e. the first head returned by the search, since this position is a prerogative of an antecedent for a trace.[116] Second, unlike the case of reflexives, the antecedent for the pronoun need not c-command the pronoun, as demonstrated in (22).

However, a search procedure that only probes heads that directly attach to the main spine, as suggested above, will miss positions that do not c-command the current element.

[116] Unless the pronoun occurs in the position in which reflexives are prohibited altogether in accordance with (21)i, i.e. the subject position in English. See the derivation for (41) in section 5.3.4.3.

(22) Becky$_i$'s dog licked her$_i$ in the cheek.

Hence, the parser must be able to search inside branches (or *sub-trees*) that attach to the main spine whenever it looks for an antecedent for a pronoun. For now, let us simply say that the parser searches sub-trees whenever it looks for an antecedent for a pronoun. For purposes of concreteness, I will assume that in searching a tree with sub-trees the parser first goes through the standard left and right edge search procedure and only then searches inside sub-trees; a discussion of how viable is this assumption will need to be postponed until a different occasion.[117]

Finally, let us discuss the status of the Binding Principles in the parser. As we mentioned earlier, in the movement approach to binding Principles A and B do not need to be postulated separately and follow from conditions d-h in (21). The remaining principle of the binding theory, Principle C, is preserved. For our purpose it suffices to state Principle C as in (21)j, i.e. as a requirement that a dependent element cannot c-command its antecedent. C-command is defined as in Chapter 2 (repeated below with insignificant modifications).[118]

(23) A c-commands B iff:
(i) any head that dominates A dominates B *and*
(ii) A does not dominate B.

If Principle C is a constraint on structure building, as a strong interpretation of our results would suggest, the parser should automatically exclude every position that is in the c-command domain of the dependent element. In the right-branching structure that we use here the c-command domain for the head that is the last word added to the structure is easy to define: it includes (at least) all heads that are subsequently incorporated into the structure by expanding the bottom of the tree. Thus, if the parser fails to find an antecedent at the step when the dependent element is incorporated into the structure, it re-checks the structure for a possible antecedent for the dependent element only when the tree gains by a new real or predicted head that does not expand the bottom of the tree.

[117] This is stated for cases where the search-triggering head is along the main spine of the tree. More discussion of sub-trees follows in section 5.3.
[118] See Hornstein (2001) for a more principled explanation of Principle C and its unbounded scope.

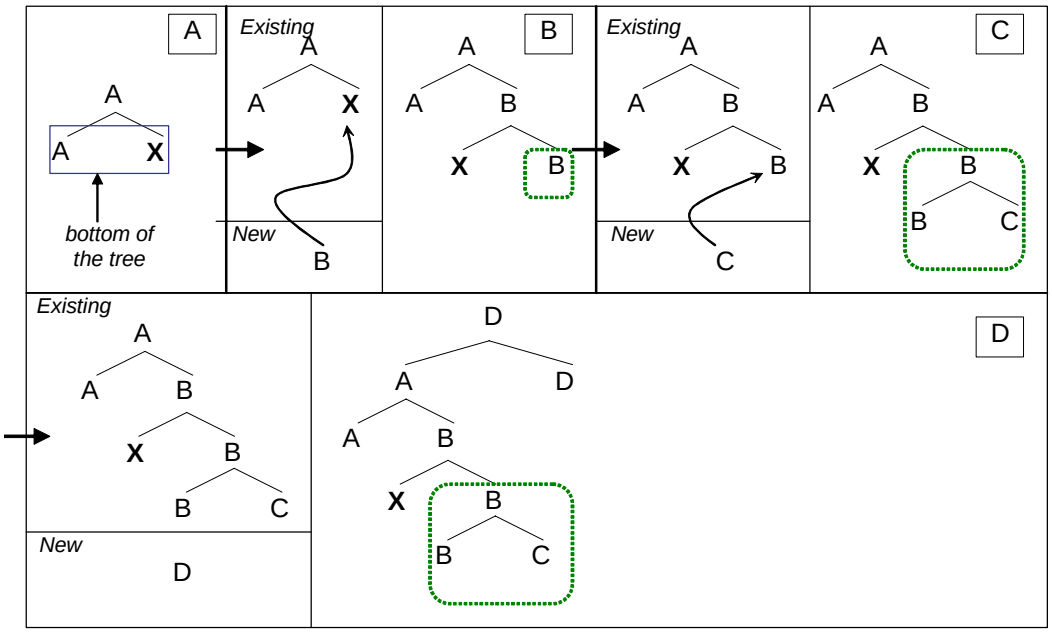


Figure 25. C-command domain for the element X.

In Figure 25, for example, if X is a dependent element, the parser fails to find an antecedent for X once it is added to the tree (panel A). It then keeps building the structure by adding new heads. Heads B and C are added by expanding the bottom of the tree, thus are in the c-command domain for X (panels B & C); the parser will not re-check for an antecedent for X at these steps. Head D does not expand the bottom of the tree, and thus the parser re-checks for an antecedent for X at this step (panel D). This procedure will be further exemplified in section 5.2.3.3.

Summarizing the proposal, target sites for movement and the locality restrictions on movement are encoded as part of the search procedure by the parser. In the following sections our proposal will be exemplified using various example sentences that involve binding dependencies.

5.2.3 Case Studies – Part 1

5.2.3.1 Forwards Anaphora, Reflexives

We will start with a case of forwards anaphora like (24).

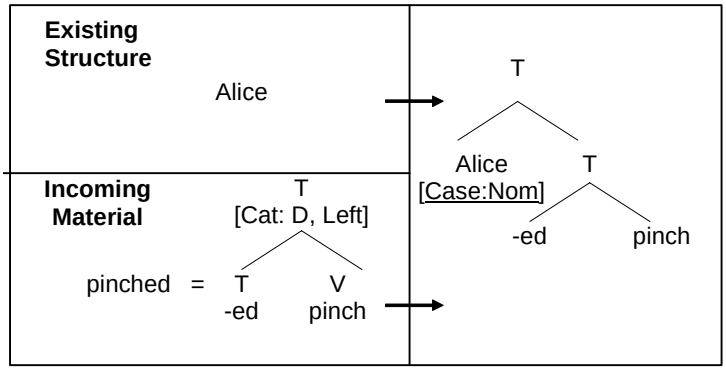
(24) Alice pinched herself.

→ *Alice*

Alice is extracted from the lexicon with (at least) the following features: [Cat: Noun], [Case: {Nom, Acc, Dat}], [Num: Sg], [Person: 3]

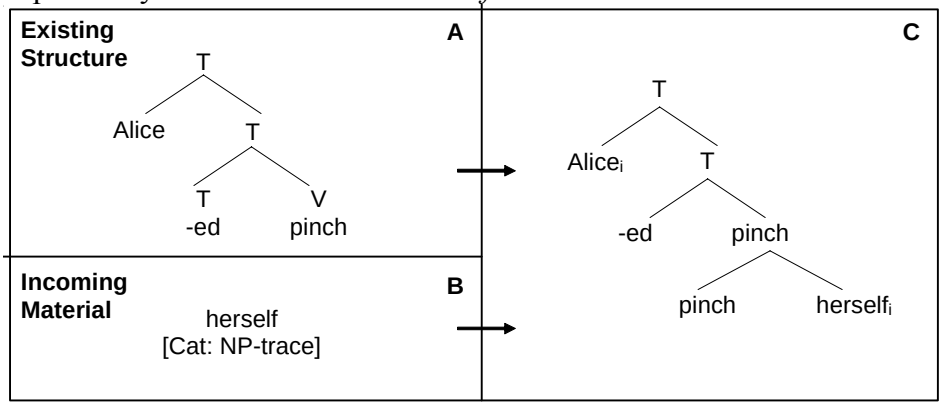
→ *pinched*

The verb *pinched* is represented as a V-T complex. The upper T-head of this complex has a licensing feature [Cat: D, Left] which it uses to attach to the right of the previous word, *Alice.* The Case feature of Alice receives the value Nominative. The agent role of *pinch* is checked off.



→ *herself*

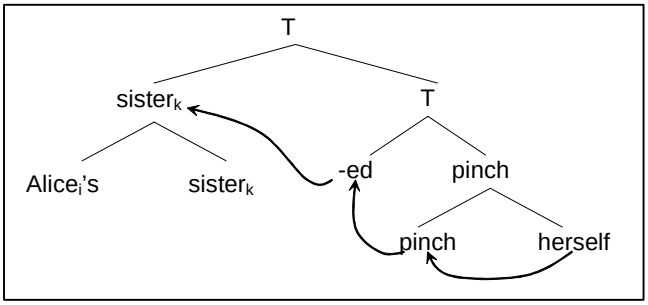
Herself can be straightforwardly added to the previous word, *pinch*, and check off its licensing features [Case: Acc, Right] and [Theta: Patient, Right]. Moreover, the parser identifies *herself* as an NP-trace and starts a search for its antecedent. The search along the left edge of the tree returns *Alice*. The parser links the two elements and checks their compatibility in morphological features. The dependency between *Alice* and *herself* is licensed.



Let us briefly discuss a modification of (24), given in (25).

(25) Alice's sister pinched herself.

In (25) the search will return the appropriate antecedent for the reflexive, i.e. *sister*, and not the incorrect *Alice*. This is because the search looks at heads that are along the main spine of the tree, but not inside those heads, as illustrated below:



That the head *Alice* will not be on the parser's search path agrees with the results of Badecker & Straub's (2002) Experiment 5, in which they found no effect of the possessor NP *Bill's/Beth's*, an inaccessible antecedent for the reflexive in sentences such as (26).

(26) Jane thought that Bill's/Beth's brother owed himself another opportunity to solve the problem.

5.2.3.2 Forwards Anaphora, Pronouns

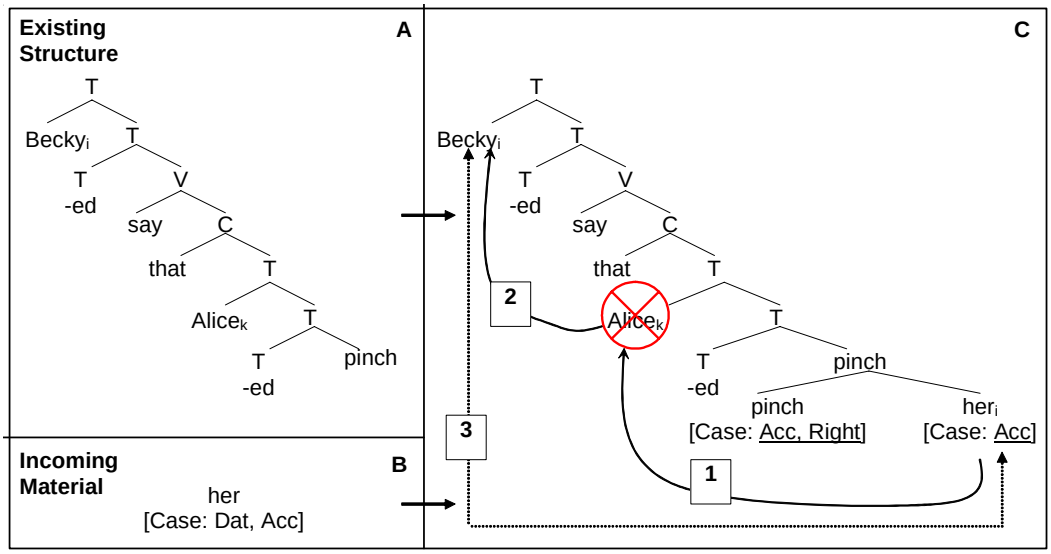
(27) Becky$_i$ said that Alice$_k$ pinched her$_i$.

I will only explain the final step of the derivation for (27). The structure built upon encountering all but the last word in (27) is given in panel A below.

→ *her*

The pronoun *her* is added to the existing tree and checks off the licensing features [Case: Acc, Right] and [Theta: Patient] for the previous word, *pinched*.

Once *her* is incorporated into the structure, a search for its antecedent along the left edge of the tree is triggered. The parser recognizes *her* as a non-trace and, hence, rather than returning the first nominal position found (*Alice*), it continues the search until the next position, *Becky*, is found. A link between *her* and *Becky* is then established, as a result of which *her* receives the same index as *Becky*, *i*.



5.2.3.3 Backwards Anaphora

Now let us consider a structure like (28). Similar structures were used as a control no-constraint condition in our Experiment 3 (Chapter 3), and yielded a GMME in the embedded subject position.

(28) His friend was smoking while John/Mary played pool.

This is the first example of how the parser builds a backwards anaphora dependency. I will use this case to illustrate the claims that were made in Chapter 3 regarding the processing of such dependencies, which are:

> *Active Search:* the parser actively looks for an antecedent for a cataphoric pronoun in the upcoming material;
>
> *Immediate respect of Principle C:* during an active search for an antecedent for a cataphoric pronoun the parser skips positions that are subject to Principle C.

The part of the algorithm that allows to implement these principles is given on page 211 and is repeated below:

> "… if the parser fails to find an antecedent at the step when the dependent element is incorporated into the structure, it re-checks the structure for a possible antecedent for the dependent element only when the tree gains by a new real or predicted head that does not expand the bottom of the tree."

→ *his*

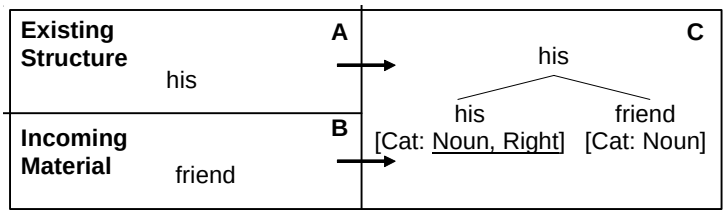
His [Cat: D] is extracted from the lexicon.

The parser identifies *his* as a non-trace dependent element. Since *his* is the first word in the input, no antecedent for it can be found at this step.

→ *friend*

Friend is added to the previous word, *his*, on the basis of the licensing feature [Cat: Noun, Right] of the determiner.

Since *friend* is added to the bottom of the previously existing tree (which was represented by the single head *his*), the parser will not re-check for an antecedent for the pronoun *his* at this step, in accordance with Principle C.



→ + *was* + *smoking*

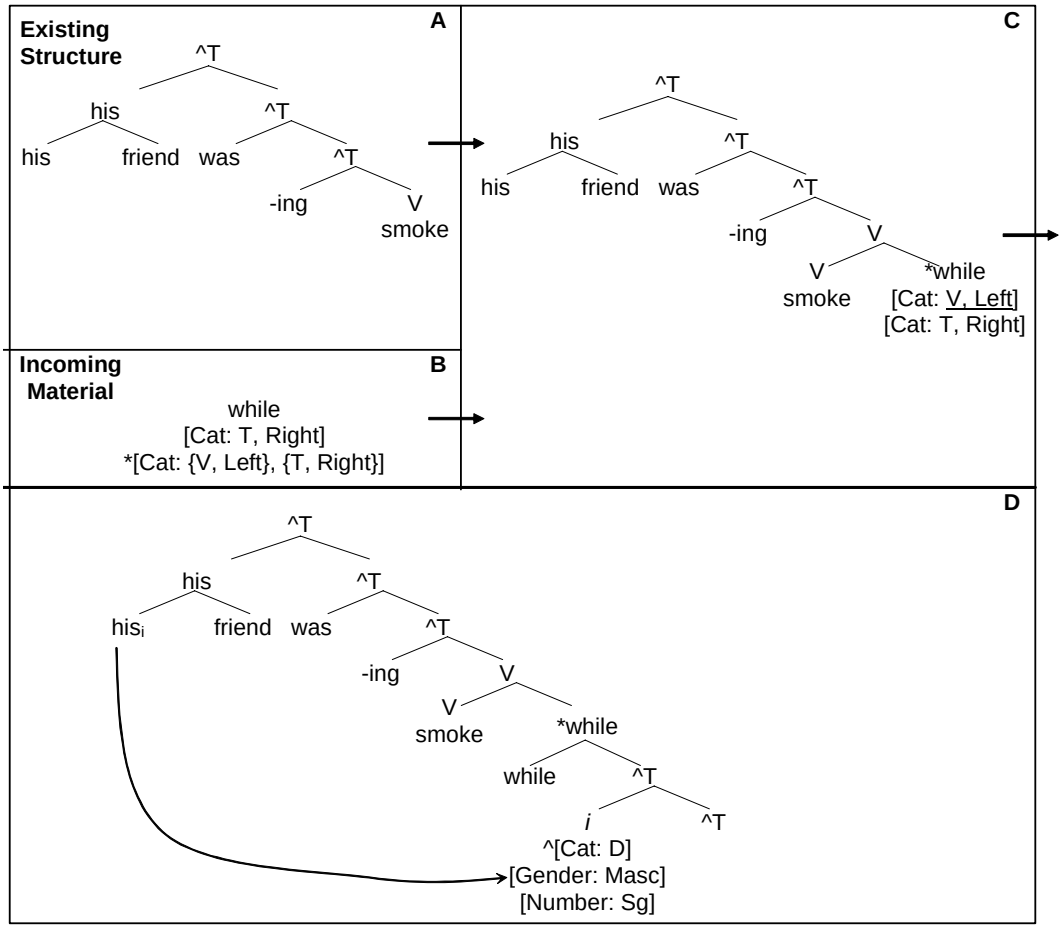
The structure built by the parser at the end of the first clause in (28) is given in panel A below.

Importantly, since *was* is added to the upper determiner head *his* in panel C, and hence does not directly expand the bottom of the tree, the parser will re-check for an antecedent for the pronoun *his* at this step and every subsequent step.

→ *while*

While adjoins to the previous word, *dance*, on the basis of its adjunct licensing feature *[Cat: V, Left] (panel B).

Once *while* is incorporated, it projects a predicted ^T-head, which in turn projects a predicted head ^[Cat: D]. The parser then rechecks whether any head can serve as an antecedent for the pronoun. ^[Cat: D] is returned and the parser projects the morphological features of the pronoun [Gender: Masc] and [Num: Sg] into that position.



→ *John/Mary*

John: John substitutes into the ^[Cat: D] position

Mary: the GMME emerges as a result from a mismatch of the morphological features of the predicted subject and the features of an NP in the input with which the parser attempts to instantiate the predicted subject.

To conclude, we have gone through a few sample sentences to illustrate how binding dependencies are established in real time. [119]

5.3 CONTROL IN SPARSE

In this section I will extend the algorithm developed for binding dependencies in section 5.2 to another class of dependencies known as *control*. The relevance of

[119] One more type of structure with a binding dependency, (i), that was used in van Gompel & Liversedge (2003) will be discussed in section 5.3.4.3.
(i) After he read the letter, John ate an apple.

control dependencies to the topic of this work is that, similar to binding, they link a pair of nominal elements one of which is referentially dependent on the other one.

In section 5.3.1 I will introduce some basic facts about control dependencies. Section 5.3.2 presents the results of an eye-tracking study by Sturt, Lombardo & Betancort (2004) on processing of control dependencies. In section 5.3.3 I extend the algorithm developed in section 5.2 for binding dependencies to cases of control dependencies and show examples of how this extended algorithm deals with cases of subject and object control into complements and cases of control into adjuncts.

5.3.1 What's Control?

Let us start by briefly introducing the phenomenon of control with the sentence in (29).

(29) Ruth forgot to turn off the alarm.

Both verbs in (29), *forget* and *turn off*, have two thematic roles to discharge. The agent role of the verb *forget* is taken by *Ruth*, and the theme role is absorbed by the complement clause. In the case of *turn off* the patient theta-role is absorbed by *the alarm*. However, the agent theta-role of *turn off* is not assigned to any overt NP. This is problematic for at least two reasons. First, (29) apparently violates the Theta-criterion which stipulates a one-to-one relation between arguments and theta-roles, as the agent role of *turn off* remains undischarged. Second, it contradicts the intuition according to which *turn off* does have an agent, *Ruth*.

A possible way to solve this contradiction is given in Chomsky (1981) who proposed that the embedded clause in (29) does in fact have an NP-element in its subject position, but that element is phonetically null. This new category was named as PRO. The structure of (29) with PRO is given in (30). PRO is controlled by the main subject *Ruth*, which explains why *Ruth* is an understood subject of the predicate *turn off the alarm*.

(30) Ruth_i forgot [PRO_i to turn off the alarm]

Note that at least in English, PRO can occur only in the subject position of infinitival and gerundive clauses, both of which are Case-less.[120]

[120] In the GB Theory the distribution of PRO was captured by means of the so-called PRO-theorem.
The PRO-theorem (Chomsky 1981): PRO is [+anaphor,+pronoun].
The PRO-theorem treats PRO both as an anaphor and as a pronoun, which makes it subject to both Principle A and Principle B of the binding theory simultaneously. The dual status of PRO amounts to saying that it is ungoverned. Were PRO governed, it would possess a governing category and

A different way of approaching these facts is presented in Hornstein (1999, 2001), who considers control dependencies (on par with binding dependencies discussed earlier) to be formed via movement. In other words, obligatorily controlled PROs are NP-traces that are related to their antecedents via A-movement. This is possible because, as we said earlier, in Hornstein's theory a single NP can check off more than one thematic role. For example, *Dan* in (31) checks the agent role of the verb *demonstrate* and moves on to check the agent role of *hope*; it leaves a trace in the embedded subject position (PRO in the traditional view). This is similar to how a complex reflexive item *Alice-self* in (32) checks off both the patient and the agent thematic role on the verb (repeated from (20)).

(31) Dan_i hopes t_i to demonstrate waltz basics.
Derivation: [$_{TP}$ Dan [$_{VP}$ Dan [hopes [$_{TP}$ Dan to [$_{VP}$ Dan demonstrate waltz basics]]]]]
(32) Alice_i pinched herself_i.
Derivation: [$_{TP}$ Alice T [$_{VP}$ [pinched [[Alice]self]]]]

Thus, both PROs and reflexives are NP traces. The distribution of these elements is determined by whether the position of a trace is [+Case]. If it is, then the trace must be instantiated as a reflexive, since in English case must be checked by an overt NP; otherwise, it is a null PRO.

An important feature of Hornstein's theory that we have not mentioned so far is an assumption of *sideward movement* (Nunes 1995, 2001), which is critical for explaining why control dependencies in structures like (33) and (34) are licit. In these cases PRO, in Hornstein's terms an NP-trace, is inside a complex subject (33) or adjunct (34) which are traditionally considered to be islands for movement. If so, an NP-trace inside them should be illicit.

(33) [t_i organizing parties] delights Caroline_i.
(34) Ilya_i never forgets to amply water his cactus [before t_i leaving the house].

In order to account for the grammaticality of (33) and (34) with the trace, Hornstein adopts the assumption of *sideward movement* (Nunes 1995, 2001). Nunes' proposal of sideward movement capitalizes on an independently motivated claim that any phrase involving a complex specifier or a complex adjunct must be assembled separately using an unconnected phrase marker (Chomsky 1993, section 3). Nunes (1995, 2001) proposes a derivational approach

---

would either satisfy Principle A (if bound inside the governing category) or satisfy Principle B (if free inside the governing category), but crucially not both principles simultaneously. Being ungoverned enables PRO to satisfy both Principles A and B simultaneously by evading both of them due to its lack of a governing category.
The fact that PRO is ungoverned also explains why it is phonetically null. The reason for it has to do with Case, which in the GB theory is assigned under government. Since PRO is ungoverned it cannot receive case. Given that every phonetically full NP must be Case-marked (the Case filter), the non-Cased-marked PRO cannot be phonetically realized.

to islands whereby a phrase is not an island for movement until after it attaches to the main tree. Consequently, movement out of complex adjuncts and subjects (that are traditionally viewed as always being islands for movement) is allowed as long as they are disconnected from the rest of the tree. Sideward movement is then an interarboreal movement from a tree to another disconnected tree. Hence, the NP-traces (PROs) in (33) & (34) are residues of a sideward movement from the subject or adjunct into the main clause. In (33) *Caroline* moves from inside the complex subject into the object position in the main clause before the subject attaches to VP, and in (34) *Ilya* moves from inside the adjunct to the main subject position before the adjunct adjoins the main VP.[121],[122]

5.3.2 An Eye-tracking Study of Control

Having discussed theoretical accounts of control dependencies, we would like to also discuss what is known about their real-time processing. Our survey will be restricted to one study that is most relevant here, an eye-tracking experiment by Sturt, Lombardo & Betancort (2004). The experiment involved sentences with control into an adjunct clause such as (35). [123]

(35) a. After (PRO) making **herself** a cup of tea, **the nurse** very carefully examined the patient.
b. After (PRO) making **himself** a cup of tea, **the nurse** very carefully examined the patient.

In (35) the reflexive is bound by PRO (Principle A), which in turn is controlled by the subject of the matrix clause. Thus, there is a mediated coreference relation between the reflexive and the matrix subject, according to which the main subject and the reflexive should match in morphological features. If the parser calculates all these relations online, an immediate surprise reaction is expected when the stereotypical features of the main subject mismatch the reflexive, as in (35)b.

Sturt et al. indeed found a GMME in early measures at the main subject *the nurse* and the following adverb *very*, i.e. first fixation durations at *the nurse* and first-pass reading times at *very* were longer in (35)b than in (35)a. Importantly, no effect was found in (36)a vs. (36)b, in which the reflexive in the adjunct was replaced with a pronoun, which cannot be bound by the PRO.

(36) a. After (PRO) making **her** a cup of tea, **the nurse** very carefully examined the patient.

[121] Note that if a sideward movement of this sort is allowed, then the requirement that the antecedent must c-command its trace cannot be sustained. In Hornstein's movement approach, this requirement on movement is not stipulated (see Kiguchi & Hornstein, ms).
[122] More details on Hornstein (1999, 2001) are given in
[123] Step-by-step structure building for sentences such as (35) is discussed in section 5.3.4.2.

b. After (PRO) making **him** a cup of tea, **the nurse** very carefully examined the patient.

The results of the experiment by Sturt and his colleagues suggest that the parser expects the controller for the PRO in the main subject position, and that this expectation is formed before any bottom-up information on the main clause becomes available.

The structures tested in Sturt et al. (2005) are similar to those from van Gompel & Liversedge's (2003) study, repeated in (37). In either case there is a fronted adjunct clause with a dependent element as a subject; the main subject is an accessible antecedent for the dependent element. The differences between the studies concerned the type of the dependent element, silent PRO in (35) vs. an overt pronoun in (37). Furthermore, whereas the main subject is the only grammatical antecedent for the dependent element in (35), in (37) it is one of many possible antecedents.

(37) When **he** was fed up, **the boy/girl** visited the girl/boy very often.

In both studies a GMME was found as a result of the manipulation of gender congruency in the main subject. This fact serves as evidence that the parser employs a similar active search mechanism for finding an antecedent for an overt pronoun (37) or a silent PRO (35). Such a similarity is not unexpected, but it is not a necessity, especially in light of the fact that only in (35) but not in (37) there is a unique grammatically available antecedent for the dependent element. Nevertheless, the parser is equally surprised to find a gender-mismatching NP in the first accessible antecedent position regardless of whether this is the only chance to find an antecedent or one of many possible chances.

In the next section I will discuss which additions to the algorithm developed in section 5.2 are necessary to bring control dependencies under its scope. I will then show how the parser establishes these dependencies incrementally using several examples, including sentences like (35) and their binding counterparts like (37).

5.3.3 More on the Algorithm for Building Referential Dependencies

Extending the parser's binding algorithm from section 5.2 to control dependencies requires considering two points. First, the dependent element in the control dependency is silent, hence its existence must be discovered via other means. Second, we have seen that in order to account for cases of control into complex subjects and adjuncts in the movement-based theory, one must appeal to the notion of sideward movement. Apart from these points, in light of the fact that the algorithm is based on Hornstein's approach in which binding and control are treated as essentially the same type dependency, the major algorithm for forming

a control dependency is the same as for binding dependencies and relies on principles listed in (21).

'Silentness' of PRO[124]

In all binding dependencies the dependent element is overt, i.e. a reflexive or a pronoun, and the parser can detect and, further, categorize it as a trace or a non-trace, based on its phonetic form. PROs cannot be detected in the same way because they are silent. The parser will postulate a trace/PRO as a result of the requirement (21)c that the [Spec TP] position in English must be filled, whenever it encounters a subject position that cannot be filled out by overt material. These are (a superset of) positions where PROs reside.[125]

Sideward movement

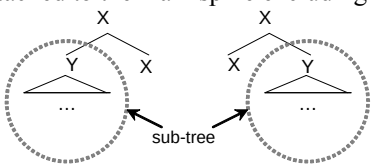
Recall that in order to explain binding and control into complex subjects and adjuncts Hornstein (1999, 2001) uses an assumption that sideward movement, i.e. an interarboreal movement between disconnected trees, is allowed.

A literal implementation of sideward movement into the parsing algorithm would require discarding a well-motivated assumption that the parser strictly incrementally builds a fully connected structure. Luckily, complex subjects and adjuncts can be identified even if they are immediately attached to the main tree: they are non-terminal nodes that 'hang off' the main spine. All information that is true of disconnected trees can be viewed as holding of sub-trees.[126] Structures with control into complex subjects or adjuncts are then structures with dependent elements inside a sub-tree. The overall procedure of searching for an antecedent for PROs inside complex adjuncts and subjects is similar to other cases. The parser finds an NP-trace inside a sub-tree (rather than along the spine of the main

[124] I will continue to use the traditional terms PRO for convenience, although it must be understood that in Hornsetein's account that our algorithm is based on PRO is nothing but an NP-trace.

[125] According to Hornstein Case-less [Spec TP] positions in fact can either be filled with a PRO (= trace) or a null pronoun *pro*, depending on whether movement out of this position is allowed. Even so, PRO will be the first option attempted, given a preference for traces over non-traces. What needs then to be explained is when the parser's initial guess, PRO = NP-trace, turns out to be inappropriate due to inability to move out of the relevant position, or in other words, how are islands defined in a derivational approach. I will briefly return to this point in section 5.3.4.4.

[126] To make it more formal, a sub-tree is any non-terminal node Y that is a sister of a node X (terminal or non-terminal) and that is dominated by a projections of X (i.e. Y does not project further up as in the diagram below). In other words, a subtree is any non-terminal projection that is attached to the main spine excluding the root of the tree.



tree). The parser first searches the left edge of the tree all way up to the root node and returns the first nominal position it finds. Otherwise, if no antecedent is returned as a result of the left edge search, the parser starts a search along the right edge.[127]

We will next go through derivations of several types of structures in SPARSE that involve subject and object control and control into an adjunct.

5.3.4 Case Studies – Part 2

5.3.4.1 Subject and Object Control

Let us start by showing how the parser builds a structure for cases of subject and object control into a complement clause such as (38) and (39) respectively.

(38) Dan hoped to leave.
= Dan_i hoped PRO_i to leave.
(39) Dan asked Silke to leave.
= Dan_i asked $Silke_k$ PRO_k to leave.

Let us start with (38).

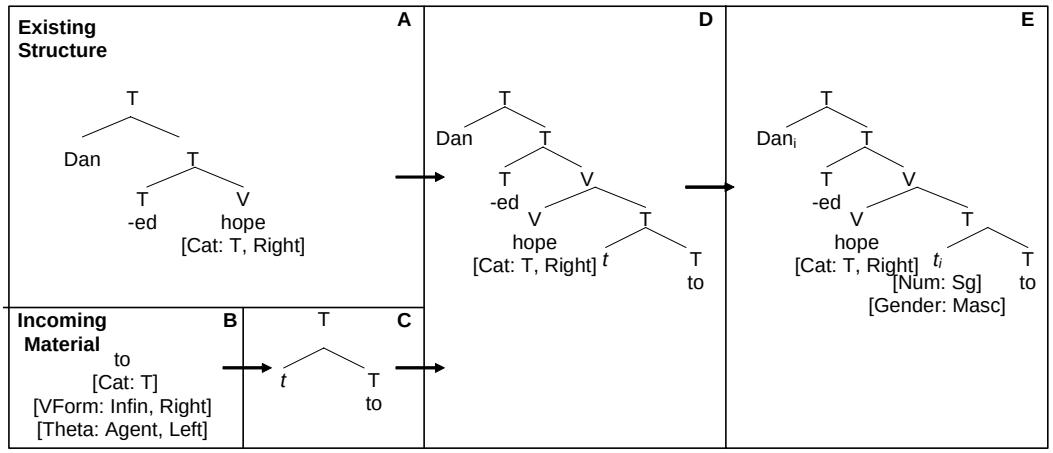
→ *Dan* + *hoped*

The structure in panel A is built.

→ *to*

To is a T-head (panel B) that can directly attach to the previous word, *hoped*, which has a licensing feature [Cat: T, Right].[128] However, if that happened, the licensing feature of the T-head [Cat: D, Left] won't have a chance to be ever checked off.[129] Hence, the parser needs to satisfy this feature before *to* is attached to *hoped*. The parser thus posits a trace to the left of the T-head, which checks off theta-feature of *to* (panel C).

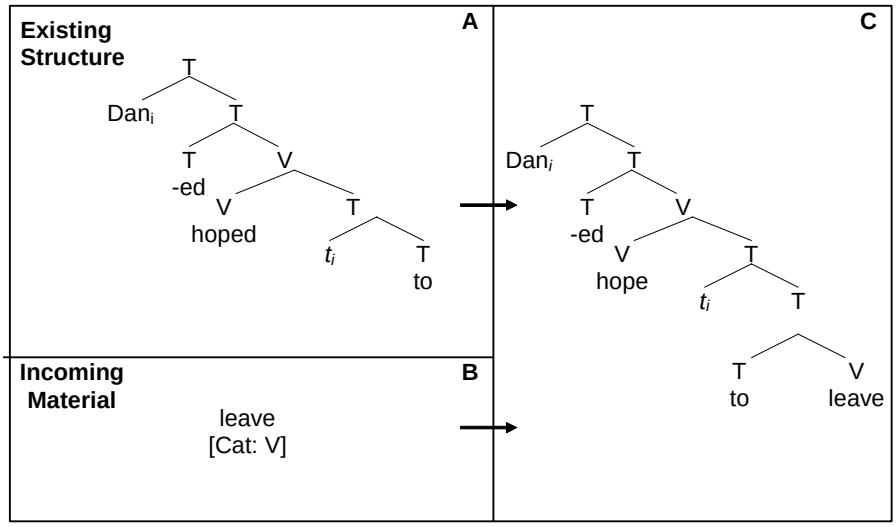
[127] As mentioned in fn. 125, the search for an antecedent for the trace is also subject to the island constraints, which we disregard at the moment.
[128] In the GB Theory it was important that the complements of control verbs were CPs rather than TPs, so that PRO is ungoverned by the verb (see fn. 120). In Hornstein's Miminalist account which dispenses with the notion of government this is no longer an issue. I will assume for simplicity that these clauses are TPs, although nothing essential hinges upon this assumption.
[129] Unlike T's of inflected verbs, T-heads introduced by infinitives and gerunds should be marked in the lexicon as weak in that they do not have a licensing feature [Case: Nom, Left].



Once the new piece of structure is attached to the tree (panel D), the parser initiates a search for an antecedent of the trace. The parser initially searches the 'left edge' of the tree, i.e. the preceding material. The first nominal position that it finds, the main subject *Dan* is appointed as the antecedent for the trace. The link between the trace and the main subject is established, and the features of the antecedent and the dependent element are matched (panel E).

→ *leave*

The verb *leave* is directly attached to the preceding word, *to*, on the basis of its licensing feature [Cat: V, Right]. The final structure for (38) is given in panel C.



The derivation for (39), which is repeated below for convenience, proceeds as follows:

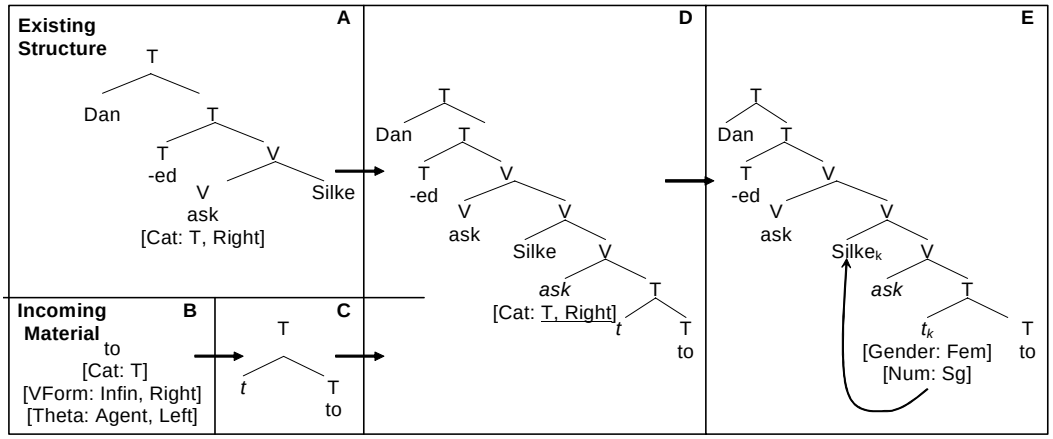
(39) Dan asked Silke to leave.
= Dan_i asked $Silke_k$ PRO_k to leave.

→ *Dan + asked + Silke*

The parser builds a main clause shown in panel A below.

→ *to*

This step is similar to the corresponding step in (38), except that the ditransitive main verb *ask* must be copied to a lower position to license its second argument in accordance with the feature [Cat: T, Right] (the copy is italicized in panel D).[130] Once the new word is added to the structure, the parser starts the left edge search for an antecedent of the trace (= the target of movement from the position of the trace). The first nominal on its way the position of the trace upwards is the object *Silke* which is appointed as the antecedent for the trace. The link between the trace and *Silke* is established, and the features of the antecedent and the dependent element are matched. Panel E shows the final structure of the sentence (39).



5.3.4.2 Preposed Adjuncts with PRO

Sentences such as (40) were shown by Sturt et al. (2004) to elicit a gender-mismatch effect at the main subject when its stereotypical gender mismatches the gender of the reflexive inside the adjunct.

(40) After making himself a cup of tea, the nurse ate an apple.
a. After *PRO* making himself a cup of tea, the nurse ate an apple.

The derivation for (40) proceeds as follows.

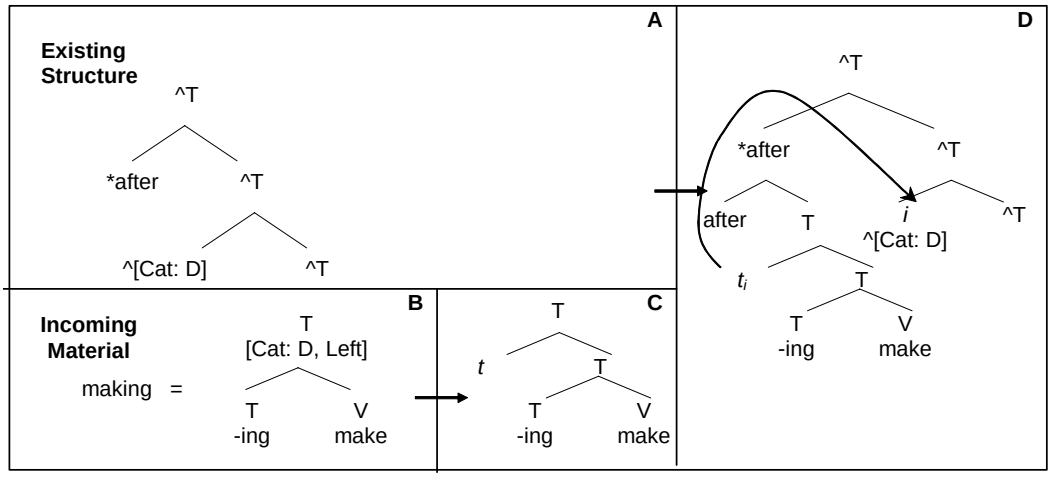
→ *after*

See panel A.

→ *making*

[130] See Phillips (1996) for a detailed discussion justifying this step from the parsing perspective. Original theoretical proposals for this structure comes are due to Larson (1988) and Aoun & Li (1989).

Making introduces a V-T complex (panel B) that can directly attach to the previous word, since *after* has a licensing feature [Cat: T, Right]. However, if that happened, the licensing feature of T-head [Cat: D, Left] would not have a chance to be ever checked off.[131] Hence, the parser needs to satisfy this feature before *making* is attached to *after*. The parser thus posits a trace to the left of the T-head, which checks off theta-feature of the verb *making*.



The parser also initiates an active search for an antecedent of the trace. The parser's initial 'left edge' search in the preceding material does not yield any results: the parser reaches the root of the tree without finding a nominal position. Hence, the parser starts a search in the predicted material along the right edge of the tree. The first nominal position that it finds, the predicted head ^[Cat: D], is appointed as the antecedent position for the trace. The link between the trace and the main subject is established. Should the trace have any features, those features would have been added to the specification of the subject in addition to the feature [Cat: D]. Since the trace is silent, no new features are added at this point.

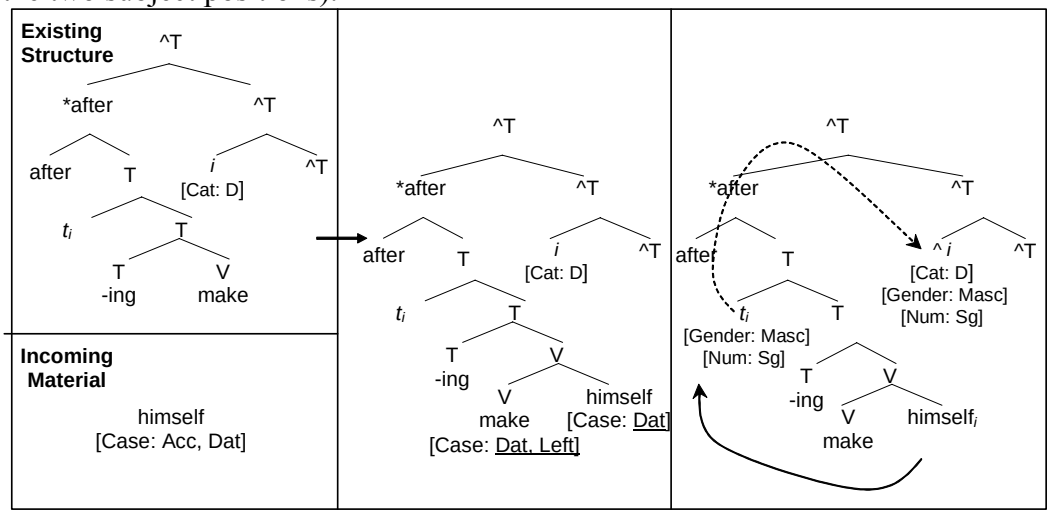
Note that in general nothing prevents the parser from reanalyzing its prediction for the antecedent for the trace if there is a closer nominal position that is not in the c-command domain for the trace. However, the parser's prediction for the antecedent will not be revised as long as new material is added to the bottom of the tree in accordance with Principle C. In this example, the parser will not re-check for an antecedent for the trace as long as more material is added to the end of the embedded clause.

→ *himself*: [Case: Acc, Dat][132]

[131] The logic is the same as in the case of attachment of the infinitival particle *to* in examples (38) & (39) discussed in section 5.3.4.1.

[132] Reflexives are represented here as morphologically simplex heads, although in principle they may be treated as introducing a pronoun-*self* complex, similar to a V-T complex introduced by a finite verb.

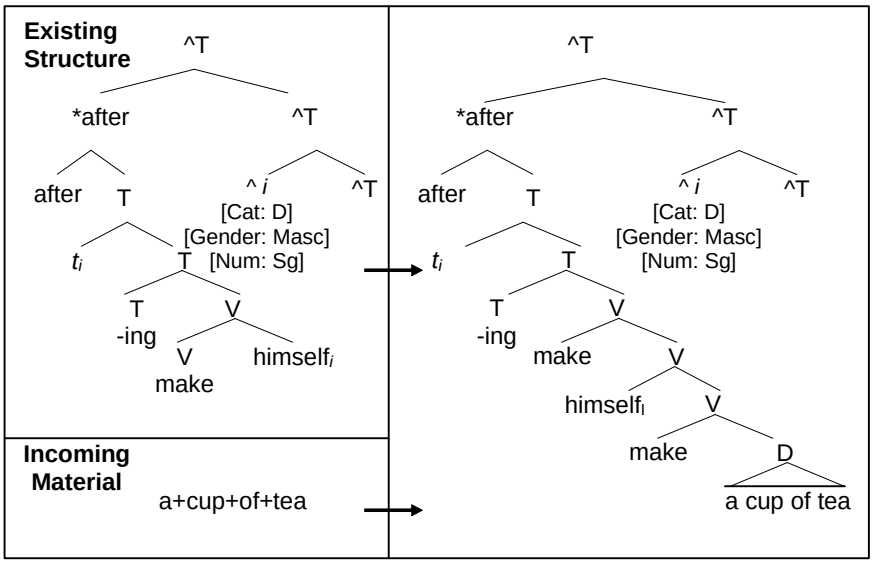
Himself is directly attached to the previous word (*making*) and checks off the dative case on the verb.[133] The parser then starts a search for an antecedent of *himself*, that must be the first nominal head found during the search. The parser's search starts at *himself* and goes upwards along the left edge in the preceding material. The embedded subject head t_i is returned and is assigned as antecedent for the reflexive. The two heads are linked, i.e. the index of the trace *i* is passed along to *himself* and the features of *himself* are passed to the trace (and, consequently, to the main subject via the previously established pointer between the two subject positions).



→ *a* + *cup* + *of* + *tea*

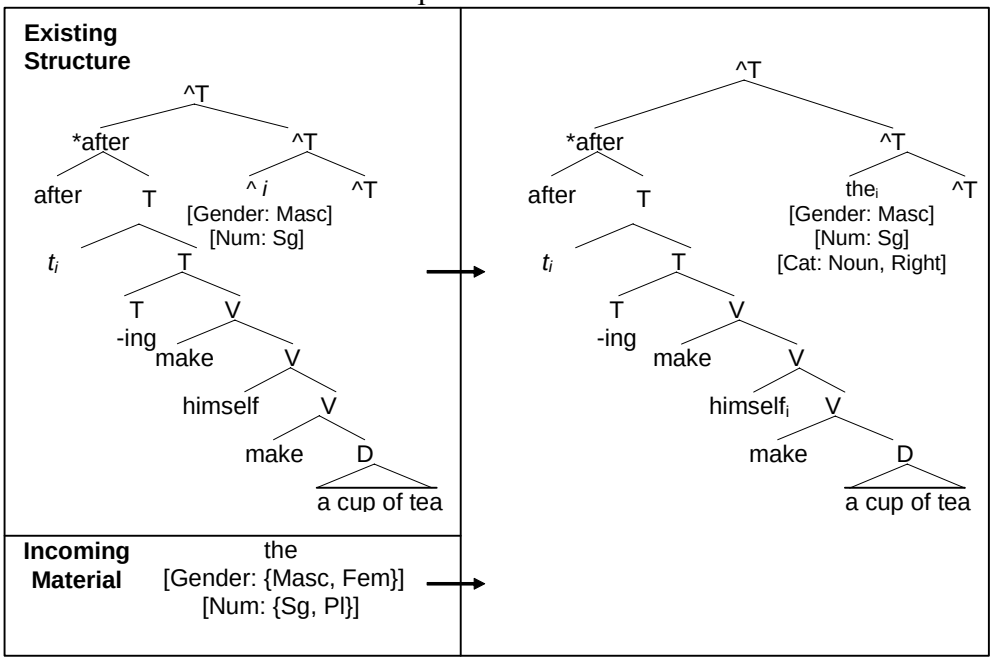
The DP *a cup of tea* is attached as a complement of *make* and checks its patient theta-role and the accusative case.

[133] I remain vague on the details of the process of establishing that in the current sentence *making* is used ditransitively. I assume for simplicity that Dative was the first choice of the parser, although presumably the default choice is Accusative and then reanalysis is started after semantic infelicity of that structure is detected or after the following material (*a cup of tea*) is encountered.



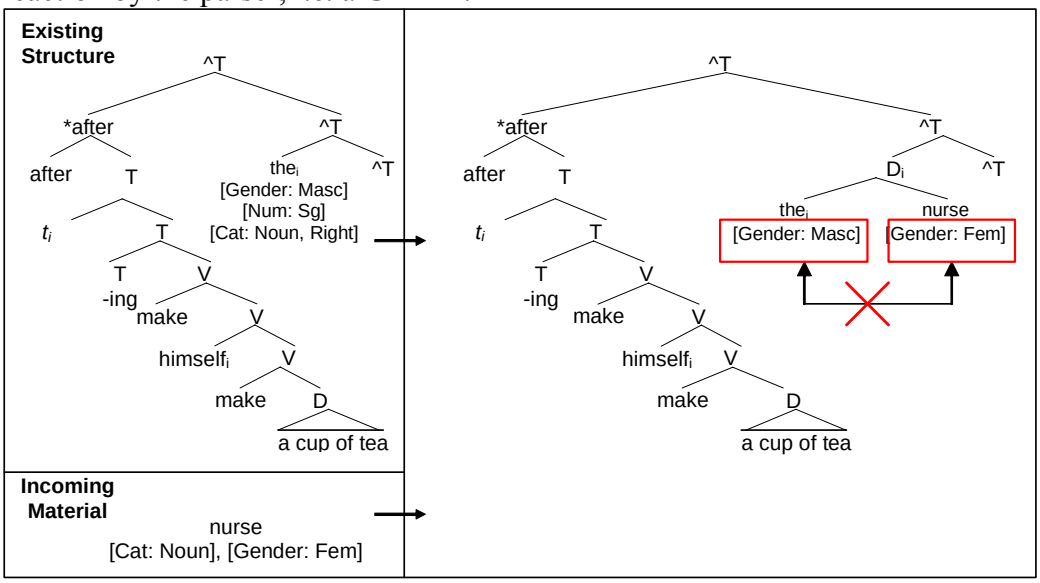
→ *the*

The cannot be directly attached to the previous word (*tea*). The parser starts searching the right edge of the tree for a possible attachment site and finds a predicted head ^[Cat: D] in the main subject position. *The* substitutes into that node; its features are intersected with the features specified in the position. As a result a non-null intersection set of features [Gender: Masc] and [Num, Sg] is found. *The* also inherits the interpretative index *i*.



→ *nurse*: [Cat: N], [Gender: Fem], [Number: Sg], [Case: {Nom, Dat, Acc}]

The parser aims to attach the new word (*nurse*) to the previous one (*the*). Although this attachment is licensed by the licensing feature [Cat: Noun, Right], a mismatch in the gender features in the determiner and the noun causes a surprise reaction by the parser, i.e. a GMME.



5.3.4.3 Binding returns: Binding inside an adjunct

Let us return to binding once again to discuss (41) that has a pronoun as a subject of the adjunct clause. Sentences of this type were shown to elicit a GMME at the main subject whenever the latter mismatches in gender with the pronoun (van Gompel & Liversedge (2003) and Chapter 3 of this dissertation).

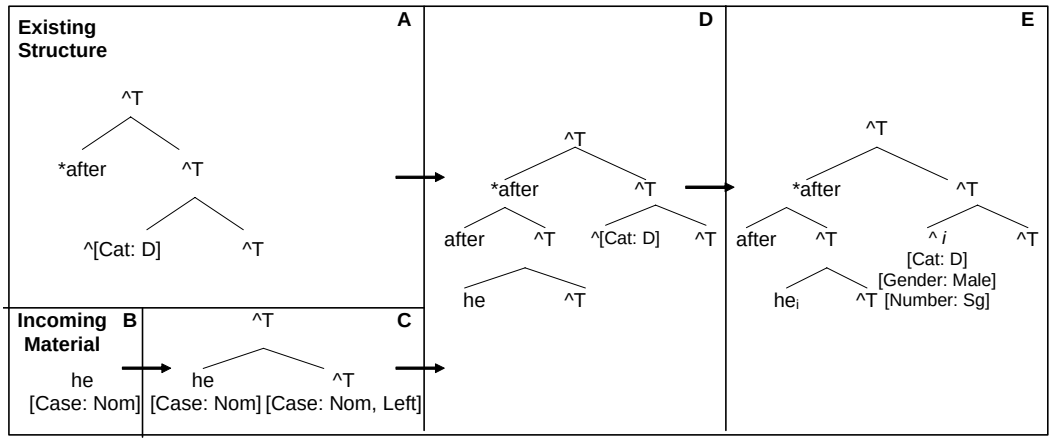
(41) After he read the letter, John ate an apple.

→ *after*

The preposition *after* is encountered first and the structure shown in panel A is projected (see the discussion of (16) for details).

→ *he*

The pronoun *he* cannot directly attach as a complement of *after*, because *he* is specified for nominative case, whereas *after* takes a dative complement. According to Schneider's algorithm (see Appendix 5-A) the parser starts searching the right edge of the tree for an attachment site for *he*. This search starts at the most recently processed word (*after*), which still needs to check off its licensing feature [Cat: T, Right] (i.e. the embedded clause is still missing). Since the search must stop whenever a head that is missing an argument on its right is encountered (see the algorithm), the parser stops searching the tree at this point.



Next, the lexicon is searched for a possible licensor for *he*. Since *he* is specified for nominative case, and since the only head that assigns nominative is T, the search returns a T-head. A predicted ^T-head is projected to the right of *he* (panel C). The tree is searched for an attachment site for the ^T-head, again starting at the most recently attached word, *after*. *After* has a licensing feature [Cat: T, Right] which licenses attachment of the predicted ^T-head as a complement of *after* (panel D).

Once *he* is attached to the tree, it initiates a search for an antecedent. Note that *he* is in the subject position, which is an illicit position for the reflexive (see (21)i). The parser therefore is allowed to return any nominal position as long as it does not violate Principle C (i.e. including the first position on the search path that is normally ruled out due to a competition with the reflexive).[134] The parser finds no candidate antecedents (i.e. no nominal positions) in the input that precedes the pronoun. It then searches along the right edge of the tree and finds the predicted head ^[Cat: D], that was projected at the previous step by the predicted main ^T-head based on its distinguished EPP feature [Cat: D, Left]. The features of *he* (i.e. its index and its phi-features, panel E) are transferred to the head ^[Cat: D]. The GMME at this subject results from a mismatch of the phi-features of the predicted subject and the features of an NP in the input with which the parser attempts to instantiate the predicted subject.

5.3.4.4 Post-posed adjuncts with PRO

(42) John kissed Mary before leaving.
a. $John_i$ kissed $Mary_k$ before PRO_i leaving.

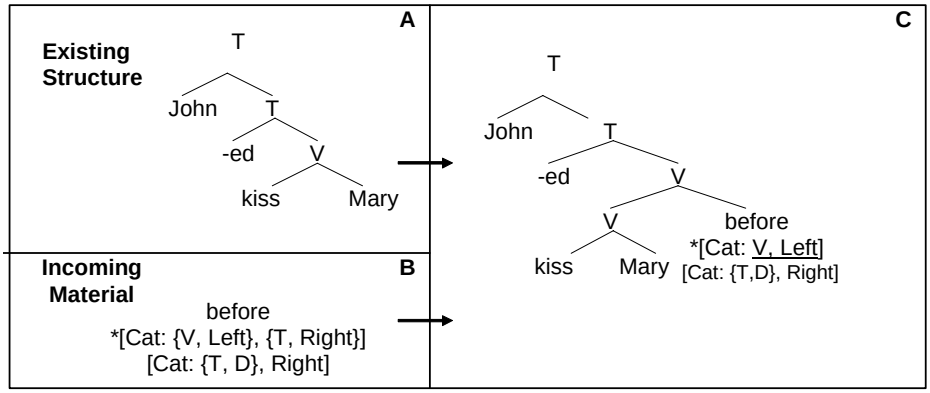
→ *John + kissed + Mary*

The structure after the parser processed the first three words is shown in panel A.

[134] Spelling out this step in detail is left for future investigation.

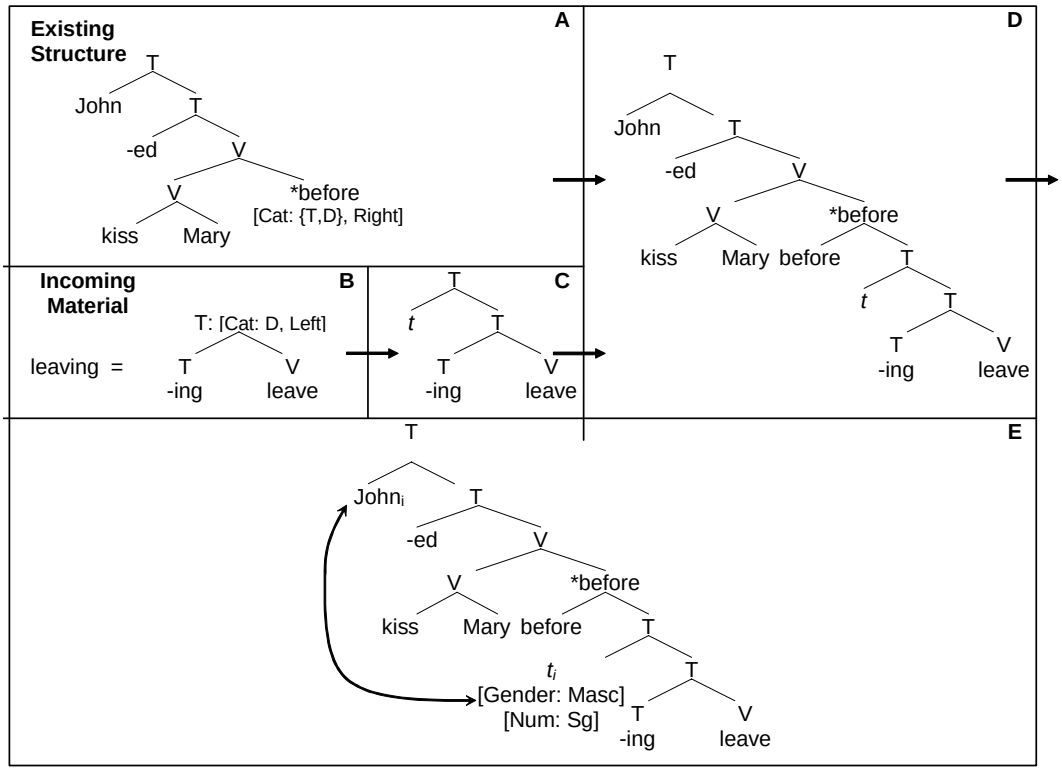
→ *before*

Before cannot be attached to the immediately preceding word, *Mary*. The parser searches along the right edge of the tree for an argument attachment of *before* (see the SPARSE algorithm, Appendix 5-A). Since this search does not return any results, the parser starts a search for an adjunct attachment site. *Before* can adjoin to the right of the V-head based on its licensing feature *[Cat: V, Left].



→ *leaving*

The verb *leaving* is incorporated via the same procedure as the verb *making* in section 5.3.4.2 (shown in panels B & C) below.



Once *leaving* is incorporated into the tree (panel D), the trace in the embedded subject position starts a search for an antecedent. The left edge of the tree is searched upwards starting at the position of the trace. The first nominal head of the search path is the main subject *John* (recall that the parser can only see nodes that are directly attached to the main spine, hence *Mary* will not be found). The parser established a link between *John* and the trace which results in a shared index *i* and feature agreement between the two nominals (panel E).

Before concluding, one final note is worth mentioning. In discussing control dependencies, I limited myself to cases where the parser's initial guess as to the type of the empty category in the [Spect TP] position, PRO, was always the correct choice. Yet, as I noted in passing, this is not always the case. Let us look at the sentence (43).

(43) John believes that washing himself would delight Mary.
a. *John$_i$ believes that PRO$_i$ washing himself would delight Mary$_k$.
b. John$_i$ believes that *pro$_i$* washing himself would delight Mary$_k$.

There are good reasons to believe that the correct structure for (43) is (43)b, but not (43)a.[135] Yet, it should be obvious that if nothing else is added to the search algorithm proposed here, the parser would build the incorrect structure (43)a. This is because while incorporating the verb *washing* into the tree the parser would postulate a trace in its [Spec TP]. Then the parser would look for an antecedent for the trace by searching up along the left edge of the tree. The main subject *John* will be the first nominal position on the search path and thus will be returned as an antecedent for the trace. This illicit outcome needs to be ruled out.

Hornstein (2001) rules out (43)b because by the time the main subject position becomes available, the adjunct *washing himself* is attached to the main tree and thus is an island for movement. In other words, a complex subject or adjunct is not an island from the perspective of its host clause, but is an island for higher clauses in the main tree. Thus, in order to correctly handle cases like (43) the parser's knowledge base must be extended to derive islands. Once this is attained, the structure (43)b will not be built because the search for an antecedent, will fail to proceed past the embedded T-node into the higher main clause.

5.4 CONCLUSION

In this chapter I developed a parsing procedure that is capable of building binding and control dependencies during incremental structure building. The algorithm follows Hornstein's (1999, 2001) theory in that binding and control are viewed as instances of movement, which in its turn is closely related to the procedure of searching the tree by the parser. The active search mechanism is instantiated via a

[135] See Appendix 5-B for an explanation what these reasons are.

requirement that the parser looks for an antecedent for a dependent element as soon as that element is incorporated into the tree and until an antecedent is found. Principle C is built into the parser's procedure of antecedent search for a dependent element: the parser disregards all nominal positions that are in the c-command domain for the dependent element, thus never considering any positions that may violate Principle C.

Appendix 5-A. SPARSE Parsing Algorithm (Schneider 1999)

1. Search unchecked features on the right edge of the existing structure for an argument attachment for the incoming material.
   - If a subsumption relation is found between all heads in the incoming item and predicted heads in the existing structure, integrate the entire new item into the existing tree.
   - If subsumption is found between some (but not all) heads in incoming item and heads in the existing tree, remove from the existing tree the portion that is compatible with the heads in the incoming item. Integrate the new item into the just-removed structure. Return to step 1 with the just-integrated new item.
2. Search the right edge of the existing structure for an adjunct attachment of the new material.
3. Build a new licenser for the new material:
   3.1. If requirements must be satisfied on the new item before structure can be built above it, build the minimum structure necessary to satisfy the requirements and continue to 3.2.
   3.2. Search the lexicon for all possible licensers of the new material. A possible licenser is either a head with appropriate left-pointing features or a null head with appropriate right-pointing features.
   3.3. Attach to the new material the intersection of all heads returned by the lexicon search.
      - If the intersection is null and the new material is headed by a predicted licenser, search for an argument attachment using all heads compatible with the head of the new constituent. If successful, make the attachment using the head of the new constituent (not the subsuming head that licensed the attachment).
      - If the intersection is null and the new material is not headed by a predicted licenser, no new licensing heads can be predicted—continue on to 4, otherwise return to step 1 with the just-built constituent.
4. Search all features (checked and unchecked) on the right edge of the existing structure for an argument attachment for the incoming material.
   4.1. If an attachment is found, remove existing element from the tree, attach it to the new material, and start over with newly-expanded constituent at step 1.

Appendix 5-B. Hornstein (1999, 2001, 2005)

Below is the gist of Hornstein (2005), which expands upon and gives answers to some problems from his earlier work (Hornstein 1999, 2001).

The hallmark of Hornstein's analysis is its parsimony and reduction of the number of different empty categories. Various silent categories that are traditionally considered to be regulated by separate modules – obligatory and non-obligatory control PRO, null pronoun *pro* and NP-traces – are all united by the notions of movement and economy. The same principles are used to derive the distribution of reflexives and bound pronouns. All these categories are viewed as grammatical formatives (rather than lexical elements), which are defined by the structural position that they appear in (in terms of features that can be checked off in a given position); those that occur in similar positions are in competition with one another. Among empty categories, traces are assumed to be the default choice by the grammar. I will dub this principle the 'Trace-first' principle.

(1) *'Trace-first' Principle of the grammar*
(Everything else being equal,) posit an NP-trace whenever possible.

The phrase 'everything else being equal' in (1) is added as a reminder that the grammar can compare derivations only if they have equivalent interpretations (logical form). For example, the 'Trace-first' principle explains why sentence (2) cannot have a representation such as (2)b, given that (2)a is licit: PRO's (NP-traces) trump *pro*'s.

(2) Dan hopes to demonstrate waltzes.
a. Dan_i hopes PRO_i to demonstrate waltzes.
b. *Dan_i hopes pro_i to demonstrate waltzes.

Categories that are not traces of movement, e.g. *pro*s and overt pronouns, are inserted as a last resort when their less costly competitors that are residues of movement, i.e. PROs and reflexives, are illicit (e.g. inside *wh*-islands).

In addition to what we said above, Hornstein's theory makes use of two parsing principles. One principle, which I dub as the 'Trace-over-pro' principle, claims that the by default the parser categorizes an empty category as an NP-trace. The second principle is the now familiar 'Active search' principle whereby the parser attempts to find an antecedent for a referentially dependent element as soon as possible. Let us start with a demonstration of why these principles are necessary with sentence (3).

(3) John persuaded Mary to leave.[136]
a. $John_i$ persuaded $Mary_k$ [PRO_k to leave].
b. *$John_i$ persuaded $Mary_k$ [PRO_i to leave].
c. *$John_i$ persuaded $Mary_k$ [pro_k to leave].
d. *$John_i$ persuaded $Mary_k$ [pro_i to leave].

(3)a is the only possible structure for (3). Hornstein (2005) explains why (3)a is the licit structure and (3)b-d are illicit as follows. The derivation in (3)b where PRO is an NP-trace anteceded by *John* does not converge as movement from the embedded subject position to the main subject position would violate minimality (*Mary* is closer to the main subject position than *John*). (3)c is illicit because *pro* is allowed to have the same index as *John* only if the same indexation pattern were not attainable through movement (i.e. (3)a trumps (3)c). Finally, the illicit structure (3)d is explained by appealing to a parsing consideration: the parser prefers to postulate an NP-trace (PRO) over *pro* whenever possible, regardless of whether the corresponding outcomes have an equivalent representation, i.e. (3)d is also trumped by (3)a.

(4) *'Trace-over-pro' Principle of the Parser*
Treat any empty category as an NP-trace.

Thus, (3)d will not ever be generated because the parser would always choose to drop a trace in the position of the empty category as it proceeds left-to-right through the sentence.[137]

However, Hornstein (2005) notes that the parsing principle whereby a trace is always preferred to *pro* is at odds with the fact that both sentences (5) and (6) are grammatical.

(5) John believes that ***e*** washing herself would delight Mary.
a. John believes that [PRO_k washing herself] would delight $Mary_k$.
(6) John believes that ***e*** washing himself would delight Mary.
a. $John_i$ believes that [pro_i washing himself] would delight Mary.

In (5) the empty category is controlled by the object *Mary* and is a PRO, the residue of sideward movement of *Mary* from inside the complex subject to the

[136] Although I will continue to use traditional PRO and *pro* in diagrams of the sentence structure, this is done only for the sake of clarity and simplicity. Every case of PRO is a trace of an NP-movement from the position of PRO to the position of its controller. *Pro* is inserted as last resort into positions from which movement cannot take place.

[137] The 'Trace-over-pro' principle of the parser should not be confused with the 'Trace-first' principle of the grammar. Both principles claim that a trace is preferred to a pronoun. However, for the grammar this preference has scope over comparable derivations (i.e. those that have an equivalent interpretation), whereas in parsing it is an across-the-board preference (which nevertheless can be in competition with other preferences, as we will discuss in a moment). Hence, (3)a trumps (3)b, but not (3)d, according to the 'Trace-first' principle. (3)d loses to (3)a as a result of application of the 'Trace-over-pro' principle of the parser.

object of the embedded clause (see (5)a). On the other hand, the antecedent for the empty category in (6) is the main subject *John*. The link between the empty category and *John* cannot be formed by movement, since the complex subject clause is an island: at the point when movement to the main clause could take place, the gerundive subject [e washing himself] is already attached to the rest of the embedded clause, which is in turn a complement of the main verb *believe* (hence attached to the main clause right away). If so, the empty category must be *pro*, as shown in (6)a. If this is correct, the empty category is PRO (i.e. an NP-trace) in (5) and *pro* in (6), which is counter-evidence to the claim that the parser always prefers to drop an NP-trace to *pro* in the place of an empty category.

Hornstein's account of these facts is the following. The grammaticality of (5) is in accordance with 'Trace-over-pro' parsing principle. What remains to be explained is why (6) is also grammatical. Hornstein suggests that the reason why the parser seems to be able to drop a *pro* in (6)is due to the preference of the parser to assign an interpretation (i.e. an antecedent) for an empty category as quickly as possible. This can be subsumed under the familiar 'active search' heading (7), whereby upon encountering a dependent element the parser starts an immediate search for an antecedent.

(7) *Active Search Principle of the Parser*
Assign an interpretation (i.e. an antecedent) to dependent element (PROs, bound pronouns (overt or silent *pro*s), reflexives) as quickly as possible.

According to Hornstein (2005), after processing an initial fragment *John believes that e washing*... the parser would both like to assign an interpretation to the empty category and to treat it as a trace rather than *pro*. However, the parser cannot do both things simultaneously, since the two preferences pull in different directions. In order for the empty category to receive an antecedent immediately it must be understood as a *pro* (and thus be related to *John*), but this will also require overriding the parser's preference for an NP-trace over a pronoun. On the other hand, if the parser does respect its preference and treats the empty category as PRO, it cannot interpret the PRO until later on. Since both options have their virtues, both parses may be available.

6 CHAPTER 6. GENERAL CONCLUSION

The chapters that precede this one introduced and investigated a specific type of a long-distance referential dependency, backward anaphora, from a theoretical standpoint (Chapters 2 & 4), a perspective of language development (Chapter 4), real-time processing (Chapters 3 & 4) and computational modeling (Chapter 5). We subjected the child and the human parser to tests involving a structural constraint on backwards anaphora, Principle C, and in either case we saw that Principle C was respected as early as possible at the respective time scale. In the child language Principle C is functional already at the earliest tested age and before other constraints on coreference become functional, which provides strong support for the innateness of the constraint (or, more precisely, of the principles that underlie the constraint). With regard to the parser, Principle C is respected at the earliest stage of processing. This fact can be added to the growing list of findings that suggest the parser is capable of building complex and grammatically accurate representations in real time (Stowe 1986, Crain & Fodor 1985, Traxler & Pickering 1996, Pablos & Phillips 2005, Yoshida 2004 among others), which undermines the view whereby all the parser has access to are oversimplified, rough-and-ready strategies (Bever 1970).

In what follows I would like to discuss some of the issues that either have been brought up earlier and deserve some more attention or have not been touched upon.

Structural vs. Non-structural Constraints

In Chapter 3 I argued that structural information may have an advantage over other types of information during parsing for reasons of efficiency, as the parser may restrict its search space based on the structural prediction sooner than or simultaneously with when it can do so based on other types of information. Upon encountering a cataphoric pronoun, for example, the parser may be able to assess whether an upcoming NP could be a valid antecedent based on the structural position of that NP before it can do so based on its morphological congruency: whether the NP is ruled out by Principle C can be predicted top-down, unlike the gender/number of that NP. (Recall also that in our long-distance context usefulness of other sources, e.g. frequency information, is largely reduced.) This position explains why structural information is used by the parser at the earliest stages of processing without stating that such information has an across-the-board architectural priority, and can be contrasted with the syntax-first approach to parsing (Frazier 1978), according to which only structural constraints can be taken into consideration at the earliest stage of processing.

Our experiments cannot distinguish between these views. In Chapter 4 we compared processing of a pair of constraints on backwards anaphora in Russian, a

structural Principle C and a discourse *poka*-constraint, and found that whereas the parser never considers an NP that is subject to Principle C as a candidate antecedent at any stage of processing, it does initially consider an NP that is later ruled out by the *poka*-constraint. Under the syntax-first approach this result is due to the fact that only structural constraints like Principle C apply at the initial stage of processing, whereas other types of information are taken into account at a later stage. For the view in which structural information need not have an architectural priority over other types of information differences in the time course of application of Principle C and the *poka*-constraint should be primarily attributed to another difference between the constraints. This difference concerns the point in the sentence at which the critical antecedent can be determined to be unambiguously subject to the constraint. In our experiment antecedents that were illicit due to Principle C could be unambiguously classified as such solely based on the left context, i.e. material that linearly precedes the antecedent. On the other hand, the left context was not sufficient to judge whether a certain antecedent was subject to the *poka*-constraint. Hence, the delay in application of the *poka*-constraint may be due to the fact that at the time when the critical antecedent is encountered it cannot be unambiguously judged as being subject to the constraint.[138] If so, it remains possible that a non-structural constraint that can be detected based on the left context would apply at the earliest stage of processing.

The Algorithm for Building Referential Dependencies in Real Time

In Chapter 5 I proposed an algorithm for constructing referential dependencies in the course of incremental structure building. The algorithm can in fact be seen as having two parts that are largely independent of each other.

One subpart of the algorithm that has a life of its own is the implementation of a 'constrained Active Search' for sentences with backwards anaphora, i.e. the ability to actively anticipate an antecedent for a cataphoric pronoun in positions that are binding-accessible, while skipping positions that are subject to Principle C. This is achieved by requiring that the parser constantly (i.e. at each new step) checks whether an antecedent for a dependent element is available, from the moment the dependent element is added to the structure until a licit antecedent is found. Immediate respect of Principle C is implemented by requiring that such re-checking is only performed in the domains that are not c-commanded by the dependent element. The only piece of grammatical knowledge that is used for this part of the algorithm is that a dependent element cannot c-command its antecedent;

[138] This alternative is especially plausible in light of findings that syntactic constraints may be temporarily 'violated' (or, perhaps more appropriately, left unnoticed at the early stage of processing) when the relevance of a constraint cannot be unambiguously detected until a later point in the sentence. In Japanese, for example, the parser may violate an island constraint by postulating a trace inside a relative clause, because at the time when a trace is posited the information available to the parser is not sufficient to classify the embedded clause as a relative clause (the parser takes the embedded clause to be a complement clause by default).

hence, this part is compatible with any approach that formulates Principle C in those terms.

The other part of the algorithm concerns implementation of the remaining constraints of the Binding Theory and constraints on control dependencies. This part is somewhat less general, since it is based on a particular view in which reflexives and PROs are considered to be traces of movement (Hornstein 1999, 2001). The knowledge base of the parser is formulated in terms that are specific to Hornstein's account, e.g. the parser does not have a separate category PRO and constraints that are specific of PRO.

In the remaining space I would like to clarify the relation between such a grammar and the parsing algorithm that is proposed in Chapter 5. First of all, the grammar in which derivations proceed right-to-left, cannot straightforwardly predict the behavior of the parser that proceeds through the sentence left-to-right.[139] This can be illustrated using the example (1).

(1) Before leaving the party John kissed Mary.
    a. Before PRO_i leaving the party $John_i$ kissed $Mary_k$.
    b. Before PRO_{*k} leaving the party $John_i$ kissed $Mary_k$.

Grammatically, in order to conclude that PRO in (1) is controlled by *John* as shown in (1)a and not by *Mary* as in (1)b one needs to compare derivations at the point when the object *Mary* merges with the main verb *kissed* (see Appendix for more details). In other words, it is possible to decide which NP is the controller for PRO only after the main verb and the object are known in (1). This is at odds with the experimental results by Sturt et al.'s (2004), which show that speakers can associate PRO with the main subject at least as soon as that subject is encountered and well in advance of when the main predicate becomes available.

Thus, a parsing model that can accurately reproduce the behavior of the human parser on structures like (1) must make use of information that is additional to what is provided by the grammar. This extra-grammatical information is a way of compensating the parser for the incompleteness of the input and serves the goal of enabling the parser to accurately and immediately establish dependencies based on a partial input.

In general, the easiest way of compensating the parser for a yet-unseen information is by making it possible for the parser to keep a list of various templates and the associated information on how different structures should be processed (Goldberg 2003). In our case, that would be to list every possible structure that has a reflexive or a PRO and specify which element is the valid

[139] Needless to say that Hornstein's theory or, for this matter, almost any other grammatical theory, was not designed to be a theory of parsing. The discussion here should not be taken as an attempt of judging grammatical theories, but rather as consideration of whether and how such grammatical theories and their insights could be related to parsing.

antecedent in each type of structure. The algorithm that is developed here for building referential dependencies does not list each structure separately. Instead, the algorithm is allowed to make use of the parser's search procedure (which is by definition extra-grammatical) for the purposes of establishing referential dependencies; the antecedent for the trace is the first element on the parser's search path. So far, however, I have not said anything about the nature of such a mechanism.

First, the observation that the antecedent for the trace is the first suitable element on the parser's search path in fact may be no more than a super-strategy that is a superficial unification of individual structure-specific strategies. If so, the availability of such descriptive generalization is merely coincidental, does not reflect any deep properties of referential dependencies and, consequently, does not have any advantage over having a list of individual structures and their properties. Under this view, our parsing algorithm belongs to the class of strategy-based parsers (Bever 1970, Kimball 1973).

Alternatively, rather than being a meaningless descriptive generalization, the extra-grammatical search procedure can be viewed as a restatement of grammatical principles in terms that are understandable to the parser. Since right-to-left derivations that are used by the grammar are inapplicable in parsing, translation of grammatical principles in parsing-friendly terms is needed. The search procedure proposed in the algorithm can be viewed as an outcome of this translation that is performed strictly in accordance with the grammatical principles. Under this view of the search procedure, our model belongs to the class of principle-based parsers (e.g. Pritchett 1992, Berwick 1991, Gorrell 1995).[140]

Finally, the search procedure that is used to derive the antecedent for the trace could turn out to be valid in explaining other parsing phenomena and be motivated by considerations that are independent of referential dependencies. If constraints on referential dependencies could be explained by an independently motivated parsing procedure, that would point in the direction that left-to-right incremental considerations play a fundamental role in 'devising' grammatical properties of such dependencies (see Hawkins 1994, Phillips 1996 for similar claims). This would provide support for the idea that the parser and the grammar are tightly connected and, possibly, that grammatical derivations proceed in the directions that are imposed by the parser, i.e. left-to-right (e.g. Phillips 1996, 2003). In our case, only some parts of the search procedure were based on independent evidence (e.g. the fact that the tree search starts from the most

[140] Given that different languages have different grammars, an appropriate 'translation' of the right-to-left grammatical principles into their left-to-right parsing equivalents would require tailoring the search procedure differently for different languages. For example, grammars of different languages may differ on whether anaphors require an antecedent for the anaphor to be the subject of the clause (the so-called subject-oriented anaphors). In a language with subject-oriented anaphors this fact will need to be reflected in the search procedure.

recently added node, rather than from the root of the tree). Other parts of the search procedure were shaped specifically to fit referential dependencies (e.g. the left-edge search only sees the nodes along the main spine of the tree). It remains to be seen whether those parts are justifiable more generally and whether search principles that are proposed on the basis of referential dependencies can account for the parser's behavior with respect to other type of phenomena.

Unsurprisingly, additional investigation is required in order to distinguish between these alternatives. I will nevertheless stop here and leave these and other issues that were brought up but not resolved in the course of this dissertation for future investigation.

REFERENCES

Allen, J. and M. S. Seidenberg (1999). The emergence of grammaticality in connectionist networks. Approaches to Language: Proceedings of the 28th Carnegie Symposium on Cognition. B. Macwhinney, Erlbaum.

Aoshima, S., C. Phillips and A. Weinberg (2003). Active filler effects and reanalysis: a study of Japanese wh- scrambling constructions. University of Maryland Working Papers in Linguistics. **12:** 26-48.

Aoshima, S., C. Phillips and M. Yoshida (2005). The source of the bias for longer filler-gap dependencies in Japanese. CUNY conference for sentence processing, University of Arizona, Tucson.

Aoun, J. and Y.-h. A. Li (1989). "Scope and constituency." Linguistic Inquiry **20**: 141–172.

Ariel, M. (1988). "Referring and Accessibility." Jounal of Linguistics **24**: 65-87.

Ariel, M. (1990). Accessing Noun-Phrase Antecedents, New York, NY: Routledge, Chapman & Hall.

Avrutin, S. (1999). Development of the syntax-discourse interface. Dordrecht, Kluwer Academic Publishers.

Avrutin, S. and E. Reuland (2003). Backward Anaphora and Tense Interpretation. St. Petersburg University Anniversary volume. L. Verbitskaya.

Avrutin, S. and K. Wexler (1999). "Children's knowledge of subjunctive clauses: obviation, binding,a dn reference." Language acquisition **8**(1): 69-102.

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

Baltin, M. and P. M. Postal (1996). "More on Reanalysis Hypothesis." Linguisic Inquiry **27**(1): 127-145.

Berwick, R. C. (1991). Principles of principle-based parsing. Principle-based Parsing: computation and Psycholinguistics. R. C. Berwick, Kluwer Academic Publishers**:** 1-37.

Bever, T. G. (1970). The cognitive basis for linguistic structures. Cognition and the Development of Language. R. Hayes. New York, Willey & Sons, Inc.

Boeckx, C., N. Hornstein and J. Nunes (2005). Overt copies in reflexive and control structures: a movement analysis.

Bolinger, D. (1979). Pronouns in discourse. Discourse and syntax: Syntax and semantics. T. Givon. New York, Academic Press. **12**.

Borik, O. (2002). Aspect and reference time, Utrecht University**.**

Brody, M. (1994). Phrase structure and dependence. UCL working papers in linguistics, Department of Phonetics and Linguistics, University College London. **6:** 1-33.

Burzio, L. (1991). "The morphological basis of anaphora." Journal of Linguistics **27**: 81-105.

Cain, B. and J. Gair (2000). Dhivehi (Maldivian). Munich, Lincom Europa.

Chafe, W. (1987). Cognitive constraints on information flow. Coherence and grounding in discourse. R. S. Tomlin. Amsterdam, John Benjamins.
Chien, Y.-C. and K. Wexler (1990). "Children's knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics." Language Acquisition **1**(3): 225-295.
Chomsky, C. (1969). The acquisition of syntax in children from 5 to 10. Cambridge, Mass.,, M.I.T. Press.
Chomsky, N. (1965). Aspects of the theory of syntax. Cambridge, Mass., MIT Press.
Chomsky, N. (1981). Lectures on government and binding. Dordrecht, Foris Publications.
Chomsky, N. (1986). Barriers. Cambridge, Mass., MIT Press.
Chomsky, N. (1995). The minimalist program. Cambridge, Mass., MIT Press.
Cowart, W. and H. S. Cairns (1987). "Evidence for an anaphoric mechanims within syntactic processing: Some reference relations defy semantic and pragmatic constraints." Memory & Cognition **15**(4): 318-331.
Crain, S. (1991). Language acquisition in the absence of experience. Behavioral & Brain Sceinces. **4:** 597-650.
Crain, S. and J. D. Fodor (1985). How can grammars help parsers. Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives. D. R. Doaety, Lauri Karttunen, & Arnold M. Zicky, Cambridge University Press**:** 94-128.
Crain, S. and C. McKee (1986). Acquisition of structural restrictions on anaphora. North Eastern Linguistic Society 16, University of Massachusetts, Amherst.
Crain, S. and R. Thornton (1998). Investigations in Universal Grammar: A Guide to Experiments in the Acquisition of Syntax and Semantics. Cambridge, MA., MIT Press.
Elbourne, P. (2005). "On the acquisition of Principle B." Linguistic Inquiry **36**(3): 333-366.
Elman, J. L. (1993). "Learning and development in neural networks: the importance of starting small." Cognition **48**: 71-99.
Enç, M. (1987). "Anchoring Conditions for Tense." Linguistic Inquiry **18**: 633-657.
Evans, G. (1980). "Pronouns." Linguistic Inquiry **11**: 337–362.
Fauconnier, G. (1985). Mental spaces: Aspects of meaning construction in natural language. Cambridge, MA, MIT Press.
Ferreira, F. and C. J. Clifton (1986). "The independence of syntactic processing." Journal of Memory and Language **25**: 348-368.
Fillmore, C. J. (1977). The case for case reopened. Grammatical Relations. P. Cole and J. M. Sadock. New York, Academic Press. **8:** 59-83.
Fodor, J. D. (1978). "Parsing strategies constraints on transformations." Linguistic Inquiry **9**(3): 427-473.

Frazier, L. (1978). On comprehending sentences: syntactic parsing strategies. PhD thesis, University of Connecticut**.**

Frazier, L. and C. J. Clifton (1989). "Successive cyclicity in the grammar and the parser." Language and Congintive Processes **4**(2): 93-126.

Gazdar, G. (1982). Phrase structure grammar. The nature of syntactic representation. P. Jacobson and G. Pullum. Dordrecht, D. Reidel Publishing Company**:** 131-186.

Gibson, E. (1991). Computational theory of human linguistic processing: Memory limitations and processing breakdown, Carnegie Mellon University**.**

Givon, T. (1983). Topic continuity in discourse: An introduction. Topic continuity in discourse: A quantitative cross-language study. T. Givon. Amsterdam, John Benjamins.

Goldberg, A. E. (2003). "Constructions: a new theoretical approach to language." Trends in Cognitive Sciences **7**(5): 219-224.

Gordon, P. (1996). The truth-value judgment task. Methods for assessing children's syntax. D. McDaniel, C. McKee and H. Cairns. Cambridge, MA, MIT Press.

Gordon, P., R. Hendrick, K. Ledoux and C. L. Yang (1999). "Processing of reference and the structure of language: An analysis of complex noun phrases." Language and Cognitive Processes **14**: 353-379.

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

Gorrell, P. (1995). Syntax and Parsing. Cambridge, UK, Cambridge University Press.

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

Guasti, M. T. and G. Chierchia (1999/2000). "Backward and forward anaphora: Reconstruction in child language." Language acquisition **8**(2): 129-170.

Haegeman, L. M. V. (1992). Theory and description in generative syntax : a case study in West Flemish. Cambridge England ; New York, NY, Cambridge University Press.

Harris, C. L. and E. A. Bates (2002). "Clausal backgrounding and pronominal reference: A functionalist approach to c-command." Language and cognitive processes **17**(3): 237–269.

Hawkins, J. A. (1994). A performance theory of order and constituency. Cambridge ; New York, Cambridge University Press.

Heim, I. (1982). The semantics of definite and indefinite noun phrases. PhD thesis, University of Massachusetts, Amherst**.**

Heim, I. (1983). File Change Semantics and the Familiarity Theory of Definiteness. Meaning, use and the interpretation of language. R. Bauerle, C. Schwarz and A. von Stechow. Berlin, Walter de Gruyter**:** 164-190.

Heim, I. (1993). Anaphora and semantic interpretation: a reinterpretation of Reinhart's approach, University of Tubingen.

Hirst, W. and G. A. Brill (1980). "Contextual aspects of pronoun assignment." Journal of Verbal Learning and Verbal Behavior **19**: 168-175.
Hornstein, N. (1998). "Movement and chains." Syntax **1**(2): 99-127.
Hornstein, N. (1999). "Movement and control." Linguistic Inquiry **30**: 69-96.
Hornstein, N. (2001). Move! : a minimalist theory of construal. Malden, Mass., Blackwell.
Inoue, A. and J. D. Fodor (1995). Information-paced parsing of Japanese. Japanese Sentence Processing. R. Mazuka and N. Nagai. Hillsdale, NJ, Lawrence Erlbaum Associates**:** 9-36.
Kam, X.-N. C., I. Stoyneshka, L. Tornyova, W. G. Sakas and J. D. Fodor (2005). Non-robustness of syntax acquisition from n-grams: A cross-linguistic perspective. CUNY conference. Phoenix, AZ.
Karttunen, L. (1976). Discourse referents. Syntax and semantics. J. D. McCawley. New York, Academic Press. **7**.
Kayne, R. S. (1984). Connectedness and binary branching. Dordrecht, Foris.
Kazanina, N. and C. Phillips (2001). Coreference in Child Russian: Distinguishing Syntactic and Discourse Constraints. 25th annual Boston University Conference for Language Development, Cascadilla Press.
Kazanina, N. and C. Phillips (2005). A Developmental Perspective on the Imperfective Paradox.
Kennison, S. M. (2003). "Comprehending the pronouns her, him, and his: Implications for theories of referential processing." Journal of Memory and Language **49**: 335-352.
Kiguchi, H. and N. Hornstein (2005). PRO gate and sideward movement. UMCP.
Kimball, J. P. (1973). "Seven principles of surface structure parsing in natural language." Cognition **2**: 15-47.
King, T. H. (1995). Configuring Topic and Focus in Russian. Stanford, CA, CSLI Publications.
Klima, E. S. (1965). Studies in diachronic transformational syntax. Thesis Ph D -- Harvard University, Klima**.**
Kuno, S. (1987). Functional syntax: Anaphora, Discourse and Emphathy. Chicago, Chicago University Press.
Lakoff, G. (1988). Cognitive Semantics. Meaning and Mental Representations. U. Eco, M. Santambrogio and P. Violi. Bloomington, Indiana University Press**:** 119-154.
Langacker, R. (1986). "An Introduction to Cognitive Grammar." Cognitive Science **10**: 1-40.
Langacker, R. W. (1987). Foundations of Cognitive Grammar, Vol 1. Stanford, CA, Stanford University Press.
Langacker, R. W. (1991). Foundations of cognitive grammar, vol 2: Descriptive application. Stanford, Standford University Press.
Larson, R. (1988). "On the double object construction." Linguistic Inquiry **19**: 335-91.
Lasnik, H. (1976). "Remarks on coreference." Linguistic Analysis **2**(1-23).

Lasnik, H. (1989). Essays on anaphora. Dordrecht, Kluwer.
Lasnik, H. (1991). On the necessity of binding conditions. Principles and Parameters in Comparative Grammar. R. Freidin, MIT Press**:** 7-28.
Lee, M.-W. E. (2004). "Another Look at the Role of Empty Categories in Sentence Processing (and Grammar)." Journal of Psycholinguistic Research **33**(1): 51-73.
Lee, M.-W. E. and J. N. Williams (2005). The role of binding condition B in intrasentential pronoun resolution. CUNY conference on human sentence processing, University of Arizona, Tucson.
Lewis, R. and M. Nakayama (2002). Syntactic and positional similarity effects in the processing of Japanese embeddings. Sentence Processing in East Asian Languages. M. Nakayama. Stanford, CSLI Publications.
Lewis, R. L. (1996). "Interference in short-term memory: the magical number two (or three) in sentence processing." Journal of Psycholinguistics Research **25**: 93-115.
Lust, B., K. Loveland and R. Kornet (1980). "The development of anaphora in first language: Syntactic and pragmatic constraints." Linguistic Analysis **6**: 359-391.
MacDonald, M. C., N. J. Pearlmutter and M. S. Seidenberg (1994). "Lexical nature of syntactic ambiguity resolution." Psychological Review **101**: 676-703.
Marslen-Wilson, W. (1975). "Sentence perception as an interactive parallel process." Science **189**: 226-228.
Matthiessen, C. and S. Thompson (1988). The structure of discourse and 'subordination'. Clause combining in grammar and discourse. J. Haiman and S. Thompson. Philadelphia, J. Benjamins.
McCray, A. (1980). The semantics of backward anaphora. The 10th annual meeting of the North Eastern Linguistic Society., Ottawa, University of Ottawa.
Moens, M. and M. Steedman (1988). "Temporal ontology and temporal reference." Computational Linguisitics **14**: 15-28.
Mortensen, D. (2004). Two types of variable elements in Hmong anaphora. Pronouns as Epiphenomena. R.-M. Déchaine and M. Wiltschko, Oxford University Press.
Neidle, C., J. Kegl, B. Bahan, D. Aarons and D. MacLaughlin (1997). Rightward Wh-Movement in American Sign Language. Rightward Movement. D. Beerman, D. LeBlanc and H. van Riemsdijk. Amsterdam, John Benjamins**:** 247-278.
Nicol, J. and D. Swinney (1989). "The role of structure in coreference assignment during sentence comprehension." Journal of Psycholinguistic Research **18**(1): 5-19.
Nicol, J. L. (1988). Coreference processing during sentence comprehension. Thesis Ph D --Massachusetts Institute of Technology Dept of Brain and Cognitive Sciences 1988**.**

O'Grady, W. (1997). Syntactic Development. Chicago, University of Chicago Press.
Otsu, Y. (1981). Universal grammar and syntactic development in children : toward a theory of syntactic development. Thesis Ph D --Massachusetts Institute of Technology Dept of Linguistics and Philosophy 1981, Massachusetts Insitute of Technology**.**
Pablos, L. and C. Phillips (2005). Rich agreement cues argument structure in on-line processing of Basque. CUNY conference for sentence processing, University of Arizona.
Pesetsky, D. and E. Torrego (2001). T-to-C movement: causes and consequences. Ken Hale: a life in language. M. Kenstowicz. Cambridge, Mass., MIT Press**:** 355-426.
Philip, W. and P. Coopmans (1996). The Role of Lexical Feature Acquisition in the Development of Pronominal Anaphora.
Phillips, C. (1996). Order and Structure. Ph.D. dissertation, MIT**.**
Piattelli-Palmarini, M. (1980). Language and learning: The debate between Jean Piaget and Noam Chomsky. Cambridge, MA, Harvard University Press.
Pickering, M. (1993). "Direct association and sentence processing: a reply to Gorrell and to Gibson and Hicock." Language and Cognitive Processes **8**(2): 163-196.
Pinker, S. (1984). Language Learnability and Language Development. Cambridge, MA, Harvard University Press.
Prince, A. S. and P. Smolensky (1993). Optimality Theory: Constraint Interaction in Generative Grammar. Piscataway, New Jersey, Rutgers University Center for Cognitive Science.
Pritchett, B. L. (1992). Grammatical Competence and Parsing Performance. Chicago, IL, The University of Chicago Press.
Pullum, G. K. and B. C. Scholz (2002). "Empirical assessment of stimulus poverty arguments." The Linguistic Review **19**(1-2): 9-50.
Reali, F. and M. H. Christiansen (2003). Reappraising Poverty of Stimulus Argument: A Corpus Analysis Approach. 28th Boston University Conference on Language Development.
Regier, T. and S. Gahl (2004). "Learning the unlearnable: the role of missing evidence." Cognition **93**: 147-155.
Reinhart, T. (1983). Anaphora and Semantic Interpretation. Chicago, University of Chicago Press.
Reinhart, T. (2000). Strategies of anaphora resolution. Interface Strategies. H. Bennis, M. Everaert and E. Reuland. Amsterdam.
Reinhart, T. and E. Reuland (1993). "Reflexivity." Linguistic Inquiry **24**(4): 657-720.
Reinhart, T. M. (1976). The syntactic domain of anaphora. Thesis 1976 Ph D -- Massachusetts Institute of Technology Dept of Foreign Literatures and Linguistics**.**
Reuland, E. (2001). "Primitives of Binding." Linguistic Inquiry **32**(3): 439-492.

Reuland, E. and S. Avrutin (2005). Binding and Beyond: Issues in Backward Anaphora. Anaphora Processing: Linguistic, cognitive and computational modelling. A. Branco, T. McEnery and R. Mitkov, John Benjamins.
Ritter, E. and M. Wiltschko (2005). Anchoring events to utterances without tense. WCCFL, Simon Fraser University.
Rizzi, L. (1982). Issues in Italian syntax. Dordrecht, D. Reidel.
Rizzi, L. (1990). Relativized minimality. Cambridge, Mass., MIT Press.
Ross, J. R. (1967). Constraints on variables in syntax. Ph.D dissertation, MIT**.**
Saffran, J., R. Aslin and E. Newport (1996). "Statistical learning by 8-month-old infants." Science **274**: 1926-1928.
Safir, K. (2004). The syntax of anaphora. Oxford, Oxford University Press.
Schneider, D. A. (1999). Parsing and incrementality. Ph.D. dissertation, University of Delaware**.**
Seidenberg, M. S. and M. C. MacDonald (1999). "A probabilistic constraints approach to language acquisition and processing." Cognitive Science **23**(4): 569-588.
Solan, L. (1983). Pronominal reference: Child language and the theory of grammar. Boston, MA, Reidel.
Spivey-Knowlton, M. and J. Sedivy (1995). "Resolving attachment ambiguities with multiple constraints." Cognition **55**: 227-267.
Stabler, E. P. (1997). Derivational minimalism. Logical Aspects of Computational Linguistics. Retoré, Springer**:** 68-95.
Stowe, L. A. (1986). "Parsing WH-constructions: evidence for on-line gap location." Language and Cognitive Processes **3**: 227-245.
Sturt, P. (2003). "A new look at the syntax-discourse interface: The use of binding principles in sentence processing." Journal of Psycholinguistic Research **32**(2): 125-139.
Sturt, P., V. Lombardo and M. Betancort (2004). Evidence against (even mild) delay models of structure-building. AMLAP conference. Aix-en-Provence.
Tavakolian, S. (1977). Structural principles in the acquisition of complex sentences, University of Massachusetts, Amherst**.**
Tenenbaum, J. B. and T. L. Griffiths (2001). "Generalization, similarity and Bayesian inference." Behavioral and Brain Sciences **24**: 629-640.
Thornton, R. and K. Wexler (1999). Principle B, VP Ellipsis, and Interpretation in Child Grammar. Cambridge, Mass, The MIT Press.
Tomasello, M. (2004). Where does grammar come from? A debate on the nature of child language acquisition. BUCLD-29, Boston University.
Travis, L. (1984). Parameters and the effects of word order variation. PhD dissertation, MIT**.**
Traxler, M. J. and M. J. Pickering (1996). "Plausibility and the processing of unbounded dependencies: An eye-tracking study." Journal of Memory and Language **35**: 454-475.

Trueswell, J. and M. Tanenhaus (1995). Sentence Comprehension. Handbook in perception and cognition, volume 11: Speech, language and communication. P. Eimas and J. Miller, Academic Press**:** 217-262.

van Gompel, R. P. G. and S. P. Liversedge (2003). "The influence of morphological information on cataphoric prounoun assignment." Journal of Experimental Psychology: Learning, Memory, and Cognition **29**(1): 128-139.

van Hoek, K. (1997). Anaphora and conceptual structure. Chicago, University of Chicago Press.

Woolford, E. (1999). "More on the anaphor agreement effect." Linguistic Inquiry **30**(22): 257-87.

Yoshida, M. (2004). Relative clause prediction in Japanese. University of Maryland, College Park.