

What Can Visual World Eye-tracking Tell Us about the Binding Theory?

Jeffrey T. Runner

Kellan D. L. Head

This paper presents the results of a visual world eye-tracking experiment that tests two claims in the literature: that the Binding Theory (BT) is a set of “linked” constraints as in the classic BT (Chomsky 1981) and HPSG’s BT (Sag, Wasow, and Bender 2003); and that the BT applies as an initial filter on processing (Nicol and Swinney 1989, Sturt 2003). Our results instead support two different claims: that the constraint on pronouns and the constraint on reflexives are separate constraints that apply differently and with different timelines, in line with “primitives of binding” theory, Reuland (2001, 2011); and that neither constraint applies as an initial filter on processing, as proposed in Badecker and Straub (2002).

Keywords: visual world eye-tracking, Binding Theory, initial filter, defeasible filter, multiple constraints

1 Introduction

A question of central importance at the interface of the grammar and the language processing system is how grammatical constraints are deployed during sentence processing. This paper focuses on how the grammatical constraints of the syntactic Binding Theory (BT)—the structural constraints on reflexives and pronouns—apply during online processing. Our study is presented against a background literature proposing a variety of models for the application of the BT during processing. The Initial Filter approach (Nicol and Swinney 1989) suggests that the BT constraints constrain from the very beginning of processing which potential antecedents people consider during processing; the Defeasible Filter approach (Sturt 2003) posits that initially people consider only potential antecedents consistent with the BT, but may at a later stage of processing consider antecedents not sanctioned by the BT; and the Multiple Constraints approach (Badecker and Straub 2002) claims, instead, that the constraints of the BT apply alongside other processing constraints throughout processing. Using a novel visual world eye-tracking method which manipulates the gender of potential antecedents visually, we find clear evidence that listeners consider gender-matching potential antecedent NPs for reflexives and pronouns that match in gender regardless of whether they are licensed structurally by the BT, consistent with the Multiple Constraints view. We also consider how our results also bear on the formulations of the BT, favoring an approach that recognizes that the constraints of the BT apply differently for reflexives and pronouns, in particular appearing to be less robust for the latter.

Many thanks to editor Christopher Piñón and an anonymous reviewer for comments that helped to improve the paper. We also thank the research assistants who helped collect the data: Kimberly Morse, Amanda Baker, Joelle Mamon, and Seth Rosenblatt. Funding was provided by NSF BCS-1150337.

1.1 Syntactic Constraints on Binding

Different types of NPs display different biases with regard to sentence-internal antecedence, which are arguably based on syntactic structure. Reflexives must find an antecedent NP in a higher position in the same clause, as in (1a), whereas pronouns resist antecedent NPs in a higher position in the same clause, as in (1b):

- (1) a. Charles_j said [that Jim_i saw himself_{i/j/*k}]
 b. Charles_j said [that Jim_i saw him_{*i/j/k}]

The classic BT (e.g. Chomsky 1981) accounted for these facts with the two constraints and the definition in (2).

- (2) Principle A. A reflexive is bound in a local domain.
 Principle B. A pronoun is free in a local domain.
 Bound (not free) = coindexed with a c-commanding noun phrase

Principle A requires a reflexive to be bound (coindexed with a c-commanding phrase) in a local domain (roughly, a clause); and Principle B requires a pronoun to be free (not bound) in that same local domain. Thus, in (1a), *himself* must be bound by *Jim*, and not by *Charles* (by Principle A); and in (1b), *him* cannot be bound by *Jim*, but may be bound by *Charles* (by Principle B). Note that the ultimate antecedent of *him* in (1b) (be it *Charles* or some other sentence-external referent) is usually thought to be determined by other, non-syntactic, considerations. In what follows we will refer to potential antecedent NPs sanctioned by the BT (e.g. *Jim* in (1a) and *Charles* in (1b)) as “BT-compatible” or “BT-accessible,” and those not sanctioned by the BT (*Charles* in (1a) and *Jim* in (1b)) as “BT-incompatible” or “BT-inaccessible.”

Most syntactic frameworks assume something like the BT, though they differ on what the relevant structural relations are (e.g. phrase structure trees, argument structures, functional structures, etc.), what count as reflexives and pronouns for the BT, and whether the principles constraining reflexives and pronouns are linked or (partially) independent.

For example, in the lexicalist framework Head-driven Phrase Structure Grammar, binding is defined on argument structures. An argument structure is an ordered list of the arguments of a head, such that an item outranks every item to its right; “outrank” corresponds to syntactic prominence and thus is similar to c-command in the classic BT. On the assumption that a verb like *see* has an argument structure that includes at least the information in (3), that it has two arguments, one more structurally prominent (NP1) than the other (NP2), the HPSG BT in (4) (based on Sag, Wasow, and Bender 2003) also predicts the binding in (1).

- (3) ARG-ST of *see*: ⟨NP1, NP2⟩
 (4) Principle A: An outranked reflexive must be outranked by a coindexed element.
 Principle B: A pronoun must not be outranked by a coindexed element.

In (1a) NP2 (*himself*) is an outranked reflexive (there is an argument to its left in the argument structure of *see*), so it must be coindexed with one of its outranking elements, here NP1 (*Jim*); and in (1b), *him* is a pronoun and must not be outranked by any coindexed element (i.e. it cannot be coindexed with NP1, *Jim*). HPSG’s BT also makes no claims about the ultimate antecedent of *him* (as in (1b), it could be *Charles* or some sentence-external NP).

The “reflexivity” approach of Reinhart and Reuland (1993), which is also employed in Reuland’s more recent (2001, 2011) “primitives of binding” approach, builds on the intuition that the BT is about licensing reflexive predicates. Essentially if the intent is for two arguments to be coindexed, that needs to be marked. The relevant conditions are in (5):

- (5) A. A reflexive-marked predicate (a predicate containing a “self” word) is reflexive (it has coindexed arguments).
 B. A reflexive predicate (a predicate with coindexed arguments) is reflexive-marked (by a “self” word).

In (1a), the predicate *see* is reflexive because its two arguments are coindexed; and it is reflexive-marked (by *himself*), satisfying A and B. In (1b), however, the assumption is that a pronoun like *him* cannot reflexive-mark the predicate, so *see* cannot be reflexive (i.e. cannot have two coindexed arguments). Coindexing *him* with *Jim* would violate B since the predicate is not reflexive-marked.

While Reflexivity rules out binding (as coindexation) in a sentence like ‘Jim adores him,’ it does not rule out coreference; that is ‘Jim’ and ‘him’ could end up referring to the same individual. It turns out that under certain circumstances a pronominal object can be coreferential with the subject, as indicated by this classic example from Reinhart (1983) (cited in Reuland 2001, p. 448):

- (6) I know what Mary and Jim have in common. Mary adores him and Jim adores him too.

To account for this possibility the Reflexivity approach is supplemented by Rule I (from Reuland 2001, p. 448):

- (7) Rule I: NP A cannot corefer with NP B if replacing A with C, C a variable A-bound by B, yields an indistinguishable interpretation.

The intuition is that in (6), *Jim adores him* does not have the same interpretation as *Jim adores himself*, and thus, by Rule I, is allowed. Unlike for reflexives, besides the Reflexivity conditions, Rule I has to be computed to properly constrain pronoun reference. The Reflexivity approach, unlike the other two approaches, allows discourse-level considerations to play a role in the ultimate fate of the interpretation of object pronouns.

Summarizing, then, the classic BT of Chomsky (1981) and the HPSG BT (Sag, Wasow, and Bender 2003) treat the intrasentential coreference possibilities of pronouns and reflexives on a par: both are determined by the binding principles. However, the Reflexivity (and Reuland’s 2001, 2011 primitives of binding) theory distinguishes the licensing of reflexive interpretations from pronoun interpretations. In particular, Rule I must apply to determine whether coreference with a pronoun is licensed. It may be worth noting that Rule I is logically independent of the choice of BT. In principle, it could be paired with the classic Chomskyan BT or the HPSG one as well. It is simply a condition stating under what conditions coreference (as opposed to binding) is (dis-)allowed.

As mentioned above at some point in coming to an interpretation for pronouns, information beyond the local syntactic context must be examined. It is clear from both psycholinguistic and computational studies of pronoun interpretation that a variety of sentence-external features (beyond the restrictions placed on it by the BT) influence the interpretation of pronouns.

1.2 The BT during Processing

In a series of cross-modal lexical priming studies, Nicol and Swinney (1989) showed that only BT-compatible antecedents were primed immediately after the proform. For the reflexive in (8), only *doctor* was primed, whereas for the pronoun, only *skier* and *boxer* were primed. The

BT-incompatible antecedents (*boxer* and *skier* for the reflexive, *team* for the pronoun) were not primed.

- (8) The boxer told the skier that the doctor for the team would blame him/himself for the recent injury.

As Nicol and Swinney summarize, "... It appears that *initial perceptual processing activates all viable candidates (here, those that conform to grammatical constraints)* and the choice among the candidates is relegated to later, perhaps nonmodular (see Fodor 1983) language processing" (p. 19) [my emphasis]. This has become known as the "Initial Filter" view of how the BT applies during processing. The BT acts like goggles on the parser, only allowing it to "see" those NPs sanctioned by the BT principles.

Based on the results of several self-paced reading experiments, Badecker and Straub (2002) showed that gender-matching BT-incompatible antecedents affected processing. In an example like (9a), participants showed a latency increase when the BT-inaccessible but gender-matching reflexive antecedent *John* was present compared to *Jane*; and in (9b), the presence of the gender-matching local subject also triggered an increase in reading latency. These effects occurred after the reflexive or pronoun.

- (9) a. John/Jane thought that Bill owed himself another opportunity to solve the problem.
b. John thought that Bill/Jane owed him another opportunity to solve the problem.

Badecker and Straub suggest that their 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" (pp. 764–765) [my emphasis]. This has come to be known as the "Multiple Constraints" view that BT constraints apply simultaneously with other constraints on processing.

Sturt (2003) argued that these previous studies (Nicol and Swinney; Badecker and Straub) used methods not sensitive enough to tap into earliest moments of processing. This argument, and his subsequent eye-tracking during reading demonstration of the early effects of BT on reflexives, have been very influential. A number of subsequent studies have built on Sturt's study, regularly finding no evidence for the effect of BT-incompatible antecedents for reflexives. We now turn to reviewing this line of studies.

Sturt (2003) used gender-stereotyped nouns like *surgeon* to examine the online application of BT principle A. In that study and all of the follow-up studies, the BT-compatible antecedent was the gender-stereotyped NP, and the BT-incompatible antecedent was either a name or a pronoun; the reflexive matched or mismatched the stereotyped gender, and matched or mismatched the gender of the inaccessible antecedent. In his experiment 1, the BT-accessible antecedent was linearly closer to the reflexive than the BT-inaccessible antecedent, as (10) illustrates. *The surgeon* is the only BT-accessible antecedent for the reflexive; the main clause sentence subject (*he/she*) is BT-inaccessible for the reflexive.

- (10) Jonathan/Jennifer was pretty worried at the City Hospital.
He/she remembered that the surgeon had pricked himself/herself with a used syringe needle.

The basic result was that first-fixation durations at the reflexive were longer if the reflexive mismatched the stereotyped gender of the accessible antecedent; that is, *surgeon...herself* >

surgeon...himself. First-fixation durations showed no effect of the gender of the inaccessible antecedent (*he/she*). However, second pass reading times, which are the sum of fixations made on a region after that region has already been exited for the first time, did show an effect of the BT-inaccessible antecedent.

Sturt suggested his results, “show that processing can indeed be *affected both by a binding-accessible antecedent, and by a (binding-inaccessible) discourse focused antecedent*,” but that “the relevant constraints become operative at temporally distinct stages, and are not both simultaneously available at the earliest point in processing, in other words, the results support a model in which *Principle A acts as an early but defeasible filter*” (p. 558) [my emphasis]. Sturt called this the Defeasible Filter view.

In Sturt’s experiment 2, which otherwise was similar to his experiment 1, the inaccessible antecedent was placed as the object of a subject relative clause; here the inaccessible antecedent was linearly closer to the reflexive than the BT-accessible one, as in (11). Here the BT-accessible antecedent (*the surgeon*) is linearly more distant from the reflexive than the BT-inaccessible antecedent (*Jonathan/Jennifer*).

(11) Jonathan/Jennifer was pretty worried at the City Hospital.

The surgeon [who treated Jonathan/Jennifer] had pricked himself/herself with a used syringe needle.

As in his experiment 1, first-fixation durations were again longer when the gender-stereotyped accessible antecedent mismatched that of the reflexive. However, unlike experiment 1, there was no early or late effect of the BT-inaccessible antecedent.

Subsequent studies using gender-stereotyped nouns in a variety of constructions all find early effects of BT condition A, but variable effects of the inaccessible antecedent. Factors that varied across experiments included whether the inaccessible antecedent was a subject and pronoun (as in Sturt’s experiment 1), was the object of a subject relative clause (as in Sturt’s experiment 2), or was the subject of an object relative clause. Several studies have placed the inaccessible NP as subject of an object relative clause, with varying degrees of BT-inaccessible antecedent effects.

Xiang, Dillon, and Phillips (2009) used ERP methods to test sentences similar to Sturt’s experiment 2, but with object relative clauses, as in (12). Here the accessible antecedent is the subject of the main clause and the inaccessible is the subject of an object relative clause.

(12) The tough soldier [that Fred/Katie treated in the military hospital] introduced himself/herself to all the nurses.

They found a significant P600 when the reflexive mismatched the stereotyped gender of the accessible antecedent, replicating the effect of the BT; they also found some marginal effects of the “intrusive” antecedent (*Katie*), as an early (250–350) central anterior negativity, and a later (800–1000) posterior negativity. They argue that these marginal effects do not actually suggest that readers were considering the intrusive antecedent, but we mention this here for completeness.

Cunnings and Felser (2012) used sentences similar to Xiang et al.’s but with a pronominal subject of an object relative clause as the resuming inaccessible antecedent, as in (13):

(13) James/Helen has worked at the army hospital for years.

The soldier [that he/she treated on the ward] wounded himself/herself while on duty in the Far East.

They divided their participants into two groups based on working memory capacity. In addition to early stereotyped gender mismatch effect, they find a robust effect of inaccessible antecedent in lower working memory participants. Though the different effects in the different working memory groups raise almost as many questions as they answer, it is worth noting that the effect of the inaccessible antecedent was strongest when it was the pronominal subject of an object relative clause, as in (13).

Summarizing, all studies employing the gender-stereotype manipulation show early effects of the BT, as the main effect of the gender-stereotype mismatch. However, studies vary on whether they find any effect of the inaccessible antecedent. Table 1 summarizes the structural differences across studies and the variable effect of the inaccessible antecedent.

Table 1

Effects of BT-accessible and -inaccessible across studies

Study	Pronoun	Subj (of RC)	Obj (of RC)	Effect of Acc	Effect of Inacc
<i>Sturt 1</i>	√	√	x	√	√
<i>Sturt 2</i>	x	x	√	√	x
<i>Xiang et al.</i>	x	√	x	√	√?
<i>C&F (low mem.)</i>	√	√	x	√	√

1.3 Some Considerations for Our Study

Examining the variations in Table 1, a generalization begins to emerge. It appears that more “topical” inaccessible antecedents exert more of an influence on binding resolution. When the inaccessible antecedent is the subject of the sentence (Sturt exp 1), or is resumed by a pronoun (Sturt exp 1, Cunnings and Felser), or is the subject of an object relative clause (Xiang et al., Cunnings and Felser), it is more likely to have an effect on the processing of reflexive binding. Sentence subjects and pronouns have long been associated with topicality; and Roland, Mauner, O’Meara, and Yun (2012) show that the subjects of object relative clauses are more likely to be “discourse old” than the objects of subject relative clauses; that is, they are likely to pick up reference to something already topical in the discourse, potentially continuing it as a topic. Our study will manipulate the “topicality” of the inaccessible antecedent by relative clause type: it will either be the subject of a relative clause (= more topical) or the object of a relative clause (= less topical).

Sturt’s (2003) study, and subsequent ones examining the processing of reflexives are typically described as providing evidence for when the BT applies. However, since they only examine reflexives, they can only reveal how Principle A applies during processing. It follows that evidence for how Principle A applies may serve as evidence for how Principle B applies if the BT actually applies as a block, as in Chomsky’s (1981) classic BT and the Sag et al. (2003) HPSG BT. However, the Reflexivity theory (Reinhart and Reuland 1993) requires examining both binding (coindexation licensed by the BT) and coreference (Rule I) for pronouns. Thus, the latter approach potentially predicts differences in the processing of pronouns compared to reflexives. Evidence that reflexives and pronouns behave differently in processing might support the latter approach, or at least an approach that recognizes the potential relevance of non-structural information in the resolution of pronoun reference. To address the question of how similar the processing of reflexives and pronouns is, our study will include both proform types in order to compare them side by side.

The main evidence for the early effect of Principle A in these previous studies comes from increased latency at (or after) the reflexive when it mismatches the stereotypical gender

of the BT-accessible antecedent (*surgeon...herself* > *surgeon...himself*). For the argument to go through that BT Principle A applies as an initial filter, we must assume that earliest representation of the noun accessed includes stereotyped gender features (like lexically marked NPs like *he* or *princess*). This is a critical assumption if the conclusion is that BT applies as an initial filter on processing (see Nicol and Swinney 1989, and discussion in Sturt 2003). This assumption may be justified, but we will instead manipulate gender without gender-stereotype.

2 Experiment

2.1 Design, Procedure, and Materials

Our experiment was designed to answer the following questions: First, does the degree of “topicality” of the BT-incompatible antecedent (operationalized as subject vs. object of relative clause) affect its influence on binding? And second, does the influence of the inaccessible antecedent hold equally and with similar timelines for both reflexives and pronouns?

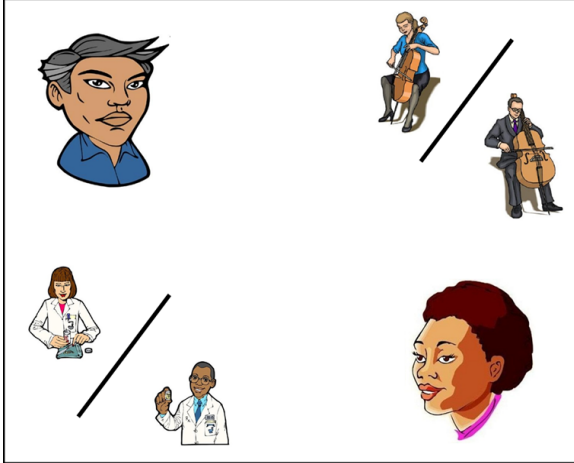
Our study employed visual world eye-tracking. Participants ($n=25$) listened to sentences containing (normed) non-gender-stereotyped occupations like ‘pharmacist’ while looking at a grid containing pictures of the mentioned characters and two distracter images. Gender was manipulated visually, using a picture of a male or a female e.g. pharmacist (see Figure 1). Half of the materials tested reflexives and the other half pronouns. The inaccessible antecedent (for reflexives) was either the subject of an object relative clause, or the object of a subject relative clause (manipulating the “topicality” of the BT-inaccessible). Participants clicked on the picture corresponding to the proform. The occupations were normed in a separate study which asked native English speakers if a particular occupation was more likely to be a male or a female (on a scale of 1-5); those occupations that scored 3 on average were included in the study. In addition, the images were normed (separately) by asking native English speakers to name each image; we used images for which the noun used in the experiment was the first choice in the norming study.

The recorded sentences were split into two parts and the border was the end of the relative clause. After the audio file containing the relative clause played, the display disappeared and a fixation cross appeared. Participants clicked on the cross and then the scene reappeared and the sentence continued. This was done to ensure that fixations when participants hear the proform were independent of their fixations while listening to the relative clause.

As is standard in visual world eye-tracking, we assume that attentional shifts to objects in the visual field are typically accompanied by a saccade; eye-movements to potential referents are closely time-locked to the input (Cooper 1974; Tanenhaus, Spivey-Knowlton, Eberhard, and Sedivy 1995); and potential referents are fixated in proportion to the likelihood of that referent being the intended target of the spoken materials (Allopenna, Magnuson, and Tanenhaus 1998). Thus, the proportion of looks to different objects provides an indication of which entities the hearer is considering as possible referents over time.

Figure 1

Sample visual display containing mentioned character (e.g. Molly in example (14a)), distractor character (e.g. Darrin), mentioned occupation (e.g. pharmacist) and distractor occupation (e.g. cellist). Scenes were constructed so that the genders of the mentioned and distractor occupations were different. Female and male versions of occupations are shown for illustration only.



The experiment manipulated three variables, each with two levels for a total of 8 conditions. Half of the experimental trials contained reflexives and half pronouns; on half of the trials the inaccessible antecedent matched and on half it mismatched the gender of the proform; and on half of the trials used subject relative clauses and half used object relative clauses. (14) and (15) illustrate the materials.

- (14) Reflexive: inaccessible (NP in RC) gender match (a) and mismatch (b):
- (a) ORC: The pharmacist(f) [that Molly met] drove herself to the party.
 - SRC: The pharmacist(f) [that met Molly] drove herself to the party.
 - (b) ORC: The pharmacist(f) [that Darrin met] drove herself to the party.
 - SRC: The pharmacist(f) [that met Darrin] drove herself to the party.
- (15) Pronoun: inaccessible (subject) gender match (a) and mismatch (b):
- (a) ORC: The pharmacist(f) [that Molly met] drove her to the party.
 - SRC: The pharmacist(f) [that met Molly] drove her to the party.
 - (b) ORC: The pharmacist(m) [that Molly met] drove her to the party.
 - SRC: The pharmacist(m) [that met Molly] drove her to the party.

In (14a) the inaccessible antecedent ('Molly') matches the gender of the reflexive, whereas in (14b) it ('Darrin') does not. And in the ORC conditions of (14) the inaccessible antecedent is the subject of the relative clause, and in the SRC conditions, it is the object. In (15) the subject of the sentence is the BT-inaccessible antecedent for the pronoun; in (a) it matches the gender of the pronoun; in (b) it mismatches. In both pronoun cases the RC-internal NP matches the gender of the pronoun, since it is the only BT-compatible antecedent in the sentence. We constructed 8 lists containing 4 lexicalizations of each condition plus 32 fillers, which were sentences of the same type as the experimental items but lacking proforms.

2.2 Results

2.2.1 Click Results We begin with the click responses. These responses indicate the final interpretation listeners assigned to the sentences. Let's first consider the predictions. On the assumption that something like the BT principles have an influence on listeners' final interpre-

tations of sentences containing pronouns and reflexives, we expect that in sentences containing reflexives (like (14)) participants should choose the subject of the sentence (and not the RC-internal NP) as antecedent; in sentences containing pronouns (like (15)) we instead expect participants to choose the RC-internal NP (and not the sentence subject) as antecedent.

Figure 2

Click responses. For reflexives, target is subject of sentence and inaccessible is RC-internal NP; for pronouns target is RC-internal NP and inaccessible is sentence subject.

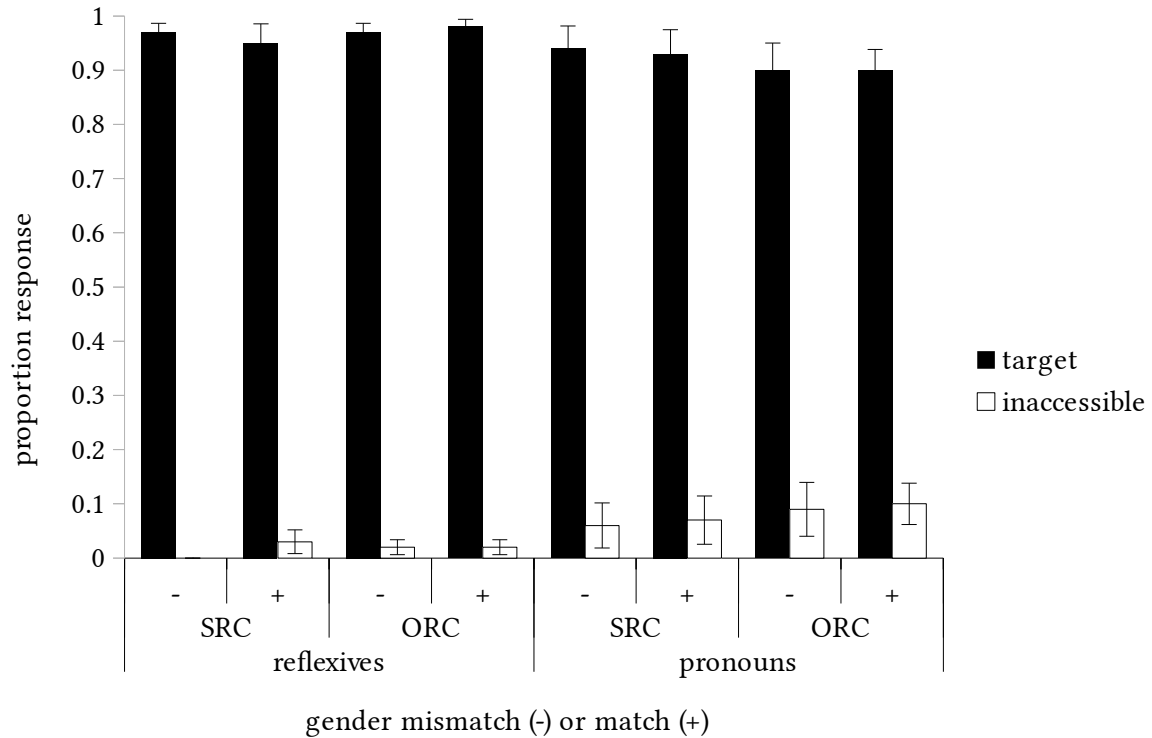


Figure 2 illustrates the proportions of BT-compatible target choices. As the figure shows there was overall a high proportion of BT-compatible responses. The mean proportions for all conditions were between .90 and .98. The only significant difference was the main effect of proform: overall participants were more likely to choose the BT-compatible target on the reflexive conditions than on the pronoun condition ($p < .03$). However, with all of the means at .90 or above, these results indicate that participants understood the sentences and that their final target choice closely followed the constraints of the BT. The fact that our participants clearly understood the sentences is relevant because in some of the previous eye-tracking studies it is difficult to know how well participants actually understood the sentences, leading to the possibility that the varying effects of the inaccessible antecedent could be partially due to the varying degrees to which participants misread or ultimately misunderstood the (fairly complex) sentences. Indeed, Sturt's experiment 1b (a small follow-up off-line study) used a subset of his experiment 1 materials to test readers' final antecedent choice; remarkably, on some conditions up to 40% of the final interpretations were inconsistent with the BT. This both suggests that the materials were complex enough to lead to error and that some of the relevant eye movements during reading may actually have come from readers not interpreting the sentences properly (see Dillon 2012). Our listeners' final interpretations were

overwhelmingly consistent with the BT, suggesting that the processing we observe through their eye movements will reflect a path to a final interpretation consistent with the grammar.

2.2.2 Fixation Data Before moving to the fixation data, let us reconsider the predictions. If the BT conditions A and B both apply from the earliest moments, there should be no gender effect, that is, no differences between trials where the BT-inaccessible matched or mismatched the gender of the proform, and this should hold for both pronouns and reflexives. However, if the BT conditions do not apply from earliest moments, or are separable we may find a gender effect early on, or a difference between how the reflexives and pronouns are processed.

In particular, the Initial Filter view predicts that there should be no early effect of the gender manipulation on participants' fixations; the Defeasible Initial Filter view also predicts no early effects of the gender manipulation, but allows for delayed effects; the Multiple Constraints view allows for the effects of both the BT principles and other factors that may be relevant for the processing of proforms. In addition, the classic and HPSG Binding Theories both predict the BT principles to be applied as a block, predicting that the structural effects should be parallel for reflexives and pronouns; however, the Reflexivity approach, which acknowledges additional constraints on pronouns, seems to predict that the processing of pronouns may be delayed or show the influence of other information in the context that reflexive binding should be insusceptible to.

Figure 3

Target advantage fixations for reflexive and pronoun trials where inaccessible mismatched or matched gender of proform

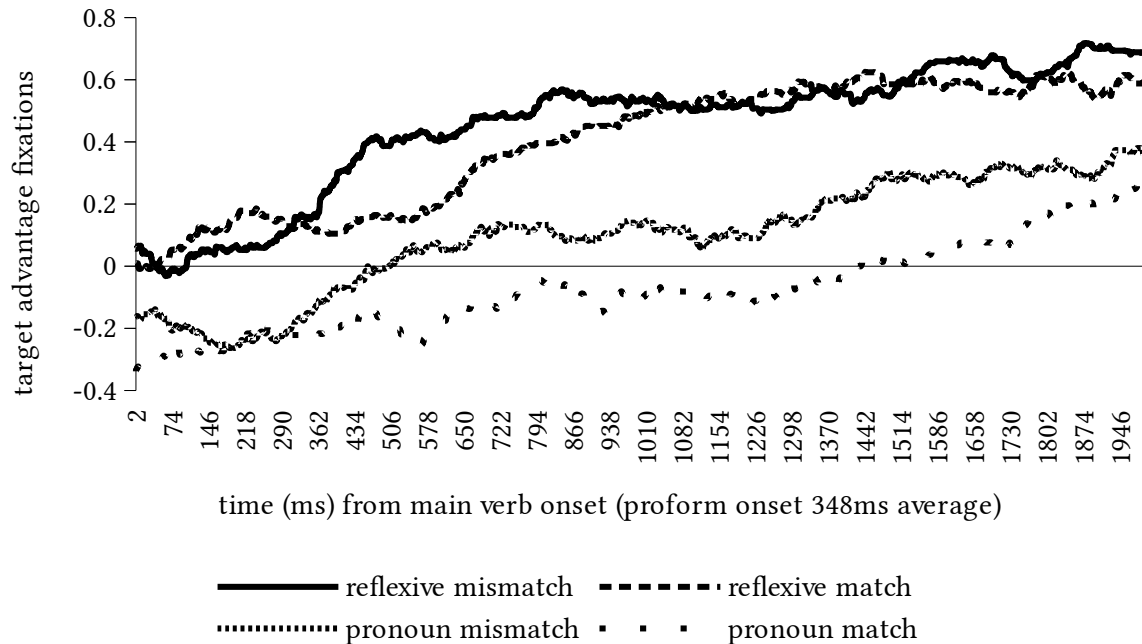


Figure 3 illustrates the “target advantage” fixations during trials in which participants chose the BT accessible target (as is clear from Figure 2, this includes over 90% of the data). Target advantage fixations are the proportion of fixations to the BT-accessible antecedent minus the proportion of fixations to the inaccessible antecedent at each 4ms time slice. If positive, it means listeners are looking at the accessible (target) more than the inaccessible; if negative, they are looking at the inaccessible more than the accessible.

It is immediately clear that reflexive and pronoun trials did not have a parallel effect on listeners' eye movements. Fixations to the target for reflexives (the blue lines in Figure 3) increased more quickly than those for pronouns (the red lines), indicating that listeners more quickly shifted their gaze to the BT-accessible antecedent (the target) when the proform was a reflexive. In addition, both reflexives and pronouns showed an early gender effect, indicated by the solid line being higher than the dotted line. That is, when the inaccessible antecedent matched the gender of the proform, listeners were slower to shift their gaze to the accessible (target) than when the inaccessible mismatched. Put another way, the presence of a gender-matching inaccessible antecedent drew fixations to that picture, reducing the target advantage. For both reflexives and pronouns the gender effect began early on, and for pronouns is rather more protracted. It is unexpected that the target advantage fixations would be negative for the pronoun trials before the onset of the proform. We return to this in the discussion.

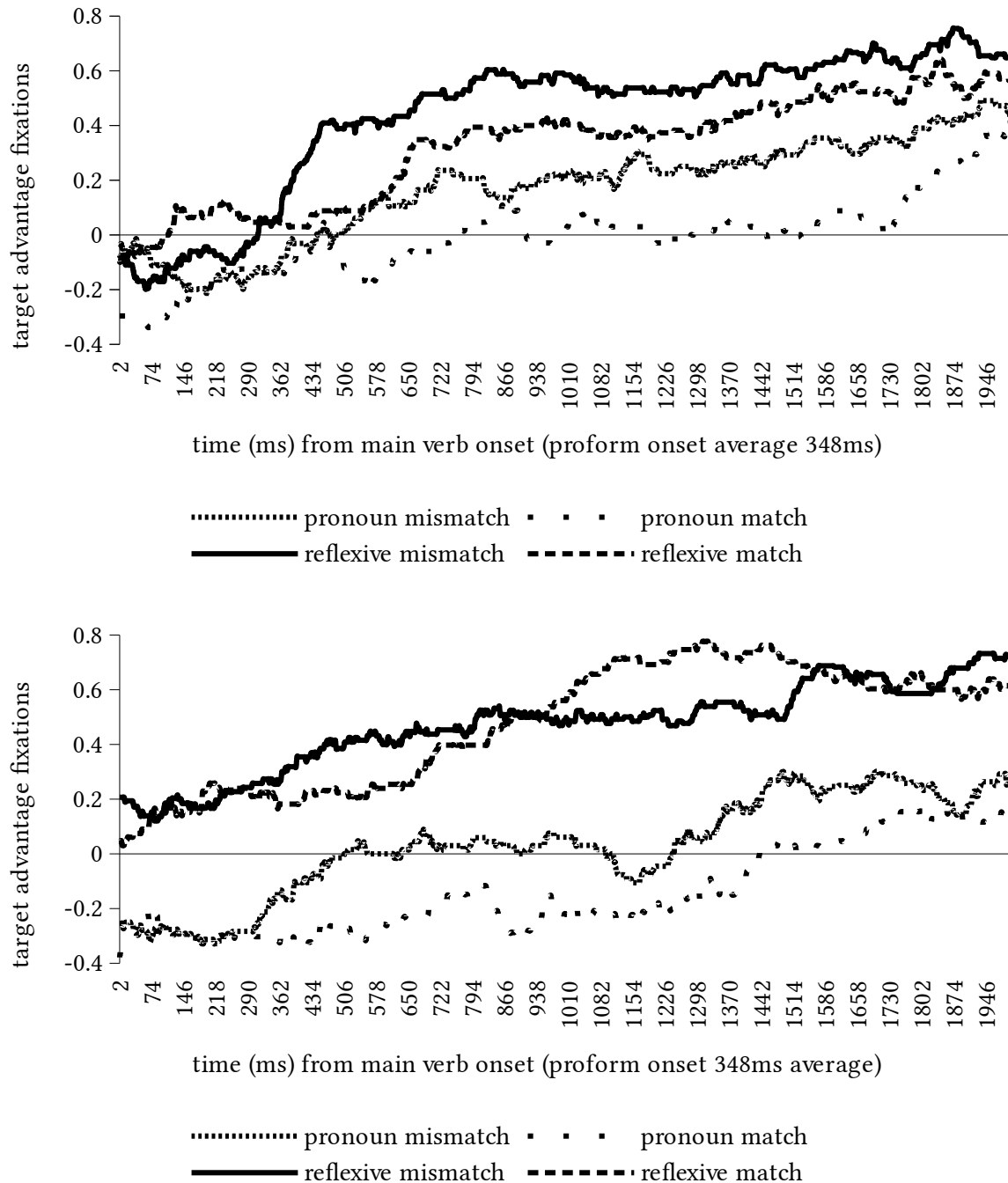
We now turn to the relative clause manipulation predictions. In object relative clauses the RC-internal NP is the subject and in subject relative clauses it is the direct object. Our prediction was that when the RC-internal NP is subject it should be more "topical" than when it is the object. Thus, we predict a larger gender effect in the object relative clause condition. No specific predictions were made for the pronoun conditions.

The top panel of Figure 4 illustrates the target advantage fixations on the subject relative clause conditions; the bottom panel illustrates those from the object relative clause conditions. Both figures show again that the target fixations on the reflexive conditions (blue lines) increased more rapidly than those on the pronoun conditions (red lines). And again, there is evidence of an early gender effect for both reflexives and pronouns (dotted line lower than solid line). With regard to the relative clause manipulation, the prediction was that on the object relative clause condition (where the inaccessible is the subject of the relative clause) there should be a larger gender effect than on the subject relative clause condition (where the inaccessible is an object). Contrary to that prediction, the gender effect appears to be more robust in the subject relative clause (for reflexives) than in the object relative clause. However, the relative clause manipulation did appear to have a large effect on fixations during the pronoun conditions. In particular, target advantage fixations were lower in the object relative clause condition than in the subject relative clause condition. We return to this in the discussion.

To statistically evaluate these effects we isolated a 500ms window starting at 500ms after main verb onset during the trials in which participants chose the BT accessible antecedent (sentence subject on reflexive trials, relative clause-internal NP on pronoun trials). This window was chosen because the average onset of the proform was at 348ms after the verb onset; thus our analysis window begins about 150ms after the average onset of the proform, about where signal-driven fixations are expected to begin to appear. The target fixations in this 500ms window were modeled in a linear mixed effects model with fixed effects of proform (reflexive, pronoun), gender (mismatch, match) and RC-type (SRC, ORC) and their interactions, and with subject and item as random intercepts; to control for over-sampling and state dependencies, we also included a control in the models that represents the fixation region on the previous time sample (Frank, Salverda, Jaeger, and Tanenhaus 2008).

Figure 4

Target advantage fixations in subject (top panel) and object (bottom panel) relative clause conditions

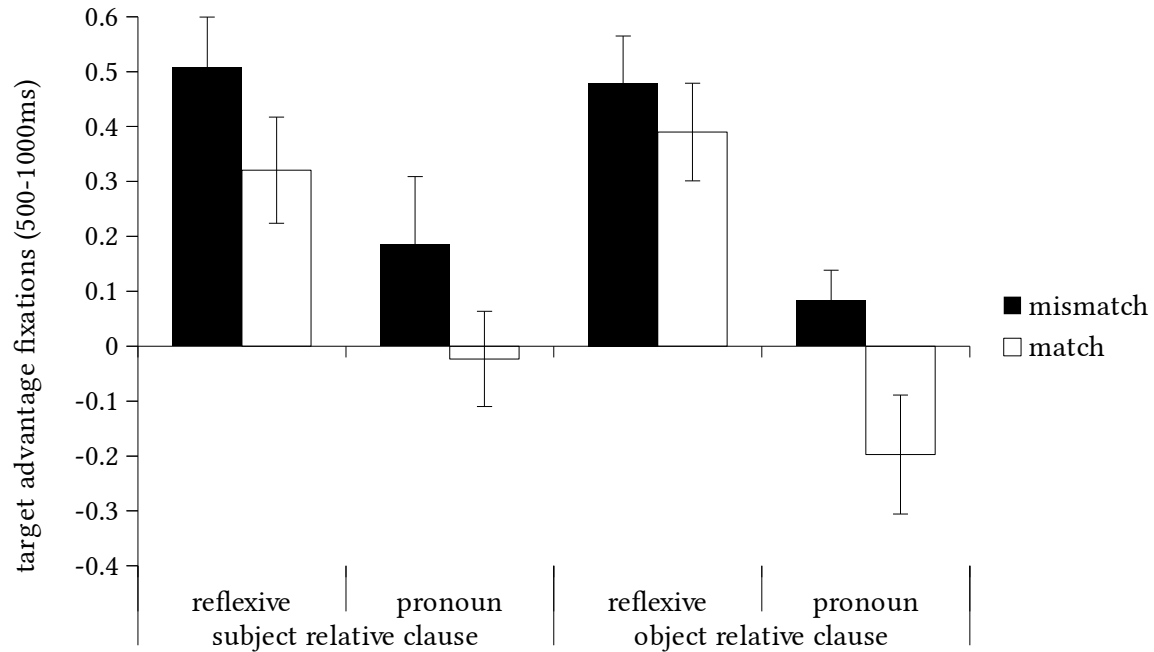


There were main effects of proform (more target fixations for reflexives than for pronouns) and gender (more target fixations when the inaccessible gender mismatched than when it matched the gender of the proform). In addition, the previous fixation control factor was highly significant. These main effects were qualified by two interactions. There was an interaction between proform and gender (the inaccessible NP's gender had a larger effect on pronoun trials than reflexive trials) and a three-way interaction among all three factors indicated that there was a gender effect in subject RCs for both reflexives and pronouns, whereas

in object RCs it is only present for pronouns and not reflexives. Figure 5 graphs target advantage fixations during this 500ms analysis window.

Figure 5

Target advantage fixations 500–1000ms after verb onset, with standard error bars



2.3 Discussion

Participants' final interpretations of the stimuli were overwhelming consistent with the predictions of all versions of the BT: on reflexive conditions participants clicked on the picture of the subject of the sentence; and on pronoun conditions they clicked on the picture of the relative clause-internal NP. All proportions were between .9 and .98, suggesting that regardless of condition, participants understood the sentences and chose the BT-compatible antecedent. There was a subtle, though statistically reliable, main effect of proform, such that on pronoun trials participants were somewhat less likely to choose the BT-accessible antecedent, though again, they still did so on 90+% of the trials. However, these interpretational differences did perhaps preview the very clear differences in fixation patterns in pronoun and reflexive trials.

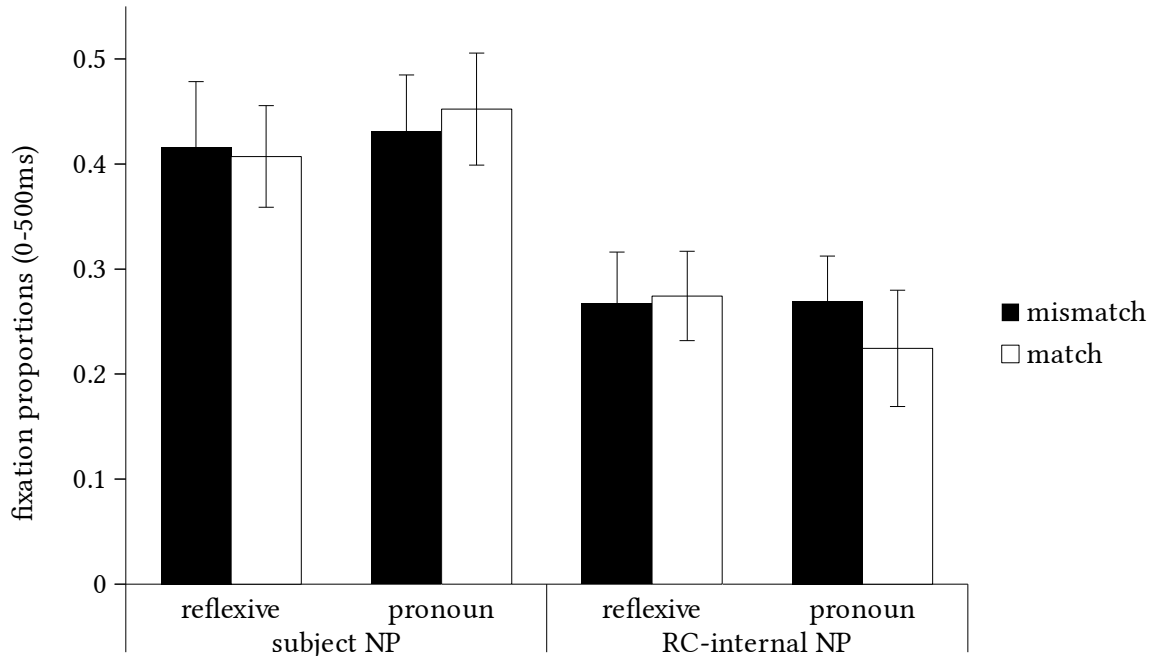
Fixations during reflexive trials indicated that participants considered the BT-accessible antecedent more than the BT-inaccessible one from soon after they heard the reflexive. Thus fixations during reflexive trials were consistent with the claim that the BT is at least one of the constraints affecting on-going reflexive interpretation. However, just as early as the effect to fixate the subject more than the RC-internal NP was the effect of the gender of that RC-internal NP: if it matched the gender of the reflexive participants were drawn to looking at it more than if it mismatched. This suggests that alongside the BT, which may account for the overall increase in looks towards the subject NP, there appears to be the consideration of other BT-incompatible NPs, from the earliest moments of reflexive processing. From the point of view of the approaches presented above, these results are inconsistent with both the Initial Filter and Defeasible Filter views, which would have predicted no gender effect, especially early on in processing. Instead, these results are consistent with the Multiple Constraints view. BT Principle A does appear to exert some influence early on, since participants

quickly started shifting their fixations to the subject; but this shift was delayed when the BT-incompatible NP matched the reflexive gender. Apparently, both BT-compatible and BT-incompatible antecedents were under consideration early on in reflexive processing.

Turning to pronouns, fixations during pronoun trials were different from those in reflexive trials in various ways. It appears that the overall effect of the BT—to resist the BT-incompatible NP as a potential antecedent—was delayed for pronouns compared to reflexives. Indeed, for some time early on participants were more likely to look at the BT-inaccessible antecedent than the BT-accessible one. This is puzzling. However, an important clue is that this effect begins before the onset of the proform, suggesting it cannot be due to pronoun processing *per se*. What could be driving this difference between pronoun and reflexive fixations? We believe the answer lies in a claim made by Koring, Mak, and Reuland (2012), that the subject of a verb is reactivated or retrieved when that verb is processed. In our materials, this would result in looks to the subject triggered by the main verb. In our previous figures, for reflexives, looks to the subject were graphed as looks to the BT-accessible antecedent; but for pronouns, looks to the subject were graphed as looks to the BT-inaccessible antecedent. To examine Koring et al.'s claim, if instead of graphing target advantage (BT-accessible fixations minus BT-inaccessible fixations), we graph looks to the subject, we would expect no difference between pronoun and reflexive fixations prior to the onset of the proform. Figure 6 illustrates fixations recast in these terms.

Figure 6

Fixations to subject NP and RC-internal NP during first 500ms after verb onset, with standard error bars



We can see in Figure 6 that before the onset of the proform participants were more likely to look at the subject of the sentence than the NP inside the relative clause. This is consistent with Koring et al.'s claim that the subject is reactivated during verb processing. Critically, the gender effect is not apparent at this early point, suggesting that this is a drive to look at the subject NP regardless of its gender. If these fixations were driven by the proform (somehow) we would expect the gender effect to appear here: more fixations to the subject NP when it

matches the gender of the proform. Given that these fixations appear prior to the onset of the proform, the lack of a gender effect is expected.

It is important to point out here that the observation that during pronoun conditions participants were more likely to look at the BT-inaccessible antecedent (the sentence subject) is not in conflict with the observation that pronouns were processed differently from reflexives. This is because the gender effect begins to appear quite early after the onset of the proform, as illustrated in Figure 3. The gender effect is a clear indicator that participants tried to interpret the proform with respect to the inaccessible antecedent, *qua* antecedent. Some of the overall depression of the pronoun fixation curves is probably due to the brief subject retrieval just mentioned, but the large and protracted gender effect strongly suggests that most of the difference is due to listeners' attempting to link the pronoun to the gender-matching BT-inaccessible antecedent.

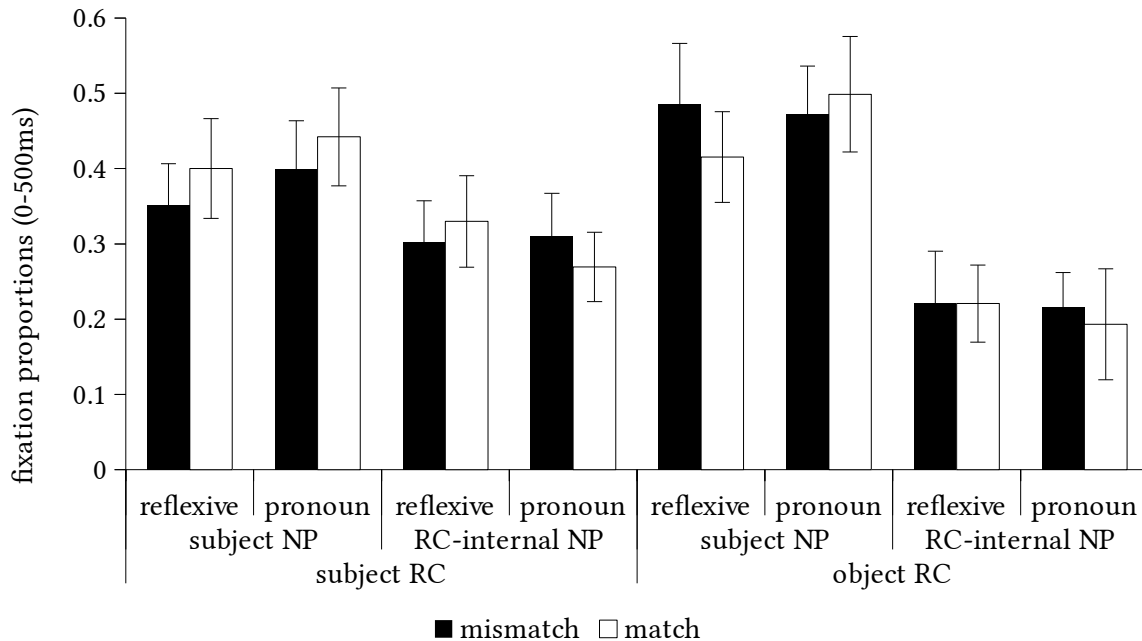
Turning now to the relative clause manipulation, given that previous studies seemed to show more of an inaccessible effect on reflexives in object rather than subject relative clauses, we had predicted something similar: that the RC-internal NP in an object relative clause, being a subject and thus potentially topical and/or expected to be discourse old (Roland et al. 2012), would have triggered a larger inaccessible gender match effect on our reflexive trials. This did not occur. Indeed the gender match effect for reflexives was somewhat smaller in object RCs than in subject RCs as is clear in Figures 4 and 5. However, the fact that the manipulation did not have the intended effect is not problematic, as the inaccessible gender effect was robust across conditions. Our goal was simply to increase the likelihood of demonstrating the inaccessible gender effect. The manipulation did not work, but the effect appeared nonetheless.

At the outset of this study, we did not have any predictions for how the relative clause manipulation would effect the pronoun conditions. The relative clause manipulation was designed to affect reflexives, based on previous studies of reflexives. However, interestingly, the relative clause manipulation did affect the pronoun fixations. We can speculate why this may have been the case, and leave for future research a more careful investigation of pronoun processing in these structures more carefully. The basic result was that in object relative clauses participants were slower to fixate the target overall and they showed a larger gender effect. One obvious possibility is the claim that object relative clauses are "harder" to process in general (Grodner and Gibson 2005). If this is the case, at least outside of a supporting context (see Roland et al. 2012), then some of the delay for pronouns may have simply been processing difficulty. For this account to work, we need to be able to assume that reflexive binding is not affected by relative clause processing complexity, since participants were not delayed at all on reflexive object relative clause conditions.

A second possibility is that the expectations for which referent was more likely to continue to be referred to may be influenced by the relative clause type. If, for example, after hearing an object relative clause, participants were more likely to expect reference to the subject NP, then reflexive trials would be comparatively easy since the subject NP turned out to be the correct NP; whereas pronoun trials would involve having to override the subject-preference of the object relative clause. Indeed, this possibility seems to be supported by the early fixations. In the main verb region, preceding the onset of the anaphor, participants fixated the subject NP more on the object RC conditions than on the subject RC conditions. This suggests that after listening to an object RC, participants were expecting reference to the subject of the sentence more than when they heard a subject RC. This is illustrated in Figure 7, which plots the fixations to the subject NP and RC-internal NP across the different relative clause types.

Figure 7

Fixations to subject NP and RC-internal NP during first 500ms after verb onset on subject relative clause (left) and object relative clause (right) conditions



It appears that after a subject relative clause, participants' expectations for ongoing reference to the sentence subject and the RC-internal NP were more balanced (though somewhat biased toward expecting reference to the sentence subject); however, after an object relative clause, there was a difference between fixations to the subject NP and fixations to the RC-internal NP, suggesting listeners were expecting ongoing reference to the subject rather than the RC-internal NP. Thus, independent of proform condition, participants expected the subject to be referred to further. On reflexive conditions this bias lines up well with the bias to consider the subject as antecedent for the reflexive; however, on pronoun conditions the bias to expect the subject to be the ongoing referent was in conflict with the expectation triggered by the BT to disregard the subject as a potential antecedent. These conflicting biases could explain the especially notable delay of pronoun reference resolution on object relative clauses.

We briefly mention two additional considerations pointed out by an anonymous reviewer that we cannot completely rule out as having contributed to the pattern of results we have presented. First, in the written language, the reflexive in English is temporarily ambiguous between being a reflexive or a pronoun, as e.g. *herself* begins with *her-*. If this ambiguity is picked up by participants in our study, it is possible that very briefly they misparsed *herself* as *her*, and thus programmed and launched fixations consistent with the pronoun interpretation. This could plausibly account for some of the fixations to the inaccessible antecedent when it matches the gender of the reflexive. However, it is worth pointing out that the materials were presented auditorily and it is less clear how auditorily ambiguous the first syllable of *herself* and *her* are. Additional analysis would be needed to rule this possibility out. Secondly, as is usual in visual world eye tracking studies we have interpreted fixations to pictures as indicating that the listeners' attention is being drawn to a particular picture because they are interpreting the linguistic form as referring to that picture. That is, they look

at the picture of the pharmacist because they have just heard the phrase *the pharmacist* or have heard a proform which they are interpreting as referring to the pharmacist. Since pronouns and reflexives in English are marked for gender, it is possible that when listeners hear a proform, its gender is “activated” and that activated features get more attention. This attention could then trigger fixations to items in the display with the same gender as the proform, even if they are not being considered as potential antecedents. The current study was not designed to test this possibility, so we must leave it to future research to tease apart differences in fixations due to attention to features like gender and attention to potential referents of the linguistic forms.

3 General Discussion

We now return to the main questions we designed our study to answer: First, does the degree of “topicality” of the BT-incompatible antecedent (operationalized as subject vs. object of relative clause) affect its influence on binding? And second, does the influence of the inaccessible antecedent hold equally and with similar timelines for both reflexives and pronouns?

The answer to Question 1 is clearly “no,” at least for reflexives, which was what the manipulation was designed for. The relative clause manipulation did affect pronoun processing, though it is not clear to what degree that had to do with topicality. The answer to Question 2 is also “no.” There was a larger and longer-lasting effect for pronouns.

In addition, our study addressed the question of the “timing” of the application of BT constraints during processing. Our results are in line with Badecker and Straub’s (2002) Multiple Constraints approach. Like their study, ours found early effects of gender-matching inaccessible antecedents on processing. This is in conflict with Sturt (2003) and Xiang et al. (2009), who used the gender-stereotype manipulation but found no early effects of inaccessible antecedents for reflexives (and did not examine pronouns).

Our study also found that the effect of gender-matching inaccessible antecedents was more robust for pronouns than reflexives. This bears on the question of whether the two BT conditions should be treated as a single constraint (applying in a block) or their effects should be treated as emerging from separate constraints (the BT + Rule I). Our results support treating them separately. In particular, our results are consistent with the Reflexivity and Primitives of Binding (Reuland 2001, 2011) approaches in which one set of conditions determines the distribution of coindexation, and an additional rule, taking into account discourse-level information, determines coreference for pronouns. That approach predicts more complexity in processing pronouns and the need to have access to additional features of the context. This is consistent with our findings, showing that pronoun reference was resolved more slowly and was more susceptible to gender-matching BT-inaccessible antecedents.

Though our results are mostly consistent with an approach like the Primitives of Binding approach, which requires an additional constraint (Rule I) to apply in order to finally determine pronoun interpretation, as pointed out above, in principle either of the other two types of BT could incorporate something like Rule I. So, our results really support an approach that recognizes additional constraints on pronoun binding, beyond those relevant to reflexive binding. That said, our results are consistent with any approach which recognizes that non-syntactic factors can have an influence on proform interpretation. Given the gender effect we found with both reflexives and pronouns it is clear that non-syntactic factors can play a role in processing proforms; and further, the influence of non-syntactic factors appears to be stronger and more ongoing during pronoun processing. Future work will have to examine what non-syntactic factors are at play, but an obvious possibility would be to examine how

the extant list of factors applying to pronouns inter-sententially may already be influencing the initial processing of pronouns, even in supposedly BT constrained positions.

References

- Alloppenna, Paul D., James S. Magnuson, and Michael K. Tanenhaus. 1998. Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language* 38.4:419–439.
- Badecker, William, and Kathleen Straub. 2002. The processing role of structural constraints on interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28.4:748–769.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Cooper, Roger M. 1974. The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology* 6.1:84–107.
- Cummings, Ian, and Claudia Felser. 2012. The role of working memory in the processing of reflexives. *Language and Cognitive Processes* 28.1-2:188–219.
- Dillon, Brian. 2012. Ungrammatical interpretations of reflexive anaphors: online or offline interference? Poster presented at the CUNY Conference on Human Sentence Processing.
- Fodor, Jerry A. 1983. *The modularity of mind: An essay on faculty psychology*. Cambridge, MA: MIT press.
- Frank, Austin, Anne Pier Salverda, T. Florian Jaeger, and Michael K. Tanenhaus. 2008. Multinomial models with “state” dependencies. Poster presented at the CUNY Conference on Human Sentence Processing.
- Grodner, Daniel, and Edward Gibson. 2005. Consequences of the serial nature of linguistic input. *Cognitive Science* 29:261–290.
- Koring, Loes, Pim Mak, and Eric Reuland. 2012. The time course of argument reactivation revealed: Using the visual world paradigm. *Cognition* 123.3:361–379.
- Nicol, Janet, and David Swinney. 1989. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18.1:5–19.
- Reinhart, Tanya, and Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24.4:657–720.
- Reuland, Eric. 2001. Primitives of binding. *Linguistics Inquiry* 32.2:439–492.
- Reuland, Eric. 2011. *Anaphora and language design*. Cambridge, MA: MIT Press.
- Roland, Douglas, Gail Mauner, Carolyn O’Meara, and Hongoak Yun. 2012. Discourse expectations and relative clause processing. *Journal of Memory and Language* 66.3:479–508.
- Sag, Ivan, Thomas Wasow, and Emily Bender. *Syntactic theory: A formal introduction*. Second edition. 2003. Stanford: CSLI.
- Sturt, Patrick. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48.3:542–562.
- Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard, and Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268.5217:1632–1634.
- Xiang, Ming, Dillon, Brian, and Colin Phillips. 2009. Illusory licensing effects across dependency types: ERP evidence. *Brain and Language* 108.1:40–55.

(Runner)
University of Rochester
jeffrey.runner@rochester.edu

(Head)
Teach For America
kellandlhead@yahoo.com