Coargumenthood and the processing of reflexives

Article

Accepted Version

Cunnings, I. ORCID: https://orcid.org/0000-0002-5318-0186 and Sturt, P. (2014) Coargumenthood and the processing of reflexives. Journal of Memory and Language, 75. pp. 117-139. ISSN 0749-596X doi: 10.1016/j.jml.2014.05.006 Available at https://centaur.reading.ac.uk/36999/

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Coargumenthood and the Processing of Reflexives

Ian Cunnings
University of Reading

&

Patrick Sturt
University of Edinburgh

Address for correspondence:

Dr Ian Cunnings
School of Psychology and Clinical Language Sciences
The University of Reading
Whiteknights Campus
Reading
RG6 7BE

Abstract

We report three eye-movement experiments and an antecedent choice task investigating the interpretation of reflexives in different syntactic contexts. This included contexts in which the reflexive and a local antecedent were coarguments of the same verbal predicate (John heard that the soldier had injured himself), and also so-called picture noun phrases, either with a possessor (John heard about the soldier's picture of himself) or without (John heard that the soldier had a picture of himself). While results from the antecedent choice task indicated that comprehenders would choose a non-local antecedent ('John' above) for reflexives in either type of picture noun phrase, the eye-movement experiments suggested that participants preferred to initially interpret the reflexive in each context as referring to the local antecedent ('the soldier'), as indexed by longer reading times when it mismatched in gender with the reflexive. We also observed a difference in the time-course of this effect. While it was observed during first-pass processing at the reflexive for coargument reflexives and those in picture noun phrases with a possessor, it was comparatively delayed for reflexives in possessorless picture noun phrases. These results suggest that locality constraints are more strongly weighted cues to retrieval than gender agreement for both coargument reflexives and those inside picture noun phrases. We interpret the observed time-course differences as indexing the relative ease of accessing the local antecedent in different syntactic contexts.

Introduction

Successful language comprehension involves forming dependencies between constituents that may span several words or sentences. A key case in point is anaphora resolution, where the ability to link reflexives and pronouns to their antecedents quickly and accurately is essential for successful sentence and discourse comprehension. Anaphor resolution thus provides an opportunity to examine the memory system that subserves language comprehension, as the interpretation of such forms relies on the retrieval of an item, the antecedent, from memory. A number of syntactic, semantic and discourse level factors are known to guide the retrieval of antecedents during anaphora resolution (see Nicol & Swinney, 2003 for review). In the current study, we focus on the role of syntactic constraints and gender agreement during the resolution of reflexives.

A growing number of studies have examined the time-course of anaphor resolution during the processing of reflexives (Badecker & Straub, 2002; Clackson, Felser, & Clahsen, 2011; Cunnings & Felser, 2013; Dillon, Mishler, Sloggett, & Phillips, 2013; Nicol & Swinney, 1989; Sturt, 2003; Xiang, Dillon & Phillips, 2009). These studies have sought to examine how syntactic constraints, specifically Principle A of the binding theory (Chomsky, 1981), and gender agreement influence antecedent retrieval during processing. However, to date, most existing research has focused on the processing of reflexives in a single type of syntactic environment, and only a few studies have examined reflexives in other constructions (e.g. Kaiser, Runner, Sussman, & Tanenhaus, 2009; Runner, Sussman, & Tanenhaus, 2003, 2006). This is unfortunate given that research in theoretical linguistics has long established that certain constructions pose problems for classic binding theory (e.g. Pollard & Sag, 1992; Reinhart & Reuland, 1993). The aim of the current study was to further examine how the syntactic environment in which a reflexive occurs influences the time-course of anaphor resolution during processing. We begin by discussing theoretical work on

the characterisation of syntactic constraints on reflexives, before discussing previous research examining how binding constraints and gender agreement guide antecedent retrieval during comprehension.

Coargumenthood and Binding Principle A

In theoretical linguistics, constraints on the interpretation of reflexives have traditionally been accounted for by binding theory (Chomsky, 1981). Binding Principles A and B were formulated to account for interpretive preferences for reflexives and pronouns as in (1).

(1)   John injured himself/him yesterday afternoon.

In (1), the reflexive *himself* can only be interpreted as referring to *John*, while the pronoun *him* must refer to some other, in this case unmentioned, antecedent. In standard binding theory, these preferences are accounted for by Principle A, which states that a reflexive must be bound to an antecedent within the local syntactic domain, in this case the same verbal predicate, and Principle B, which states that pronouns must be free within the local domain. We will refer to local antecedents of reflexives such as *John* in (1), which are predicted to be preferred according to standard binding theory as *accessible antecedents*, and other antecedents in a piece of discourse as binding theory *inaccessible antecedents*.

A key prediction of binding theory as originally formulated is that interpretive preferences for reflexives and pronouns should be in complementary distribution. Whilst this prediction is true for cases such as (1), a number of counter-examples have long been noted (see e.g. Pollard & Sag, 1992; Reinhart & Reuland, 1993). One particularly well-discussed case is that of the so-called referential or picture noun phrase (henceforth PNP), as in (2).

(2a)   John saw the picture of himself/him by the fireplace.

(2b)   John saw that there was a picture of himself/him by the fireplace.

While in (2a), the predictions of binding theory superficially may appear to hold, it is clear from example (2b) that reflexives inside PNPs can be bound to an antecedent outside of the local domain, in this case to an antecedent in a higher clause. Additionally, the complementary distribution of reflexives and pronouns appears to have broken down. In (2b) in particular, the acceptability of use of either a reflexive or pronoun to refer to *John* is not nearly as marked as in (1). The interpretation of reflexives in PNPs is also sensitive to non-syntactic factors, including pragmatic properties of the text such as 'point of view' or the 'source' of information (Kaiser et al., 2009; Kuno, 1987).

Considerations such as these led to a reformulation of the constraints on binding. Chomsky (1986) provided an account that maintained the core claims of structural binding theory, but in which the binding domain for reflexives can be enlarged if no binder is available within the most local domain. In this way, binding to the higher clause in (2b) is possible, as no antecedent is available in the reflexive's most local domain. Others have proposed more radical departures from standard binding theory, and have in particular highlighted the importance of the *coargument* relationship (Pollard & Sag, 1992; Reinhart, 1983; Reinhart & Reuland, 1993; Reuland, 2001, 2011). Although there are differences in the precise characterisation of these different theories, they share in common the importance of coargumenthood as an explanation for the contrast between (1), where the original predictions of binding theory seem to hold, and exceptional cases as in (2). Coargumenthood refers to a relationship between the core arguments of a predicate. In (1), the reflexive and antecedent are both core coarguments (subject and object) of the same verbal predicate. In

such coargument contexts the predictions of binding theory are predicted to hold, such that the reflexive must be bound to one of its coarguments (and a pronoun must be free from it). Thus, complementarity is predicted to hold, and the reflexive must be bound, in this case, to *John*. In cases such as (2), the PNP itself is argued to form a nominal predicate. Here, the reflexive has no coarguments within the local domain (the nominal PNP predicate) that it can bind to and as such, under this account, binding constraints simply cannot apply. In this case, a breakdown in complementarity is expected. Pollard and Sag (1992) refer to such reflexives as being binding theory exempt anaphors, while others (e.g. Reuland 2001, 2011) restrict use of the term *anaphor* to coargument reflexives, and refer to other types of reflexives as *logophors*. In both cases, it is claimed that it is only coargument reflexives that are syntactically bound.

A related construction is that of the PNP with a possessor, or possessed picture noun phrase (henceforth PPNP), as in (3).

(3)   John's picture of himself/him was found by the fireplace.

In contrast to debate regarding reflexives inside PNPs without a possessor, it has generally been assumed in the linguistics literature that PPNPs are restricted by binding constraints, such that a reflexive must bind to the possessor, while a pronoun must be free from it (Pollard & Sag, 1992; Reinhart & Reuland, 1993). If PNPs form nominal predicates, as argued by Pollard and Sag, an antecedent within this predicate (the possessor) must be bound to a reflexive in much the same way that reflexives have to be bound to an antecedent in the same verbal predicate in standard coargument contexts. In this way, the possessor of the PNP in (3) behaves just as the subject of the verb does in (1). However, in contrast to this account of binding in PPNPs, some have found that PPNP reflexives can indeed take

antecedents other than the possessor (Keller & Asudeh, 2001; Runner et al., 2003, 2006). Based on evidence from a series of visual wold paradigm studies (see below), Runner et al. claim that *all* reflexives inside PNPs, irrespective of whether or not they have a possessor, should be considered as being exempt from binding theory.

Antecedent retrieval during sentence processing

The formation of anaphoric dependencies during language processing clearly implicates the memory mechanisms that underlie language comprehension. A growing body of research has claimed that language comprehension is subserved by a content-addressable memory architecture (for review see Lewis, Vasishth, & Van Dyke, 2006; Van Dyke & Johns, 2012). Such theories assume that there is a severely limited focus of attention that is only capable of holding up to four items (Cowan, 2000), and potentially only one item (McElree, 2001), in the focus at one time. Retrieval is required for all other previously encountered items. Interpreting an anaphor will thus likely involve retrieval of the antecedent from memory, as it is unlikely to still be in the focus of attention when the anaphor is encountered.

A hallmark of content-addressable memory systems is that of direct access to items in memory without the need to search through irrelevant items. As items are accessed directly, this predicts that there should be no differences in retrieval speed as dependency length increases. In content-addressable architectures, retrieval occurs when an item in memory matches a set of retrieval cues. As retrieval cues are matched against all items in memory in parallel, all items that (partially) match the cues receive some amount of activation. This leads to the possibility of similarity-based interference, if a (partially) content-matching but ultimately incorrect item is retrieved. Key evidence that implicates such a memory architecture during language processing has been the demonstration of invariable access

times to items in memory irrespective of dependency length (Martin & McElree, 2008; McElree, 2000; McElree, Foraker, & Dyer, 2003), and of similarity-based interference from competitors (Van Dyke, 2007, Van Dyke & McElree, 2006; 2011).

An important issue that such models must address relates to the question of what sources of information constitute cues to retrieval. Across different dependencies, a number of potential information sources could act as cues to memory retrieval during language processing. In the case of anaphora, a number of cues could potentially guide retrieval, including gender/number agreement, discourse prominence and syntactic constraints. That agreement might cue memory retrieval during language comprehension is most clearly demonstrated by 'attraction' effects in the processing of subject-verb agreement (Pearlmutter, Garnsey, & Bock, 1999; Wagers, Lau, & Phillips, 2009). For example, Wagers et al. examined sentences such as (4).

(4a)   The key to the cell unsurprisingly were rusty from many years of disuse.

(4b)   The key to the cells unsurprisingly were rusty from many years of disuse.

Both (4a,b) are ungrammatical as the verb *were* mismatches in number agreement with the head noun *the key*. Wagers et al. found longer reading times shortly after the critical verb for ungrammatical in comparison to grammatical control sentences. However, the size of this effect was reliably attenuated when the local 'attractor' noun *cells* matched in number agreement with the verb. Wagers et al. concluded that this attraction effect strongly implicates a direct access content-addressable memory architecture with highly ranked agreement cues. When the correct subject head noun does not fully match with the retrieval cues of the verb (which cue for a number matching head noun), a grammatically illicit but number matching noun is sometimes retrieved.

In contrast to subject-verb agreement, reliable attraction effects have not been reported for reflexives in coargument contexts (Dillon et al., 2013; Sturt, 2003; Xiang et al., 2009). In two experiments, Sturt (2003) recorded participants' eye-movements as they read sentences similar to (5).

(5a)   Jonathan/Jennifer remembered that the surgeon had pricked himself/herself with a used syringe needle.

(5b)   The surgeon who treated Jonathan/Jennifer had pricked himself/herself with a used syringe needle.

Sturt manipulated gender agreement between the reflexive and both the binding-theory accessible antecedent *the surgeon* and the inaccessible antecedent (*Jonathan/Jennifer*). In his Experiment 1, the accessible antecedent was linearly closer to the reflexive than the inaccessible antecedent, as in (5a), while in his Experiment 2, the order was reversed as in (5b). Across both experiments, reading times were longer when the binding theory accessible antecedent *the surgeon* mismatched in stereotypical gender with the reflexive, during first-pass processing at the reflexive. Some effects of the nonlocal inaccessible antecedent were observed in Experiment 1, but they were restricted to second-pass processing measures, and were not in the direction predicted by attraction. At the sentence prefinal region (*used syringe* in (5)), in conditions when the local accessible antecedent matched the gender of the reflexive, reading times were longer when the nonlocal inaccessible antecedent mismatched the reflexive's gender. The inaccessible antecedent did not reliably affect processing at any point in time in Experiment 2 however. As such, Sturt hypothesised that binding constraints act as an initial filter to anaphor resolution, which nevertheless may be defeasible during later stages of processing. In this way, binding constraints may guide the initial retrieval, but this

initial retrieval can be subsequently reanalysed. Indeed, Sturt originally reasoned that one possible reason for a defeasible but not absolute filter would be to allow for the interpretation of reflexives that are not governed by binding constraints.

More recently, Dillon et al. (2013) directly compared attraction effects for both subject-verb agreement and reflexive binding. While they replicated the results of Wagers et al. (2009) for agreement, they failed to observe reliable attraction effects for reflexives in sentences similar to (5b), but which contained number agreement rather than gender agreement manipulations. They concluded that, at least for coargument reflexives, Principle A acts as a 'hard constraint' that solely guides antecedent retrieval during anaphor resolution.

The hypothesis that certain types of information might take priority over others during antecedent retrieval raises questions with regard to how different cues to retrieval are combined during processing. Some cue-based models (e.g. Lewis & Vasishth, 2005) assume that all sources of information, and thus all cues, are weighted equally. More recently however it has been claimed that at least for some dependencies, some cues might take priority over others. Van Dyke and McElree (2011), for example, claimed that certain cues may 'gate' access to other types of information. In their examination of constituent retrieval across relative clause boundaries they claimed that syntactic cues, in this case subjecthood, may gate access to semantic cues such as animacy. While attraction effects in subject-verb agreement suggest that agreement is a highly weighted cue in such cases, the results of Sturt (2003) and Dillon et al. (2013) suggest that for coargument reflexives, binding constraints may be more heavily weighted cues to retrieval than gender/number agreement.

Whether or not binding constraints *uniquely* cue memory access for coargument reflexives has however been the subject of considerable debate. In addition to the delayed effects observed by Sturt (2003) in his Experiment 1, in a self-paced reading experiment Badecker and Straub (2002) observed longer reading times when a nonlocal inaccessible

antecedent matched in gender with the reflexive in comparison to when it did not, in sentences such as *John/Jane thought that Bill owed himself another opportunity to solve the problem*. Badecker and Straub interpreted this result as indicating competition between antecedents when multiple gender matching antecedents are available in the discourse. Cunnings & Felser (2013) also recently reported an eye-movement experiment which used materials similar to (5b), where the accessible antecedent was linearly more distant to the reflexive than the inaccessible one. However, in their experiment the inaccessible antecedent appeared as the subject of the relative clause, while in the materials tested by Sturt, as in (5b), it was the object. In comparison to a 'double match' condition in which both antecedents matched in gender with the reflexive (e.g. '*the surgeon ... Jonathan ... himself...*'), Cunnings & Felser observed longer reading times during first-pass processing at the reflexive in conditions when either antecedent mismatched in gender with the reflexive (e.g. '*the surgeon ... Jennifer ... himself...*' and '*the surgeon ... Jonathan ... herself...*'). This effect however, was restricted to participants who scored comparatively lower on a standard reading span test (Daneman & Carpenter, 1980). These results indicate that individual differences may affect anaphor resolution, and suggest that for readers with a lower working memory span, a binding theory inaccessible antecedent may be retrieved if it intervenes between the reflexive and accessible antecedent, and is a discourse prominent subject.

In summary, the results of a number of studies suggest that binding constraints constitute highly weighted cues to antecedent retrieval during the processing of coargument reflexives. Whether such structural constraints uniquely cue retrieval has been debated. While effects of structurally inaccessible antecedents have been reported, they have not been observed consistently across experiments, and numerical trends are not always in the same direction. Badecker and Straub (2002), for example, observed longer reading times following multiple gender *matching* antecedents, whereas a different effect was observed by Sturt

(2003) in his Experiment 1, and for lower working memory span readers by Cunnings and Felser (2013) in their Experiment 2. In contrast to these studies examining coargument reflexives, far fewer researchers have investigated the processing of reflexives inside PNPs.

We are aware of only one study to date that has examined the time-course of antecedent retrieval for reflexives inside PNPs without possessors. Kaiser at el. (2009) conducted a visual world paradigm experiment using sentences as in (6), in which the pragmatic properties of the discourse were manipulated. Participants in this task both viewed a visual scene of the participants described in the sentences while their eye-movements were monitored, and additionally had to provide an overt response as to who they thought the reflexive referred to.

(6a)   Peter told Andrew about the picture of himself on the wall.

(6b)   Peter heard from Andrew about the picture of himself on the wall.

In (6a), *Peter* is the 'source' of information and *Andrew* the 'perceiver', while in (6b) these roles are reversed. Kaiser et al. hypothesised that reflexives may prefer antecedents that are 'sources' of information (see Kuno, 1987). They observed that while participants generally preferred the subject as the antecedent of the reflexive over 80% of the time, this preference was stronger when the subject was a source of information. Participants' eye-movements across a visual display also exhibited a subject advantage, which was stronger when the subject was the source of information, during an early time window starting 200ms after the onset of the reflexive. Kaiser et al. concluded that the resolution of PNP reflexives is immediately sensitive to both structural (in this case subjecthood) and pragmatic (source vs. perceiver) properties of the discourse, a finding potentially compatible with the hypothesis that reflexives inside PNPs are exempt from binding constraints.

Keller and Asudeh (2001) presented results from an acceptability judgement study using magnitude estimation examining PPNP reflexives. They found that participants did not rate sentences such as 'Hanna found Peter's picture of herself' as being fully unacceptable. This finding suggests that, at least in an untimed task, reflexives in PPNPs do not necessarily have to be bound to the possessor, contrary to the predictions of both binding theory (Chomsky, 1981) and subsequently revised proposals (Pollard & Sag, 1992; Reinhart & Reuland, 1993). Runner et al. (2003, 2006) also examined PPNP reflexives in a series of visual world paradigm experiments. Participants in their studies acted out instructions as in (7) while their eye-movements were monitored.

(7)   Look at Ken. Have Joe touch Harry's picture of himself.

Across their experiments, Runner et al. observed that participants chose an antecedent other than the possessor up to approximately 30% of the time. Additionally, looks to antecedents other than the possessor were observed in a time-window starting 300ms after the onset of the reflexive. Here, they observed that participants were equally likely to fixate upon either the possessor or subject of the critical sentence, but not the antecedent in the lead-in sentences. Runner et al. concluded that antecedents other than the possessor are immediately considered during processing, and as such claimed that reflexives inside PPNPs should be treated as exempt from binding theory.

Kaiser et al. (2009) also examined PPNP reflexives in sentences such as (8) in a visual world paradigm experiment. As in their experiment on PNP reflexives without a possessor, they manipulated the source/perceiver properties of potential antecedents.

(8a)   Peter told Andrew about Greg's picture of himself.

(8b)   Peter heard from Andrew about Greg's picture of himself.

In contrast to their findings for PNP reflexives without a possessor, in this experiment Kaiser et al. observed no reliable effect of their pragmatic manipulation, and instead participants chose the possessor as the antecedent for the reflexive over 90% of the time. Participants' eye-movements across a visual display were also focused on the possessor, with no reliable influence of the pragmatic manipulation at any point in time.

In summary, previous studies examining coargument reflexives have suggested that binding constraints constitute a highly weighted cue to antecedent retrieval. Kaiser et al. (2009) showed that the resolution of PNP reflexives is sensitive to pragmatic properties of the text, suggesting that the resolution of reflexives inside PNPs is not guided solely by structural cues. Research examining PPNP reflexives has shown that antecedents other than the possessor are sometimes considered immediately during processing (Runner et al, 2003, 2006), suggesting that such reflexives should be considered as being exempt from binding theory. However, this finding was not replicated in a subsequent study, in which there was no evidence that antecedents other than the possessor were being considered during early stages of processing (Kaiser et al., 2009).

The present study

Against this background, the aim of the current study was to further investigate how the syntactic context in which a reflexive appears influences the extent to which binding theory accessible and inaccessible antecedents are considered. In order to investigate both the time-course of anaphor resolution and the final interpretations given to reflexives in different contexts, we report three eye-movement experiments and an antecedent choice task. The eye-movement experiments were conducted to examine initial preferences and the relative time-

course of antecedent retrieval in coargument contexts (Experiment 1), PNP contexts (Experiment 2) and PPNP contexts (Experiment 3). The antecedent choice task (Experiment 4) was designed to gauge the extent to which comprehenders are willing to ultimately consider local and nonlocal antecedents for coargument and (P)PNP reflexives. We are unaware of any previous studies that have compared coargument reflexives to reflexives in (P)PNPs with maximally similar materials. In contrast to Runner et al. (2003, 2006) and Kaiser et al. (2009), our eye-movement experiments adopted a reading paradigm rather than the visual world paradigm to investigate anaphora resolution. We are particularly interested in the early stages of anaphora resolution to investigate what cues are used to guide the earliest stages of antecedent retrieval in different structural contexts. To this end, we used the gender mismatch paradigm (Cunnings & Felser, 2013; 2013; Sturt, 2003; Xiang et al., 2009) to examine the extent to which binding theory accessible and inaccessible antecedents are preferentially retrieved during comprehension.

Experiment 1

To examine the time-course of anaphora resolution, we monitored participants' eye-movements as they read a series of texts as shown in (9) below. The aim of Experiment 1 was to replicate Sturt (2003) and provide a 'yardstick' of the time-course of antecedent retrieval during anaphora resolution for coargument reflexives, to contrast with reflexives inside PNPs and PPNPs in Experiments 2 and 3 respectively. Gender congruence between a reflexive and two antecedents in the discourse was manipulated as in (9).

(9a)   *Local antecedent match, nonlocal antecedent match*

Jonathan was walking through the military barracks. He heard that

the soldier had positioned himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(9b)   *Local antecedent match, nonlocal antecedent mismatch*

Jennifer was walking through the military barracks. She heard that

the soldier had positioned himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(9c)   *Local antecedent mismatch, nonlocal antecedent match*

Jennifer was walking through the military barracks. She heard that

the soldier had positioned herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(9d)   *Local antecedent mismatch, nonlocal antecedent mismatch*

Jonathan was walking through the military barracks. He heard that

the soldier had positioned herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

In (9a-d), the local antecedent *the soldier* is the only antecedent that is accessible according to both standard binding theory (Chomsky, 1981) and accounts reformulated in terms of coargumenthood (Pollard & Sag, 1992; Reinhart & Reuland, 1993). In (9a,b) this local antecedent matches in stereotypical gender with the reflexive, while in (9c,d) there is a stereotypical gender mismatch. Additionally, the gender of the non-local antecedent (Jonathan/Jennifer), predicted to be inaccessible according to binding theory, has also been manipulated. In (9a,c), the nonlocal antecedent matches in gender with the reflexive, while in (9b,d) there is a gender mismatch.

If antecedent retrieval is guided solely by binding constraints (Dillon et al., 2013), such that coargument reflexives must be bound to an antecedent in the same clause, we should observe reliable effects of the gender of the local antecedent only. In this case, reading

times in (9c,d), when the local antecedent mismatches in stereotypical gender with the reflexive, should be longer than in (9a,b), when there is a gender match. Reading times should not differ as a result of the gender of the nonlocal antecedent. If binding theory acts as an initial but nevertheless defeasible filter (Sturt, 2003), effects of the nonlocal antecedent might be observed but these should be relatively delayed in comparison to effects of the local antecedent. There are different ways in which the nonlocal antecedent might influence reading times, in either early or later stages of processing.

One possibility could be that both binding constraints and gender agreement combine to cue antecedent retrieval during processing. The strongest evidence for this would be an attenuation of the local antecedent gender mismatch effect when a gender matching but structurally inaccessible nonlocal antecedent is available in the discourse, as is observed in agreement attraction (Dillon et al., 2013; Wagers et al., 2009). In this case, the local antecedent gender mismatch effect in (9c), when the nonlocal antecedent matches in gender with the reflexive, should be attenuated in comparison to (9d), when neither antecedent match the reflexive's gender. Alternatively, if gender is not a strongly weighted cue to retrieval for reflexives, there is the possibility that other potential cues (e.g. subjecthood) might combine with binding constraints such that nonlocal antecedents are sometimes retrieved. If discourse-prominent but binding theory inaccessible subjects are sometimes retrieved irrespective of gender (*Jonathan/Jennifer* in our examples), we might observe main effects of gender mismatch for both the local and nonlocal antecedents. Yet another possibility could be that reading times are longer in double match condition (9a) compared to the local antecedent match, nonlocal antecedent mismatch condition (9b). This could be taken as evidence of competition between the two antecedents when both match in gender with the reflexive (Badecker & Straub, 2002).

Method

Participants

28 native English speakers (9 males, mean age 23) were paid a small fee to participate in the experiment. All participants had normal or corrected to normal vision and were recruited from the University of Edinburgh community.

Materials

32 sets of experimental items were constructed as in (9). Gender congruence between the local antecedent and the reflexive was manipulated using highly gender biased nouns, all of which had previously been rated for gender stereotypicality in previous studies (Cunnings & Felser, 2013; Kennison & Trofe, 2003; Kreiner, Sturt, & Garrod, 2008). The gender of the nonlocal antecedent was manipulated using proper names. For each item, the nonlocal antecedent was introduced in the first introduction sentence, and then referred to with a pronoun at the start of the second, critical sentence to place this character strongly into discourse focus (Sanford & Garrod, 1988). The second sentence always included the local antecedent and critical reflexive, while a third final 'wrap-up' sentence was included to avoid the influence of end-of-trial artefacts from affecting reading times of the critical second sentence. The full set of experimental items is provided in the Appendix.

In addition to the experimental items, 64 filler texts were also constructed. These included distractor items that were structurally similar to the experimental items but did not contain reflexives.

Procedure

The experimental and filler items were pseudo-randomised such that no two experimental items appeared adjacent to each other and were spread across four presentation

lists in a Latin-square design. The experiment was divided into four blocks at which point participants could take a break if required. Forward and reverse orders within each block were constructed and the ordering of each block was different for each participant. The experiment began with some practice items to familiarize participants with the procedure. All items were presented in Consolas fixed width font and displayed across up to three lines of text onscreen. Critical items were always spread across three lines, with line breaks after the complementiser 'that' and after the end of the critical sentence.

Eye-movements were recorded using the EYELINK 2000 system, with eye-movements being recorded at a rate of 1000Hz. While viewing was binocular, the eye-movement record was recorded from the right eye only. Each experimental session began with calibration of the eye-tracker on a nine-point grid, and any drift in calibration was compensated for via recalibration between trials if required. Before each trial, participants fixated on a fixation marker above the first word of the trial to be displayed. Upon fixation on this marker, the trial text appeared. Participants read each text silently at their normal reading rate, pressing a button on a control pad once completed. Content questions requiring a yes-no push button response followed all critical trials and half of the fillers. Half of the questions required a 'yes' response and half a 'no' response. The entire experiment lasted approximately 30-45 minutes in total.

Data analysis

To examine the time-course of anaphor resolution we calculated reading times for four regions of text. The critical *reflexive region* consisted of the critical reflexive (*himself/herself*), while the *spillover region* consisted of the two words directly after it. The *prefinal region* consisted of the two words following the spillover region, and finally the *final region* consisted of the rest of the sentence. Three reading time measures are reported for

each region of text. *First pass reading time* is the summed duration of fixations within a region during its first inspection, until it is exited to the left or right, while *regression path duration* is calculated by summing the duration of each fixation, starting with the first fixation when a region is entered from the left, up until but not including the first fixation in a region to the right. In addition to these two *first-pass* processing measures, we also calculated *second pass times*, which included all fixations within a region *after* it has been exited following the first-pass.

All trials in which track loss occurred were discarded, and regions which were initially skipped during reading were treated as missing data in the two first-pass measures. For second pass times, trials in which a region was not fixated after the first-pass contributed a second pass time of zero to the calculation of averages. Following Sturt (2003), to increase the probability of a first pass fixation at the critical reflexive region, a leftward-shifting procedure was used in calculation of the first pass and regression path times at the reflexive. If the reflexive was not fixated during the first pass, we included fixations that landed up to 4 characters to the left of the region boundary (see Sturt 2003: 548). Prior to the calculation of all reading time measures an automatic procedure merged short fixations of 80ms or below that were within one degree of visual arc of another fixation. All other fixations of 80ms or below, as well as those above 800ms, were removed before further analysis.

Results

Overall accuracy to the comprehension questions was 95% (all above 85%), indicating that participants paid attention to the content of the sentences. Track loss accounted for 0.22% of the data, and skipping rates for the reflexive region (after leftward-shifting) was 3.24%. Skipping rates for the spillover, prefinal and final regions were 19.98%, 8.26% and 7.14% respectively.

Analysis was conducted using linear-mixed effects models with crossed random effects for subjects and items (see Baayen, 2008; Baayen, Davidson, & Bates, 2008) using the lme4 package in R. For each reading time measure the analysis included sum coded, fixed main effects of 'local antecedent' (match vs. mismatch), 'nonlocal antecedent' (match vs. mismatch) and the 'local antecedent' by 'nonlocal antecedent' interaction. Subject and item random intercepts, as well as subject and item random slopes for each fixed effect, were included using a 'maximal' random effects structure (Barr, Levy, Scheepers, & Tily, 2013). Models were fit using restricted maximum likelihood. If the maximal model failed to converge, the random effect parameter that accounted for the least variance in the data was removed, and the model refit until convergence was achieved. For each measure, p values for each fixed effect were calculated from the t distribution (Baayen, 2008: 248).

Summaries of the reading time data and statistical analysis for Experiment 1 are presented in Tables 1 and 2 respectively.

(TABLE 1 ABOUT HERE)

At the reflexive region, first pass reading times were marginally longer in conditions (9c,d), when the stereotypical gender of the local antecedent mismatched in gender with the reflexive, in comparison to conditions (9a,b), when there was a stereotypical gender match (235ms vs 222ms). The same stereotypical gender mismatch effect was significant in both the regression path times (295ms vs. 249ms), and the second pass times (193ms vs, 135ms). In contrast to these local antecedent effects, the nonlocal antecedent did not significantly affect reading times at the reflexive region.

(TABLE 2 ABOUT HERE)

At the spillover region, there was a trend for first pass reading times to be longer in conditions (9a,c), when the nonlocal antecedent matched in gender with the reflexive, in comparison to conditions (9b,d), when it mismatched (278ms vs. 255ms), but the main effect of the nonlocal antecedent was only marginally significant. In regression path times and second pass reading times there were significant main effects of the local antecedent. In both measures, reading times were longer when the local antecedent mismatched in stereotypical gender with the reflexive in comparison to when there was a gender match (for regression path times, 431ms vs. 345ms; for second pass times, 251 vs 190ms).

At the prefinal region, there was a marginally significant main effect of the local antecedent in regression path times, with reading times again tending to be longer following stereotypical gender mismatches (513ms vs. 434ms) in the absence of any reliable effects in either first pass or second pass times. No reliable effects were observed at the sentence final region.

Discussion

The results of Experiment 1 replicate previous studies showing relatively early effects of the gender of the local, structurally accessible antecedent at a point in time when no reliable influence of the nonlocal inaccessible antecedent was observed (Sturt, 2003). Reading times at the reflexive, in both first- and second-pass measures, were reliably influenced by the stereotypical gender of the local antecedent, with reading times being longer following stereotypical gender violations. The only hint of an effect of the nonlocal antecedent was observed in first-pass times at the spillover region. We did not replicate the delayed nonlocal inaccessible antecedent effect observed by Sturt (2003), but instead found

marginally significant trends similar to Badecker & Straub (2002), in that reading times were longer when the structurally inaccessible antecedent matched in gender with the reflexive.

In sum, these results are compatible with the hypothesis that binding constraints cue antecedent retrieval during the processing of coargument reflexives (Clackson et al., 2011; Dillon et al., 2013; Nicol & Swinney, 1989; Sturt, 2003; Xiang et al., 2009). Experiment 2 sought to examine PNP reflexives with similar materials.

Experiment 2

In Experiment 2, we examined the time-course of reference resolution for reflexives inside PNPs. The items used were identical to those in Experiment 1, except that the reflexive now appeared inside a picture noun phrase. The items from Experiment 1 were adapted as in (10).

(10a)   *Local antecedent match, nonlocal antecedent match*

Jonathan was walking through the military barracks. He heard that

the soldier had a picture of himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

 (10b)   *Local antecedent match, nonlocal antecedent mismatch*

Jennifer was walking through the military barracks. She heard that

the soldier had a picture of himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

 (10c)   *Local antecedent mismatch, nonlocal antecedent match*

Jennifer was walking through the military barracks. She heard that

the soldier had a picture of herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(10d) *Local antecedent mismatch, nonlocal antecedent mismatch*

       Jonathan was walking through the military barracks. He heard that

       the soldier had a picture of herself in the middle of the mess hall.

       The food being served for dinner did not look very appetising.

As in Experiment 1, *the soldier* is the only antecedent accessible according to standard binding theory (Chomsky, 1981, 1986), as it is the only antecedent in the same clause as the reflexive. Accordingly, *Jonathan/Jennifer* should be inaccessible. If this is indeed the case, then we should observe similar results in Experiment 2 as in Experiment 1. In this case, reading times should be reliably longer at or soon after the reflexive in conditions (10c,d), when the local antecedent mismatches in stereotypical gender with the reflexive, in comparison to conditions (10a,b), when there is a gender match.

An alternative set of predictions can be considered if reflexives inside PNPs are exempt from binding theory. If this is the case, then it may be that reflexives in such contexts may optionally take a nonlocal antecedent. As in Experiment 1, there are different ways in which this effect might be observed. If PNP reflexives cue retrieval of any gender matching antecedent, we might find processing difficulty in 'double mismatch' condition (10d) only. In all other conditions, the reflexive can find a gender-matching antecedent by either retrieving *the soldier* or *Jonathan/Jennifer*. Another possibility could be that discourse-salience and/or subjecthood, rather than gender agreement, guide antecedent retrieval for PNP reflexives. In this case, we might observe main effects of *either* the local or nonlocal antecedent, as both are subjects, such that reading times might be longer in (10c,d) in comparison to (10a,b) and also in conditions (10b,d) in comparison to (10a,c). This pattern of results would suggest that either antecedent was equally likely to be retrieved. Longer reading times in double match

condition (10a) compared to the local antecedent match, nonlocal antecedent mismatch condition (10b) could also be taken as evidence of competition between the two antecedents.

Note that although Pollard and Sag (1992) claim that PNP reflexives are exempt from binding constraints, interpretation of such reflexives is argued to be subject to additional factors, including an 'intervention' constraint. This constraint states that reflexives can only take a long distance antecedent as long as there is no other potential binder in a more local domain. Consider sentences such as (11).

(11a)   James said that John published a picture of himself yesterday afternoon.

(11b)   James said that the newspaper published a picture of himself yesterday afternoon.

Pollard and Sag claim that the reflexive is unlikely to refer to *James* in (11a) as a more local antecedent (*John*) intervenes. However, in cases when there is no local *animate* intervener, it is possible for the reflexive to take a long distance antecedent (e.g. *James* in (11b)). Pollard and Sag take this contrast as evidence that there is no absolute *grammatical* constraint on binding reflexives in PNPs, but that their interpretation is subject to additional factors. Assuming this account, then the predictions for Experiment 2 would again be similar to those in Experiment 1, such that we should observe gender mismatch effects for the local antecedent only.

Method

Participants

28 native English speakers (9 males, mean age 22), none of whom took part in Experiment 1, were paid to participate. All had normal or corrected to normal vision and were recruited from the University of Edinburgh community.

Materials

The 32 sets of experimental items from Experiment 1 were adapted as in (10). Apart from the reflexive now appearing inside a PNP, all other aspects of the experimental items were identical to those used in Experiment 1. Note that as in Experiment 1, the nonlocal antecedent is in a discourse-prominent position. Additionally note that in (10), the local antecedent is the subject of a verb that does *not* specifically entail that the local antecedent was the producer/creator of the picture. Creation verbs of this type (e.g. *took* a picture) were not used as it has been claimed that they in particular force the reflexive to be obligatorily bound to (and a pronoun free from) the local antecedent (Keller & Asudeh, 2001). We avoided such verbs to thus further increase the chances of the nonlocal antecedent being retrieved. The full set of experimental items can be found in the Appendix.

As in Experiment 1, 64 filler texts were also constructed which included distractor items that were structurally similar to the experimental items but did not contain reflexives.

Procedure and Data Analysis

The procedure and data analysis was the same as in Experiment 1.

Results

Overall accuracy to the comprehension questions was 95% (all above 89%), indicating that participants paid attention to the content of the sentences. There was no track loss, and skipping rates for the reflexive, spillover, prefinal and final regions were 7.59%, 22.11%, 14.89% and 12.21% respectively. Summaries of the reading time data and statistical analysis for Experiment 2 are presented in Tables 3 and 4 respectively.

(TABLE 3 ABOUT HERE)

In contrast to Experiment 1, we did not observe any significant effects in either first pass reading times or regression path times at the reflexive. There was however a significant main effect of the local antecedent in second pass times. Here, reading times were longer in conditions (10c,d), when the local antecedent mismatched in stereotypical gender with the reflexive, in comparison to conditions (10a,b), when there was a gender match (120ms vs. 88ms). No reliable effects of the nonlocal antecedent were observed at the reflexive region.

(TABLE 4 ABOUT HERE)

At the spillover region, we again observed significant main effects of the local antecedent. In each reading time measure at this region, we observed longer reading times when the local antecedent mismatched in stereotypical gender with the reflexive in comparison to when there was a gender match (for first pass reading times, 279ms vs. 253ms; for regression path times, 415ms vs. 321ms; for second pass times, 148ms vs. 113ms).

At the prefinal region, we again observed a significant stereotypical gender mismatch effect for the local antecedent in the regression path times (497ms vs. 420ms). In second pass times at this region, we observed a marginally significant main effect of the nonlocal antecedent. In this measure, reading times tended to be longer in conditions (10a,c), when the nonlocal antecedent matched in gender with the reflexive, in comparison to conditions (10c,d), when it did not (148ms vs. 116ms). No reliable effects were observed at the final region.

Discussion

The results of Experiment 2 can be summarised as follows. In second pass reading times at the reflexive and both first and second pass measures at subsequent regions of text, we observed reliably longer reading times when the local antecedent mismatched in stereotypical gender with the reflexive in comparison to when there was a gender match. The only trend of an effect of the nonlocal antecedent was comparatively delayed until second-pass times at the prefinal region where, similar to Experiment 1, we observed longer reading times when the nonlocal antecedent matched in gender with the reflexive in comparison to when it mismatched.

The results of Experiment 2 are similar to those of Experiment 1 in that reading times were reliably influenced by the gender of the local antecedent. However, one difference between Experiments 1 and 2 appears to be the relative time-course of the local antecedent gender mismatch effect. While we observed this mismatch effect during first pass processing at the reflexive in Experiment 1, it was delayed until second pass reading times at the reflexive, and first pass times at the spillover region, in Experiment 2. To statistically evaluate this potential time-course difference between coargument and PNP reflexives, we conducted an additional between-experiment analysis for the two critical first pass measures at the reflexive region where this difference in time-course was most apparent. This additional 2x2x2 analysis contained sum coded, fixed effects for the 'local antecedent' (match vs. mismatch), 'nonlocal antecedent' (match vs. mismatch), 'structure' (coargument reflexive vs. PNP reflexive), and their interactions. The 'maximal' random effects structure that converged was again used (Barr et al., 2013), with 'structure' treated as a between subjects but within items variable. For first pass reading times, this analysis revealed no significant main effects or interactions (all $t < 1.65$, all $p > .101$). For regression path times however, there was a significant main effect of structure (estimate = 52, SE = 25, $t = 2.06$, $p = .040$) that was qualified by a significant structure by local antecedent interaction (estimate

= 60, SE = 28, t = 2.18, p = .030). No other main effects or interactions were significant (all other t < 1.13, all other p > .260). The reliable structure by local antecedent interaction in regression path times at the reflexive confirms that the local antecedent stereotypical gender mismatch effect was indeed observed at an earlier point in time in Experiment 1 than Experiment 2. One possible account of this is that while the reflexive can access an antecedent in the most local domain (the same verbal predicate) in Experiment 1, if picture noun phrases form a type of nominal predicate (Chomsky, 1981, 1986; Pollard & Sag, 1992), reflexives in Experiment 2 cannot access an antecedent within their most local domain. This may suggest that dependency formation is comparatively easier or quicker when an antecedent is available in the most local domain. We return to discussion of these time-course differences in the General Discussion.

The results of Experiment 2 suggest that reflexives inside PNPs preferentially cue retrieval of an antecedent in the same clause. Whether or not the reflexives in PNPs as tested in Experiment 2 are ever interpreted as referring to a nonlocal antecedent is further examined in Experiment 4, which tested interpretive preferences in an offline task. We first report the results of a third eye-movement experiment examining reflexives in PNPs with a possessor.

Experiment 3

In Experiment 3, we adapted the materials from Experiments 1 and 2 to investigate the processing of reflexives in PPNP contexts. The items were adapted as in (12).

(12a)   *Local antecedent match, nonlocal antecedent match*

Jonathan was walking through the military barracks. He heard about

the soldier's picture of himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(12b) *Local antecedent match, nonlocal antecedent mismatch*

Jennifer was walking through the military barracks. She heard about

the soldier's picture of himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(12c) *Local antecedent mismatch, nonlocal antecedent match*

Jennifer was walking through the military barracks. She heard about

the soldier's picture of herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(12d) *Local antecedent mismatch, nonlocal antecedent mismatch*

Jonathan was walking through the military barracks. He heard about

the soldier's picture of herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

As in Experiments 1 and 2, *the soldier* is the only antecedent accessible according to standard binding theory (Chomsky, 1981, 1986), whereas *Jonathan/Jennifer* should be inaccessible. The same predictions also hold for revised theories (Pollard & Sag, 1992; Reinhart & Reuland, 1993) under the assumption that PNPs form nominal predicates which thus require obligatory binding to the possessor (Pollard & Sag, 1992; Reinhart & Reuland, 1993). Such accounts would also predict that *the soldier* is the only accessible antecedent. In this case, we should find effects of the stereotypical gender of this local antecedent only, with reading times being longer in conditions (12c,d) in comparison to (12a,b).

If antecedents other than the possessor are immediately considered during processing for PPNP reflexives (Runner et al, 2003, 2006), then this would predict that antecedents other than the possessor might sometimes be retrieved upon encountering the reflexive. In this case, we might find that only 'double mismatch' condition (12d) causes processing difficulty,

as this is the only condition in which no gender matching antecedent is available. Alternatively, based on evidence from their visual world experiments, Runner et al. claimed that both the possessor and a sentential subject are initially considered as the antecedent for a PPNP reflexives. If PPNP reflexives thus cue retrieval of discourse-prominent subjects, we may observe main effects of the gender of either (or both) antecedents, such that reading times may be longer in conditions (12c,d) in comparison to (12a,b), and also (12b,d) in comparison to (12a,c). As in Experiments 1 and 2, longer reading times in double match condition (12a) compared to condition (12b) might also index competition between the two antecedents when both match in gender.

In light of the results of Experiments 1 and 2, additional predictions can also be made with regard to the time-course of observed effects. Recall that in Experiment 2, the local antecedent gender mismatch effect was observed at a relatively later point in time in comparison to Experiment 1. One possible explanation of this is that in Experiment 2 there is no antecedent in the most local syntactic domain, the nominal picture noun phrase predicate. In Experiment 1 however, the antecedent and reflexive are both within the same verbal predicate. If picture noun phrases form a type of complex predicate (Pollard & Sag 1992), then the possessor in Experiment 3 can be considered a predicate-internal antecedent. Under this analysis, as local antecedent retrieval in Experiment 1 involves retrieval of a predicate-internal antecedent, so too does retrieval of the possessor in Experiment 3. If this is indeed the case, then we should observe no relative difference in the time-course of local antecedent gender mismatch effects in Experiment 3 in comparison to Experiment 1.

Method

Participants

28 native English speakers (4 males, mean age 21) with normal or corrected to normal vision from the University of Edinburgh community were paid to participate. None had taken part in either Experiments 1 or 2.

Materials

The 32 sets of experimental items from Experiments 1 and 2 were adapted as in (12). While the reflexive now appeared inside a possessed picture noun phrase, all other aspects of the experimental items were identical to those used in Experiments 1 and 2, and a full set of experimental items can be found in the Appendix. Note that in this experiment, the nonlocal inaccessible antecedent is in the same clause as the reflexive and accessible local antecedent. As in the previous experiments, the nonlocal antecedent is also highly discourse-salient. The materials were constructed with these factors in mind in an attempt to increase the possibility of the nonlocal antecedent being retrieved. As in the other experiments, 64 filler texts were also included which included distractor items that were structurally similar to the experimental items but did not contain reflexives.

Procedure and Data Analysis

The procedure and data analysis was the same as in Experiment 1.

Results

Overall accuracy to the comprehension questions was 94% (all above 84%), indicating that participants paid attention to the content of the sentences. Track loss occurred in 0.22% of trials, and skipping rates for the reflexive, spillover, prefinal and final regions were 3.21%, 24.01%, 13.30% and 11.50% respectively. Summaries of the reading time data and statistical analysis for Experiment 3 are presented in Tables 5 and 6 respectively.

(TABLE 5 ABOUT HERE)

At the reflexive region, there were marginally significant main effects of the local antecedent in both first pass and regression path times. In both measures, reading times tended to be longer in conditions (12c,d), when the local antecedent mismatched in stereotypical gender with the reflexive, in comparison to conditions (12a,b), when there was a gender match (for first pass times, 259ms vs. 240ms; for regression path times, 330ms vs. 297ms). In the second pass times at the reflexive, the stereotypical gender mismatch effect for the local antecedent was significant (114ms vs. 72ms). No reliable effects of the nonlocal antecedent were observed at the reflexive region.

(TABLE 6 ABOUT HERE)

At the spillover region, there was a significant main effect of the local antecedent in regression path times. Here, reading times were longer when the local antecedent mismatched in stereotypical gender with the reflexive in comparison to when there was a gender match (371ms vs. 316ms). In second pass times, there was a significant main effect of the nonlocal antecedent. In this measure, reading times were longer in conditions (12a,c), when the nonlocal antecedent matched in gender with the reflexive, in comparison to conditions (12b,d), when there was a gender mismatch (139ms vs. 100ms).

No reliable effects were observed at the prefinal or final regions in Experiment 3.

Discussion

In Experiment 3, the only reliable effects we observed at the reflexive region were of the local antecedent, with reading times being longer following stereotypical gender violations. We also observed a similar effect in regression path times at the spillover region. In this experiment we did observe a comparatively delayed significant main effect of the nonlocal antecedent. This effect was restricted to second-pass times at the spillover region and, as was observed numerically in Experiments 1 and 2, the direction of the effect indicated longer reading times when the nonlocal antecedent matched in gender with the reflexive, as also observed by Badecker & Straub (2002).

In contrast to Experiment 2, there did not seem to be any noticeable difference in the time-course of the local antecedent stereotypical gender mismatch effect observed in Experiment 3 in comparison to Experiment 1. To statistically analyse the time-course of this effect across Experiments 1 and 3, we again conducted a between experiment analysis containing 'local antecedent' (match vs. mismatch), 'nonlocal antecedent' (match vs. mismatch) and 'structure' (coargument reflexive vs. PPNP reflexive) as independent variables for the two first pass measures at the reflexive, where effects of the local antecedent were first observed in Experiment 1. For first pass reading times, this analysis revealed a significant main effect of the local antecedent (estimate = 15, SE = 7, t = 2.36, p = .018) in the absence of any other reliable main effects or interactions (all other t < 1.41, all other p > .161). For regression path times there was a marginally significant main effect of structure (estimate = 20, SE = 10, t = 1.92, p = .055), with reading times tending to be longer for reflexives inside PPNPs than those in coargument contexts. The main effect of the local antecedent was significant (estimate = 37, SE = 11, t = 3.36, p < .001) in the absence of any other main effects or interactions (all other t < 1, all other p > .364). This additional analysis

confirms that we observed no evidence of any delay in the local antecedent gender mismatch effect for reflexives inside PPNPs in comparison to coargument contexts[1].

The results of Experiment 3 suggest that participants initially preferred the possessor as an antecedent for the reflexive. This interpretation of our results contrasts with Runner et al. (2003, 2006), who observed equally early effects of both the possessor and a sentential subject in their visual world paradigm studies. We return to differences between the results of Experiment 3 and Runner et al. in the General Discussion.

The local antecedent stereotypical gender mismatch effects observed in Experiments 1 – 3 suggest that the local antecedent was preferentially retrieved. In Experiment 4, we further tested how absolute these preferences are by investigating the extent to which local and nonlocal antecedents are considered by comprehenders in an untimed antecedent-choice task.

Experiment 4

While Experiments 1 – 3 investigated initial retrieval preferences, the aim of Experiment 4 was to investigate the final interpretations that comprehenders are willing to give to reflexives in coargument, PNP and PPNP contexts. A subset of six conditions from the eye-movement experiments was tested as in (13). The nonlocal antecedent always matched in gender with the reflexive, while we manipulated gender congruence between the

---

[1] We also conducted an analysis that combined the data from Experiments 1 - 3. As comparisons of the PPNP and coargument reflexives did not reveal reliable local antecedent by structure interactions in first pass measures at the reflexive, we lumped the data from Experiments 1 and 3 together and compared them to the PNP reflexives from Experiment 2. For first-pass times at the reflexive, this analysis did not reveal any significant differences (all $t < 1.63$, all $p > .105$). For regression path times, the local antecedent by structure interaction was significant (estimate = 51, SE = 24, $t = 2.13$, $p = .033$), in the absence of any other reliable main effects or interactions (all other $t < 1.44$, all other $p > .150$). This additional analysis suggests that both PPNP and coargument reflexives together patterned differently to PNP reflexives in showing reliable local antecedent stereotypical gender mismatch effects at a point in time when they were not observed for reflexives inside PNPs.

reflexive and local antecedent across the three construction types tested in Experiments 1 – 3.

Participants were asked to choose who they thought the reflexive most likely referred to.

(13a) *Coargument reflexive, local antecedent match*

Jonathan was walking through the military barracks. He heard that

the soldier had positioned himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(13b) *Coargument reflexive, local antecedent mismatch*

Jennifer was walking through the military barracks. She heard that

the soldier had positioned herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(13c) *PNP reflexive, local antecedent match*

Jonathan was walking through the military barracks. He heard that

the soldier had a picture of himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(13d) *PNP reflexive, local antecedent mismatch*

Jennifer was walking through the military barracks. She heard that

the soldier had a picture of herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(13e) *PPNP reflexive, local antecedent match*

Jonathan was walking through the military barracks. He heard about

the soldier's picture of himself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

(13f) *PPNP reflexive, local antecedent mismatch*

Jennifer was walking through the military barracks. She heard about

the soldier's picture of herself in the middle of the mess hall.

The food being served for dinner did not look very appetising.

The different constructions tested in Experiments 1 – 3 have been manipulated such that (13a/b) tests coargument reflexives, (13c/d) PNP reflexives and (13e/f) PPNP reflexives. In (13a/c/e) both antecedents match in gender with the reflexive. In (13b/d/f) the local antecedent mismatches in stereotypical gender with the reflexive, while the nonlocal antecedent matches. As such, we are able to gauge the extent to which the nonlocal antecedent is considered when the reflexive either matches or mismatches in gender with the reflexive, and also the extent to which preferences for the nonlocal antecedent may differ across construction types.

If binding acts as a 'hard' constraint to retrieval in coargument contexts (Dillon et al., 2013), participants should always choose the accessible antecedent, irrespective of the gender manipulation in (13a/b). The classic account of binding (Chomsky, 1981, 1986) would predict that reflexives in PNPs should prefer a local antecedent in the same clause, if one is available. This would predict that participants in Experiment 4 should choose the local antecedent for PNPs in (13c/d) to the same extent as coarguments reflexives in (13a/b). Fewer local antecedent responses for PNPs in (13c/d) than the coarguments in (13a/b) would be compatible with accounts in which PNP reflexives are exempt from binding constraints (Pollard & Sag, 1992; Reinhart & Reuland, 1993). Both classic accounts of binding (Chomsky, 1981, 1986) and revised accounts in terms of coargument (Pollard & Sag, 1992; Reinhart & Reuland, 1993) would predict that PPNP reflexives should bind to the possessor. In this case, participants should choose the local antecedent equally as often in (13e/f) as (13a/b). Previous experiments investigating PPNP reflexives have however shown that participants are willing to consider an antecedent other than the possessor when asked to

make a conscious judgement (Keller & Asudeh, 2001; Runner et al. 2003, 2006). If these results replicate here, we would expect fewer local antecedent choices for PPNP reflexives in (13e/f) than for coargument reflexives in (13a/b).

Method

Participants

36 native English speakers (9 males, mean age 23.2) volunteered to take part, none of whom took part in any of the eye-tracking experiments.

Materials

Materials consisted of the 32 items in the two nonlocal antecedent match conditions tested in Experiments 1 − 3. As these 32 items do not divide equally into the six conditions in Experiment 4, an additional 4 items were constructed giving a total of 36 items (see Appendix). 54 fillers were also constructed that contained reflexives and pronouns in a variety of different constructions that had either a single or multiple gender matching antecedents available to them in the discourse.

Procedure

The experimental items were distributed across six lists in a latin-square design and pseudo-randomised with the fillers such that no two critical items appeared next to each other. Both forward and reverse pseudo-randomised orders were presented to participants.

The questionnaire was administered over email as a Word document. Critical items were displayed across three lines of text as in the eye-movement experiments, with the reflexives and filler pronouns being contained within a box. Below each piece of discourse, two antecedents appeared as choices (A) and (B) across two lines (e.g. in (13a) the choices

were (A) The soldier or (B) Jonathan). Participants were given instructions to choose who they thought the boxed reflexive or pronoun most likely referred to. Possible responses were either person (A), person (B) or either of them. Participants responded by choosing either '(A)', '(B)' or 'Either' from a drop-down list that appeared in a column to the right of each critical sentence. The number of local and nonlocal antecedent responses that referred to either person (A) or (B) was counterbalanced across the 36 critical items.

Results

The percentage of local antecedent, nonlocal antecedent and 'either' responses for each condition are provided in Table 7.

(TABLE 7 ABOUT HERE)

Table 7 suggests that participants almost always chose the local antecedent in coargument contexts. There was a slightly weaker preference for the local antecedent in PNP contexts, which was weaker again in PPNP contexts. Additionally, a gender mismatching local antecedent appears to decrease the proportion of local antecedent choices.

For analysis, we coded responses binomially as either 'local antecedent' or 'other response'. We collapsed nonlocal antecedent and 'either' responses into a single category as choosing either the nonlocal antecedent or 'either antecedent' corresponds to an interpretation that is not predicted by classic binding theory. The data were analysed using a mixed logit model (see Jaeger, 2008) with the maximal random effects structure. The 3-level fixed effect 'construction type' (coargument, PNP, PPNP) was Helmert coded, with one contrast comparing coargument reflexives to the average response for PNP and PPNP reflexives

combined, and a second contrast comparing PNP to PPNP reflexives. The 2-level fixed effect 'local antecedent gender' (match vs. mismatch) was sum coded.

The contrast between coargument reflexives and those inside picture noun phrases was significant (estimate = 0.91, SE = 0.17, z = 5.22, p < .001), indicating there were more local antecedent responses for coargument reflexives (96%) than those in either type of picture noun phrase (79%). The contrast between PNP and PPNP reflexives was also significant (estimate = 0.54, SE = 0.12, z = 4.39, p < .001). There were more local antecedent responses for reflexives inside PNPs (86%) than PPNPs (72%). The main effect of local antecedent gender was also significant (estimate = 1.40, SE = 0.35, z = 4.03, p < .001), with more local antecedent responses when the gender of the local antecedent matched the gender of the reflexive in comparison to when it mismatched (89% vs. 80% respectively). Although the difference between local antecedent gender match and mismatch conditions was smaller for coargument reflexives (4%) than reflexives in either PNPs (11%) or PPNPs (13%), the two interactions between construction type and local antecedent gender were not significant (both z < 1, both p > .579).

Discussion

The results of Experiment 4 contrast with the eye-movement experiments reported above. For coargument reflexives, we observed a strong local antecedent preference in both Experiments 1 and 4. For PNP and PPNP reflexives in Experiments 2 and 3, we argued that our results indicated that local antecedents were preferentially retrieved. The results of Experiment 4 indicate that comprehenders are sometimes willing to ultimately consider a nonlocal antecedent in these contexts when consciously asked to make a choice between two antecedents. Perhaps the most surprising result from Experiment 4 is that PPNPs received the fewest local antecedent choices overall. This would not be predicted by both classic and

revised accounts of binding (Chomsky, 1981, 1986; Reinhart & Reuland. 1993; Pollard & Sag, 1992) which claim that reflexives in PPNPs must be bound to the possessor. We discuss the results of Experiment 4, along with Experiments 1 – 3, in more detail below.

General Discussion

The primary aim of this study was to investigate the time-course of anaphora resolution for reflexives in different structural contexts. We were particularly interested in examining the extent to which antecedents hypothesised to be accessible and inaccessible according to binding theory are retrieved during processing for coargument reflexives in comparison to those inside picture noun phrases. The results of Experiment 4 indicated that comprehenders are willing to interpret a reflexive inside either a PNP or PPNP as referring to a nonlocal antecedent, while this interpretation was almost never considered for coargument reflexives. In three eye-movement experiments investigating coargument, PNP and PPNP reflexives however, we found reliable evidence that local antecedents were retrieved at a point in time when no reliable influence of the nonlocal antecedent was observed. Where we did observe significant differences between reflexives in different contexts relates to the time-course of the local antecedent gender mismatch effect. This was observed during first-pass processing of the reflexive in coargument contexts. There appeared to be no delay in the onset of this effect for PPNP reflexives relative to coargument reflexives. However, there was a delay in onset of the local antecedent gender mismatch effect for PNP reflexives compared to coargument reflexives. We discuss the implications of these findings for models of memory retrieval during language processing, and discuss how they relate to existing studies of reflexives in picture noun phrases and different accounts of binding constraints in turn below.

Memory retrieval during anaphora resolution

In each eye-movement experiment, we observed longer reading times at or shortly after the reflexive when it mismatched in stereotypical gender with the local antecedent. For coargument reflexives, this replicates previous results suggesting that binding constraints provide a highly weighted cue to retrieval (Clackson et al., 2011; Dillon et al., 2013; Nicol & Swinney, 1989; Sturt, 2003; Xiang et al., 2009). We suggest that the results of PNP and PPNP reflexives in Experiments 2 and 3 respectively indicate that reflexives in such contexts also preferentially cue retrieval of an antecedent within the most local domain.

More generally, we believe our findings are compatible with recent claims that retrieval cues during language processing may not be equally weighted (Van Dyke & McElree, 2011; Dillon et al., 2013). We suggest that for reflexives in both coargument and (P)PNP contexts, locality constraints are more highly weighted retrieval cues than gender agreement. We hypothesised that the strongest evidence of gender being a highly weighted retrieval cue would have been from an 'attraction' effect as is observed in subject-verb agreement (Wagers et al., 2009; Dillon et al., 2013). This would predict that the size of the local antecedent gender mismatch effect should be attenuated when the nonlocal antecedent matched in gender with the reflexive. We did not observe a significant effect of this type. To the extent that reading times in the local antecedent mismatch conditions were influenced by the nonlocal antecedent, the observed effects actually went in the *opposite* direction to this prediction. We acknowledge that we cannot conclude that gender cues are not used at all during antecedent retrieval. However, we maintain that our results indicate that, unlike closely related phenomena such as subject-verb agreement (Wagers et al., 2009; Dillon et al., 2013), agreement features do not constitute highly weighted cues to antecedent retrieval during the resolution of reflexives. Importantly, our results suggest this finding holds for both reflexives in 'standard' coargument contexts and those inside picture noun phrases.

Our lack of attraction effects could potentially be because the structurally inaccessible antecedent was linearly more distant to the reflexive than the accessible antecedent in our experiments. We think this is unlikely to be the case however. Wagers et al. (2009) reported attraction effects for subject-verb agreement when the 'attractor' both intervened between the target noun and verb, and also when the attractor was linearly further away.

In addition to attraction, we discussed other ways in which the nonlocal antecedent could have affected antecedent retrieval. For example, we also considered the possibility that we might observe main effects of *both* antecedents, such that reading times could have been longer when either the local or nonlocal antecedent mismatched in gender with the reflexive. Even if gender is not a highly weighted cue to retrieval, such an effect would be compatible with the hypothesis that highly salient but binding theory inaccessible subjects can occasionally be retrieved (Cunnings & Felser, 2013). We however again failed to observe a significant effect of this type for either coargument or (P)PNP reflexives. Note that Cunnings & Felser only observed interference of this sort when the structurally inaccessible antecedent intervened between the reflexive and accessible antecedent. They did not observe such effects when the structurally accessible antecedent was linearly closer to the reflexive, as was the case in the experiments reported here.

Although found in different reading time measures in each experiment, we did observe numerical trends in Experiments 1 and 2, that were reliable in Experiment 3, for longer reading times when the nonlocal antecedent matched in gender with the reflexive. Different accounts of such effects have been proposed. In a self-paced reading study, Badecker and Straub (2002) reported longer reading times when a local and nonlocal antecedent *both* matched in gender with a reflexive. They interpreted this finding as indexing competition between the two antecedents. An alternative account of such *multiple match* effects is provided by Dillon et al., who note that in feature-overwriting models (e.g. Nairne,

1990), the fidelity of a representation in memory may degrade if it matches in content with other representations. In this way, multiple match effects might indicate that the target representation has degraded in quality, with reading time slowdowns in such cases indicating impeded access to a degraded representation (see Dillon 2011: 93-94). As such, longer reading times in the local antecedent match, nonlocal antecedent match conditions (9a, 10a, 12a) compared to the local antecedent match, nonlocal antecedent mismatch conditions (9b, 10b, 12b) may indicate relative difficulty in accessing a degraded target representation rather than necessarily implicating active consideration of the nonlocal antecedent.

Note that this feature overwrite account only provides an explanation for longer reading times in conditions (9a, 10a, 12a) when both the local and nonlocal antecedents share the same gender features. In local antecedent mismatch, nonlocal antecedent match conditions (9c, 10c, 12c), the two antecedents do not both match with regard to the relevant gender features for overwrite to occur. We however observed trends for main effects of gender matching nonlocal antecedents, in both local antecedent match and local antecedent mismatch conditions. How to account for this finding is unclear. It could be that these trends do indeed index some consideration of the nonlocal antecedent. The observed longer reading times in the nonlocal antecedent match conditions could perhaps index coercion of an otherwise dispreferred interpretation in which the reflexive refers to the gender-matching nonlocal antecedent. The fact that these trends were delayed in comparison to the local antecedent mismatch effects, might suggest that they index reanalysis following an initial retrieval of the local antecedent. We emphasise however that different nonlocal antecedent effects have been reported previously which have not always replicated. This account thus remains speculative until further research confirms the precise nature of this possible effect.

Irrespective of how these nonlocal antecedent effects are interpreted, the results of Experiment 4 clearly show that comprehenders are willing to interpret reflexives as referring

to nonlocal antecedents in PNP and PPNP contexts. We suggest that the relative time-course of local and nonlocal antecedent effects observed in the eye-movement experiments, along with the offline results of Experiment 4, may suggest a time-course difference such that local antecedents are initially retrieved, but that this initial retrieval can potentially be overriden. We suggest that if certain retrieval cues are more highly weighted than others (Van Dyke & McElree, 2011), it naturally follows that there may be instances in which an initial retrieval needs to be revised. Thus, we suggest that highly weighted locality constraints, which will lead to successful interpretation the vast majority of the time, are so strongly weighted that they favour an initial retrieval of the local antecedent, that may subsequently be revised, for reflexives in both coargument and non-coargument contexts.

One remaining question relates to how locality constraints are implemented. Dillon et al. (2013) claimed that syntactic cues, such as 'clause-mate subject', gate access to the local subject for coargument reflexives. The eye-movement results for coargument and PNP reflexives reported here could also be explained in terms of such a syntactic cue. As the nonlocal antecedent in Experiments 1 and 2 was not in the same clause as the reflexive, the cue 'clause-mate subject' would uniquely cue retrieval of the local antecedent. The eye-movement results for PPNP reflexives are less clearly explained in this way however, as in Experiment 3 the nonlocal antecedent was also a clause-mate subject in the same clause as the reflexive and local antecedent. Yet, we still observed a preference for the possessor to be initially retrieved. It might be that the locality requirement of reflexives is better explained not in terms of constraints cueing a clause-mate antecedent but instead a predicate-mate antecedent, if one is available. This possibility is discussed in more detail in the next section.

Another factor that may influence the locality preference is that the local antecedent is the most recently encoded item in memory at the point when the reflexive is encountered. As such, this most recent item might have a particularly high level of activation, strongly biasing

its retrieval. Even in cases when the accessible antecedent is linearly more distant to the reflexive than an inaccessible antecedent, as in (5b), retrieval of the main clause subject at the verb preceding the reflexive may increase activation levels of the structurally accessible antecedent. Dillon et al. (2013) reported that computational simulations suggested that this reactivation increases the likelihood of the accessible antecedent being retrieved. Whether the latent activation levels of a local antecedent can explain interpretive preferences for reflexives across contexts is currently not well understood, but it remains possible that the locality preferences observed in Experiments 1 – 3 are at least partially attributable to the activation levels of the local antecedent in memory at the point of retrieval.

Note that our prediction of an initial preference for local antecedents, even in the case of PNP and PPNP reflexives, is incompatible with results from the visual world paradigm (Runner et al., 2003, 2006). We discuss how our data relate to previous studies of reflexives in (P)PNPs and different theoretical accounts of binding constraints, in turn below.

Anaphora resolution in picture noun phrases

Recall that Kaiser et al. (2009) examined reflexives inside PNPs in sentences such as 'Peter told/heard from John about the picture of himself on the wall', and found that pragmatic properties of the text influenced which antecedent was retrieved. Note that in this context, *both* antecedents are in the same clause as the reflexive. Even standard reflexives without picture noun phrases (e.g. Peter told/heard from John about himself the other day) can be bound to either antecedent in such cases (Domínguez, Hicks, & Song, 2012; Hicks, 2009; Pollard & Sag, 1992). In the current Experiment 2, only one antecedent was in the same clause as the reflexive, and we have argued that the eye-movement data reported here suggest an initial preference for this antecedent. Together, the results from both our Experiment 2 and from Kaiser et al. are compatible with the hypothesis that PNP reflexives

prefer an antecedent in the same clause. Other factors, including pragmatic notions such as the 'source' vs. 'perceiver' of information, may provide additional retrieval cues when multiple antecedents are available in the same clause.

Our results for PNPs with a possessor are however different to a previous study by Runner et al. (2003, 2006). We claimed that the possessor is initially retrieved. Recall that for instructions such as 'Look at Ken. Have Joe touch Harry's picture of himself', Runner et al. found that participants' eye-movements across a visual display shortly after the onset of the reflexive fixated on both the possessor *Harry* and the subject of the sentence *Joe* but not the antecedent in the lead-in sentence, *Ken.* Runner et al. claimed that these results indicated that antecedents other than the possessor were immediately accessed. We note that a second visual paradigm study failed to replicate this result, and instead observed a preference for the possessor only during early stages of processing (Kaiser et al., 2009). However, the results of Runner et al. (2003, 2006) merit further discussion.

One difference between our Experiment 3 and those reported by Runner et al. is that we did not require participants to make an overt response regarding their interpretive preferences for the reflexive. It could be that task demands influence the extent to which antecedents other than the possessor are considered, such that in tasks where intuitions on anaphor resolution are not overtly probed, readers tend to interpret reflexives inside PPNPs as referring to the possessor. Task demands could also potentially account for the strong local antecedent preferences we observed in Experiments 1 – 3 versus the relative consideration of the nonlocal antecedent, particularly in (P)PNP constructions, in Experiment 4. Task demands have been shown to influence the extent to which 'good enough' strategies are employed during syntactic ambiguity resolution (Swets, Desmet, Clifton & Ferreira, 2008). The extent to which tasks demands may influence antecedent retrieval strategies during the interpretation of reflexives is thus a possible avenue of future research.

Two other properties of the visual world paradigm might also explain the discrepancy in early effects reported here and those reported by Runner et al. (2003, 2006). It could be that the visual displays used by Runner et al. meant that both the local and nonlocal antecedents were more prominent than is usually the case in reading paradigms. In this way, the early effects of antecedents other than the possessor observed by Runner et al. might indicate that the nonlocal antecedents were more salient in their studies than in the current experiments. We note that this interpretation of Runner et al.'s early nonlocal antecedent effects would still implicate consideration of both local and nonlocal antecedents during early stages of processing, contra our claims, if the nonlocal antecedent is prominent enough.

We do however believe that another interpretation of results from the visual world paradigm is possible that does not necessarily implicate antecedent retrieval during anaphor resolution. In a visual world paradigm study, Dahan and Tanenhaus (2005) report that upon hearing a word such as *snake*, participants' eye-movements across a visual display fixated both on the correct referent (a snake) and a visual competitor (a coil of rope). These results indicate that even in cases of *unambiguous* reference, *visual* competitors in a display can attract fixations. We suggest that this leaves open the possibility that in studies of anaphora resolution using the visual world paradigm, *some proportion* of looks to competitor antecedents *may* indicate competition during visual search, rather than necessarily implicating that comprehenders are actively considering the particular object in the visual display as an antecedent. Thus, although Runner et al.'s results *may* indicate early retrieval of nonlocal antecedents, without *unambiguous reference* controls (e.g. 'Have Joe touch Harry's picture of Ken'), we maintain that it is difficult to disentangle effects of competition during visual search from active consideration of nonlocal antecedents during memory retrieval.

In summary, we believe the results of Experiment 2 reported here, which examined PNP reflexives, are compatible with previous findings reported by Kaiser et al. (2009). For

PPNP reflexives, we note that while previous studies and our own Experiment 4 show that participants may ultimately be willing to choose an antecedent other than the possessor, we argue that the results of Experiment 3 suggest an initial preference for the local antecedent.

We now turn to discussion of our findings in relation to different theoretical accounts of binding. In particular, we focus here on the results of Experiment 4 and the observed differences in time-course of the local antecedent gender mismatch effect in the eye-movement experiments. This difference in time-course is illustrated in Figure 1, which plots a measure known as cumulative progression (see Kreiner et al. 2008). This figure is included as a graphical illustration of the time-course differences in the local antecedent gender mismatch effect observed across Experiments 1 − 3, which is backed up statistically by the between-experiment analyses on the first-pass measures reported above. Cumulative progression was measured by taking the character position of the first fixation on the critical reflexive, and then calculating how far forward (in characters) the reader progressed in the sentence over the next 800ms in a series of 10ms intervals. The Y-axis of the graph shows the *difference* in cumulative progression between the local antecedent match and mismatch conditions. Thus, a positive value on the Y-axis indicates that the reader has progressed further through the sentence in the local antecedent match condition relative to the mismatch condition. The X-axis indicates time, in milliseconds, from the onset of the first fixation on the critical reflexive. As such, the onset of the mismatch effect in each experiment can be estimated by examining the point where the progression difference begins to exceed zero. It can be seen that coargument and PPNP reflexives show a similar profile, with a relatively early onset of the mismatch effect around 200msec following the first fixation on the reflexive, while PNP reflexives show a later onset, at around 400msec.

(FIGURE 1 ABOUT HERE)

We first compare our results for standard coargument reflexives to reflexives inside PNPs without a possessor. The results of Experiment 4 showed that comprehenders are more willing to consider a nonlocal antecedent for a reflexive inside a PNP than for coargument reflexives. This finding is compatible with theoretical accounts which assume that reflexives inside PNPs are exempt from binding constraints (Pollard & Sag, 1992; Reinhart & Reuland, 1993), but is difficult to reconcile with classic binding theory (Chomsky, 1981, 1986) which predicts that a reflexive inside a PNP must bind to the structurally most local antecedent. Note also that Pollard and Sag's (1992) revised account assumes that reflexives in PNPs are subject to an 'intervention' constraint which states that they must refer to the most local animate antecedent. The results of Experiment 4 would not be predicted on these grounds, and instead suggest that however the locality requirement of reflexives in PNPs is best characterised, it is not an absolute preference.

We also observed a difference in the time-course of the local antecedent gender mismatch effect for reflexives in coargument and PNP contexts. While this effect was observed during first-pass processing at the reflexive in coargument contexts, it was comparatively delayed until second pass processing of the reflexive, and first and second pass processing at the spillover region, for PNPs. In standard theoretical accounts, coargument reflexives are bound to an antecedent within the most local syntactic domain. In Experiment 1, this amounts to the reflexive cueing retrieval of a predicate-mate antecedent, in this case the local subject of the same verbal predicate as the reflexive. The different theoretical accounts of PNPs discussed above assume that picture noun phrases themselves form a type of nominal predicate. Under this analysis, the reflexive inside a possessorless PNP has no coargument that it can bind to in the most local domain (the nominal PNP predicate). The relative delay in the time-course of the local antecedent gender mismatch effect for PNPs

may thus indicate relative difficulty in retrieving an antecedent that is not within the most local syntactic domain.

Turning to PNPs with a possessor, Experiment 4 indicated that participants were willing to consider a nonlocal antecedent for reflexives in PPNPs approximately 28% of the time. This finding is incompatible with both classic binding theory (Chomsky, 1981, 1986) and revised accounts which predict that reflexives in PPNPs must be bound to the possessor (Pollard & Sag, 1992). These results are however compatible with the claim that both PNPs and PPNPs should be considered exempt from binding theory (Runner et al., 2003, 2006). One surprising finding is that nonlocal antecedent choices were more frequent for reflexives in PPNPs than in PNPs without a possessor. One possible reason for this difference is that the nonlocal antecedent in our PPNP stimuli, as in (13e/f), was in the same clause as the local antecedent, while in our PNP stimuli, as in (13c/d), the nonlocal antecedent was in a higher clause. This was necessarily to ensure that our stimuli were maximally similar at and after the reflexive in the three eye-movement experiments, but leaves open the possibility that clause structure may influence the extent to which nonlocal antecedents are considered.

While Experiments 1 and 2 indicated a difference in time-course of the local antecedent gender mismatch effect for reflexives in coargument contexts compared to those in PNPs, the results of Experiments 1 and 3 suggested no relative delay in the onset of this effect for PPNP compared to coargument reflexives. If picture noun phrases form a nominal predicate (Chomsky 1981, 1986; Pollard & Sag, 1992), the possessor can be considered as an antecedent within this most local domain. Under this analysis, both coargument and PPNP reflexives can cue retrieval of a predicate-mate antecedent, and as such the ease with which this antecedent is retrieved during processing is similar in both cases. Note however, that even if this is indeed the case, the results of Experiment 4 still clearly show that the local antecedent preference is less absolute in PPNP than coargument contexts.

The theoretical consequences of Experiments 1 – 4 could potentially be explained in different ways. One possibility could be to maintain that something akin to binding theory applies but only *initially* across contexts, and that binding constraints are more easily overridden in PNPs and PPNPs. Another possibility could be that binding constraints strictly only apply to coargument reflexives and that other contexts, such as PNPs and PPNPs, are exempt from binding theory but an analogous locality constraint strongly, but not uniquely, biases the local antecedent. A further hypothesis could be that there is no strict dichotomy between coargument and non-coargument relations, but that preferences are graded across contexts. It is difficult to choose between these proposals based on eye-movement data alone. We do note however that if picture noun phrases are completely exempt from binding constraints, this would predict that reflexives and pronouns should be equally acceptable in such cases, as has been claimed by some authors (see Reinhart & Reuland, 1993: 661). However, Keller and Asudeh (2001) showed that in sentences such as 'Hanna found a picture of her/herself' the reflexive was considered more acceptable. Runner et al. (2003, 2006) found that reflexives prefer local antecedents and pronouns prefer nonlocal antecedents in PPNPs, although these preferences were not as strong as the strict complementarity found in coargument contexts. Although in-depth discussion of these issues is beyond the scope of this paper, the finding that pronouns *do not* prefer a local antecedent in non-coargument contexts is unexpected under accounts which predict a strict dichotomy between coargument and non-coargument relations (Pollard & Sag, 1992; Reinhart & Reuland, 1993).

Finally, we note that in a previous study Cunnings and Felser (2013) reported that individual differences in the processing of coargument reflexives correlated with performance on a reading span test (Daneman & Carpenter, 1980). We did not collect reading span scores for the participants in the experiments reported here. We acknowledge that this leaves open the possibility that individual differences in reading span may affect anaphora resolution

differently in coargument and non-coargument contexts. As we did not collect individual differences measures in our eye-movement studies, we cannot tell whether individual differences may have affected our results. We note however, that closer inspection of the results of the antecedent choice task may suggest differences between participants. For the two coargument conditions, 34 out of 36 participants chose the local antecedent over 90% of the time, while 20 and 14 participants chose it over 90% of the time in PNP and PPNP contexts respectively. Thus, while some participants showed near unanimous preferences for the local antecedent, even in PNP and PPNP contexts, others were more willing to entertain nonlocal antecedent interpretations. The extent to which individual differences scores such as the reading span test can explain these differences, as well as potential differences during processing, may thus be an interesting avenue of future research.

Conclusion

The results of three eye-movement experiments showed that readers initially preferred to resolve a reflexive anaphor as referring to a structurally local antecedent. These results are compatible with claims that different cues to memory retrieval during language processing may not be equally weighted (Dillon et al., 2013; Van Dyke & McElree, 2011), and that locality constraints are more heavily weighted retrieval cues than gender for reflexives across different syntactic contexts. We also observed a difference in the time-course of local antecedent gender mismatch effects across different syntactic environments. For coargument reflexives, this gender mismatch effect occurred at a point in time earlier than for reflexives inside picture noun phrases. No time-course differences were observed between coargument reflexives and reflexives inside possessed picture noun phrases. We suggest these differences index the relative ease of accessing the local antecedent across different syntactic contexts.

References

Baayen, H. (2008). *Analyzing linguistic data. A practical introduction to statistics using R*. Cambridge: Cambridge University Press.

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

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

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

Büring, D. (2005). *Binding theory*. Cambridge: Cambridge University Press.

Chomsky, N. (1981). *Lectures on government and binding*. Dordrecht: Foris.

Chomsky, N. (1986). *Knowledge of language. Its nature, origin, and use*. New York: Praeger.

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

Cowan, N. (2000). The magic number 4 in short-term memory: A reconsideration of mental storage capacity. *Behavioral and Brain Sciences*, 24, 87-185.

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

Dahan, D., & Tanenhaus, M. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychological Bulletin & Review*, 12, 455-459.

Daneman, M., & Carpenter, P. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior*, 19, 456-466.

Dillon, B. (2011). *Structured access in sentence comprehension*. Unpublished PhD thesis, University of Maryland.

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

Domínguez, L., Hicks, G., & Song, H. (2012). Untangling locality and orientation constraints in the L2 acquisition of anaphoric binding: a feature based approach. *Language Acquisition*, 19, 266-300.

Foraker, S. & McElree, B. (2007). The role of prominence in pronoun resolution: Availability vs. accessibility. *Journal of Memory and Language*, 56, 357-383.

Hicks, G. (2009). *The derivation of anaphoric dependencies*. Amsterdam: John Benjamins.

Jaeger, F. (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. *Journal of Memory and Language*, 59, 434-446.

Kaiser, E., Runner, J., Sussman, R., & Tanenhaus, M. (2009). Structural and semantic constraints on the resolution of pronouns and reflexives. *Cognition*, 112, 55-80.

Keller, F., & Asudeh, A. (2001). Constraints on linguistic coreference: Structural vs. pragmatic factors. In Moore, J., & Stenning, K. (eds.), *Proceedings of the 23rd annual conference of the cognitive science society* (pp. 483-488). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Kennison, S., & Trofe, J. (2003). Comprehending pronouns: A role for word-specific gender stereotype information. *Journal of Psycholinguistic Research*, 32, 355-378.

Kreiner, H., Sturt, P., & Garrod, S. (2008). Processing definitional and stereotypical gender in reference resolution: Evidence from eye-movements. *Journal of Memory and Language*, 58, 239-261.

Kuno, S. (1987). *Functional syntax: Anaphora, discourse and empathy*. Chicago: University of Chicago Press.

Lewis, R. L., Vasishth, S., and Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10, 44-54.

Lewis, R. L. & Vasishth, S. (2007). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science*, 29, 1-45.

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

McElree, B. & Dosher, B.A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General,* 122, 291-315.

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

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

McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 817-835.

McElree, B. (2006). Accessing recent events. In Ross, B. (ed.), *The psychology of learning and motivation Vol. 46* (pp. 155-200). San Diego: Academic Press.

Nairne, J. (1990). A feature model of immediate memory. *Memory & Cognition*, 18, 251-269.

Nicol, J., & Swinney, D. (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research,* 18, 5-20.

Nicol, J., & Swinney, D. (2003). The psycholinguistics of anaphora. In Barss, A. (ed.), *Anaphora: A reference guide* (pp. 72-104). Oxford: Blackwell.

Pearlmutter, N., Garnsey, S., & Bock, K. (1999). Agreement processes in sentence comprehension. *Journal of Memory and Language*, 41, 427-456.

Pollard, C., & Sag, I. (1992). Anaphors in English and the scope of binding theory. *Linguistic Inquiry*, 23, 261-303.

Reinhart, T. (1983). *Anaphora and semantic interpretation*. London: Croom Helm.

Reinhart, T., & Reuland, E. (1993). Reflexivity. *Linguistic Inquiry*, 24, 657-720.

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

Reuland, E. (2011). *Anaphora and language design*. Cambridge, Massachusetts: The MIT Press.

Sanford, A., & Garrod, S. (1988). Proper names as controllers of discourse focus. *Language and Speech*, 31, 43-56.

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

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

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

Van Dyke, J. & McElree, B. (2006). Retrieval interference in sentence processing. *Journal of Memory and Language*, 55, 157-166.

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

Van Dyke, J. & Johns, C. (2012). Memory interference as a determinant of language comprehension. *Language and Linguistics Compass*, 6, 193-211.

Wagers, M., Lau, E. & Phillips, C. (2009). Attraction in comprehension: representation and processes. *Journal of Memory and Language*, 61, 206-237.

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

Appendix

The materials for Experiments 1 - 4 are given below. Items 1 – 32 were used in each experiment, while items 33 – 36 were additionally tested in Experiment 4. The local and nonlocal antecedent manipulations are shown in parentheses. Square brackets denote the manipulations across coargument, picture noun phrase and possessed picture noun phrase constructions, delimited with a forward slash (/).

1   Jonathan (Jennifer) was walking through the military barracks. He (She) [heard that the soldier had positioned / heard that the soldier had a picture of / heard about the soldier's picture of] himself (herself) in the middle of the mess hall. The food being served for dinner did not look very appetizing.

2   Timothy (Miranda) was confused by the size of the depot. He (She) [knew that the builder had lost / knew that the builder kept a photo of / knew about the builder's photo of] himself (herself) near the back of the store. It was easy to get lost amongst all the supplies.

3   Nicholas (Victoria) worked in the fire station in town. He (She) [knew that the firefighter had injured / knew that the firefighter left a portrait of / knew about the firefighter's portrait of] himself (herself) by the table in the room. It was a difficult job that required being on call for many hours.

4   Michael (Barbara) enjoyed the trip in the small aircraft. He (She) [remembered that the pilot had readied / remembered that the pilot kept a painting of / remembered about the pilot's painting of] himself (herself) near the front of the plane. The view up in the sky was certainly something to behold.

5   Richard (Jessica) remembered the tour of the national stadium. He (She) [recalled that the footballer had prepared /recalled that the footballer hung a picture of / recalled

about the footballer's picture of] himself (herself) in the brand new locker room. It had been a very exciting day out indeed.

6   Steven (Joanna) worked out after a long day at the office. He (She) [learnt that the boxer had coached / learnt that the boxer kept a photo of / learnt about the boxer's photo of] himself (herself) towards the back of the gym. It took a lot of hard work to stay fit and healthy.

7   Joshua (Rachel) tries to support the small shops in town. He (She) [remembered that the butcher had cut / remembered that the butcher hung a portrait of / remembered about the butcher's portrait of] himself (herself) behind a table in the shop. Despite everything, business was doing well during the recession.

8   William (Susanne) remembered how bad the boat trip had been. He (She) [recalled that the sailor had positioned / recalled that the sailor had a painting of / recalled about the sailor's painting of] himself (herself) next to a wheel on the bridge. It was a shame that the water had been so rough.

9   Benjamin (Beverley) visited the garage in the city centre. He (She) [thought that the mechanic had hurt / thought that the mechanic kept a video of / thought about the mechanic's video of] himself (herself) next to some old power tools. It was probably time that some new equipment was bought.

10   Gregory (Hillary) had lived in the mining village for many years. He (She) [heard that the miner had wounded / heard that the miner left a tape of / heard about the miner's tape of] himself (herself) underground in the old mine shaft. It was a difficult job but somebody had to do it.

11   Clarence (Caroline) reminisced about the country walk. He (She) [remembered that the farmer had stopped / remembered that the farmer kept a sketch of / remembered

about the farmer's sketch of] himself (herself) near the tractor in the field. The weather had certainly been beautiful that day.

12   Anthony (Abigail) was out on the road learning the trade. He (She) [remembered that the electrician had seated / remembered that the electrician lost a drawing of / remembered about the electrician's drawing of] himself (herself) toward the back of the van. There was a lot to learn on the new job.

13   Bernard (Bethann) had watched the operation in the hospital. He (She) [knew that the surgeon had trained / knew that the surgeon had a video of / heard about the surgeon's video of] himself (herself) to help get through the surgery. Luckily everything went well in the end.

14   Derrick (Deirdre) had been having trouble with a leaky sink. He (She) [recalled that the plumber had sprayed / recalled that the plumber had a tape of / recalled about the plumber's tape of] himself (herself) while trying to fix the taps. Luckily it didn't take long to find the source of the leak.

15   Desmond (Dorothy) didn't like being in the wrong part of town. He (She) [recalled that the criminal had disguised / recalled that he criminal dropped a sketch of / recalled about the criminal's sketch] himself (herself) beside the bins in the alley. It really was time to get back home.

16   Charles (Kathryn) enjoyed the charity event at city hall. He (She) [remembered that the bodyguard had positioned / remembered that the bodyguard had a drawing of / remembered about the bodyguard's drawing of] himself (herself) towards the back of the room. There were lots of important people there that night.

17   Jennifer (Jonathan) was working late at the police station. She (He) [thought that the prostitute had hurt / thought that the prostitute held a picture of / thought about the

prostitute's picture of] herself (himself) beside a chair by the cells. It had been a troublesome night so far and would likely get worse.

18   Miranda (Timothy) learnt about the future at the magic shop. She (He) [recalled that the fortune teller had prepared / recalled that the fortune teller had a photo of / recalled about the fortune teller's photo of] herself (himself) at the back of the room. The expensive trip turned out to be a waste of money in the end.

19   Victoria (Nicholas) worked for a nanny agency in town. She (He) [remembered that the babysitter had amused / remembered that the babysitter had a portrait of / remembered about the babysitter's portrait of] herself (himself) at the back of the office. It was a busy but friendly place to work.

20   Barbara (Michael) had some business to finish at the office. She (He) [learnt that the typist had prepared / learnt that the typist left a painting of / learnt about the typist painting of] herself (himself) next to a very untidy desk. The mess was going to be cleared up soon.

21   Jessica (Richard) had not worked for the company for very long. She (He) [thought that the secretary had organised / thought that the secretary hid a picture of / thought about the secretary's picture of] herself (himself) behind a big old work desk. It was going to take a little while to get used to the office.

22   Joanna (Steven) knew how tiresome it was to keep the flat tidy. She (He) [learnt that the cleaner had amused / learnt that the cleaner kept a photo of / learnt about the cleaner's photo of] herself (himself) near the table in the lounge. Sometimes the job could get a little boring.

23   Rachel (Joshua) had enjoyed the day at the health spa. She (He) [learnt that the beautician had manicured / learnt that the beautician kept a portrait of / learnt about

the beautician's portrait] herself (himself) beside a table in the salon. It had certainly been a relaxing day indeed.

24   Susanne (William) did not like having to go to the hospital. She (He) [recalled that the nurse had prepared / recalled that the nurse put a painting of / recalled about the nurse's painting of] herself (himself) near the staff schedule that day. It looked like it was going to be a very busy shift on the ward.

25   Beverley (Benjamin) was enjoying the glamorous clothes show. She (He) [knew that the fashion model had dressed / knew that the fashion model kept a video of / knew about the fashion model's video of] herself (himself) in the busy room next door. It obviously took a lot of work to look that good.

26   Hillary (Gregory) had a good time at the American football match. She (He) [knew that the cheerleader had embarrassed / knew that the cheerleader kept a tape of / knew about the cheerleader's tape of] herself (himself) by the edge of the field. The half time dance routine needed a lot of practice.

27   Caroline (Clarence) was never able to find time to tidy up. She (He) [thought that the housekeeper had busied / thought that the housekeeper left a sketch of / thought about the housekeeper's sketch of] herself (himself) by the sofa in the lounge. There was certainly a lot of tidying up to do.

28   Abigail (Anthony) had been temping with the agency for some time. She (He) [heard that the receptionist had positioned / heard that the receptionist kept a drawing of / heard about the receptionist's drawing of] herself (himself) by the old pair of shelves. The working hours in the office were long and the pay was poor.

29   Bethann (Bernard) was getting anxious ahead of the big day. She (He) [remembered that the florist had prepared / remembered that the florist watched a video of /

remembered about the florist's video of] herself (himself) ahead of the wedding next week. There was still much to get ready.

30   Deirdre (Derrick) worked for the large clothing manufacturer. She (He) [heard that the dressmaker had pricked / heard that the dressmaker found a tape of / heard about the dressmaker's tape of] herself (himself) while at work late one day. The job wasn't very interesting but the overtime pay was worth it.

31   Dorothy (Desmond) found it hard to work long hours as a parent. She (He) [knew that the childminder had strained / knew that the childminder left a sketch of / knew about the childminder's sketch of] herself (himself) next to the new lounge sofa. It was a difficult job to look after all the children.

32   Kathryn (Charles) was waiting in the beauty parlour in town. She (He) [remembered that the manicurist had entertained / remembered that the manicurist put a drawing of / remembered about the manicurist's drawing of] herself (himself) near the back of the salon. It had not been a very busy afternoon in town that day.

33   Mark (Mary) knew that something serious had happened. He (She) recalled [that the bank robber had concealed / that the banker robber hid a picture of / about the bank robber's picture of] himself (herself) behind a bin down an alley. The evidence would be very incriminating.

34   Jason (Nancy) had driven to the local petrol station. He (She) learnt [that the trucker had embarrassed / that the trucker saw a video of / about the trucker's video of] himself (herself) on the new CCTV camera system. It was quite funny.

35   Mary (Mark) worked long hours at the local hospital. She (he) heard [that the midwife had prepared / that the midwife kept a photo of / about the midwife's photo of] herself (himself) in the staffroom near the ward. It was going to be a very long and tiring shift.

36   Nancy (Jason) had really enjoyed the programme immensely. She (He) remembered [that the ballet dancer had excelled / that the ballet dancer showed a tape of / about the ballet dancer's tape of] herself (himself) on the popular new TV show. It was going to be on every Saturday night.

Table 1. *Reading times for three eye-movement measures at four regions of texts in Experiment 1 (SDs in parentheses)*

	First Pass Reading Time	Regression Path Time	Second Pass Time
Reflexive Region			
Local Match, Nonlocal Match	221 (93)	241 (133)	132 (222)
Local Match, Nonlocal Mismatch	222 (87)	256 (158)	138 (197)
Local Mismatch, Nonlocal Match	236 (122)	295 (248)	207 (265)
Local Mismatch, Nonlocal Mismatch	233 (108)	294 (212)	178 (228)
Spillover Region			
Local Match, Nonlocal Match	261 (130)	357 (334)	191 (253)
Local Match, Nonlocal Mismatch	254 (129)	332 (253)	188 (240)
Local Mismatch, Nonlocal Match	294 (163)	461 (384)	257 (310)
Local Mismatch, Nonlocal Mismatch	255 (113)	401 (363)	245 (295)
Prefinal Region			
Local Match, Nonlocal Match	240 (132)	439 (460)	193 (264)
Local Match, Nonlocal Mismatch	256 (137)	428 (377)	182 (280)
Local Mismatch, Nonlocal Match	259 (171)	531 (568)	220 (268)
Local Mismatch, Nonlocal Mismatch	253 (124)	495 (558)	207 (279)
Final Region			
Local Match, Nonlocal Match	337 (216)	840 (1119)	112 (216)
Local Match, Nonlocal Mismatch	307 (218)	660 (560)	108 (219)
Local Mismatch, Nonlocal Match	336 (263)	834 (969)	142 (287)
Local Mismatch, Nonlocal Mismatch	342 (251)	884 (1059)	113 (218)

Local = Local antecedent, Nonlocal = Nonlocal antecedent

Table 2. *Summary of statistical analyses for three eye-movement measures at four regions of texts in Experiment 1*

	First Pass Reading Time		Regression Path Time		Second Pass Time	
	Estimate	*t*	*Estimate*	*t*	*Estimate*	*t*
Reflexive Region						
Local	13 (8)	1.75 [+]	45 (15)	2.96 *	57 (20)	2.88 *
Nonlocal	1 (8)	0.07	7 (15)	0.45	12 (14)	0.81
Local * Nonlocal	3 (18)	0.18	15 (35)	0.44	34 (29)	1.17
Spillover Region						
Local	15 (12)	1.28	87 (25)	3.53 **	61 (21)	2.97 *
Nonlocal	20 (12)	1.66 [+]	36 (29)	1.22	7 (18)	0.41
Local * Nonlocal	26 (22)	1.21	35 (64)	0.55	9 (34)	0.27
Prefinal Region						
Local	7 (11)	0.65	80 (41)	1.94 [+]	26 (20)	1.28
Nonlocal	6 (10)	0.59	27 (39)	0.68	12 (21)	0.57
Local * Nonlocal	19 (23)	0.80	24 (68)	0.35	2 (37)	0.06
Final Region						
Local	16 (15)	1.05	104 (76)	1.37	18 (19)	0.91
Nonlocal	10 (17)	0.56	67 (73)	0.91	17 (15)	1.12
Local * Nonlocal	42 (33)	1.27	225 (139)	1.62	26 (31)	0.83

Local = Local antecedent, Nonlocal = Nonlocal antecedent.
Estimate = Model Estimate (SE in brackets). [+] *= p < .10, * = p < .05, ** = p < .001*

Table 3. *Reading times for three eye-movement measures at four regions of texts in Experiment 2 (SDs in parentheses)*

	First Pass Reading Time	Regression Path Time	Second Pass Time
Reflexive Region			
Local Match, Nonlocal Match	250 (104)	321 (263)	82 (157)
Local Match, Nonlocal Mismatch	252 (116)	345 (424)	94 (191)
Local Mismatch, Nonlocal Match	246 (109)	310 (215)	121 (200)
Local Mismatch, Nonlocal Mismatch	249 (115)	319 (221)	118 (219)
Spillover Region			
Local Match, Nonlocal Match	257 (123)	335 (337)	124 (202)
Local Match, Nonlocal Mismatch	249 (124)	307 (255)	101 (170)
Local Mismatch, Nonlocal Match	280 (161)	437 (559)	150 (210)
Local Mismatch, Nonlocal Mismatch	277 (145)	392 (346)	146 (234)
Prefinal Region			
Local Match, Nonlocal Match	283 (157)	417 (367)	135 (270)
Local Match, Nonlocal Mismatch	280 (155)	422 (383)	111 (205)
Local Mismatch, Nonlocal Match	274 (140)	535 (606)	160 (254)
Local Mismatch, Nonlocal Mismatch	291 (158)	458 (410)	121 (217)
Final Region			
Local Match, Nonlocal Match	324 (213)	704 (840)	89 (256)
Local Match, Nonlocal Mismatch	359 (274)	608 (610)	75 (198)
Local Mismatch, Nonlocal Match	337 (269)	659 (733)	91 (262)
Local Mismatch, Nonlocal Mismatch	359 (256)	702 (1022)	85 (237)

Local = Local antecedent, Nonlocal = Nonlocal antecedent

Table 4. *Summary of statistical analyses for three eye-movement measures at four regions of texts in Experiment 2*

	First Pass Reading Time		Regression Path Time		Second Pass Time	
	Estimate	*t*	*Estimate*	*t*	*Estimate*	*t*
Reflexive Region						
Local	2 (9)	0.17	14 (26)	0.55	32 (15)	2.16 *
Nonlocal	5 (8)	0.56	16 (23)	0.70	4 (14)	0.32
Local * Nonlocal	1 (15)	0.09	14 (46)	0.29	14 (25)	0.58
Spillover Region						
Local	23 (11)	2.05 *	89 (34)	2.66 *	36 (13)	2.80 *
Nonlocal	2 (12)	0.18	40 (37)	1.09	14 (13)	1.03
Local * Nonlocal	6 (22)	0.27	12 (63)	0.19	19 (26)	0.72
Prefinal Region						
Local	2 (10)	0.16	75 (33)	2.29 *	18 (19)	0.92
Nonlocal	7 (11)	0.62	36 (32)	1.12	32 (19)	1.66 +
Local * Nonlocal	21 (26)	0.83	77 (70)	1.10	16 (37)	0.42
Final Region						
Local	6 (17)	0.37	12 (63)	0.20	6 (19)	0.31
Nonlocal	26 (16)	1.61	39 (68)	0.58	11 (16)	0.68
Local * Nonlocal	10 (38)	0.27	106 (107)	0.99	8 (32)	0.25

Local = Local antecedent, Nonlocal = Nonlocal antecedent.
Estimate = Model Estimate (SE in brackets). [+] *= p < .10, * = p < .05, ** = p < .001*

Table 5. *Reading times for three eye-movement measures at four regions of texts in Experiment 3 (SDs in parentheses)*

	First Pass Reading Time	Regression Path Time	Second Pass Time
Reflexive Region			
Local Match, Nonlocal Match	237 (115)	296 (244)	72 (152)
Local Match, Nonlocal Mismatch	242 (121)	298 (210)	71 (158)
Local Mismatch, Nonlocal Match	260 (148)	315 (245)	108 (197)
Local Mismatch, Nonlocal Mismatch	254 (128)	334 (228)	119 (214)
Spillover Region			
Local Match, Nonlocal Match	263 (148)	333 (266)	126 (225)
Local Match, Nonlocal Mismatch	249 (137)	299 (257)	103 (180)
Local Mismatch, Nonlocal Match	264 (150)	388 (395)	152 (223)
Local Mismatch, Nonlocal Mismatch	268 (153)	354 (285)	96 (179)
Prefinal Region			
Local Match, Nonlocal Match	276 (199)	405 (383)	125 (199)
Local Match, Nonlocal Mismatch	286 (155)	388 (265)	111 (192)
Local Mismatch, Nonlocal Match	270 (142)	428 (440)	120 (202)
Local Mismatch, Nonlocal Mismatch	294 (191)	463 (525)	123 (205)
Final Region			
Local Match, Nonlocal Match	344 (238)	665 (752)	84 (191)
Local Match, Nonlocal Mismatch	337 (216)	603 (585)	65 (164)
Local Mismatch, Nonlocal Match	342 (217)	680 (654)	82 (180)
Local Mismatch, Nonlocal Mismatch	369 (265)	627 (583)	93 (232)

Local = Local antecedent, Nonlocal = Nonlocal antecedent

Table 6. *Summary of statistical analyses for three eye-movement measures at four regions of texts in Experiment 3*

	First Pass Reading Time		Regression Path Time		Second Pass Time	
	Estimate	*t*	*Estimate*	*t*	*Estimate*	*t*
Reflexive Region						
Local	17 (9)	1.86 [+]	27 (15)	1.78 [+]	42 (14)	3.08 *
Nonlocal	0 (8)	0.05	9 (18)	0.50	5 (12)	0.40
Local * Nonlocal	11 (17)	0.65	16 (31)	0.51	11 (24)	0.47
Spillover Region						
Local	11 (11)	0.97	51 (24)	2.12 *	10 (14)	0.67
Nonlocal	3 (11)	0.24	26 (27)	0.96	39 (14)	2.79 *
Local * Nonlocal	18 (24)	0.77	11 (49)	0.23	33 (26)	1.24
Prefinal Region						
Local	2 (15)	0.17	48 (47)	1.03	4 (15)	0.24
Nonlocal	15 (13)	1.12	4 (30)	0.13	5 (12)	0.43
Local * Nonlocal	15 (28)	0.54	46 (68)	0.68	17 (26)	0.66
Final Region						
Local	13 (16)	0.82	26 (58)	0.45	13 (15)	0.85
Nonlocal	9 (16)	0.58	45 (53)	0.85	4 (13)	0.30
Local * Nonlocal	29 (29)	1.00	11 (112)	0.10	31 (28)	1.09

Local = Local antecedent, Nonlocal = Nonlocal antecedent.
Estimate = Model Estimate (SE in brackets). [+] *= p < .10, * = p < .05, ** = p < .001*

Table 7. Percentage of responses in six conditions in Experiment 4

	Coargument Reflexives		PNP Reflexives		PPNP Reflexives	
	Local Antecedent Match	Local Antecedent Mismatch	Local Antecedent Match	Local Antecedent Mismatch	Local Antecedent Match	Local Antecedent Mismatch
Percentage of local antecedent responses	97.7%	94.4%	91.2%	80.1%	78.2%	65.3%
Percentage of nonlocal antecedent responses	1.9%	4.2%	2.3%	6.9%	5.6%	14.8%
Percentage of either antecedent responses	0.5%	1.4%	6.5%	13.0%	16.2%	19.9%

Figure 1. *Progression difference as a function of time*