

Covariation in processing: grammar vs. context

Abstract

In addition to referential uses, pronouns can have covarying interpretations, i.e., exhibit the behavior of a bound variable. The grammatical mechanism(s) behind such readings have been subject to longstanding debates: some authors argue for a fairly flexible but unified semantic mechanism that is not tied closely to syntactic configurations, while others distinguish a core class of *bona-fide* binding with tight syntactic constraints from other mechanisms that give rise to ultimately parallel effects, but do so more indirectly. Psycholinguistic work has started to uncover the processing mechanisms in evaluating dependencies between covarying pronouns and (candidate) antecedents. Moulton and Han (2018) leverage the processing perspective to try to shed light on the theoretical question of what mechanism is at play for a given covarying pronoun, arguing that so-called Gender Mismatch Effects only arise for cases of *bona-fide* binding, supporting the existence of distinct mechanisms. However, Kush and Eik (2019), looking at another construction involving the relevant other covariation mechanisms, do find Gender Mismatch Effects. They suggest that various contextual factors can make a covarying interpretation harder to obtain, and that suitable adjustments to Moulton and Han's stimuli would lead to Gender Mismatch Effects emerging for these types of sentences even when no *bona-fide* binding is involved. A series of self-paced reading experiments replicate the results from Moulton and Han, and then extend the paradigm to a variation along the lines suggested by Kush and Eik. The adjustment of contextual factors indeed results in Gender Mismatch Effects for both environments. We discuss how the processing evidence informs the theoretical issues.

Keywords: covariation, binding, bound variable, pronoun, c-command, processing

1. Introduction

In addition to referential uses, pronouns can have covarying interpretations, i.e., exhibit the behavior of a bound variable. The grammatical mechanism(s) behind such readings have been subject to longstanding debates: some authors argue for a fairly flexible but unified semantic mechanism that is not tied closely to syntactic configurations, while others distinguish a core class of *bona-fide* binding with tight syntactic constraints from (a) separate grammatical mechanism(s) that give rise to ultimately parallel effects, but do so more indirectly. Psycholinguistic work has started to uncover the cognitive processes in evaluating dependencies between covarying pronouns and (candidate) antecedents. Moulton and Han (2018) leverage the processing perspective to try to shed light on the theoretical question of what grammatical mechanism is at play for a given covarying pronoun, arguing that so-called Gender Mismatch Effects (GMMs) in processing only arise for cases of *bona-fide* binding, and in favor of the existence of distinct grammatical mechanisms. However, Kush and Eik (2019), looking at another construction involving the putative other covariation mechanisms, do find GMMs. They suggest that various contextual factors can make a covarying interpretation harder to obtain, and that suitable adjustments to Moulton and Han’s stimuli would lead to GMMs emerging for these types of sentences as well, despite the absence of *bona-fide* binding.

We report a series of self-paced reading experiments to further inform these debates. Our first goal is empirical, namely, to determine whether the types of sentences without *bona-fide* binding investigated by Moulton and Han (2018) do exhibit GMMs once properly modified along the lines suggested by Kush and Eik (2019). We first replicate the results from Moulton and Han, and then extend the paradigm to a variation along the lines suggested by Kush and Eik. The adjustment of contextual factors indeed results in GMMs for both environments. Our second goal is to assess the theoretical implications of the overall body of empirical evidence. We argue that our data provide clear evidence against the notion that in early processing, the search for quantificational antecedents is restricted to c-commanding noun phrases, e.g., as put forward by Moulton and Han. At the same time, more needs to be said about the apparent difference between c-commanding and non-c-commanding quantificational antecedents, since immediate accessibility of the latter, but not the former, seems to require contextual support. Furthermore, the question remains open as to just how such support facilitates access. We sketch a possible explanation of this, as well as of the independence of c-commanding antecedents from such support. While our data clearly undermine Moulton and Han’s argument from pronoun processing in favor of a two-mechanism grammar, we argue them to be compatible with both unified and two-mechanism theoretical perspectives, i.e., the present processing considerations do not settle the question of what mechanisms for covariation exist in the grammar. Finally, our results confirming the importance of contextual motivation of the relevant antecedent choices for pronouns make clear that even in early processing, antecedent accessibility and retrieval is guided by deeper interpretive factors, and not just dependent on purely structural considerations.

The remainder of the introduction section introduces more detailed background on theories of covariation and prior experimental work on related issues. The following sections present three experiments investigating GMMs for non-c-commanding quantificational antecedents in several variations and with different controls. Section 5 presents a general discussion of the theoretical implications of our findings.

1.1 Requirements for covariation: c-command vs. semantic scope

According to the classic account by Reinhart (1983), binding requires the antecedent to c-command the pronoun.¹ This view captures contrasts such as the following, where the antecedent noun phrase (**bold**) occurs inside of a relative clause within the subject, thus not c-commanding the pronoun in the verb phrase (*italicized*; adapted from Barker, 2012):

- (1) The [man who traveled with **[the woman]_i**] denied that *she_i* met the shah.
- (2) # The [man who traveled with **[each woman]_i**] denied that *she_i* met the shah.

While the referential determiner phrase (DP) in (1) allows a *co-referential* interpretation of the pronoun, which—in contrast to binding—does not require c-command, the minimal variant with a quantifier phrase (QP) antecedent in (2) does not allow a *covarying* reading due to the unavailability of binding.² However, Barker (2012) catalogues a variety of both previously known and novel exceptions to the c-command requirement. For example, non-c-commanding QPs in possessive DPs (3a), inverse linking constructions (3b), PPs (3c), and VPs (3d) (all taken from Barker) can serve as antecedents for a pronoun receiving a covarying interpretation:

- (3)
 - a. [No_i one's mother-in-law] fully approves of her_i.
 - b. [Someone from every_i city] hates it_i.
 - c. John gave [to each_i participant] a framed picture of his_i mother.
 - d. John [visited each_i student] on his_i birthday.

Various types of reactions to considerations of this type have been offered in the literature, which fall into two broad camps: Unified accounts pursue a reconceptualization of the notions of binding and covarying interpretations that covers all of these cases. In contrast, two-mechanism accounts argue that cases of non-c-commanding quantificational antecedents do not come about via *bona-fide* binding, for which a c-command requirement is maintained, but are instead to be derived using another interpretive mechanism.

The most relevant instance of a unified account for our purposes comes from Barker (2012), who argues, building on Safir (2004), that there simply is no c-command requirement for binding.³ Instead, he proposes a weaker scope requirement, according to which a necessary condition for a QP binding a pronoun is that it can take semantic scope over the position of the pronoun. This still excludes (2), as shown in (4), but allows (3a-d), as illustrated here for (3b) and (3d) by the availability of interpretations where the relevant QP takes scope over an indefinite appearing in the position of a pronoun:

¹ While precise definitions of c-command vary in the literature, they are generally based on the notion that for X to c-command Y, Y has to be properly contained in the syntactic sister-node of X.

² We adopt the DP/QP labels for consistency with prior work in this area; no suggestion of different syntactic categories for referential and quantificational noun phrases is intended.

³ Other contributions towards a unified account, often considered relative to just one specific type of exception, have proposed to adjust the definition of c-command. The detailed variations are of no concern for our purposes; see, Barker (2012) for comprehensive references.

- (4) The [man who traveled with each woman] denied that a student met the shah.
 CANNOT MEAN: *For each woman x , it holds that the man y who traveled with x denied that there is a student that met the shah.* * (“each woman” > “a student”)
- (5)
- a. [Someone from every city] hates a city regulation.
 CAN MEAN: *For every city x there is some person y , such that there exists a city regulation z and y hates z .* (“every city” > “a city regulation”)
 - b. John [visited each student] on a Monday.
 CAN MEAN: *For each student x there is some Monday y , such that John visited x on y .* (“each student” > “a Monday”)

Various grammatical mechanisms have been proposed by two-mechanism accounts in order to derive covarying interpretations relative to non-c-commanding quantificational antecedents without invoking regular binding, most prominently involving quantification over situations. Crucially, these mechanisms, unlike genuine binding, do not require a c-command relationship between QP and pronoun. The technical details here get complex (see, e.g., Büring, 2004 for a formally detailed account of *binding out of possessors* exemplified by (3a)), and the generality of this move across cases is far from clear, but see Section 5 for some illustrations relative to cases at play in the experiments discussed here.

What is most relevant to us in more general terms is how the general configuration of the theoretical landscape can be related to processing data: On the one hand, if we find processing differences, e.g., with respect to the presence of GMMs, between cases of covarying pronouns that are derived through different means according to two-mechanism accounts (because they occur in c-commanding vs. non-c-commanding environments), this offers potential support to the relevant two-mechanism proposal insofar as it can readily attribute these to the inherent theoretical contrast that it posits—this is the line of argument pursued by Moulton and Han (2018). In contrast, a unified perspective comes with no inherent explanation of such a contrast. Importantly, as we’ll elaborate in the general discussion below, the converse does not hold, i.e., while the absence (or lack of full generality) of such potential processing effects undermines the argument in favor of two-mechanism approaches just sketched, it does not necessarily in and of itself argue against two-mechanism grammatical accounts (as different theoretical derivations of an interpretation need not correspond to processing differences).

On the other hand, processing accounts—like the one based on a two grammatical-mechanisms account we’ll be considering—that tie the presence (or absence) of GMMs to purely structural factors, e.g., whether or not the antecedent QP c-commands the pronoun, don’t predict there to be any variation in GMMs based on non-structural factors, e.g., based on contextual or content variations within the same structural configuration, unless supplemented in some way to integrate such factors. In contrast, accounts that relate the presence of GMMs to other, non-structural factors, such as the availability of a given scope interpretation, may have a more direct way of explaining such factors playing a crucial role for GMMs.

1.2 Previous experimental work on covarying pronouns

Several lines of recent psycholinguistic work on the processing of covarying pronouns have investigated the role of c-command in online processing, using so-called Gender Mismatch Effects (GMMs; Sturt, 2003): relative to cases where a pronoun is anteceded by a noun phrase of matching gender, reading times on the pronoun (and following region(s)) increase when there’s a

mismatch in gender between the two. This is taken to indicate that while, with the gender-matching antecedent, interpretation of the pronoun in processing progresses smoothly, there is a disruption of some sort when the antecedent comes with a gender mismatch.⁴ This perspective makes the presence of GMMs a highly useful diagnostic for whether the processor is attempting to establish a dependency between a given pronoun and potential antecedent, as the only way such an effect could be absent is if the potential antecedent in either form is not considered at all to begin with (the original study by Sturt used this to test reflexives for effects of condition A of the binding theory). Cunnings et al. (2015) and Kush et al. (2015) leverage this to test whether c-command constitutes such a factor, i.e., whether upon encountering a pronoun, the processor only considers a candidate quantificational noun phrase antecedent if it c-commands the pronoun. And indeed, Cunnings et al. find GMMs for sentences like (6b), (where “every old man” c-commands the pronoun: “CC”) but not (6a) (where it doesn’t: “NoCC”), which crucially manifests in a statistical interaction of the gender and c-command factors.

- (6)
- a. The surgeon who every old man on the emergency ward saw silently wished that {(i) he/ (ii) she} could go a little bit faster. (*NoCC* (i) *Match* / (ii) *Mismatch*)
 - b. The surgeon saw that every old man on the emergency ward silently wished that {(i) he/ (ii) she} could go a little bit faster. (*CC* (i) *Match* / (ii) *Mismatch*)

Further experiments by Kush et al. (2015) corroborate these findings, and taken together, these results suggest that the processor does not attempt to establish a dependency with a QP antecedent when the pronoun is not in its c-command domain.

Results of this sort inform models of pronoun resolution in processing in important ways. One of the most prominent approaches to this are cue-based models, which assume that the search for an antecedent is guided by item-specific features intrinsic to the items being retrieved, like morphological features (Lewis et al., 2006). Each word encountered by the parser triggers specific retrieval cues, which guide the parser to rapidly form dependencies with items matching in appropriate features (e.g., gender or number; McElree, 2000). What Cunnings et al. (2015) and Kush et al. (2015) show, along with a growing body of research on the impact of other structural constraints in online antecedent retrieval (e.g., Sturt, 2003 on Principle A; Chow et al., 2014 on Principle B), is that relations between items, rather than item-specific information alone, are relevant for the retrieval of the QP antecedent. Kush et al., building on Kush (2013), propose to capture such structural effects in a cue-based framework by positing an ACCESSIBLE feature whose value can be dynamically updated as the parse unfolds. In the studies in question, the idea would be that the switch from a given candidate QP antecedent being accessible to being inaccessible would be made as the parse reaches a stage that is outside of the QP’s c-command domain, leading to that QP no longer being considered as an antecedent if a pronoun is subsequently encountered.

However, in light of the theoretical discussion above, it is important to acknowledge that the experiments in question do not specifically tease out c-command as the relevant structural

⁴ The precise nature of this disruption can in principle be construed in various ways. The perhaps most prominent view, adapted by cue-based retrieval approaches (discussed below), is that the mismatching antecedent is not considered for pronoun resolution, and that the delays are due to difficulties in interpreting the pronoun relative to a suitable antecedent, whether it is because there isn’t one at all or because it is less easily accessible than in the gender-matching variant. Alternatively, one could posit that the mismatching antecedent is temporarily considered for interpreting the pronoun, which then leads to a clash based on the feature mismatch. Our discussion is largely independent of this, but see footnote 5 and the discussion of Experiment 2 for some further relevant considerations.

constraint at play. Given the primary interest in testing for relational cues in general, these authors intentionally chose constructions where c-command and the scope constraint make the same predictions, as a QP in a relative clause cannot take scope over the main clause verb phrase in (6). This leaves open the possibility that what governs the setting of the ACCESSIBLE feature is a matter of scope, not c-command. In an attempt to differentiate between these two possibilities, and to potentially garner evidence for a two-mechanism approach to covariation, Moulton and Han's (2018) Experiment 2 presents a variant of the general Cummings et al. (2015) experiment, using stimuli with a QP in sentence-initial temporal adjunct clauses:

- (7)
- a. After each boy brought fresh water from the kitchen quickly it seems that {(i) he / (ii) she} went on an early break. (*QP & NoCC* (i) *Match* / (ii) *Mismatch*)
 - b. It seems each boy brought fresh water from the kitchen quickly right before {(i) he / (ii) she} went on an early break. (*CC* (i) *Match* / (ii) *Mismatch*)

Crucially, and unlike in Cummings et al. (2015), the no-c-command (NoCC) condition here allows the quantifier to take scope over the pronoun in the absence of c-command (as confirmed in an offline judgement task). This manipulation thus makes it possible to distinguish the processor's sensitivity to scope and c-command, unlike Cummings et al. and Kush et al. (2015): if scope is decisive, GMMEs are expected in both conditions. If c-command is what matters, we'd only expect a GMME in the CC condition. In a self-paced reading (SPR) task, Moulton and Han (2018) find an initial interaction indicative of GMMEs in the c-command (CC) condition but not the NoCC condition. In line with the reasoning above, they interpret this to show that in contrast to the CC condition, the processor was not attempting to establish a dependency between the pronoun and the potential QP-antecedent in the NoCC condition. A second SPR experiment by Moulton and Han (Experiment 3) compares exceptionally covarying sentences with QP antecedents like (7a) to identically structured sentences with DP antecedents, as in (8), yielding a parallel interaction due to the presence of a GMME in the DP condition (where the referential interpretation is not dependent on c-command, as in (1) above) but not the QP condition.

- (8) After the boy brought fresh water from the kitchen quickly it seems that {(i) he / (ii) she} went on an early break. (*DP* (i) *Match* / (ii) *Mismatch*)

In theoretical terms, Moulton and Han (2018) see their results as supporting a two-mechanism view, where covariation due to standard binding is subject to a c-command requirement, whereas the mechanism at play in other cases (such as (7a)) does not require this structural configuration. With regards to processing, it furthermore needs to be explained why the latter cases do not give rise to GMMEs. Moulton (2017) spells out a specific proposal for such a mechanism, according to which exceptionally covarying pronouns are interpreted as D-type pronouns, (Postal, 1966; Elbourne, 2005)—essentially, a definite description with a phonologically null noun phrase. The interpretation of the D-type pronoun is crucially mediated by a situation variable that is in turn quantified over, which results in covariation (adapting the analysis of temporal adjunct clauses in Artstein, 2005). Illustrating informally for (7a), for any given situation *s* containing a boy who brought water, there is a temporally later situation in which the unique boy in *s* went on a break. Importantly, at a formal level, this only involves binding of, and quantification over, situation pronouns, which are not directly associated with the gender features of the “antecedent” noun

phrase. The pronoun and the candidate antecedent are not formally related to one another, so that an evaluation of whether or not the gender features match does not take place (at least not initially). On this account, the interaction found by Moulton and Han suggesting an absence of GMMEs in the NoCC condition is then explained under the assumption that, at least initially, feature match only plays a role in processing for pronominal variables in a binding configuration (presumably reflected both in the syntax and the semantics), and not for covarying pronouns in other structural environments. The semantic misalignment ultimately matters, of course, since the Mismatch version of (7a) does not allow for a covarying reading, but, the story must go, this doesn't happen until later (the question of just how much later being an open question), and it may not give rise to the same effect as gender mismatch on a (potentially) bound pronoun even then, depending on how the nature of GMMEs is construed.⁵

A straightforward prediction of this proposal based on a situation semantic account of covariation is that other cases subject to such an analysis should behave similarly. Two studies have looked at related issues: earlier work by Carminati et al. (2002) compared so called telescoping sentences (such as “Every British soldier aimed and then he killed an enemy soldier.”) with parallel variants allowing for binding under c-command, with both DP and QP antecedents (but not testing for GMMEs), and found no processing costs associated with a non-c-commanding antecedent. This suggests that the relevant two types of covariation mechanisms do not necessarily differ in their processing time-course, but this is consistent with Moulton and Han's (2018) proposal, since GMMEs are not at play. Kush and Eik (2019), however, directly test the prediction by looking at donkey pronouns (9a), one of the most prominent cases for which situation semantic D-type analyses have been proposed (Heim, 1990; Elbourne, 2005):

- (9) English paraphrases of Kush and Eik's (2019) Norwegian sample stimuli:
 - a. Every father who had a daughter in a soccer league drove {her/him} to the games.
 - b. The father who had a daughter in a soccer league drove {her/him} to the games.

Using referential pronouns (9b) for a baseline comparison, Kush and Eik (2019) conducted an SPR study in Norwegian, in a standard GMME design with gender matching or mismatching pronouns. They find GMMEs in both cases, and no interaction, suggesting that the non-c-commanding indefinite antecedent and its particular gender-features were accessible in early processing not only for referential but also for covarying interpretations. Assuming a situation semantic analysis of donkey sentences, this directly contradicts the key prediction from Moulton and Han (2018) spelled out above.

The contrast in GMME findings for the NoCC stimuli in Moulton and Han (2018) and Kush and Eik (2019) raises a new question about the sources of GMMEs under covariation. In theoretical terms, there certainly are differences between the constructions in play that could have repercussions for processing. While the cases with QPs in temporal adjuncts, as in (7a), are among those that are standardly considered to fall under the scope constraint according to Barker (2012),

⁵ See footnote 4. If one construes delays for mismatching antecedents as involving an initial attempt at forming an anaphoric dependency which then falters in light of the incompatible gender features, then perhaps ultimately the Moulton and Han account predicts effects that are parallel in nature but arise later, as the feature mismatch doesn't arise in the initial phase of considering pronoun and antecedent, when only the situation pronoun and its abstract binder are considered.

donkey sentences are not (and neither are cases of telescoping).⁶ This is because they do not involve the antecedent QP taking scope over the position of the pronoun. In contrast, QPs in temporal adjuncts can take scope over the position of the pronoun. Just why this difference should lead to the pattern of GMMes reviewed here remains open at this point, but it's worth noting these differences of potential relevance. Kush and Eik, extending the cue-based approach to GMMes in Kush et al. (2015), suggest an alternative processing proposal, which maintains that a single processing mechanism uniformly governs antecedent retrieval in covarying cases (as well as referential ones), with or without c-command, while allowing for possible variation between specific types of cases. In particular, they suggest that the presence of an ACCESSIBLE feature is affected not just by structural factors like c-command, but also by contextual factors at play in settling on an overall interpretation, specifically with regards to scope. The idea that such factors can have an impact is not new and has been discussed for telescoping in some detail (Poesio & Zucchi, 1992; Anderssen, 2011). Leaving details for later, the general direction of Kush and Eik's take on the Moulton and Han results is that the relevant contextual pressures for a covarying interpretation of the stimuli in Moulton and Han were not sufficiently strong to make the QP immediately accessible, and that offline covariation judgments reflected later stages in the comprehension process. They suggest a number of specific changes that they speculate should increase the availability of covarying readings in the Moulton and Han type stimuli, which should lead to GMMes parallel to other constructions. The main empirical contribution of this paper is to test stimuli implementing these adjustments to assess what effect, if any, this has on the presence of GMMes in the SPR paradigm. This will be done in Experiment 2. Before proceeding to this, we first report a replication of the original Moulton and Han study to ensure that we are starting from the same baseline.

2. Experiment 1

Our first experiment is a replication of Experiment 3 from Moulton and Han (2018). The critical experimental stimuli were identical to the original ones, crossing antecedent type (QP vs. DP) with gender match vs. mismatch, as illustrated in (10). However, we were only able to include 20 of the original study's 36 filler sentences. The replication ensures that this and other potential minor deviations in methods did not affect the GMME pattern, to provide a sound comparison with the results for the modified stimuli in Experiment 2.

(10)

- a. After $\frac{1}{2}$ each boy $\frac{2}{3}$ brought fresh water $\frac{3}{4}$ from the kitchen $\frac{4}{5}$ quickly $\frac{5}{6}$ it seems $\frac{6}{7}$ that he $\frac{7}{8}$ went $\frac{8}{9}$ on an early $\frac{9}{10}$ break. (*QP Match*)
- b. After $\frac{1}{2}$ each boy $\frac{2}{3}$ brought fresh water $\frac{3}{4}$ from the kitchen $\frac{4}{5}$ quickly $\frac{5}{6}$ it seems $\frac{6}{7}$ that she $\frac{7}{8}$ went $\frac{8}{9}$ on an early $\frac{9}{10}$ break. (*QP Mismatch*)
- c. After $\frac{1}{2}$ the boy $\frac{2}{3}$ brought fresh water $\frac{3}{4}$ from the kitchen $\frac{4}{5}$ quickly $\frac{5}{6}$ it seems $\frac{6}{7}$ that he $\frac{7}{8}$ went $\frac{8}{9}$ on an early $\frac{9}{10}$ break. (*DP Match*)
- d. After $\frac{1}{2}$ the boy $\frac{2}{3}$ brought fresh water $\frac{3}{4}$ from the kitchen $\frac{4}{5}$ quickly $\frac{5}{6}$ it seems $\frac{6}{7}$ that she $\frac{7}{8}$ went $\frac{8}{9}$ on an early $\frac{9}{10}$ break. (*DP Mismatch*)

⁶ Though this ultimately depends on your analysis of these sentences, and the details of how scope is construed in a given semantic framework. Note that Barker and Shan (2008) analyze donkey pronouns as involving in-scope binding in continuation semantics.

As a reminder, a replication of the Moulton and Han (2018) results crucially should involve an interaction between antecedent type and gender match in the critical and/or spillover regions, with GMMEs only arising in the DP conditions.

2.1 Materials and Procedure

The 20 critical items from Experiment 3 in Moulton and Han (2018), varying by the four conditions in (10), were distributed across four lists in a Latin-square design, with individual participants only seeing each item in one condition. In addition, 20 of the 36 fillers from Moulton and Han with structures unrelated to the manipulation of interest were included, so that participants saw a total of 40 items (see Appendix B for a full list of materials). These were presented in randomized order, with an alternating pattern of a critical item followed by a filler. Each sentence was split into ten regions in a moving-window self-paced reading paradigm; starting from a set of dashes replacing each character on the screen, participants advanced to the next region by pressing the space bar (previous regions turned back to dashes). Each trial was followed by a yes/no comprehension question. The questions asked about the content of the sentences, but they were orthogonal to the manipulation so as to not interfere with the data. The study was hosted on Ibex, an online experiment hosting service, and took place remotely. Participants were instructed to be in a quiet place without distraction. They received explicit instructions on doing the task, as well as three practice trials with feedback on the accuracy of their comprehension questions. The total experiment lasted about 10 minutes. A link to a copy of the experiment can be found in (20) in Appendix A.

2.2 Participants

Eighty-three undergraduate students, who self-identified as native speakers of English, were recruited through our university's subject pool and received course credit for their participation. In line with subject pool policy, they saw a debriefing about the main research questions addressed by the experiment at the end.

2.3 Results

Prior to statistical analysis, data from participants with an accuracy rate for comprehension questions below 70% (across all conditions and fillers) or an average reading time (RT) below 300ms (across all regions) were removed.⁷ This eliminated two participants, leaving data from 81 participants to analyze.

Mean accuracy on the comprehension questions after removal across all conditions and fillers was 0.90 ($SE = .008$). Table 1 shows the mean accuracy rate broken down by condition, which shows no major effect on comprehension.

	Match	Mismatch
QP	.89 (.022)	.89 (.018)
DP	.92 (.016)	.90 (.018)

⁷ In their Experiment 2, Moulton and Han (2018) report removing two participants with an average RT below 400ms, who also had the two lowest comprehension rates (mean 62%); after removal, the lowest comprehension rate was 75%. In their Experiment 3, they removed no participants at all, and the lowest comprehension rate was 70%. We adjusted the RT-based cut-off to averages lower than 300ms, as our participants seem to have been somewhat faster, though there also is no straightforward comparison, since we were missing 16 of Moulton and Han's fillers. Generally, even our faster readers had high comprehension question accuracy rates, suggesting sufficient engagement with, and comprehension of, the stimuli.

Table 1: Mean accuracy rates of comprehension question responses (*SE*) in Experiment 1

Individual trials were also removed if any one region’s RT during that trial was above 3000ms, eliminating 60 experimental trials (4%). Mean RTs by condition and region after removal are shown in Table 2. The graphs in Figure 1 show natural log-transformed mean RTs for each region. Region 7 is the critical one containing the pronoun. Regions 8 and 9 are considered spillover regions, as effects from the manipulation in region 7 may emerge here as well.

		Region									
		1	2	3	4	5	6	7	8	9	10
QP	Match	407	518	663	666	604	543	486	560	578	620
	Mismatch	425	527	662	654	662	533	476	601	622	628
DP	Match	411	487	613	639	611	527	459	504	571	601
	Mismatch	431	486	662	659	646	560	490	604	643	665

Table 2: Mean RTs (ms) by region in Experiment 1

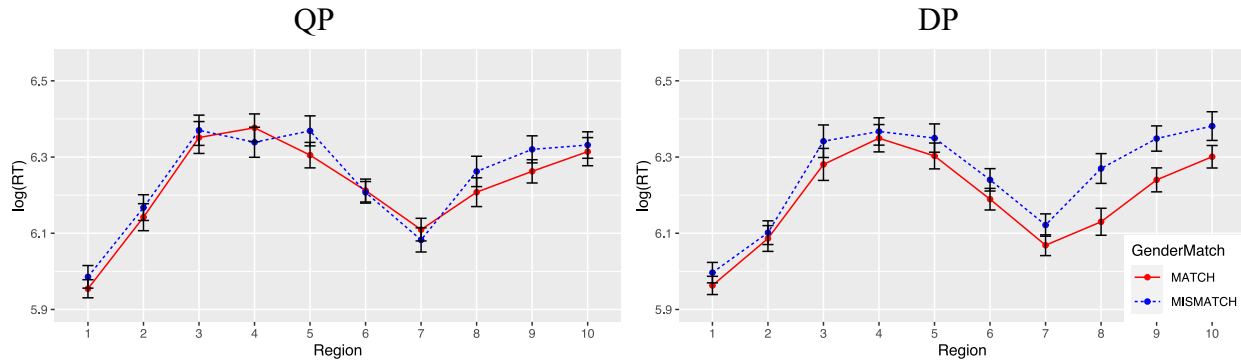


Figure 1: Log-transformed mean RTs by region in Experiment 1⁸

Statistical analyses used the natural log-transformed RTs as the dependent variable (as in Moulton & Han, 2018). For each region, a linear mixed-effects model analysis using the lme4 package (Bates et al., 2015) was conducted in R (version 4.2.2). For this initial replication, we report two analysis variants: first, one only including the manipulated factors as predictors, as reported by Moulton and Han, to have a fully parallel point of reference. Second, since it has become standard practice to include reading times for the previous region as an additional predictor in analyses of self-paced reading data, we also report an analysis with this added in.⁹

The first model included fixed effects of antecedent type, gender match, and their interaction. Antecedent type and gender match were sum-coded with one level as -1 and the other as 1. The initial model used a maximal random-effects structure (Barr et al., 2013), with random intercepts and random slopes and interactions for participants and items. In case of convergence issues or random effect correlation issues, the random effect structure was gradually simplified by

⁸ In the DP condition, it can be seen on the graph that the gender mismatch condition is consistently slightly higher than the gender match condition even before the critical region 7. This can be safely ignored, as the stimuli were identical up to that point, and further analysis showed no significant interactions ($p > 0.05$) until the critical region.

⁹ Thanks to the handling editor, Ming Xiang, for suggesting to include the latter analysis.

removing individual random effect slopes and by removing correlations with random slopes. For region 7, this resulted in a model with only random intercepts; for region 8, a model with (uncorrelated) by-participants slopes for antecedent type and gender match and a by-item slope for gender match; and for region 9, a by-participants slope for antecedent type and a random intercept for items. Planned comparisons to measure GMMs per antecedent type were computed using the emmeans package. *P*-values were determined using the lmerTest package via the Satterthwaite method (Kuznetsova et al., 2017). An overview of the interaction analyses can be found in Table 3, and results by region are summarized below.

	Region 7 (pronoun)			Region 8 (spillover)			Region 9 (spillover)		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
Antecedent Type	0.000	0.007	-0.022	-0.018	0.009	-2.028 *	0.003	0.009	0.379
Gender Match	-0.008	0.007	-1.084	-0.047	0.013	-3.539 **	-0.042	0.008	-4.959 ***
Type × Match	-0.020	0.007	-2.726 **	-0.020	0.009	-2.255 *	-0.009	0.008	-1.077

Table 3: Summary of statistical analysis for Experiment 1 (without previous region RT as a predictor)
. $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Region 7: The analysis revealed a significant interaction of antecedent type and gender match. Planned comparisons found a significant simple effect of gender match in the DP condition (Est. = -0.056, $SE = 0.021$, $t = -2.700$, $p < 0.01$) but not in the QP condition (Est. = 0.024, $SE = 0.021$, $t = 1.159$, $p = 0.247$).

Region 8: The analysis revealed a significant main effect of antecedent type as well as of gender match, and a significant interaction of antecedent type and gender match. Planned comparisons found a significant simple effect of gender match in the DP condition (Est. = -0.133, $SE = 0.032$, $t = -4.200$, $p < 0.001$) and a marginally significant one in the QP condition (Est. = -0.054, $SE = 0.032$, $t = -1.710$, $p = 0.093$).

Region 9: The analysis revealed a significant main effect of gender match. Planned comparisons found a significant simple effect of gender match in both the DP condition (Est. = -0.102, $SE = 0.024$, $t = -4.274$, $p < 0.001$) and the QP condition (Est. = -0.066, $SE = 0.024$, $t = -2.731$, $p < 0.01$).

As we'll discuss in more detail below, these results, based on analyses parallel to those reported in Moulton and Han (2018), overall align with those in the original paper, confirming the comparability of our methods and data. But before discussing the interpretation of the data in more detail, let us offer a second set of analyses, which add the reading time in the preceding region as a further factor to the various models (the random effect structures here were the same as above, with the exception of having to remove the by-participants random effect correlation for region 9). Unsurprisingly, the reading time of the previous region was a highly significant predictor, as can be seen in Table 4, but we will not comment further below on this, given that it's not of direct theoretical interest.

	Region 7 (pronoun)			Region 8 (spillover)			Region 9 (spillover)		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
Antecedent Type	-0.001	0.007	-0.229	-0.018	0.008	-2.214 *	0.011	0.009	1.428
Gender Match	-0.003	0.007	-0.426	-0.043	0.012	-3.551 **	-0.022	0.008	-2.808 **
Type \times Match	-0.014	0.007	-2.127 *	-0.009	0.008	-1.083	-0.001	0.008	-0.102

Table 4: Summary of statistical analysis for Experiment 1 (with previous region RT as a predictor)
. $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Region 7: The analysis revealed a significant interaction of antecedent type and gender match. Planned comparisons found a marginally significant simple effect of gender match in the DP condition (Est. = -0.033, $SE = 0.019$, $t = -1.807$, $p = 0.071$) but none in the QP condition (Est. = 0.022, $SE = 0.019$, $t = 1.201$, $p = 0.230$).

Region 8: The analysis revealed significant main effects of antecedent type and gender match. Planned comparisons found a significant simple effect of gender match in both the DP condition (Est. = -0.102, $SE = 0.029$, $t = -3.553$, $p < 0.001$) and QP condition (Est. = -0.068, $SE = 0.029$, $t = -2.342$, $p < 0.05$).

Region 9: The analysis revealed a significant main effect of gender match. Planned comparisons found a significant simple effect of gender match in the DP condition (Est. = -0.047, $SE = 0.023$, $t = -2.058$, $p < 0.05$) and a marginally significant simple effect of gender match in the QP condition (Est. = -0.043, $SE = 0.023$, $t = -1.919$, $p = 0.055$).

The impact of including previous region reading times as a predictor is shifting the effects in the original analysis, making the key outcome pattern somewhat more subtle: There still is an interaction in region 7, but the GMME in the DP conditions is now only marginally significant. Additionally, the interaction disappears in region 8, where we now also find a fully significant GMME for the QP conditions. (But note that our marginally significant simple effect in the QP condition in the first analysis here already contrasts with Moulton and Han’s findings, as they found no GMME for QP conditions.) Nonetheless, at a minimum, it still holds that there is a difference between QP and DP conditions, given the interaction in region 7. We can’t know whether the original data would exhibit these same shifts if previous region reading times were included there. For purposes of the discussion to follow here, we’ll take the new analyses including the additional predictor as our main focus, though we’ll summarize outcomes for analyses without that predictor in footnotes when reporting subsequent data analyses.

2.4 Discussion

Experiment 1 was conducted for the purpose of replicating Moulton and Han (2018) to ensure a sound baseline for the variations in Experiment 2. Overall, the data do generally replicate the effects in Moulton and Han, though in a more subtle way once we include the additional predictor of previous region reading times in the model. Even so, the interaction in region 7 indicates a difference in the impact of the gender manipulation based on antecedent type. This is further confirmed by the marginally significant simple effect of gender in the DP condition and the absence thereof in the QP condition. However, note that in our data, the interaction disappears in spillover regions 8 and 9, and the gender manipulation has a significant or marginally significant simple

effect there for both DP and QP conditions. This contrasts with Moulton and Han,¹⁰ but is not in principle incompatible with their overall generalizations, in that it still reflects an initial phase where the referential condition, with no c-command requirement, involves immediate establishment of the dependency on the antecedent in a way that is sensitive to gender features, whereas effects of gender don't arise until later in the QP condition with a covarying pronoun. The difference is that we do find an effect in the reading times downstream, but since conceptually, the gender mismatch has to matter at some point down the line (given the interpretation of the sentence), whether or not that has repercussions in a given processing measure at some later point is largely an orthogonal question.¹¹

While the present results are compatible with Moulton and Han's (2018) preferred interpretation, they, just like Moulton and Han's, are subject to a potential alternative interpretation, as they discuss: While it could be that the dependency is immediately established in the QP condition, with any potential impact of gender mismatch not unfolding until later, it also could be that the dependency itself is not established until later. If that were the case, another effect might be expected in both the Match and Mismatch conditions, namely a (temporary) unheralded pronoun effect (Greene et al., 1994), due to encountering a pronoun without an explicit antecedent within the sentence. This would predict initial slow-downs in the Match condition for QPs relative to DPs. Moulton and Han discuss and discard this possibility in the context of their Experiment 2, where a c-command condition (rather than replacing the QP with a DP) serves as a control, and they do not find any difference in processing time upon encountering the pronoun. However, the comparison here is less than ideal due to structural differences between NoCC and CC sentences—the structurally parallel DP and QP sentences make for a more telling comparison. And indeed, additional post-hoc analyses on the match conditions in our data, again using the emmeans package and applying it to the model that includes the additional previous region reading time predictor, show a significant simple effect of antecedent type in the match condition in region 8 (Est. = -0.054, $SE = 0.023$, $t = -2.339$, $p < 0.05$), due to slower reading times in the QP match condition.¹² This may reflect an unheralded pronoun effect, and therefore be evidence against the proposal that a dependency is immediately established, perhaps because wide scope has not been robustly computed in early processing for non-c-commanding QP antecedents in these materials, as suggested by Kush and Eik (2019). The results of Experiment 2 will shed more light on this issue, so we return to it in their discussion (also see the general discussion). For the moment, the main takeaway from this initial replication is that we can detect differences between the DP and QP conditions in terms of the emergence of GMMEs. This sets the stage for our second experiment, where we manipulate the original stimuli along the lines suggested by Kush and Eik.

3. Experiment 2

¹⁰ Although, we should note that the Moulton and Han (2018) experiment also contains a slight hint of a parallel effect, with a marginally significant effect ($p = 0.09$) of increased difficulty in the QP mismatch condition compared to the QP match condition in that same region.

¹¹ On the most extreme construal of the Moulton (2017) proposal, one might not expect any GMME at all, since there is no direct link between the pronoun and the candidate antecedent. But one could also imagine other ways in which infelicity causing a processing delay could come into play later on in this proposal that would be compatible with the present finding. Since the varied stimuli in Experiment 2 yield immediate GMMEs for the QP condition, we won't dwell on this issue here.

¹² A parallel trend seems to be present at least numerically in the graphs for the original version of this experiment in Moulton and Han (2018).

In light of their finding of immediate GMMs in donkey sentences, where the explanation of the Moulton and Han (2018) data detailed in Moulton (2017) predicts the absence of GMMs parallel to the NoCC condition in their work, Kush and Eik (2019) argue for a uniform processing mechanism that retrieves the antecedent in the presence of c-command or other contextual factors facilitating an anaphoric dependency. But while Kush and Eik's results clearly establish that some non-c-commanding antecedents for covarying pronouns can give rise to immediate GMMs, it's far from clear whether this will generalize to other cases. Recall, among other things, that donkey sentences do not actually fall under Barker's (2012) scope constraint proposal, since the indefinite antecedents in donkey sentences do not take scope over the relevant pronoun sites. Thus, it remains a genuinely open question whether immediate GMMs can arise in sentences such as those in the NoCC condition, and in particular whether this can be brought about by manipulating the sorts of factors suggested to be at play by Kush and Eik.

3.1 *Adjustments to Experiment 1*

Kush and Eik (2019, p. 12) speculate that a key factor for how quickly anaphoric dependencies are established in the relevant sentences is “how easy it is to adopt a *quantificational*, *distributive*, or *multi-event* reading of the fronted adjunct,” and the Moulton and Han (2018) stimuli arguably do not facilitate such readings. The pragmatics of producing such a reading may be informed by the literature on telescoping. Telescoping, as mentioned, is usually described as semantic binding across sentential boundaries. Certain factors have been identified to promote felicitous telescoping (Poesio & Zucchi, 1992; Anderssen, 2011). Among these is a *scripted*, *non-accidental*, and/or *generic* relationship between sentences, such that there is an appearance of regular relatedness—perhaps causation or expected succession—between the events of the first sentence and the second sentence. For example, (11b) demonstrates the change in felicity of a telescoping interpretation of (11a) when a context is provided to promote a scripted reading.¹³

(11)

- a. # Every_i dog came in. It_i lay down under the table.
- b. I went to the circus last night. They had a number involving dogs that went like this: The circus performers put a table on some supports. Then, every dog came in. It lay down under the table, stood on its back paws, and lifted the table with its front paws.

As these types of cues seem to promote covariation for telescoping examples, Kush and Eik (2019) propose that they will do so as well for the stimuli in Moulton and Han (2018). They propose four specific adjustments in this vein in particular:

(12)

- a. Change from past tense to present tense
- b. Add an indefinite DP to the adjunct clause
- c. Remove the intervening raising predicate *it seems/it appears*
- d. Appear generally scripted in nature

For Experiment 2, we implemented adjustments along these lines for the Moulton and Han (2018) stimuli used in Experiment 1. While (12a)–(12c) involved fairly straightforward alterations, (12d) was more open-ended, and involved modifying the content of each clause such that the events in

¹³ The examples in (11) are from Poesio and Zucchi (1992).

the second clause were more closely related to those in the first. The outcome was a set of stimuli that was completely parallel to the stimuli in Experiment 1 in terms of the overall syntactic configuration, but whose salient interpretation had the features above to have the relevant scripted nature. (13) illustrates the resulting variation of the original stimuli in (10).

(13)

- a. After $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ he $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*QP Match*)
- b. After $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ she $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*QP Mismatch*)
- c. After $\frac{1}{2}$ the boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ he $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*DP Match*)
- d. After $\frac{1}{2}$ the boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ she $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*DP Mismatch*)

Kush and Eik (2019) argue that the changes outlined in (12) and implemented in (13) intuitively make a scripted interpretation of these sentences more easily available. The change to present tense is of particular importance, we argue (and elaborate on in Section 5.3), as the sentences are now more plausibly interpreted as being part of a general script of what happens whenever a boy fetches a bucket of water: that same boy goes to clean the barn and stables. This is in contrast to the past tense sentences in the original stimuli in (10), which seem to be easily construed as a simple recount of events that have occurred, without any scripted notions. We find that the raising predicate *it seems/it appears* also promotes a circumstantial, non-scripted reading, motivating its removal. Finally, by thematically connecting the events in the subordinate and matrix clauses, and using an indefinite which can quantify over the events in the matrix clause, we reinforce the availability of the scripted interpretation and thus, in theory, the covarying interpretation.

Experiment 2 keeps constant the syntactic structure of Experiment 1 while contextually facilitating a scripted interpretation. If the absence or delay of a GMME is due to the general syntactic structure of the sentences and the underlying semantic mechanisms of covariation (via quantification over situations), as on the Moulton and Han (2018) account, then this variation should have no effect. In contrast, on the Kush and Eik (2019) proposal, where contextual factors matter for how likely a relevant scopal interpretation and corresponding anaphoric dependency is, we should see comparable GMMEs in both antecedent type conditions and no interaction (assuming our manipulation is sufficient and successful).

3.2 Materials and Procedure

The procedure remained the same as in Experiment 1, using the adjusted stimuli. A full list of materials can be found in Appendix C. The materials were directly adapted from those in Experiment 1, with twenty test items in four conditions, as exemplified in (13). The same twenty fillers from Experiment 1 were used, with one region removed (typically a modifying adjunct) to match the number of regions in the adjusted stimuli. A link to the experiment can be found in (21) in Appendix A.

3.3 Participants

Sixty-seven undergraduates self-identifying as native-English speakers were recruited through our university's subject pool, none of whom had participated in Experiment 1.

3.4 Results

The same data removal criteria as for Experiment 1 were applied. This eliminated three participants, leaving data from a total of 64 participants.

Across all conditions and fillers, participants answered the comprehension questions with a mean accuracy of about 0.89 ($SE = .008$). Table 5 shows the mean proportion of correct responses in each condition. There appears to be no major effect on comprehension from the manipulations.

	Match	Mismatch
QP	.89 (.018)	.88 (.024)
DP	.90 (.017)	.88 (.022)

Table 5: Mean accuracy rates of comprehension question responses (SE) in Experiment 2

Forty-seven individual experimental trials (4%) were removed following the same removal criteria in Experiment 1. Table 6 provides the mean reading times (RTs) for each region. Figure 2 provides a graph of natural log-transformed mean RTs by region. In this experiment, region 6 is the critical one, but since it only included the pronoun in this version, the spillover regions 7 and 8 are especially important to consider.

		Region								
		1	2	3	4	5	6	7	8	9
QP	Match	431	573	707	655	644	508	512	545	648
	Mismatch	447	609	705	665	661	520	568	634	695
DP	Match	449	547	707	671	679	516	540	601	633
	Mismatch	451	544	672	665	658	534	597	636	705

Table 6: Mean RTs (ms) by region in Experiment 2

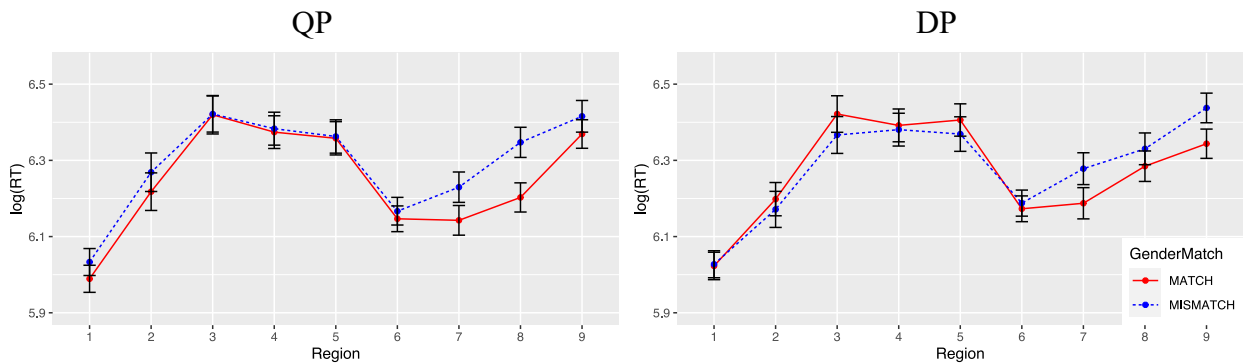


Figure 2: Log-transformed mean RTs by region in Experiment 2

RT data were natural log-transformed and analyzed with a linear mixed-effects model following the same approach as for Experiment 1, using the maximal random effect structures that would converge. For all three regions, this was a model with a random slope for antecedent type by-

participants. We focused on models that included previous region reading times as a predictor.¹⁴ The results of the analysis are summarized in Table 7. The effects of previous region reading times were highly significant throughout, but are not reported here in detail.

	Region 6 (pronoun)			Region 7 (spillover)			Region 8 (spillover)		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
Antecedent Type	0.007	0.009	0.769	0.017	0.009	1.838 .	0.005	0.010	0.472
Gender Match	-0.011	0.009	-1.328	-0.041	0.009	-4.473 ***	-0.035	0.010	-3.523 ***
Type × Match	-0.003	0.009	-0.309	-0.004	0.009	-0.419	0.022	0.010	2.235 *

Table 7: Summary of statistical analysis for Experiment 2
. $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Region 6: There are no significant effects.

Region 7: The analysis revealed a marginally significant main effect of antecedent type and a significant main effect of gender match. Planned comparisons found significant simple effects of gender match in both the QP condition (Est. = -0.074, $SE = 0.026$, $t = -2.857$, $p < 0.01$) and DP condition (Est. = -0.089, $SE = 0.026$, $t = -3.464$, $p < 0.001$).

Region 8: The analysis revealed a significant main effect of gender match and a significant interaction of antecedent type and gender. Planned comparisons found a significant simple effect of gender match in the QP condition (Est. = -0.114, $SE = 0.028$, $t = -4.067$, $p < 0.001$) but not in the DP condition (Est. = -0.026, $SE = 0.028$, $t = -0.924$, $p = 0.356$).

3.5 Discussion

The adjustments implemented for the stimuli from Experiment 1 led to immediate GMMs in both the QP and DP conditions, and no interaction of antecedent type and gender match of the sort found before (see below on the interaction in region 8). These results suggest that the QP antecedents, including their gender information, were as immediately and robustly accessed as the DP antecedents. The changes aiming to create a more script-like interpretation, while leaving the syntactic structure as before, thus did induce an interpretation where an anaphoric dependency of the pronoun on the QP antecedent was established without any delay, as predicted by Kush and Eik (2019).

Unlike in Experiment 1 (but parallel to Moulton & Han's, 2018 Experiment 3), GMMs did not emerge until the first spillover region. Note, however, that unlike there, the critical region in the present study only contained the pronoun, and thus was very short, making spillover effects more likely. Moreover, prior findings by Cunnings et al. (2015) that found GMMs beginning in the pronoun region in a DP condition in their eye-tracking data suggest that the slightly delayed effect in the spillover region here is due to the nature of the self-paced reading paradigm (though also note that we do find GMMs on the region containing only the pronoun in Experiment 3).

Another aspect of the data worth commenting on is the interaction between antecedent type and gender match in region 8, with a greater effect of gender mismatch in the QP condition. Note

¹⁴ The results without such a predictor were essentially identical in terms of significance patterns; the only minor divergences were not crucial to our interpretation of the data, namely: (i) in region 7, the main effect of antecedent type was significant ($p < 0.05$) rather than merely marginally significant. (ii) in region 8, the simple effect indicative of a GMM in the DP condition was still significant ($p < 0.05$). (iii) the simple effect of antecedent type in the match condition was significant ($p < 0.05$) rather than merely marginally significant.

that this seems to mainly be driven by slower reading times in the DP match condition (post-hoc analysis of region 8 shows a marginally significant simple effect of antecedent type in the match condition; $Est. = 0.053$, $SE = 0.028$, $t = 1.897$, $p = 0.059$), which may be slightly less easily compatible with the scripted nature of the new stimuli. While further aspects of this tentative finding—if fully substantiated statistically—need to be explored in future work, it’s clear that, if anything, the direction of this effect in the second spillover region goes directly counter to the Moulton and Han (2018) predictions of an absence of GMMes for QP-anteceded pronouns in the NoCC condition; i.e., if anything, the GMME for the QP conditions seems to be more pronounced and long-lasting than in the DP conditions.

Returning to the finding that is key to the main question we are pursuing, Experiment 2 crucially establishes that syntactic structure alone cannot be blamed for the delayed (or entirely absent, in Moulton & Han’s, 2018 data) GMMes for the stimuli in Experiment 1. Experiment 2 employed exactly the same syntactic configuration, with a QP in a temporal adjunct clause, and yet, with the modifications of the stimuli aiming for a more scripted and non-accidental relationship between the two clauses, we get immediate GMMes. Taking the standard stance that GMMes reflect the establishment of an anaphoric dependency, these data thus show that non-c-commanding QPs are immediately considered as antecedents, and their gender is evaluated, as the relevant pronoun (that covaries with the quantifier in the match condition) is encountered (or in any case, this happens just as fast as in the DP condition, which doesn’t require c-command).

There is another point worth noting in relating the findings for Experiment 2 to Moulton and Han’s (2018) account and their interpretation of their data. Their proposal explicitly argued that in NoCC configurations with QPs, the dependency between the situation pronoun and its binder is indeed established immediately, but since there is no gender information encoded at this level, gender mismatch effects go unnoticed, as it were (until the D-Type pronoun interpretation relative to the situation quantified over is fully considered). Correspondingly, and as noted in the discussion of Experiment 1 above, they argued against an alternative interpretation of their effect in terms of an unheralded pronoun effect. The results of Experiment 2 speak against both of these takes. First, we do find immediate GMMes in the QP condition, suggesting full and immediate consideration of the information encoded on the pronoun, including gender features, rather than an initial phase of merely considering the anaphoric dependency of the situation pronoun. Secondly, while there was what can plausibly be taken as an unheralded pronoun effect in Experiment 1, in the form of a significant simple effect in the Match conditions with faster reading times in the first spillover region for DP than QP sentences, no such effect is present in Experiment 2 (in fact, we find a marginally significant effect in the opposite direction in the second spillover region). This, in turn, suggests that GMMes are indeed tied to whether the relevant potential antecedent is considered in interpreting the pronoun (see our comment in footnote 4 on possible different takes on the nature of GMMes).

The present results align with the predictions by Kush and Eik (2019) and are in principle compatible with their proposal of a mechanism that uniformly governs the retrieval of QP antecedents. However, the details of just how the scripted interpretation and the overall manipulation of the contextual factors feed into such a mechanism are yet to be spelled out. Furthermore, the fact remains that c-command and non-c-command configurations seem to differ in that only the latter are affected by the present contextual manipulations. We will return to these more general issues about how to interpret the data in the general discussion. But first, it is worth testing for any potential remaining differences empirically, by following suit with the previous literature and not only comparing referential and quantifying noun phrases as antecedents in identical structures, but

also c-commanding and non-c-commanding QP antecedents in as minimally varied sentences as possible, as in Experiment 2 from Moulton and Han (2018). Using the modified stimuli from Experiment 2 (with additional c-command variations), Experiment 3 will allow us to assess more directly whether the non-c-commanding antecedents there exhibit any differences in processing time-course from c-commanding ones, as this is not directly ruled out yet by our findings so far.

4. Experiment 3

While the previous experiment has shown that syntactic structure alone cannot be held responsible for processing delays in general, the possibility remains that c-command does play a role for how easily QP antecedents are retrieved. By adding certain contextual pressures to exceptionally covarying sentences, QP antecedents become more readily accessible, according to Experiment 2. Are they, however, as readily accessible as *c-commanding* QP antecedents?

Experiment 3 utilizes the QP sentences (match and mismatch conditions) from Experiment 2, labeled NoCC in (14), and adds c-commanding (CC) condition variants. Parallel to Moulton and Han's (2018) Experiment 2, these were implemented by moving the temporal conjunction ("After/Before/When") to the region immediately preceding the pronoun. To maintain a parallel event structure and overall interpretation between the CC and NoCC conditions, "before" was changed to "after" (and vice versa) in relevant variants. One challenge in creating stimuli varying the main clause and adjunct clause role for the two clauses in play is keeping the quantifier and pronoun in place in terms of their surface position, which is important to maintain a constant distance between them. Moulton and Han's stimuli achieve this by adding the embedding with "it seems" that can appear in either clause. However, this embedding is hard to integrate with the manipulations in our Experiment 2. As an alternative solution, we add a fronted adjunct, usually an adverb or a PP, to the CC sentences. The same adjunct is also added to the NoCC condition in the position immediately preceding the pronoun. The goals guiding the particular adjunct choices were chiefly the following: 1) maintain the scripted nature of the sentence to ensure the contextual cues remained intact; 2) make adjunction to the verb phrase in the NoCC condition plausible; 3) limit the number of syllables of the adjunct to three to as closely as possible match pronoun-antecedent distance between conditions. To meet 2) in particular and prevent potential garden-pathing, the adjunct was followed by a comma. The resulting sentences are exemplified in (14).¹⁵

(14)

- a. After $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ on foot, $\frac{6}{7}$ he $\frac{7}{8}$ goes $\frac{8}{9}$ to clean the $\frac{9}{10}$ barn and stables. (*NoCC Match*)
- b. After $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ on foot, $\frac{6}{7}$ she $\frac{7}{8}$ goes $\frac{8}{9}$ to clean the $\frac{9}{10}$ barn and stables. (*NoCC Mismatch*)
- c. On foot, $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ before $\frac{6}{7}$ he $\frac{7}{8}$ goes $\frac{8}{9}$ to clean the $\frac{9}{10}$ barn and stables. (*CC Match*)

¹⁵ It is possible that introducing the additional adjunct, which can make the sentences somewhat cumbersome in certain cases, introduces a confound of its own. Appendix E summarizes a version of this experiment, not reported here, that utilizes the same stimuli with the adjunct removed (and antecedent-pronoun distance correspondingly varying slightly). The results are comparable to those in Experiment 3, suggesting that neither the antecedent distance nor the adjunct addition crucially contributes to the relevant aspects of the results. Note that in this variant, the NoCC did not have a comma preceding the pronoun, matching the CC condition in this regard, suggesting that the comma in the NoCC sentences in (14) did not crucially affect the results.

- d. On foot, $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ before $\frac{6}{7}$ she $\frac{7}{8}$ goes $\frac{8}{9}$ to clean the $\frac{9}{10}$ barn and stables. (*CC Mismatch*)

If c-command has an independent role to play in the accessibility of QP antecedents in covarying constructions, we should expect to see the NoCC conditions exhibit reliably greater retrieval costs, such that there is an interaction between the structure type and gender match conditions, at a minimum in early regions. Alternatively, if c-command has no privileged role in processing (at least in the present sentence variants), we should see no such interaction, but rather consistent early GMMEs across conditions, parallel to Experiment 2.

4.1 Materials and procedure

The procedure followed that of the previous experiments. A full list of materials can be found in Appendix D. There again were twenty test items in four conditions, as exemplified in (14). Each condition had either a NoCC structure or a CC structure, and either a gender match or gender mismatch between the pronoun and its possible antecedent. The same twenty fillers from Experiment 1 were used. A link to the experiment can be found in (22) in Appendix A.

4.2 Participants

Seventy-five undergraduates self-identifying as native-English speakers took part in the experiment, none of whom had participated in Experiments 1 or 2 before.

4.3 Results

Following the same data removal criteria as for the previous experiments, five participants' data were removed, leaving a total of 70 participants for analysis.

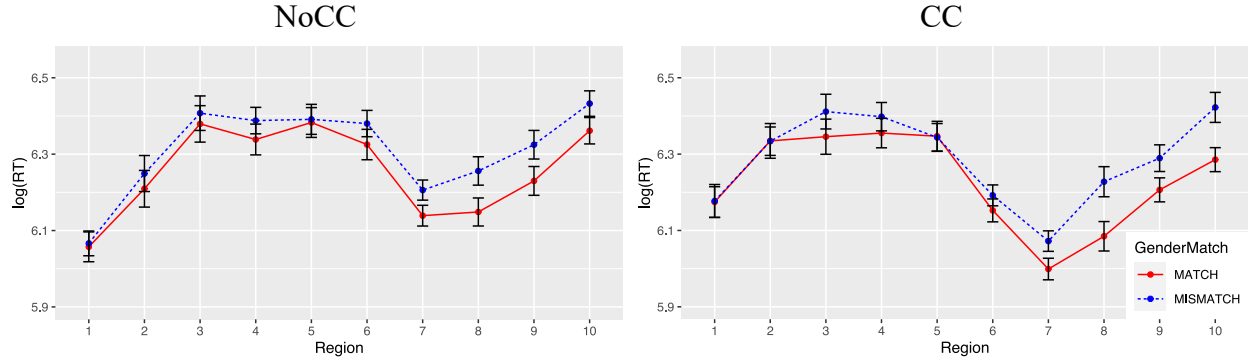
Across all conditions and fillers, participants answered the comprehension questions with a mean accuracy of 0.92 ($SE = 0.007$). Table 8 shows the mean proportion of correct responses by condition, with no major effect of condition on comprehension.

	Match	Mismatch
NoCC	.91 (.019)	.89 (.017)
CC	.92 (.014)	.89 (.019)

Table 8: Mean accuracy rates of comprehension question responses (SE) in Experiment 3

Forty-two individual experimental trials (3%) were removed following the same removal criteria in the previous experiments. Table 9 provides the mean reading times (RTs) for each region. Figure 3 provides a graph of natural log-transformed mean RTs for each region. The critical region is region 7 (again just containing the pronoun), and regions 8 and 9 are possible spillover regions.

		Region									
		1	2	3	4	5	6	7	8	9	10
NoCC	Match	472	572	678	629	664	619	487	510	551	637
	Mismatch	462	594	689	657	672	647	528	579	619	688
CC	Match	541	640	656	638	633	500	425	483	540	587
	Mismatch	546	621	692	668	627	517	457	563	589	691

Table 9: Mean RTs (ms) by region in Experiment 3**Figure 3:** Log-transformed mean RTs by region in Experiment 3

Analysis was carried out on natural log-transformed RT data, using a linear mixed-effects model with the maximal random-effect structure that would converge, as before. The model for region 7 had by-participant random slopes for gender match and structure, and an uncorrelated by-items slope for gender match; the one for region 8 had the same by-participant random slopes and a random intercept only for items; the only random slope for region 9 was for gender match by participants. We focused on models that included previous region reading times as a predictor.¹⁶ The results of the analysis are summarized in Table 10. The effects of previous region reading times were highly significant throughout, but are not reported here in detail.

	Region 7 (pronoun)			Region 8 (spillover)			Region 9 (spillover)		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
Structure Type	-0.044	0.008	-5.622 ***	0.012	0.011	1.103	-0.007	0.008	-0.829
Gender Match	-0.028	0.008	-3.435 **	-0.047	0.009	-5.124 ***	-0.023	0.008	-2.696 **
Type × Match	-0.002	0.007	-0.282	-0.014	0.008	-1.781 .	0.005	0.008	0.550

Table 10: Summary of statistical analysis for Experiment 3
. $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Region 7: The analysis revealed significant main effects of structure type and gender match. Planned comparisons found significant simple effects of gender match in both the NoCC condition (Est. = -0.053, $SE = 0.021$, $t = -2.467$, $p < 0.05$) and CC condition (Est. = -0.060, $SE = 0.022$, $t = -2.811$, $p < 0.01$).

Region 8: The analysis revealed a significant main effect of gender match and a marginally significant interaction of structure type and gender match. Planned comparisons found significant simple effects of gender match in both the NoCC condition (Est. = -0.065, $SE = 0.024$, $t = -2.709$, $p < 0.01$) and CC condition (Est. = -0.122, $SE = 0.024$, $t = -5.015$, $p < 0.001$).

¹⁶ The results without such a predictor were essentially identical in terms of significance patterns; the only minor divergences were not crucial to our interpretation of the data, namely: (i) in region 8, there was a marginally significant main effect of structure ($p = 0.057$), as opposed to no effect; (ii) in region 9, there still was a significant simple effect indicating a GMME in the CC conditions ($p < 0.001$), as opposed to no effect.

Region 9: The analysis revealed a significant main effect of gender match. Planned comparisons found a significant simple effect of gender match in the NoCC condition (Est. = -0.055, $SE = 0.023$, $t = -2.329$, $p < 0.05$) but none in the CC condition (Est. = -0.037, $SE = 0.024$, $t = -1.531$, $p = 0.127$).

4.4 Discussion

We find GMMs in both the NoCC and CC conditions in the critical and first spillover region (and in the second spillover region for NoCC). In contrast to Moulton and Han's (2018) Experiment 2, with syntactic structures overall parallel to ours here, our modified stimuli aiming to set up a more scripted relation between events do not give rise to an interaction of structure type and gender match in any of the regions. As such, these results show that QP antecedents in exceptionally covarying constructions with a QP in a temporal adjunct clause can be as easily and as quickly processed as QP antecedents that c-command a pronoun, once the overall interpretation is more supportive of a covarying interpretation.

Parallel to what we saw in Experiment 2, we find early significant GMMs in the NoCC condition (here even starting in the critical region containing just the pronoun, in contrast to Experiment 1, where GMMs only emerged in the spillover region across conditions), with no indication of a delay in GMMs relative to controls (here the CC condition; DP antecedents in Experiment 2). Thus, the presence or absence of c-command between a potential QP antecedent and a pronoun does not in general determine processing time-course of effects based on gender features. Both configurations can exhibit immediate effects, although in the NoCC condition, this is further modulated by contextual and general interpretive properties of the sentences.

Before discussing the broader theoretical repercussions of the present set of findings, also in comparison to previous work, a caveat is in order that the absence of initial interactions here is of course not fully conclusive, in that it is limited to the still relatively course-grained method of SPR at hand, and there could be smaller scale timing discrepancies between the CC and NoCC conditions that are too fine-grained to be captured here. Some potential hints for this are in the data, e.g., in region 8, where a marginally significant interaction appears to be due to a numerical difference in the size of the GMMs, slightly larger in the CC condition, with 80ms, than in the NoCC condition, with 69ms. At the same time, the fact that we already find GMMs on the region containing only the pronoun in Experiment 3 (an effect that replicates in the variant of the experiment reported in Appendix E) points to rather rapid effects. Furthermore, in some cases we find more extended or pronounced GMMs for the QP and NoCC conditions. On balance, and in the absence of positive evidence for differences due to the presence or absence of c-command, we will proceed to our general discussion taking the present result patterns at face value.

5. General Discussion

Our experiments look at variations of the stimuli from Moulton and Han (2018), which implement suggestions by Kush and Eik (2019) for making covarying interpretations of pronouns with QP antecedents in temporal adjunct clauses more easily accessible. Our main empirical question was whether these pronouns would then give rise to GMMs. In theoretical terms, we were interested in how the empirical findings inform theories of pronoun processing and antecedent retrieval, as well as the grammatical mechanisms giving rise to covarying interpretations in different configurations. In this section, we sum up our findings and discuss them in light of the broader questions at hand.

5.1 Can non-c-commanding QP antecedents be accessed early in processing?

In the present context, this question boils down to whether, relative to controls, such as DP antecedents or c-commanding QP antecedents, we find GMMs for non-c-commanding QP antecedents, or whether these are absent or delayed. Moulton and Han's (2018) original findings, as well as our replication in Experiment 1, suggest the latter. Moulton and Han found no GMMs in such cases. Our initial experiment replicates this, with an initial interaction due to GMMs for DP antecedents but not QP antecedents. However, we do find an effect of gender in the spillover regions for QPs, pointing to a delay rather than a complete absence of GMMs. Nonetheless, the pattern in Experiment 1 on its own is compatible with the generalization Moulton and Han build on, in that there is an initial phase where GMMs for pronouns with non-c-commanding antecedents fail to arise.

The data from Kush and Eik (2019) on GMMs in donkey sentences provided a first contrast to this, as GMMs with non-c-commanding QP antecedents in that construction arise as quickly as for DP antecedents. Our data from Experiments 2 and 3 on sentences structurally parallel to ones from Moulton and Han's (2018) experiments extend this to temporal adjunct clauses: in Experiment 2, DP and QP antecedents give rise to GMMs in the same regions, and there's no interaction between them. Experiment 3 further confirms that non-c-commanding QP antecedents lead to GMMs as quickly as c-commanding ones.

The shift in GMMs between the original and modified stimuli appears to be due to the changes based on suggestions from Kush and Eik (2019), which aimed to give the stimuli a more scripted and non-accidental interpretation. While suggesting that non-c-commanding QP antecedents *can* be accessed quickly, the overall data show that this is subject to variation based on non-structural aspects of the stimuli.

5.2 What are the implications for c-command in theories of pronoun processing?

One class of theories on processing pronouns relative to QP antecedents posits a special role for c-command. The probably most straightforward variant of this is that the processor's search space for antecedents is strictly restricted in purely structural terms, such that only noun phrases in syntactic positions that c-command the pronoun are considered (at least during an initial phase). In particular, we considered Moulton and Han's (2018) proposal along these lines, spelled out in technical detail in Moulton (2017). Here, covariation in non-c-commanding environments, analyzed as involving binding of situation variables, is claimed to not (or at least not immediately) involve evaluation of gender features, because the pronoun for which an antecedent is sought is a situation pronoun without such features. Our findings from Experiments 2 and 3 are inconsistent with this general type of approach, as non-c-commanding QP-antecedents there do give rise to GMMs as quickly as the relevant controls. This shows that the QP antecedents in question are accessed early on, and speaks against an initial processing phase where only c-commanding QP antecedents are considered, adding to the previous evidence from donkey sentences in Kush and Eik (2019). More generally, this drives home the point that the presence or absence of GMMs is not determined by purely structural properties of the constructions involved alone, and correspondingly, that theories about the search space for antecedent retrieval can't be defined in purely structural terms.

To some extent, this conclusion contrasts with other findings for conditions of the binding theory, especially for reflexives, where it has been argued that structurally illicit antecedents are categorically ignored in early processing (e.g., Chow et al., 2014). This work in part inspired Kush et al. (2015) and Cummings et al. (2015) to test for similar effects due to the c-command constraint

in QP binding proposed by Reinhart (1983). Whether or not the former findings turn out to hold in full generality (non-standard uses of reflexives could muddy the waters), it is clear from the present results that the range of potential noun phrase antecedents for bound pronouns is not in general limited to ones that c-command the relevant pronoun, not even during an initial processing phase.¹⁷ Further confirmation for this finding should be sought in future work, e.g., using eye-tracking, a tool with higher temporal resolution than self-paced reading, to investigate the processing of stimuli like those in Experiments 2 and 3. Alternative research paradigms (e.g., see Badecker & Straub, 2002) may also shed more light on the way in which c-command and contextual pressures interact in retrieval, by measuring the time-course of processing multiple candidate antecedents.

While our data clearly show that non-c-commanding QP antecedents *can* be accessed quickly, there still remains a difference from the c-commanding ones: for the former, GMMEs only arose right away once adjustments to the original stimuli were made, while the latter showed the effects regardless of the version of the stimuli. This variation in when the effect is present then leaves open the question of whether c-command has a special role in processing pronouns with QP antecedents after all, and if so, what that role might be. Considering what might modulate GMMEs in non-c-commanding cases, as we'll do in more detail in Section 5.3, such a role most likely is at best a very indirect one. The most plausible possibility is that whether or not a QP antecedent is accessed at a given point in time depends on whether or not it is interpreted as having semantic scope over the pronoun.¹⁸ If this is on the right track, then what's special about c-commanding QP antecedents might simply be that they are interpreted to take scope over expressions in their c-command domain by default: c-command does generally align with surface scope interpretations, and these, in turn, have been argued to be preferred in processing (Anderson, 2004). Thus, while semantic scope does not imply c-command (as in the temporal adjunct clauses in our experiments), c-command does quite generally imply (the easy availability of) corresponding semantic scope. The early GMMEs in Moulton and Han's (2018) original c-command conditions and our Experiment 1 (with no enhancements to promote more scripted interpretations) would then be explained in terms of the relevant scope being available by default, based on independent processing principles, without any dependence on contextual support. In contrast, the QPs in temporal adjunct clauses in the no-c-command conditions seem to require some additional support to make an interpretation where they take scope over the pronoun available easily and quickly. The broader prediction that follows from this is that the presence of GMMEs should correlate with the availability of relevant scopal interpretations (as independently measured). Testing this more generally seems like a formidable task to be taken up in future work.

As the present study implemented the changes to the Moulton and Han (2018) materials suggested by Kush and Eik (2019), and our findings directly align with the latter authors' prediction for them, the emerging empirical picture is entirely consistent with the uniform antecedent retrieval mechanism these authors propose. This involves an ACCESSIBLE feature for possible antecedent noun phrases whose setting can be affected by contextual considerations relevant to determining the scope of potential antecedent QPs. Our finding of GMMEs with non-c-

¹⁷ Another early finding along parallel lines comes from research on VP-ellipsis, which argued that binding in that construction can occur in the absence of c-command with the presence of certain intonational signals (Hirschberg & Ward, 1991).

¹⁸ Note that we're considering this here only from the perspective of what constraints might guide the processor in accessing antecedents. This is a separate question from whether or not a grammatical account in terms of scope, without any role for c-command, to all covarying pronouns is theoretically warranted. See Section 5.4 for more details.

commanding QPs adds to their results on donkey sentences, further informing what factors impact the setting of this ACCESSIBLE feature. Since the former, but not the latter, are a case of the QP taking scope over the position of the pronoun (at least on standard views of donkey sentences), our results broaden our understanding of antecedent retrieval in processing across different cases of exceptional covariation. To the extent that GMMs in donkey sentences are not subject to the same type of contextual variation that we found in temporal adjunct clauses, this further aligns with the notion that the availability of the relevant scopal interpretation in the latter is associated with the presence of GMMs there. Future research should extend this approach to further types of exceptionally covarying constructions, such as the various others documented in Barker (2012).

5.3 *What modulates the availability of antecedents and the presence of GMMs?*

While the present data are indeed compatible with a uniform antecedent retrieval mechanism, we hasten to note that the details of how this retrieval mechanism works, and in particular the way in which a host of different factors of quite different nature affect the setting of the ACCESSIBLE feature, have yet to be spelled out. The key questions in the present context are (i) why there is variation of accessibility of QP antecedents in the temporal adjunct configuration, but not in c-command configurations, and (ii) exactly what the nature of the variation in the former is. As already noted briefly above, it is plausible that in general, variation in QP antecedent accessibility is modulated by the relative availability of an interpretation where the QP takes scope over the pronoun. With regards to (i), the source of the variation does not seem to lie in differences in global plausibility of the relevant scopal interpretation, because the CC and NoCC sentence variants in Moulton and Han's (2018) study (illustrated in (7) above) essentially convey the same propositional meaning (given the switch between *before* and *after*).¹⁹ Rather, there does seem to be an interplay of the structural configuration with how easily a given scope interpretation is available. Specifically, when the potential antecedent c-commands the pronoun, this corresponds to a surface scope interpretation, which in turn has been argued to be more easily available in general (Anderson, 2004). In contrast, in non-c-commanding configurations, as with the temporal adjunct clauses in our stimuli, the relevant scope may well be available in principle, but does not in general constitute the default choice. Instead, its availability is modulated by other factors.

This brings us to (ii), i.e., the question of exactly how our modification of the stimuli made an interpretation of the QP taking scope over the pronoun more easily accessible. While we aren't in a position to present a fully fleshed out formal analysis, we offer some more detailed speculations of a plausible-seeming account here, in particular with regards to the semantics involved in the change from past to present tense. Building on Kush and Eik's (2019) intuition that covarying readings are more easily available for quantificational, multi-event readings of these types of sentences, we suggest that this (at least in large part) results from the impact of the quantificational force of the tense operator on the relative availability of an interpretation where the quantifier in the temporal adjunct takes scope over the position of the pronoun. Consider the sketch of an analysis in (15b) of the meaning of the sentence in (15a) from Moulton (2017), which roughly follows the semantic analysis of temporal adjunct clauses in Artstein (2005):²⁰

¹⁹ This is complicated, among other things, by the inclusion and varying location of "it seems," though at least intuitively, this does not seem to affect the interpretation in ways relevant for our purposes.

²⁰ Note that modeling tense in terms of quantification over situation here makes it possible to consider a standard D-Type analysis of the pronoun, which is interpreted as "the boy" relative to the situations temporally specified by tense.

(15)

- a. After each boy came home, he practiced piano
- b. $\forall x[\text{boy}(x) \rightarrow \exists s[\text{came.home}(x)(s) \ \& \ \exists s'[\text{practiced}(\text{hes}')(s') \ \& \ \text{after}(s)(s') \ \& \ M(x)(s')]]]$

While omitting various details, this captures an episodic meaning such that for each boy x , there exists some situation s (in the past, though that's not marked here) in which x came home, and that there is a matching subsequent situation s' in which the relevant boy practices piano.²¹ On a situation-based D-Type analysis, “hes'” here stands for a covert definite description, effectively “the boy in s' ,” thus allowing for covariation without the pronoun being directly bound by “each boy.”²² This reading is a generalization about boys, such that for each of them, a certain sequence of events is said to have occurred once. What changes when we switch to present tense is that the relevant temporal quantification becomes universal as well, which changes the logical configuration.²³

(16)

- a. After each boy comes home, he practices piano
- b. $\forall x \forall s [[\text{boy}(x) \ \& \ \text{comes.home}(x)(s)] \rightarrow \exists s'[\text{practice}(\text{hes}')(s') \ \& \ \text{after}(s)(s') \ \& \ M(x)(s')]]]$

With the two universals—quantifying over boys and situations—taking highest scope together, this now becomes a generalization over what happens when boys come home: all boy-home-coming situations are said to be followed by a situation of the relevant boy practicing piano. We think it's plausible that such a generalization is more natural and cognitively more easily accessible, and that this plays a crucial role in facilitating fast access to a covarying interpretation in our stimuli in Experiment 2. Note that any difference in accessibility across these variants has to be seen in relation to alternative scopings, which we assume the grammar makes equivalently available for both configurations. To illustrate, consider the following variations (without pronouns, but with an indefinite, to explore the different scope interpretations more directly):

(17)

- a. After each boy came home, a snack was served.
- b. $\exists s[\forall x[\text{boy}(x) \rightarrow \text{came.home}(x)(s)] \ \& \ \exists s' \exists y [\text{snack}(y) \ \& \ \text{served}(y)(s')]]$

²¹ “M” here represents a matching function in the sense of Rothstein (1995), which she argues to be required even for simple temporally quantified sentences such as (i):

(i) Each time the door bell rang, Sue opened the door.

This sentence conveys that Sue opened the door as many times as the door bell rang. But if one modeled its truth conditions simply as “for all door bell ringings x , there exists a door opening by Sue y ,” this would be too weak in that they would be met by a situation where, say, after a total of 15 door bell ringings, Sue opened the door one time. The matching function “M” ensures that there is a different door opening for each time the door bell rang.

²² This, of course, is another instance of spelling out a two-grammatical-mechanisms view of covariation, which allows one to maintain a c-command requirement for *bona fide* binding. Without such a requirement, the pronoun here could simply be represented as “ x ” and be directly bound by the universal quantifier “each boy.” See Section 5.4 for more discussion.

²³ At least if we simplify the generic, habitual interpretation of the English present tense, which seems to allow for exceptions, and thus is not fully universal, but we leave it at that for present purposes. As noted by an anonymous reviewer, it may be less clear how exactly this facilitates the relevant scope if we adopt a more fleshed out analysis of the present tense here; since our proposal here is relatively speculative to begin with, we leave further exploration of this issue for future discussion.

- $$\begin{aligned} & \& \text{after}(s)(s') \& M(x)(s'))]] \\ \text{c. } & \forall x [\text{boy}(x) \rightarrow \exists s [\text{came.home}(x)(s) \& \exists s' \exists y [\text{snack}(y) \& \text{served}(y)(s') \\ & \& \text{after}(s)(s') \& M(x)(s'))]]] \\ (18) \\ \text{a. } & \text{After each boy comes home, a snack is served} \\ \text{b. } & \forall s [\forall x [\text{boy}(x) \rightarrow \text{comes.home}(x)(s)] \rightarrow \exists s' \exists y [\text{snack}(y) \& \text{serve}(y)(s') \\ & \& \text{after}(s)(s') \& M(x)(s'))]] \\ \text{c. } & \forall x \forall s [[\text{boy}(x) \& \text{comes.home}(x)(s)] \rightarrow \exists s' \exists y [\text{snack}(y) \& \text{serve}(y)(s') \\ & \& \text{after}(s)(s') \& M(x)(s'))]] \end{aligned}$$

The (b) variants here, where “each boy” does not take scope over “a snack,” require there to be one snack to be served once all the boys are home. The (c) variants, in contrast, require there to be different individual snack servings upon the arrival of each single boy. The key claim based on Kush and Eik's (2019) suggestion and our proposed explanation of the effect of the change to present tense is that in (17), interpretation (c) is less prominent and accessible relative to (b), as compared to (18), where interpretation (c) is more readily available. This naturally is subject to more rigorous empirical assessment, but it strikes us as intuitively plausible. As to a potential explanation of this contrast, should it indeed be confirmed, we'll offer the speculation that this has to do with a preference for letting two universal quantifiers scope together, whether it's because the overall semantic representation winds up simpler in some regard, or whether there is an advantage in conceptual simplicity or naturalness for such a configuration. Naturally, this line of reasoning is subject to further development and investigation, but we hope this is a useful first step in that direction.

5.4 Does our GMME processing data differentiate grammatical theories of covariation?

The final question is what the current empirical picture means for theories about the grammatical mechanisms underlying covariation. The results and their interpretation from Moulton and Han (2018) owed their intrigue to the notion that they provide processing evidence supporting the two grammatical mechanisms of covariation view, with a privileged role for c-commanding antecedents in processing. This was in line with a contrast in the grammar between *bona fide* binding at play in c-command configurations and an alternative, situation-based mechanism at play in non-c-command configurations. But the fuller empirical picture that is now emerging provides a more nuanced and complex picture. What our new data clearly show is that there is not a general restriction in early phases of processing to only consider c-commanding antecedents (with the standard caveat about potential limitations due to the relatively course-grained nature of SPR measures). Once the interpretation of a quantifier in a temporal adjunct clause where it scopes over the matrix clause is made easily available (by a variety of factors, though probably most importantly the switch to present tense), GMMEs arise as early as for c-command and referential DP control variants. Thus, access to the relevant scope interpretation does seem to play a crucial role for the availability of a covarying interpretation.

One might take the apparently central role of scope to align well with Barker's (2012) proposal that a scope constraint is all that is needed in terms of grammatical restrictions on covariation. However, the absence of a general timing difference in accessing c-commanding and non-c-commanding antecedents does not in and of itself speak in favor of positing a single grammatical mechanism to be at play. Oversimplifying somewhat, the choice between the two theories comes

down to variants of the following semantic representations (where “hes” in (19b) again is understood to stand for “the boy in s”):

(19)

- a. After each boy comes home, he practices piano
- b. $\forall s \forall x [[\text{boy}(x) \ \& \ \text{comes.home}(x)(s)] \rightarrow \exists s' [\text{practice}(\text{hes}')(s') \ \& \ \text{after}(s)(s') \ \& \ M(x)(s')]]$
- c. $\forall s \forall x [[\text{boy}(x) \ \& \ \text{comes.home}(x)(s)] \rightarrow \exists s' [\text{practice}(\mathbf{x})(s') \ \& \ \text{after}(s)(s') \ \& \ M(x)(s')]]$

There does not seem to be any principled reason leading us to expect that one of these should be linked to a slower or fundamentally different cognitive process in comprehension.²⁴ But then the fact that in this version of temporal adjunct clauses, we get GMMs as quickly as in DP-antecedent and c-commanding QP controls does not speak for or against either one of these possible theoretical analyses. Thus, we see no general reason to favor either a one-mechanism or two-mechanism grammatical theory of covariation based on the reading time data considered here.

5.5 Conclusion

The perhaps clearest lesson from the present enterprise is that the processing mechanisms of retrieving candidate quantificational antecedents that a given pronoun covaries with are guided by rather deep and subtle aspects of the sentence’s interpretation as it unfolds incrementally. In particular, the semantically subtle, and structurally innocuous, variations between Moulton and Han’s (2018) stimuli and our Experiments 2 and 3 seem to directly and immediately affect how accessible a potential QP antecedent is, arguably due to the availability of an interpretation where the QP takes scope over the pronoun. Thus, the processes involved are not merely formal and mechanical in terms of considering specific syntactic domains and checking for formal features; rather, they engage deeply with the compositional semantic interpretation, including whatever grammatical mechanisms one favors to deal with deviations from surface scope. We therefore submit that work on GMMs in pronoun processing and theories of antecedent retrieval more generally should embrace the intricacies of the semantics and explore the full richness of the hypothesis space that emerges as we consider theoretical and processing questions in this domain in a fully integrated perspective, including the subtleties of both structure and meaning.

Data accessibility statement

Data and analysis code are being made available as part of the submission and will be made accessible for readers on the OSF in case of publication.

Ethics and consent

These studies were conducted within an approved IRB protocol at our university [details omitted for peer review], and participants provided consent for their participation.

²⁴ Moulton and Han (2018) were of course in a very different position, given their data, in that they found differences and offered to ground them in subtle theoretical distinctions. The data become much harder to interpret in such a way once we’re no longer dealing with a general processing pattern correlated with the relevant structural configurations.

Competing interests

The author(s) has/have no competing interests to declare.

Authors' contributions

[Omitted for peer review.]

References

- Anderson, C. (2004). *The structure and real-time comprehension of quantifier scope ambiguity*. [Doctoral dissertation, Northwestern University].
- Anderssen, J. (2011). *Quantification, misc*. [Doctoral dissertation, University of Massachusetts, Amherst]. https://scholarworks.umass.edu/open_access_dissertations/430/
- Artstein, R. (2005). Quantificational Arguments in Temporal Adjunct Clauses. *Linguistics and Philosophy*, 28(5), 541–597. <https://doi.org/10.1007/s10988-005-6921-6>
- Badecker, W., & Straub, K. (2002). The processing role of structural constraints on interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28(4), 748–769. <https://doi.org/10.1037/0278-7393.28.4.748>
- Barker, C., & Shan, C. (2008). Donkey anaphora is in-scope binding. *Semantics and Pragmatics*, 1.1, 1–42. <https://doi.org/10.3765/sp.1.1>
- Barker, C. (2012). Quantificational Binding Does Not Require C-Command. *Linguistic Inquiry*, 43(4), 614–633. https://doi.org/10.1162/ling_a_00108
- Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language*, 68(3), 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, D., Mächler, M., Bolker, B., & Walker, S. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, 67, 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Büring, D. (2004). Crossover Situations. *Natural Language Semantics*, 12(1), 23–62. <https://doi.org/10.1023/B:NALS.0000011144.81075.a8>
- Carminati, M. N., Frazier, L., & Rayner, K. (2002). Bound Variables and C-Command. *Journal of Semantics*, 19(1), 1–34. <https://doi.org/10.1093/jos/19.1.1>
- Chow, W.-Y., Lewis, S., & Phillips, C. (2014). Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology*, 5, 630. <https://doi.org/10.3389/fpsyg.2014.00630>
- Cunnings, I., Patterson, C., & Felser, C. (2015). Structural constraints on pronoun binding and coreference: Evidence from eye movements during reading. *Frontiers in Psychology*, 6. <https://doi.org/10.3389/fpsyg.2015.00840>
- Elbourne, P. D. (2005). *Situations and Individuals*. MIT Press.
- Greene, S. B., Gerrig, R. J., Mckoon, G., & Ratcliff, R. (1994). Unheralded Pronouns and Management by Common Ground. *Journal of Memory and Language*, 33(4), 511–526. <https://doi.org/10.1006/jmla.1994.1024>
- Heim, I. (1990). E-Type pronouns and donkey anaphora. *Linguistics and Philosophy*, 13(2), 137–177. <https://doi.org/10.1007/BF00630732>
- Hirschberg, J., & Ward, G. (1991). Accent and bound anaphora. *Cognitive Linguistics*, 2(2), 101–122. <https://doi.org/10.1515/cogl.1991.2.2.101>

- Kush, D., & Eik, R. (2019). Antecedent accessibility and exceptional covariation: Evidence from Norwegian Donkey Pronouns. *Glossa*, 4(1), 1–17. <http://dx.doi.org/10.5334/gjgl.930>
- Kush, D., Lidz, J., & Phillips, C. (2015). Relation-sensitive retrieval: Evidence from bound variable pronouns. *Journal of Memory and Language*, 82, 18–40. <https://doi.org/10.1016/j.jml.2015.02.003>
- Kush, D. W. (2013). *Respecting Relations: Memory Access and Antecedent Retrieval in Incremental Sentence Processing* [Doctoral dissertation, University of Maryland]. <https://drum.lib.umd.edu/handle/1903/14589>
- Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). lmerTest Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, 82, 1–26. <https://doi.org/10.18637/jss.v082.i13>
- Lewis, R. L., Vasishth, S., & Van Dyke, J. A. (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Sciences*, 10(10), 447–454. <https://doi.org/10.1016/j.tics.2006.08.007>
- McElree, B. (2000). Sentence Comprehension Is Mediated by Content-Addressable Memory Structures. *Journal of Psycholinguistic Research*, 29(2), 111–123. <https://doi.org/10.1023/A:1005184709695>
- Moulton, K. (2017). Retrieving Antecedents in Processing: Binder Indices and ϕ -Features. In W. G. Bennett, L. Hracs, & D. R. Storoshenko (Eds.), *Proceedings of the 35th West Coast Conference on Formal Linguistics* (pp. 30–40). Cascadia Proceedings Project. <http://www.lingref.com/cpp/wccfl/35/abstract3372.html>
- Moulton, K., & Han, C. (2018). C-command vs. scope: An experimental assessment of bound-variable pronouns. *Language*, 94(1), 191–219. <https://doi.org/10.1353/lan.2018.0005>
- Poesio, M., & Zucchi, A. (1992). On Telescoping. *Semantics and Linguistic Theory*, 2(0), 347–366. <https://doi.org/10.3765/salt.v2i0.3034>
- Postal, P. (1966). On so-called “pronouns” in English. *Report on the Seventeenth Annual Round Table Meeting on Linguistics and Language Studies*, 177–206.
- Reinhart, T. (1983). *Anaphora and Semantic Interpretation*. Routledge. <https://doi.org/10.4324/9781315536965>
- Rothstein, S. (1995). Small Clauses and Copular Constructions. In A. Cardinaletti & M. T. Guasti (Eds.), *Small Clauses* (Vol. 28, pp. 25–48). Brill. https://doi.org/10.1163/9780585492209_003
- Safir, K. (2004). *The Syntax of Anaphora*. Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195166132.001.0001>
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48(3), 542–562. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)

Appendix A: Experiment Demos

- (20) Experiment 1: <https://farm.pcibex.net/r/WRsDZO/>
 (21) Experiment 2: <https://farm.pcibex.net/r/BooSRB/>
 (22) Experiment 3: <https://farm.pcibex.net/r/ngjyma/>

Appendix B: Experiment 1 Materials

1	QP	M	After - each fireman - carried the hose - from the firetruck - to the house - it appears - that he - called for - backup to help - quickly.	Did the hose come from the house?	n
1	QP	MM	After - each fireman - carried the hose - from the firetruck - to the house - it appears - that she - called for - backup to help - quickly.	Did the hose come from the house?	n
1	DP	M	After - the fireman - carried the hose - from the firetruck - to the house - it appears - that he - called for - backup to help - quickly.	Did the hose come from the house?	n
1	DP	MM	After - the fireman - carried the hose - from the firetruck - to the house - it appears - that she - called for - backup to help - quickly.	Did the hose come from the house?	n
2	QP	M	Before - each ballerina - fainted on stage - dramatically - during the recital - it appears - that she - practiced - the dance steps - carefully.	Was the recital on stage?	y
2	QP	MM	Before - each ballerina - fainted on stage - dramatically - during the recital - it appears - that he - practiced - the dance steps - carefully.	Was the recital on stage?	y
2	DP	M	Before - the ballerina - fainted on stage - dramatically - during the recital - it appears - that she - practiced - the dance steps - carefully.	Was the recital on stage?	y
2	DP	MM	Before - the ballerina - fainted on stage - dramatically - during the recital - it appears - that he - practiced - the dance steps - carefully.	Was the recital on stage?	y
3	QP	M	Before - each prince - went hunting - in the woods - for a stag - it seems - that he - poisoned - the food - with cyanide.	Was the food poisoned with arsenic?	n
3	QP	MM	Before - each prince - went hunting - in the woods - for a stag - it seems - that she - poisoned - the food - with cyanide.	Was the food poisoned with arsenic?	n
3	DP	M	Before - the prince - went hunting - in the woods - for a stag - it seems - that he - poisoned - the food - with cyanide.	Was the food poisoned with arsenic?	n
3	DP	MM	Before - the prince - went hunting - in the woods - for a stag - it seems - that she - poisoned - the food - with cyanide.	Was the food poisoned with arsenic?	n
4	QP	M	After - each policeman - went on duty - at the event - downtown - it appears - that he - investigated the - crime scene - in the dark.	Was the event downtown?	y
4	QP	MM	After - each policeman - went on duty - at the event - downtown - it appears - that she - investigated the - crime scene - in the dark.	Was the event downtown?	y
4	DP	M	After - the policeman - went on duty - at the event - downtown - it appears - that he - investigated the - crime scene - in the dark.	Was the event downtown?	y
4	DP	MM	After - the policeman - went on duty - at the event - downtown - it appears - that she - investigated the - crime scene - in the dark.	Was the event downtown?	y
5	QP	M	After - each salesman - sold a brand new Mercedes - to the rich man - for a big profit - it appears - that he - got a big bonus - this year - from the boss.	Was a brand new Toyota sold to the rich man?	n
5	QP	MM	After - each salesman - sold a brand new Mercedes - to the rich man - for a big profit - it appears - that she - got a big bonus - this year - from the boss.	Was a brand new Toyota sold to the rich man?	n
5	DP	M	After - the salesman - sold a brand new Mercedes - to the rich man - for a big profit - it appears - that he - got a big bonus - this year - from the boss.	Was a brand new Toyota sold to the rich man?	n
5	DP	MM	After - the salesman - sold a brand new Mercedes - to the rich man - for a big profit - it appears - that she - got a big bonus - this year - from the boss.	Was a brand new Toyota sold to the rich man?	n
6	QP	M	Before - each woman - got worried - about proper - safety procedures - it seems - that she - checked on - the operating room - several times.	Was the room an operating room?	y
6	QP	MM	Before - each woman - got worried - about proper - safety procedures - it seems - that he - checked on - the operating room - several times.	Was the room an operating room?	y
6	DP	M	Before - the woman - got worried - about proper - safety procedures - it seems - that she - checked on - the operating room - several times.	Was the room an operating room?	y
6	DP	MM	Before - the woman - got worried - about proper - safety procedures - it seems - that he - checked on - the operating room - several times.	Was the room an operating room?	y
7	QP	M	When - each waitress - spilled the drinks - all over the floor - of the restaurant - it seems - that she - went quickly - back to - the kitchen.	Did the drinks spill all over the floor of the hospital?	n
7	QP	MM	When - each waitress - spilled the drinks - all over the floor - of the restaurant - it seems - that he - went quickly - back to - the kitchen.	Did the drinks spill all over the floor of the hospital?	n

7	DP	M	When - the waitress - spilled the drinks - all over the floor - of the restaurant - it seems - that she - went quickly - back to - the kitchen.	Did the drinks spill all over the floor of the hospital?	n
7	DP	MM	When - the waitress - spilled the drinks - all over the floor - of the restaurant - it seems - that he - went quickly - back to - the kitchen.	Did the drinks spill all over the floor of the hospital?	n
8	QP	M	When - each groomsman - was on the dance - floor having fun - with friends - it appears - that he - asked one of - the little flower girls - to dance.	Were the flower girls little?	y
8	QP	MM	When - each groomsman - was on the dance - floor having fun - with friends - it appears - that she - asked one of - the little flower girls - to dance.	Were the flower girls little?	y
8	DP	M	When - the groomsman - was on the dance - floor having fun - with friends - it appears - that he - asked one of - the little flower girls - to dance.	Were the flower girls little?	y
8	DP	MM	When - the groomsman - was on the dance - floor having fun - with friends - it appears - that she - asked one of - the little flower girls - to dance.	Were the flower girls little?	y
9	QP	M	Before - each mother - arrived at school - for a meeting - early in the morning - it seems - that she - phoned the principal - to complain - about bullies.	Was the meeting at night?	n
9	QP	MM	Before - each mother - arrived at school - for a meeting - early in the morning - it seems - that he - phoned the principal - to complain - about bullies.	Was the meeting at night?	n
9	DP	M	Before - the mother - arrived at school - for a meeting - early in the morning - it seems - that she - phoned the principal - to complain - about bullies.	Was the meeting at night?	n
9	DP	MM	Before - the mother - arrived at school - for a meeting - early in the morning - it seems - that he - phoned the principal - to complain - about bullies.	Was the meeting at night?	n
10	QP	M	When - each actress - was practicing - lines for the play - at home - it seems - that she - learned that the - play would appear - on Broadway.	Would the play appear on Broadway?	y
10	QP	MM	When - each actress - was practicing - lines for the play - at home - it seems - that he - learned that the - play would appear - on Broadway.	Would the play appear on Broadway?	y
10	DP	M	When - the actress - was practicing - lines for the play - at home - it seems - that she - learned that the - play would appear - on Broadway.	Would the play appear on Broadway?	y
10	DP	MM	When - the actress - was practicing - lines for the play - at home - it seems - that he - learned that the - play would appear - on Broadway.	Would the play appear on Broadway?	y
11	QP	M	When - each little girl - was at the park - playing with friends - from school - it appears - that she - got a strawberry - ice-cream cone - for free.	Was the ice-cream chocolate-flavored?	n
11	QP	MM	When - each little girl - was at the park - playing with friends - from school - it appears - that he - got a strawberry - ice-cream cone - for free.	Was the ice-cream chocolate-flavored?	n
11	DP	M	When - the little girl - was at the park - playing with friends - from school - it appears - that she - got a strawberry - ice-cream cone - for free.	Was the ice-cream chocolate-flavored?	n
11	DP	MM	When - the little girl - was at the park - playing with friends - from school - it appears - that he - got a strawberry - ice-cream cone - for free.	Was the ice-cream chocolate-flavored?	n
12	QP	M	After - each little boy - came home - late - from a long day at school - it seems - that he - practiced playing - the piano - for half an hour.	Was it a long day at school?	y
12	QP	MM	After - each little boy - came home - late - from a long day at school - it seems - that she - practiced playing - the piano - for half an hour.	Was it a long day at school?	y
12	DP	M	After - the little boy - came home - late - from a long day at school - it seems - that he - practiced playing - the piano - for half an hour.	Was it a long day at school?	y
12	DP	MM	After - the little boy - came home - late - from a long day at school - it seems - that she - practiced playing - the piano - for half an hour.	Was it a long day at school?	y
13	QP	M	When - each businessman - spoke at the - city council meeting - last night - it appears - that he - was angry about - high taxes - and new bylaws.	Was the city council meeting yesterday morning?	n
13	QP	MM	When - each businessman - spoke at the - city council meeting - last night - it appears - that she - was angry about - high taxes - and new bylaws.	Was the city council meeting yesterday morning?	n
13	DP	M	When - the businessman - spoke at the - city council meeting - last night - it appears - that he - was angry about - high taxes - and new bylaws.	Was the city council meeting yesterday morning?	n
13	DP	MM	When - the businessman - spoke at the - city council meeting - last night - it appears - that she - was angry about - high taxes - and new bylaws.	Was the city council meeting yesterday morning?	n
14	QP	M	When - each old woman - came quickly - out of the house - down the road - it seems - that she - yelled mean - things loudly - for hours.	Was the house down the road?	y
14	QP	MM	When - each old woman - came quickly - out of the house - down the road - it seems - that he - yelled mean - things loudly - for hours.	Was the house down the road?	y
14	DP	M	When - the old woman - came quickly - out of the house - down the road - it seems - that she - yelled mean - things loudly - for hours.	Was the house down the road?	y
14	DP	MM	When - the old woman - came quickly - out of the house - down the road - it seems - that he - yelled mean - things loudly - for hours.	Was the house down the road?	y
15	QP	M	After - each secretary - typed up the minutes - in the office - from the meeting yesterday - it seems - that she - organized the shelves - in the photocopying room - very neatly.	Were the shelves in the storage room?	n

15	QP	MM	After - each secretary - typed up the minutes - in the office - from the meeting yesterday - it seems - that he - organized the shelves - in the photocopying room - very neatly.	Were the shelves in the storage room?	n
15	DP	M	After - the secretary - typed up the minutes - in the office - from the meeting yesterday - it seems - that she - organized the shelves - in the photocopying room - very neatly.	Were the shelves in the storage room?	n
15	DP	MM	After - the secretary - typed up the minutes - in the office - from the meeting yesterday - it seems - that he - organized the shelves - in the photocopying room - very neatly.	Were the shelves in the storage room?	n
16	QP	M	After - each hockey player - changed into skates - carefully - at the rink - it appears - that he - skated on the - fresh ice - for several hours.	Was the ice at the rink initially fresh?	y
16	QP	MM	After - each hockey player - changed into skates - carefully - at the rink - it appears - that she - skated on the - fresh ice - for several hours.	Was the ice at the rink initially fresh?	y
16	DP	M	After - the hockey player - changed into skates - carefully - at the rink - it appears - that he - skated on the - fresh ice - for several hours.	Was the ice at the rink initially fresh?	y
16	DP	MM	After - the hockey player - changed into skates - carefully - at the rink - it appears - that she - skated on the - fresh ice - for several hours.	Was the ice at the rink initially fresh?	y
17	QP	M	After - each plumber - fixed a leaky sink - in the kitchen - in the morning - it appears - that he - picked out a new - faucet for the bathroom - in the afternoon.	Was the sink in the laundry room?	n
17	QP	MM	After - each plumber - fixed a leaky sink - in the kitchen - in the morning - it appears - that she - picked out a new - faucet for the bathroom - in the afternoon.	Was the sink in the laundry room?	n
17	DP	M	After - the plumber - fixed a leaky sink - in the kitchen - in the morning - it appears - that he - picked out a new - faucet for the bathroom - in the afternoon.	Was the sink in the laundry room?	n
17	DP	MM	After - the plumber - fixed a leaky sink - in the kitchen - in the morning - it appears - that she - picked out a new - faucet for the bathroom - in the afternoon.	Was the sink in the laundry room?	n
18	QP	M	After - each bridesmaid - had her hair - done up fancy - for the wedding - it appears - that she - helped the bride - get to the church - on time.	Did the bride get to church on time?	y
18	QP	MM	After - each bridesmaid - had her hair - done up fancy - for the wedding - it appears - that he - helped the bride - get to the church - on time.	Did the bride get to church on time?	y
18	DP	M	After - the bridesmaid - had her hair - done up fancy - for the wedding - it appears - that she - helped the bride - get to the church - on time.	Did the bride get to church on time?	y
18	DP	MM	After - the bridesmaid - had her hair - done up fancy - for the wedding - it appears - that he - helped the bride - get to the church - on time.	Did the bride get to church on time?	y
19	QP	M	When - each new father - arrived at the - hospital in a rush - from work - it seems - that he - was told by - the doctor that - everything would be fine.	Did the doctor say there was a tragedy?	n
19	QP	MM	When - each new father - arrived at the - hospital in a rush - from work - it seems - that she - was told by - the doctor that - everything would be fine.	Did the doctor say there was a tragedy?	n
19	DP	M	When - the new father - arrived at the - hospital in a rush - from work - it seems - that he - was told by - the doctor that - everything would be fine.	Did the doctor say there was a tragedy?	n
19	DP	MM	When - the new father - arrived at the - hospital in a rush - from work - it seems - that she - was told by - the doctor that - everything would be fine.	Did the doctor say there was a tragedy?	n
20	QP	M	After - each boy - brought fresh water - from the kitchen - quickly - it seems - that he - went - on an early - break.	Did the water come from the kitchen?	y
20	QP	MM	After - each boy - brought fresh water - from the kitchen - quickly - it seems - that she - went - on an early - break.	Did the water come from the kitchen?	y
20	DP	M	After - the boy - brought fresh water - from the kitchen - quickly - it seems - that he - went - on an early - break.	Did the water come from the kitchen?	y
20	DP	MM	After - the boy - brought fresh water - from the kitchen - quickly - it seems - that she - went - on an early - break.	Did the water come from the kitchen?	y
21	filler	filler	The waiter - said - he was getting - really - annoyed with - the apparent - lack of effort - of the kitchen - staff - tonight.	Were the kitchen staff working hard?	n
22	filler	filler	The firefighter - thought - he was - certainly - going to faint - from exhaustion - as the fire - in the barn - was blazing - so strongly.	Was the firefighter feeling faint?	y
23	filler	filler	The chairman - of the board - solemnly - announced that - he would retire - next - year - because - of chronic - illness.	Was the chairman retiring next week?	n
24	filler	filler	The waiter - flirtatiously - said - he could - offer - Susan - a strong - margarita - for free - that evening.	Did the waiter offer a margarita?	y
25	filler	filler	The landlord - said - he would - happily throw - a welcome party - for new residents - on the block - next Friday - if the weather - held up.	Will the party happen rain or shine?	n
26	filler	filler	The best man - at the wedding - in Boston - last weekend - thought - she was - always - the best dressed - at every - occasion.	Was the wedding in Boston?	y
27	filler	filler	The policeman - said - she could - easily figure out - the details - of the robbery - last night - given the evidence - left behind - at the scene.	Was there a robbery that morning?	n
28	filler	filler	The waiter - at the bar - said - she wore - a big smile - for the whole - time - that evening - at the surprise - birthday party.	Does the waiter work at a bar?	y

29	filler	filler	The boy - said - she was - very - disappointed at - the result - of the game - since - the Eagles - lost again.	Did the Eagles win the game?	n
30	filler	filler	President Obama - said - she would - be given - an opportunity - to ask just - one question - right - after the press - conference finishes.	Did President Obama allow for a question?	y
31	filler	filler	Kate - wished - he would - sing Uptown Funk - and - Gangnam Style - karaoke - to break the ice - at the staff - orientation.	Was Kate at a going away party?	n
32	filler	filler	Susan - commented that - he liked - the mascots - and their - costumes - a lot - but not - the tie-dyed team - uniforms.	Did someone like the mascots?	y
33	filler	filler	The nurse - said - he was - surprised at - how much - the tiny wound - on his finger - hurt the varsity - basketball - athlete.	Was the wound big?	n
34	filler	filler	The saleswoman - said - she loved - the style - and the fit - of the shirt - the model - wore - from this year's - winter collection.	Did the shirt fit properly?	y
35	filler	filler	The girl - said - she marveled at - the large - selection of - fresh fruits - and local vegetables - at the farmers' - market - last Saturday.	Did the girl marvel at the selection of bread?	n
36	filler	filler	The waiter - cleaned up - the spoons - before - he served - the decadent - cheesecake - to the newly - engaged - couple.	Did the waiter serve cheesecake?	y
37	filler	filler	The housemaid - put away - the stacks of - magazines - before - he left - the room - to prepare - for the guests - to arrive.	Did the housemaid put away the dishes?	n
38	filler	filler	The policewoman went - to the coffee shop - for a quick - lunch break - before - she finished - filling out - the document - concerning - the recent crime.	Did the policewoman fill out a document?	y
39	filler	filler	The hairstylist - consulted - a few current - magazines - before - he decided - on a hairstyle - for the upcoming - graduation - ceremonies.	Was the hairstyle chosen right away?	n
40	filler	filler	The landlady - barbecued - huge amounts of - sausage and ham - for her - European tenants - before - she baked - some fresh - banana bread.	Did the landlady bake banana bread?	y
41	practice	practice	This is - a practice - sentence - to get you - used to - reading sentences - like this - during the - experiment - today.	N/A	N/A
42	practice	practice	The monkey - swung from - tree to tree - through - the jungle - as the - angry lion - pursued him - desperately - close behind.	Was the lion following the monkey?	y
43	practice	practice	The motorist - failed to make - a complete stop - prior to - entering - the intersection - which - angered the - nearby - pedestrian.	Did the motorist make a complete stop?	n

Appendix C: Experiment 2 Materials

1	QP	M	After - each boy - fetches a bucket - of water - from the well - he - goes - to clean the - barn and stables.	Does the bucket contain milk?	n
1	QP	MM	After - each boy - fetches a bucket - of water - from the well - she - goes - to clean the - barn and stables.	Does the bucket contain milk?	n
1	DP	M	After - the boy - fetches a bucket - of water - from the well - he - goes - to clean the - barn and stables.	Does the bucket contain milk?	n
1	DP	MM	After - the boy - fetches a bucket - of water - from the well - she - goes - to clean the - barn and stables.	Does the bucket contain milk?	n
2	QP	M	Before - each ballerina - performs a routine - to audition - for the troupe - she - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
2	QP	MM	Before - each ballerina - performs a routine - to audition - for the troupe - he - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
2	DP	M	Before - the ballerina - performs a routine - to audition - for the troupe - she - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
2	DP	MM	Before - the ballerina - performs a routine - to audition - for the troupe - he - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
3	QP	M	Before - each prince - steps forward - to receive a sword - for bravery - he - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
3	QP	MM	Before - each prince - steps forward - to receive a sword - for bravery - she - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
3	DP	M	Before - the prince - steps forward - to receive a sword - for bravery - he - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
3	DP	MM	Before - the prince - steps forward - to receive a sword - for bravery - she - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
4	QP	M	After - each policeman - investigates - an incident and - interviews witnesses - he - files a - police report - detailing the event.	Does the police report detail the incident?	y
4	QP	MM	After - each policeman - investigates - an incident and - interviews witnesses - she - files a - police report - detailing the event.	Does the police report detail the incident?	y
4	DP	M	After - the policeman - investigates - an incident and - interviews witnesses - he - files a - police report - detailing the event.	Does the police report detail the incident?	y

4	DP	MM	After - the policeman - investigates - an incident and - interviews witnesses - she - files a - police report - detailing the event.	Does the police report detail the incident?	y
5	QP	M	Before - each woman - chooses a perfume - to wear to - the banquet - she - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
5	QP	MM	Before - each woman - chooses a perfume - to wear to - the banquet - he - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
5	DP	M	Before - the woman - chooses a perfume - to wear to - the banquet - she - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
5	DP	MM	Before - the woman - chooses a perfume - to wear to - the banquet - he - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
6	QP	M	After - each fireman - retrieves an axe - and gas mask - from the firetruck - he - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
6	QP	MM	After - each fireman - retrieves an axe - and gas mask - from the firetruck - she - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
6	DP	M	After - the fireman - retrieves an axe - and gas mask - from the firetruck - he - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
6	DP	MM	After - the fireman - retrieves an axe - and gas mask - from the firetruck - she - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
7	QP	M	When - each waitress - prepares a table - for customers - at the restaurant - she - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
7	QP	MM	When - each waitress - prepares a table - for customers - at the restaurant - he - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
7	DP	M	When - the waitress - prepares a table - for customers - at the restaurant - she - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
7	DP	MM	When - the waitress - prepares a table - for customers - at the restaurant - he - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
8	QP	M	When - each groomsman - stands up - to deliver a speech - on stage - he - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
8	QP	MM	When - each groomsman - stands up - to deliver a speech - on stage - she - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
8	DP	M	When - the groomsman - stands up - to deliver a speech - on stage - he - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
8	DP	MM	When - the groomsman - stands up - to deliver a speech - on stage - she - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
9	QP	M	After - each salesman - makes a sale - while traveling - on the job - he - returns to - the office to - report the sale.	Is the sale made while at the office?	n
9	QP	MM	After - each salesman - makes a sale - while traveling - on the job - she - returns to - the office to - report the sale.	Is the sale made while at the office?	n
9	DP	M	After - the salesman - makes a sale - while traveling - on the job - he - returns to - the office to - report the sale.	Is the sale made while at the office?	n
9	DP	MM	After - the salesman - makes a sale - while traveling - on the job - she - returns to - the office to - report the sale.	Is the sale made while at the office?	n
10	QP	M	When - each actress - recites a line - while preparing - for the new play - she - holds the - script nearby - for reference.	Is the play a new play?	y
10	QP	MM	When - each actress - recites a line - while preparing - for the new play - he - holds the - script nearby - for reference.	Is the play a new play?	y
10	DP	M	When - the actress - recites a line - while preparing - for the new play - she - holds the - script nearby - for reference.	Is the play a new play?	y
10	DP	MM	When - the actress - recites a line - while preparing - for the new play - he - holds the - script nearby - for reference.	Is the play a new play?	y
11	QP	M	After - each bridesmaid - picks out - a dress from - the bridal boutique - she - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
11	QP	MM	After - each bridesmaid - picks out - a dress from - the bridal boutique - he - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
11	DP	M	After - the bridesmaid - picks out - a dress from - the bridal boutique - she - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
11	DP	MM	After - the bridesmaid - picks out - a dress from - the bridal boutique - he - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
12	QP	M	When - each little girl - is given a - word to spell at - the spelling bee - she - is told - the definition - and etymology.	Is the etymology of the word provided?	y
12	QP	MM	When - each little girl - is given a - word to spell at - the spelling bee - he - is told - the definition - and etymology.	Is the etymology of the word provided?	y
12	DP	M	When - the little girl - is given a - word to spell at - the spelling bee - she - is told - the definition - and etymology.	Is the etymology of the word provided?	y
12	DP	MM	When - the little girl - is given a - word to spell at - the spelling bee - he - is told - the definition - and etymology.	Is the etymology of the word provided?	y
13	QP	M	After - each plumber - installs a pipe - in the bathroom - of the new house - he - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n

13	QP	MM	After - each plumber - installs a pipe - in the bathroom - of the new house - she - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
13	DP	M	After - the plumber - installs a pipe - in the bathroom - of the new house - he - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
13	DP	MM	After - the plumber - installs a pipe - in the bathroom - of the new house - she - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
14	QP	M	After - each secretary - types up a memo - for the employees - to read - she - emails the boss - with relevant - important information.	Is the memo for employees?	y
14	QP	MM	After - each secretary - types up a memo - for the employees - to read - he - emails the boss - with relevant - important information.	Is the memo for employees?	y
14	DP	M	After - the secretary - types up a memo - for the employees - to read - she - emails the boss - with relevant - important information.	Is the memo for employees?	y
14	DP	MM	After - the secretary - types up a memo - for the employees - to read - he - emails the boss - with relevant - important information.	Is the memo for employees?	y
15	QP	M	After - each hockey player - buys a pair - of skates - at the arena - he - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
15	QP	MM	After - each hockey player - buys a pair - of skates - at the arena - she - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
15	DP	M	After - the hockey player - buys a pair - of skates - at the arena - he - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
15	DP	MM	After - the hockey player - buys a pair - of skates - at the arena - she - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
16	QP	M	After - each little boy - takes a few swings - at the plate - during the baseball game - he - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
16	QP	MM	After - each little boy - takes a few swings - at the plate - during the baseball game - she - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
16	DP	M	After - the little boy - takes a few swings - at the plate - during the baseball game - he - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
16	DP	MM	After - the little boy - takes a few swings - at the plate - during the baseball game - she - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
17	QP	M	When - each father - plans a - family trip - for a weekday - he - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
17	QP	MM	When - each father - plans a - family trip - for a weekday - she - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
17	DP	M	When - the father - plans a - family trip - for a weekday - he - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
17	DP	MM	When - the father - plans a - family trip - for a weekday - she - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
18	QP	M	When - each old woman - attends a bingo night - at the rec center - downtown - she - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
18	QP	MM	When - each old woman - attends a bingo night - at the rec center - downtown - he - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
18	DP	M	When - the old woman - attends a bingo night - at the rec center - downtown - she - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
18	DP	MM	When - the old woman - attends a bingo night - at the rec center - downtown - he - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
19	QP	M	When - each businessman - files a tax return - at the end - of the quarter - he - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
19	QP	MM	When - each businessman - files a tax return - at the end - of the quarter - she - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
19	DP	M	When - the businessman - files a tax return - at the end - of the quarter - he - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
19	DP	MM	When - the businessman - files a tax return - at the end - of the quarter - she - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
20	QP	M	Before - each mother - sets up - a playdate - for the kids - she - checks to - make sure the date - is available.	Is the playdate for the kids?	y
20	QP	MM	Before - each mother - sets up - a playdate - for the kids - he - checks to - make sure the date - is available.	Is the playdate for the kids?	y
20	DP	M	Before - the mother - sets up - a playdate - for the kids - she - checks to - make sure the date - is available.	Is the playdate for the kids?	y
20	DP	MM	Before - the mother - sets up - a playdate - for the kids - he - checks to - make sure the date - is available.	Is the playdate for the kids?	y
21	filler	filler	The waiter - said - he was getting - annoyed with - the apparent - lack of effort - of the kitchen - staff - tonight.	Were the kitchen staff working hard?	n
22	filler	filler	The firefighter - thought - he was - going to faint - from exhaustion - as the fire - in the barn - was blazing - so strongly.	Was the firefighter feeling faint?	y

23	filler	filler	The chairman - of the board - announced that - he would retire - next - year - because - of chronic - illness.	Was the chairman re- tiring next week?	n
24	filler	filler	The waiter - said - he could - offer - Susan - a strong - margarita - for free - that evening.	Did the waiter offer a margarita?	y
25	filler	filler	The landlord - said - he would - happily throw - a welcome party - for new residents - next Friday - if the weather - held up.	Will the party happen rain or shine?	n
26	filler	filler	The best man - at the wedding - in Boston - thought - she was - always - the best dressed - at every - occasion.	Was the wedding in Boston?	y
27	filler	filler	The policeman - said - she could - easily figure out - the details - of the robbery - last night - given the evidence - left behind.	Was there a robbery that morning?	n
28	filler	filler	The waiter - at the bar - said - she wore - a big smile - for the whole - time - at the surprise - birthday party.	Does the waiter work at a bar?	y
29	filler	filler	The boy - said - she was - disappointed at - the result - of the game - since - the Eagles - lost again.	Did the Eagles win the game?	n
30	filler	filler	President Obama - said - she would - be given - an opportunity - to ask just - one question - after the press - conference finishes.	Did President Obama allow for a question?	y
31	filler	filler	Kate - wished - he would - sing Uptown Funk - and - Gangnam Style - kara- oke - at the staff - orientation.	Was Kate at a going away party?	n
32	filler	filler	Susan - commented that - he liked - the mascots - and their - costumes - but not - the tie-dyed team - uniforms.	Did someone like the mascots?	y
33	filler	filler	The nurse - said - he was - surprised at - how much - the tiny wound - hurt the varsity - basketball - athlete.	Was the wound big?	n
34	filler	filler	The saleswoman - said - she loved - the fit - of the shirt - the model - wore - from this year's - winter collection.	Did the shirt fit properly?	y
35	filler	filler	The girl - said - she marveled at - the large - selection of - fresh fruits - at the farmers' - market - last Saturday.	Did the girl marvel at the selection of bread?	n
36	filler	filler	The waiter - cleaned up - the spoons - before - he served - the cheesecake - to the newly - engaged - couple.	Did the waiter serve cheesecake?	y
37	filler	filler	The housemaid - put away - the stacks of - magazines - before - he left - the room - to prepare - for the guests.	Did the housemaid put away the dishes?	n
38	filler	filler	The policewoman went - for a quick - lunch break - before - she finished - filling out - the document - concerning - the recent crime.	Did the policewoman fill out a document?	y
39	filler	filler	The hairstylist - consulted - a few current - magazines - before - he decided - on a hairstyle - for the upcoming - graduation.	Was the hairstyle chosen right away?	n
40	filler	filler	The landlady - barbecued - huge amounts of - sausage and ham - for her - European tenants - before - she baked - banana bread.	Did the landlady bake banana bread?	y
41	practice	practice	This is - a practice - sentence - to get you - used to - reading sentences - like this - during the - experiment.	N/A	N/A
42	practice	practice	The monkey - swung from - tree to tree - through - the jungle - as the - an- gry lion - pursued him - desperately.	Was the lion follow- ing the monkey?	y
43	practice	practice	The motorist - failed to make - a complete stop - prior to - entering - the in- tersection - which - angered the - pedestrian.	Did the motorist come to a complete stop?	n

Appendix D: Experiment 3 Materials

1	CC	M	On foot, - each boy - fetches a bucket - of water - from the well - before - he - goes - to clean the - barn and stables.	Does the bucket con- tain milk?	n
1	CC	MM	On foot, - each boy - fetches a bucket - of water - from the well - before - she - goes - to clean the - barn and stables.	Does the bucket con- tain milk?	n
1	NoCC	M	After - each boy - fetches a bucket - of water - from the well - on foot, - he - goes - to clean the - barn and stables.	Does the bucket con- tain milk?	n
1	NoCC	MM	After - each boy - fetches a bucket - of water - from the well - on foot, - she - goes - to clean the - barn and stables.	Does the bucket con- tain milk?	n
2	CC	M	With grace, - each ballerina - performs a routine - to audition - for the troupe - after - she - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
2	CC	MM	With grace, - each ballerina - performs a routine - to audition - for the troupe - after - he - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
2	NoCC	M	Before - each ballerina - performs a routine - to audition - for the troupe - with grace, - she - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
2	NoCC	MM	Before - each ballerina - performs a routine - to audition - for the troupe - with grace, - he - practices the - dance steps - many times.	Is the audition for a ballet troupe?	y
3	CC	M	With pride, - each prince - steps forward - to receive a sword - for bravery - after - he - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
3	CC	MM	With pride, - each prince - steps forward - to receive a sword - for bravery - after - she - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
3	NoCC	M	Before - each prince - steps forward - to receive a sword - for bravery - with pride, - he - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n

3	NoCC	MM	Before - each prince - steps forward - to receive a sword - for bravery - with pride, - she - takes a - bow in front of - the king.	Is the sword being given for chivalry?	n
4	CC	M	Thoroughly, - each policeman - investigates - an incident and - interviews witnesses - before - he - files a - police report - detailing the event.	Does the police report detail the incident?	y
4	CC	MM	Thoroughly, - each policeman - investigates - an incident and - interviews witnesses - before - she - files a - police report - detailing the event.	Does the police report detail the incident?	y
4	NoCC	M	After - each policeman - investigates - an incident and - interviews witnesses - thoroughly, - he - files a - police report - detailing the event.	Does the police report detail the incident?	y
4	NoCC	MM	After - each policeman - investigates - an incident and - interviews witnesses - thoroughly, - she - files a - police report - detailing the event.	Does the police report detail the incident?	y
5	CC	M	At last, - each woman - chooses a perfume - to wear to - the banquet - after - she - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
5	CC	MM	At last, - each woman - chooses a perfume - to wear to - the banquet - after - he - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
5	NoCC	M	Before - each woman - chooses a perfume - to wear to - the banquet - at last, - she - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
5	NoCC	MM	Before - each woman - chooses a perfume - to wear to - the banquet - at last, - he - samples several - different scents - at the store.	Is the perfume for a business meeting?	n
6	CC	M	With strength, - each fireman - retrieves an axe - and gas mask - from the firetruck - before - he - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
6	CC	MM	With strength, - each fireman - retrieves an axe - and gas mask - from the firetruck - before - she - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
6	NoCC	M	After - each fireman - retrieves an axe - and gas mask - from the firetruck - with strength, - he - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
6	NoCC	MM	After - each fireman - retrieves an axe - and gas mask - from the firetruck - with strength, - she - searches the - scene for - trapped victims.	Does the axe come from the firetruck?	y
7	CC	M	At mealtime, - each waitress - prepares a table - for customers - at the restaurant - after - she - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
7	CC	MM	At mealtime, - each waitress - prepares a table - for customers - at the restaurant - after - he - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
7	NoCC	M	Before - each waitress - prepares a table - for customers - at the restaurant - at mealtime, - she - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
7	NoCC	MM	Before - each waitress - prepares a table - for customers - at the restaurant - at mealtime, - he - polishes the - nice silverware - meticulously.	Is the table for hotel patrons?	n
8	CC	M	One-by-one, - each groomsman - stands up - to deliver a speech - on stage - before - he - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
8	CC	MM	One-by-one, - each groomsman - stands up - to deliver a speech - on stage - before - she - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
8	NoCC	M	After - each groomsman - stands up - to deliver a speech - on stage - one-by-one, - he - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
8	NoCC	MM	After - each groomsman - stands up - to deliver a speech - on stage - one-by-one, - she - raises a - toast to the - bride and groom.	Is the toast to the bride and groom?	y
9	CC	M	Through effort, - each salesman - makes a sale - while traveling - on the job - before - he - returns to - the office to - report the sale.	Is the sale made while at the office?	n
9	CC	MM	Through effort, - each salesman - makes a sale - while traveling - on the job - before - she - returns to - the office to - report the sale.	Is the sale made while at the office?	n
9	NoCC	M	After - each salesman - makes a sale - while traveling - on the job - through effort, - he - returns to - the office to - report the sale.	Is the sale made while at the office?	n
9	NoCC	MM	After - each salesman - makes a sale - while traveling - on the job - through effort, - she - returns to - the office to - report the sale.	Is the sale made while at the office?	n
10	CC	M	With grace, - each actress - recites a line - to prepare - for the new play - while - she - holds the - script nearby - for reference.	Is the play a new play?	y
10	CC	MM	With grace, - each actress - recites a line - to prepare - for the new play - while - he - holds the - script nearby - for reference.	Is the play a new play?	y
10	NoCC	M	While - each actress - recites a line - to prepare - for the new play - with grace, - she - holds the - script nearby - for reference.	Is the play a new play?	y
10	NoCC	MM	While - each actress - recites a line - to prepare - for the new play - with grace, - he - holds the - script nearby - for reference.	Is the play a new play?	y
11	CC	M	With care, - each bridesmaid - picks out - a dress from - the bridal boutique - before - she - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
11	CC	MM	With care, - each bridesmaid - picks out - a dress from - the bridal boutique - before - he - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
11	NoCC	M	After - each bridesmaid - picks out - a dress from - the bridal boutique - with care, - she - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n
11	NoCC	MM	After - each bridesmaid - picks out - a dress from - the bridal boutique - with care, - he - texts a picture - to the bride - for approval.	Does the picture get emailed to the bride?	n

12	CC	M	At first, - each little girl - is given a - word to spell at - the spelling bee - before - she - is told - the definition - and etymology.	Is the etymology of the word provided?	y
12	CC	MM	At first, - each little girl - is given a - word to spell at - the spelling bee - before - he - is told - the definition - and etymology.	Is the etymology of the word provided?	y
12	NoCC	M	After - each little girl - is given a - word to spell at - the spelling bee - at first, - she - is told - the definition - and etymology.	Is the etymology of the word provided?	y
12	NoCC	MM	After - each little girl - is given a - word to spell at - the spelling bee - at first, - he - is told - the definition - and etymology.	Is the etymology of the word provided?	y
13	CC	M	With a wrench, - each plumber - installs a pipe - in the bathroom - of the new house - before - he - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
13	CC	MM	With a wrench, - each plumber - installs a pipe - in the bathroom - of the new house - before - she - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
13	NoCC	M	After - each plumber - installs a pipe - in the bathroom - of the new house - with a wrench, - he - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
13	NoCC	MM	After - each plumber - installs a pipe - in the bathroom - of the new house - with a wrench, - she - inspects the - joints with - a special tool.	Is the pipe installed in the kitchen?	n
14	CC	M	With speed, - each secretary - types up a memo - for the employees - to read - before - she - emails the boss - with relevant - important information.	Is the memo for employees?	y
14	CC	MM	With speed, - each secretary - types up a memo - for the employees - to read - before - he - emails the boss - with relevant - important information.	Is the memo for employees?	y
14	NoCC	M	After - each secretary - types up a memo - for the employees - to read - with speed, - she - emails the boss - with relevant - important information.	Is the memo for employees?	y
14	NoCC	MM	After - each secretary - types up a memo - for the employees - to read - with speed, - he - emails the boss - with relevant - important information.	Is the memo for employees?	y
15	CC	M	With cash, - each hockey player - buys a pair - of skates - at the arena - before - he - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
15	CC	MM	With cash, - each hockey player - buys a pair - of skates - at the arena - before - she - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
15	NoCC	M	After - each hockey player - buys a pair - of skates - at the arena - with cash, - he - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
15	NoCC	MM	After - each hockey player - buys a pair - of skates - at the arena - with cash, - she - spends time - on the rink - wearing them in.	Are the skates bought from the internet?	n
16	CC	M	For practice, - each little boy - takes a few swings - at the plate - during the baseball game - before - he - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
16	CC	MM	For practice, - each little boy - takes a few swings - at the plate - during the baseball game - before - she - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
16	NoCC	M	After - each little boy - takes a few swings - at the plate - during the baseball game - for practice, - he - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
16	NoCC	MM	After - each little boy - takes a few swings - at the plate - during the baseball game - for practice, - she - gets ready - to aim - and hit the ball.	Are the swings taken at the plate?	y
17	CC	M	With joy, - each father - plans a - family trip - for a weekday - when - he - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
17	CC	MM	With joy, - each father - plans a - family trip - for a weekday - when - she - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
17	NoCC	M	When - each father - plans a - family trip - for a weekday - with joy, - he - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
17	NoCC	MM	When - each father - plans a - family trip - for a weekday - with joy, - she - sets aside - vacation days - from work.	Is the family trip on a weekend?	n
18	CC	M	For pleasure, - each old woman - attends a bingo night - at the rec center - downtown - after - she - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
18	CC	MM	For pleasure, - each old woman - attends a bingo night - at the rec center - downtown - after - he - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
18	NoCC	M	Before - each old woman - attends a bingo night - at the rec center - downtown - for pleasure, - she - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
18	NoCC	MM	Before - each old woman - attends a bingo night - at the rec center - downtown - for pleasure, - he - gets multiple boards - to increase - potential wins.	Is the bingo night downtown?	y
19	CC	M	With care, - each businessman - files a tax return - at the end - of the quarter - before - he - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
19	CC	MM	With care, - each businessman - files a tax return - at the end - of the quarter - before - she - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
19	NoCC	M	After - each businessman - files a tax return - at the end - of the quarter - with care, - he - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n
19	NoCC	MM	After - each businessman - files a tax return - at the end - of the quarter - with care, - she - faces review - and scrutiny - from the IRS.	Are the tax returns filed at the beginning of the quarter?	n

20	CC	M	With joy, - each mother - sets up - a playdate - for the kids - after - she - checks to - make sure the date - is available.	Is the playdate for the kids?	y
20	CC	MM	With joy, - each mother - sets up - a playdate - for the kids - after - he - checks to - make sure the date - is available.	Is the playdate for the kids?	y
20	NoCC	M	Before - each mother - sets up - a playdate - for the kids - with joy, - she - checks to - make sure the date - is available.	Is the playdate for the kids?	y
20	NoCC	MM	Before - each mother - sets up - a playdate - for the kids - with joy, - he - checks to - make sure the date - is available.	Is the playdate for the kids?	y
21	filler	filler	The waiter - said - he was getting - really - annoyed with - the apparent - lack of effort - of the kitchen - staff - tonight.	Were the kitchen staff working hard?	n
22	filler	filler	The firefighter - thought - he was - certainly - going to faint - from exhaustion - as the fire - in the barn - was blazing - so strongly.	Was the firefighter feeling faint?	y
23	filler	filler	The chairman - of the board - solemnly - announced that - he would retire - next - year - because - of chronic - illness.	Was the chairman retiring next week?	n
24	filler	filler	The waiter - flirtatiously - said - he could - offer - Susan - a strong - margarita - for free - that evening.	Did the waiter offer a margarita?	y
25	filler	filler	The landlord - said - he would - happily throw - a welcome party - for new residents - on the block - next Friday - if the weather - held up.	Will the party happen rain or shine?	n
26	filler	filler	The best man - at the wedding - in Boston - last weekend - thought - she was - always - the best dressed - at every - occasion.	Was the wedding in Boston?	y
27	filler	filler	The policeman - said - she could - easily figure out - the details - of the robbery - last night - given the evidence - left behind - at the scene.	Was there a robbery that morning?	n
28	filler	filler	The waiter - at the bar - said - she wore - a big smile - for the whole - time - that evening - at the surprise - birthday party.	Does the waiter work at a bar?	y
29	filler	filler	The boy - said - she was - very - disappointed at - the result - of the game - since - the Eagles - lost again.	Did the Eagles win the game?	n
30	filler	filler	President Obama - said - she would - be given - an opportunity - to ask just - one question - right - after the press - conference finishes.	Did President Obama allow for a question?	y
31	filler	filler	Kate - wished - he would - sing Uptown Funk - and - Gangnam Style - karaoke - to break the ice - at the staff - orientation.	Was Kate at a going away party?	n
32	filler	filler	Susan - commented that - he liked - the mascots - and their - costumes - a lot - but not - the tie-dyed team - uniforms.	Did someone like the mascots?	y
33	filler	filler	The nurse - said - he was - surprised at - how much - the tiny wound - on his finger - hurt the varsity - basketball - athlete.	Was the wound big?	n
34	filler	filler	The saleswoman - said - she loved - the style - and the fit - of the shirt - the model - wore - from this year's - winter collection.	Did the shirt fit properly?	y
35	filler	filler	The girl - said - she marveled at - the large - selection of - fresh fruits - and local vegetables - at the farmers' - market - last Saturday.	Did the girl marvel at the selection of bread?	n
36	filler	filler	The waiter - cleaned up - the spoons - before - he served - the decadent - cheesecake - to the newly - engaged - couple.	Did the waiter serve cheesecake?	y
37	filler	filler	The housemaid - put away - the stacks of - magazines - before - he left - the room - to prepare - for the guests - to arrive.	Did the housemaid put away the dishes?	n
38	filler	filler	The policewoman went - to the coffee shop - for a quick - lunch break - before - she finished - filling out - the document - concerning - the recent crime.	Did the policewoman fill out a document?	y
39	filler	filler	The hairstylist - consulted - a few current - magazines - before - he decided - on a hairstyle - for the upcoming - graduation - ceremonies.	Was the hairstyle chosen right away?	n
40	filler	filler	The landlady - barbecued - huge amounts of - sausage and ham - for her - European tenants - before - she baked - some fresh - banana bread.	Did the landlady bake banana bread?	y
41	practice	practice	This is - a practice - sentence - to get you - used to - reading sentences - like this - during the - experiment.	N/A	N/A
42	practice	practice	The monkey - swung from - tree to tree - through - the jungle - as the - angry lion - pursued him - desperately.	Was the lion following the monkey?	y
43	practice	practice	The motorist - failed to make - a complete stop - prior to - entering - the intersection - which - angered the - pedestrian.	Did the motorist come to a complete stop?	n

Appendix E: Experiment 3 Variant

A variant of Experiment 3 was run using stimuli that lacked the additional adjunct that had been added to the Experiment 3 stimuli to maintain a constant antecedent-pronoun distance between conditions. The QP sentences in Experiment 2, shown in (13a) and (13b) above, are left unchanged in the stimuli for this experiment variant, labeled NoCC in (23). In these sentences, the clausal role manipulation leads to a one-region difference in antecedent-pronoun distance between the NoCC

and CC conditions. The analysis nevertheless yields overall parallel patterns to those reported in Experiment 3. These are summarized below.

(23)

- After $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ he $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*NoCC Match*)
- After $\frac{1}{2}$ each boy $\frac{2}{3}$ fetches a bucket $\frac{3}{4}$ of water $\frac{4}{5}$ from the well $\frac{5}{6}$ she $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*NoCC Mismatch*)
- Each boy $\frac{1}{2}$ fetches a bucket $\frac{2}{3}$ of water $\frac{3}{4}$ from the well $\frac{4}{5}$ before $\frac{5}{6}$ he $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*CC Match*)
- Each boy $\frac{1}{2}$ fetches a bucket $\frac{2}{3}$ of water $\frac{3}{4}$ from the well $\frac{4}{5}$ before $\frac{5}{6}$ she $\frac{6}{7}$ goes $\frac{7}{8}$ to clean the $\frac{8}{9}$ barn and stables. (*CC Mismatch*)

Participants ran: 76

Participants analyzed after removal: 73

Overall comprehension accuracy: 0.92 ($SE = .007$)

	Match	Mismatch
NoCC	.92 (.017)	.89 (.022)
CC	.90 (.015)	.89 (.021)

Table 11: Mean accuracy rates of comprehension question responses (SE)

Experimental trials removed: 36 (2%)

		Region								
		1	2	3	4	5	6	7	8	9
NoCC	Match	460	572	695	658	672	498	526	580	663
	Mismatch	464	552	698	652	663	555	593	628	730
CC	Match	552	693	676	665	516	438	506	560	654
	Mismatch	583	683	661	680	533	469	560	611	704

Table 12: Mean RTs (ms) by region

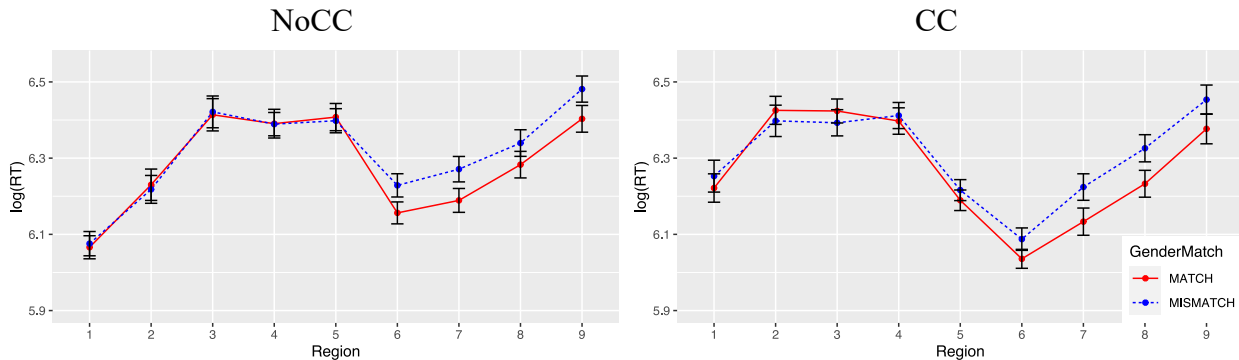


Figure 4: Log-transformed mean RTs by region

Regions 6 used a model with by-participant random slopes for structure and gender match, and a by-item random slope for gender match; for region 7, there was an uncorrelated by-participant random slope for structure and an uncorrelated by-item random slope for gender match; and for region 8, there were by-participant random slopes for structure and gender match as well as their interaction, and an uncorrelated by-item random slope for structure.

We focused on models that included previous region reading times as a predictor.²⁵ The results of the analysis are summarized in Table 13. The effects of previous region reading times were highly significant throughout, but are not reported here in detail.

	Region 6 (pronoun)			Region 7 (spillover)			Region 8 (spillover)		
	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>	Est.	SE	<i>t</i>
Structure Type	-0.049	0.008	-6.086 ***	-0.005	0.009	-0.593	-0.008	0.010	-0.790
Gender Match	-0.030	0.009	-3.355 **	-0.033	0.009	-3.602 ***	-0.023	0.009	-2.642 **
Type × Match	0.007	0.007	0.928	-0.001	0.008	-0.159	-0.007	0.009	-0.752

Table 13: Summary of statistical analysis
. $p < 0.10$; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Region 6: The analysis revealed significant main effects of structure type, with longer reading times in the NoCC condition, and of gender match (in the expected direction). Planned comparisons found significant simple effects of gender match in the NoCC condition (Est. = -0.073, $SE = 0.023$, $t = -3.210$, $p < 0.01$) and in the CC condition (Est. = -0.047, $SE = 0.023$, $t = -2.043$, $p < 0.05$).

Region 7: The analysis revealed a significant main effect of gender match. Planned comparisons found significant simple effects of gender match in the NoCC condition (Est. = -0.064, $SE = 0.025$, $t = -2.561$, $p < 0.05$) and in the CC condition (Est. = -0.069, $SE = 0.025$, $t = -2.764$, $p < 0.01$).

Region 8: The analysis revealed a significant main effect of gender match. Planned comparisons revealed a significant simple effect of gender match in the CC condition (Est. = -0.061, $SE = 0.026$, $t = -2.373$, $p < 0.05$) but not in the NoCC condition (Est. = -0.033, $SE = 0.025$, $t = -1.309$, $p = 0.195$).

²⁵ The results without such a predictor were essentially identical in terms of significance patterns; the only minor divergences were not crucial to our interpretation of the data, namely: (i) in region 7, there was a significant main effect of structure ($p < 0.01$); (ii) in region 8, there was a marginally significant main effect of structure ($p = 0.088$); (iii) in region 8, there was a marginally significant simple effect indicating a GMME in the NoCC condition ($p = 0.057$), as opposed to no effect.