Covariation in processing: grammar vs. context

Nikhil Lakhani & Florian Schwarz
Manuscript, March 2022

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 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 would lead to Gender Mismatch Effects emerging for these types of sentences, too. 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
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, thus not c-commanding the pronoun (italicized; adapted from Barker, 2012):

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

While the referential determiner phrase (DP) in (1) allows a co-referential interpretation of the pronoun (not requiring c-command), the minimal variant with a quantifier phrase (QP) antecedent

---

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.
in (2) does not allow a *covarying* reading. However, Barker (2012) catalogues a variety of both previously known and novel exceptions, where a pronoun receives a covarying interpretation relative to a non-c-commanding QP antecedent, embedded in various environments, including possessive DPs (3a), inverse linking (3b), PPs (3c), and VPs (3d) (from Barker):

(3)
   a. [No1 one’s mother-in-law] fully approves of heri.
   b. [Someone from everyi city] hates iti.
   c. John gave [to eachi participant] a framed picture of hisi mother.

One reaction in the literature, often considered relative to just one specific type of exception, has been to adjust the definition of c-command. Taking the whole of the evidence hecatalogues together, Barker (2012) argues, building on Safir (2004), that there simply is no c-command requirement for binding, but rather 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. A crucial shared feature of these accounts is that they aim to maintain a unified account of the various cases of covarying interpretations. An alternative stance is to say that these exceptions are in fact not cases of *bona-fide* binding, but rather instances of covariation brought about more indirectly by alternative mechanisms whose nature does not require a c-command relationship between QP and pronoun, most prominently involving quantification over situations. While the technical details are complex (see Büring, 2004 for such an account of *binding out of possessors* exemplified by (3a)), and the generality of this move across cases is far from clear, what matters for our purposes is the general configuration of the theoretical landscape and its relation to processing data: for a given pair of cases of covarying pronouns differentiated theoretically by such *two-mechanism* approaches, differences in processing effects can be seen as lending empirical support to the theoretical distinction, which, in contrast to a unified perspective, provides inherent theoretical grounds for explaining such differences. As we’ll elaborate in the general discussion below, the absence (or lack of full generality) of such processing effects, however, does not necessarily argue against two-mechanism theories.

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 (GMME; 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 an attempt at establishing a dependency between the two is made upon encountering the pronoun, which then falters due to the gender mismatch. This perspective positions the presence of GMME’s as a diagnostic for whether the processor is attempting to establish such a dependency, which makes for a highly useful tool for investigating the factors affecting whether or not such attempts are made (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,

---

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 The detailed variations are of no concern for our purposes; see, Barker (2012) for comprehensive references.
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 GMMEs for sentences like (4a), (where every old man c-commands the pronoun) but not (4b) (where it doesn’t, i.e. NoCC), which crucially manifests in a statistical interaction of the gender and c-command factors.

\[(4)\]
\[
a. \text{The surgeon saw that every old man on the emergency ward silently wished that } \{\text{he/she}\} \text{ could go a little bit faster. } (\text{CC (i) Match / (ii) Mismatch})
b. \text{The surgeon who every old man on the emergency ward saw silently wished that } \{\text{he/she}\} \text{ could go a little bit faster. } (\text{NoCC (i) Match / (ii) Mismatch})
\]

Further experiments by Kush et al. (2015) corroborate these findings, and taken together, these results suggest that the processor does not initially 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 QPs 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 constraint at play. Being primarily interested in testing for relational cues, these authors intentionally chose constructions where c-command and the scope constraint make the same predictions, leaving open the possibility that what governs the setting of the ACCESSIBLE feature is a matter of scope. 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 Cunnings et al. (2015) experiment, using stimuli with a QP in sentence-initial temporal adjunct clauses:

\[(5)\]
\[
a. \text{After each boy brought fresh water from the kitchen quickly it seems that } \{\text{(i) he / (ii) she}\} \text{ went on an early break. } (Q\text{P & NoCC (i) Match / (ii) Mismatch})
b. \text{It seems each boy brought fresh water from the kitchen quickly right before } \{\text{(i) he / (ii) she}\} \text{ went on an early break. } (C\text{C (i) Match / (ii) Mismatch})
\]
Crucially, and unlike in Cunnings et al. (2015), the 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 Cunnings et al. and Kush et al. (2015). In a self-paced reading (SPR) task, Moulton and Han (2018) find an initial interaction such that GMMEs were present in the 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 in the NoCC condition. A second SPR experiment by Moulton and Han (Experiment 3) compares exceptionally covarying sentences with QP antecedents like (5a) to identically structured sentences with DP antecedents, as in (6), 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.

(6) 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 (5a)) does not require this structural configuration. With regards to processing, it furthermore needs to be explained why these 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 (5a), 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. Formally, this only involves binding of, and quantification over, situation pronouns, which are not directly associated with the gender features of the “antecedent” noun phrase. Thus, the interaction found by Moulton and Han suggesting an absence of GMMEs in the NoCC condition is explained if we assume that, at least initially, feature match only plays a role in processing for variables in a binding configuration (presumably reflected both in the syntax and the semantics). The semantic misalignment ultimately matters, of course, since the Mismatch version of (5a) 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 if the way it is interpreted never allows for consideration of a reading where this pronoun is related to the relevant QP.

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 (7a), one of the most prominent cases for which situation semantic D-type analyses have been proposed (Heim, 1990; Elbourne, 2005):

(7) English paraphrases 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 (7b) 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 GMME effects under covariation. In theoretical terms, there certainly are differences between the constructions that could have repercussions for processing. While the cases with QPs in temporal adjuncts, as in (5a), are among those that fall under the scope constraint according to Barker (2012), donkey sentences do not (and neither do cases of telescoping). 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 GMME effects 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 mechanism uniformly governs antecedent retrieval in covarying cases, 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 GMME effects 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 GMME effects 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 (8). 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.

(8)

a. After 1/2 each boy 2/3 brought fresh water 3/4 from the kitchen 4/5 quickly 5/6 it seems 6/7 that he 7/8 went 8/9 on an early 9/10 break. *(QP Match)*

b. After 1/2 each boy 2/3 brought fresh water 3/4 from the kitchen 4/5 quickly 5/6 it seems 6/7 that she 7/8 went 8/9 on an early 9/10 break. *(QP Mismatch)*

c. After 1/2 the boy 2/3 brought fresh water 3/4 from the kitchen 4/5 quickly 5/6 it seems 6/7 that he 7/8 went 8/9 on an early 9/10 break. *(DP Match)*

d. After 1/2 the boy 2/3 brought fresh water 3/4 from the kitchen 4/5 quickly 5/6 it seems 6/7 that she 7/8 went 8/9 on an early 9/10 break. *(DP Mismatch)*

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 (8), 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 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 the experiment can be found in (18) 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 a comprehension rate below 70% correct (across all conditions and fillers) and/or an average reading time below 400ms (across all regions)
were removed, as closely as possible reflecting the removal criteria used by Moulton and Han (2018).\(^4\) This eliminated 10 participants, leaving data from 73 participants to analyze.

Mean accuracy on the comprehension questions after removal across all conditions and fillers was 0.91 (SE = .009). Table 1 shows the mean accuracy rate broken down by condition, which shows no major effect on comprehension.

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>QP</td>
<td>.89 (.024)</td>
<td>.89 (.019)</td>
</tr>
<tr>
<td>DP</td>
<td>.92 (.016)</td>
<td>.90 (.019)</td>
</tr>
</tbody>
</table>

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

Individual trials were also removed if any one region’s reading time (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.

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>QP Match</td>
<td>416</td>
<td>543</td>
<td>694</td>
<td>689</td>
<td>630</td>
<td>560</td>
<td>502</td>
<td>585</td>
<td>594</td>
<td>635</td>
</tr>
<tr>
<td>QP Mismatch</td>
<td>439</td>
<td>550</td>
<td>696</td>
<td>687</td>
<td>692</td>
<td>552</td>
<td>496</td>
<td>629</td>
<td>643</td>
<td>647</td>
</tr>
<tr>
<td>DP Match</td>
<td>424</td>
<td>510</td>
<td>647</td>
<td>664</td>
<td>640</td>
<td>544</td>
<td>476</td>
<td>527</td>
<td>592</td>
<td>620</td>
</tr>
<tr>
<td>DP Mismatch</td>
<td>447</td>
<td>508</td>
<td>697</td>
<td>691</td>
<td>675</td>
<td>581</td>
<td>509</td>
<td>632</td>
<td>669</td>
<td>688</td>
</tr>
</tbody>
</table>

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

\(^4\) In their Experiment 2, Moulton and Han (2018) removed two participants with a sub 400ms avg RT, who also had the two lowest comprehension rates, and 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%.

\(^5\) 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 (\(p > 0.10\)) until the critical region.
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 4.1.1). The 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. If this did not converge, the random slope parameter for the interaction for either participants or items was removed for another attempt to fit the model. For regions 7 and 8, the maximal model converged. In region 9, for participants, the random slope for the interaction was removed. Planned comparisons to measure GMMEs per antecedent type were calculated by dummy-coding antecedent type and gender match to find simple effects of gender match. The maximal model method was again utilized with the dummy-coded factors. 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.

<table>
<thead>
<tr>
<th>Region 7 (pronoun)</th>
<th>Region 8 (spillover)</th>
<th>Region 9 (spillover)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Antecedent Type</strong></td>
<td><strong>Gender Match</strong></td>
<td><strong>Type × Match</strong></td>
</tr>
<tr>
<td>Est.</td>
<td>SE</td>
<td>t</td>
</tr>
<tr>
<td>0.001</td>
<td>0.008</td>
<td>0.158</td>
</tr>
<tr>
<td>-0.012</td>
<td>0.008</td>
<td>-1.441</td>
</tr>
<tr>
<td>-0.017</td>
<td>0.008</td>
<td>-2.187 *</td>
</tr>
</tbody>
</table>

Table 3: Summary of statistical analysis for Experiment 1

.p < 0.10; * p < 0.05; ** p < 0.01; *** p < 0.001

**Region 7:** The analysis revealed a significant interaction of gender match and antecedent type. Planned comparisons found a significant simple effect of gender match in the DP condition (Est. = 0.058, SE = 0.023, t = 2.558, p < 0.05) but not in the QP condition (Est. = -0.012, SE = 0.023, t = -0.526, p = 0.599).

**Region 8:** The analysis revealed a significant main effect of gender match and a marginally significant interaction of gender match and antecedent type. Planned comparisons found a significant simple effect of gender match in the DP condition (Est. = 0.131, SE = 0.036, t = 3.598, p < 0.01) but not in the QP condition (Est. = 0.055, SE = 0.034, t = 1.620, p = 0.114).

**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 QP condition (Est. = 0.071, SE = 0.029, t = 2.458, p < 0.05) and DP condition (Est. = 0.108, SE = 0.027, t = 4.057, p < 0.001).

**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. The interaction in region 7 (and marginal interaction in region 8) indicates a difference in the impact of the gender manipulation based on antecedent type. This is further confirmed by the 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 region 9, and the gender manipulation has a significant simple effect there for both DP and QP conditions. This contrasts with Moulton and Han, 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.

While the present results are compatible with Moulton and Han (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 and any potential impact of gender mismatch does not unfold 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, 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 show a significant simple effect of antecedent type in the match condition in region 8 (Est. = 0.073, SE = 0.027, t = 2.695, p < 0.01), due to slower reading times in the QP match condition. 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). We’ll return to this issue in the general discussion. For the moment, the main take-away 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

In light of their finding of immediate GMMEs in donkey sentences, where the explanation of the Moulton and Han (2018) data detailed in Moulton (2017) predicts the absence of GMMEs parallel to the NoCC condition in their work, Kush and Eik (2019) argue for a uniform processing

---

6 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.

7 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 GMME effects for the QP condition, we won’t dwell on this issue here.

8 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).
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 GMMEs, 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 GMMEs 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, (9b) demonstrates the change in felicity of a telescoping interpretation of (9a) when a context is provided to promote a scripted reading.⁹

(9)

a. # Every dog came in. It 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 identified four specific adjustments in this vein:

(10)

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 (10a)–(10c) involved fairly straightforward alterations, (10d) was more open-ended, and involved modifying the content of each clause such that the events in 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

⁹ The examples in (9) are from Poesio and Zucchi (1992).
configuration, but whose salient interpretation had the features above to have the relevant scripted nature. (11) illustrates the resulting variation of the original stimuli in (8).

(11)

a. After 1/2 each boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 he 6/7 goes 7/8 to clean the 8/9 barn and stables. *(QP Match)*

b. After 1/2 each boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 she 6/7 goes 7/8 to clean the 8/9 barn and stables. *(QP Mismatch)*

c. After 1/2 the boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 he 6/7 goes 7/8 to clean the 8/9 barn and stables. *(DP Match)*

d. After 1/2 the boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 she 6/7 goes 7/8 to clean the 8/9 barn and stables. *(DP Mismatch)*

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 (11). 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 (19) 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 eight participants, leaving data from a total of 59 participants.

Across all conditions and fillers, participants answered the comprehension questions with a mean accuracy of about 0.90 (SE = .009). Table 4 shows the mean proportion of correct responses in each condition. There appears to be no major effect on comprehension from the manipulations.

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>QP</td>
<td>.89 (.020)</td>
<td>.88 (.026)</td>
</tr>
<tr>
<td>DP</td>
<td>.91 (.018)</td>
<td>.89 (.023)</td>
</tr>
</tbody>
</table>

Table 4: 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 5 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.

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>QP</td>
<td>Match</td>
<td>443</td>
<td>597</td>
<td>734</td>
<td>683</td>
<td>670</td>
<td>522</td>
<td>528</td>
<td>562</td>
</tr>
<tr>
<td></td>
<td>Mismatch</td>
<td>460</td>
<td>636</td>
<td>736</td>
<td>694</td>
<td>688</td>
<td>537</td>
<td>588</td>
<td>660</td>
</tr>
<tr>
<td>DP</td>
<td>Match</td>
<td>463</td>
<td>569</td>
<td>738</td>
<td>699</td>
<td>709</td>
<td>531</td>
<td>558</td>
<td>623</td>
</tr>
<tr>
<td></td>
<td>Mismatch</td>
<td>463</td>
<td>563</td>
<td>695</td>
<td>688</td>
<td>683</td>
<td>547</td>
<td>616</td>
<td>657</td>
</tr>
</tbody>
</table>

**Table 5:** Mean RTs (ms) by region in Experiment 2

![Log-transformed mean RTs by region in Experiment 2](image)

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. All three regions converged on the maximal model.

The results of the analysis are summarized in Table 6.

<table>
<thead>
<tr>
<th>Region 6 (pronoun)</th>
<th>Region 7 (spillover)</th>
<th>Region 8 (spillover)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Est.</td>
<td>SE</td>
<td>t</td>
</tr>
<tr>
<td>Antecedent Type</td>
<td>0.006</td>
<td>0.011</td>
</tr>
<tr>
<td>Gender Match</td>
<td>-0.010</td>
<td>0.009</td>
</tr>
<tr>
<td>Type × Match</td>
<td>0.002</td>
<td>0.010</td>
</tr>
</tbody>
</table>

**Table 6:** 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 significant main effect of gender match. Planned comparisons found significant simple effects of gender match in the QP condition (Est. = 0.085,
Region 8: The analysis revealed a significant main effect of gender match and a marginally significant interaction of antecedent type and gender, due to a greater effect of gender mismatch in the QP condition. Planned comparisons found significant simple effects of gender match in the QP condition (Est. = 0.159, SE = 0.031, t = 5.205, p < 0.001) and in the DP condition (Est. = 0.063, SE = 0.030, t = 2.098, p < 0.05).

3.5 Discussion
The adjustments implemented for the stimuli from Experiment 1 led to immediate GMMEs 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 marginal 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, and the changes aiming to create a more script-like interpretation, while leaving the syntactic structure as before, were able to induce an interpretation where an anaphoric dependency 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), GMMEs 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. Also note that Cunnings et al. (2015) found GMMEs beginning in the pronoun region in a DP condition in their eye-tracking data, so the variation in when the effects emerge may largely be due to the nature of the self-paced reading paradigm.

Another aspect of the data worth commenting on is the marginal interaction between antecedent type and gender match in region 8, with a greater effect of gender mismatch in the QP condition. However, 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 (Est. = -0.070, SE = 0.036, t = -1.916, p < 0.10) simple effect of antecedent type in the match condition), 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 goes directly counter to the Moulton and Han (2018) predictions of an absence of GMMEs for QP-anteceded pronouns in the NoCC condition.

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). These 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 interact with 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. 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 (12), 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 (12).\[^{10}\]

\[^{10}\] 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.

\begin{align*}
(12) & \\
& a. \text{After} \ 1/2 \text{ each boy} \ 2/3 \text{ fetches a bucket} \ 3/4 \text{ of water} \ 4/5 \text{ from the well} \ 5/6 \text{ on foot, 6/7 he} \ 7/8 \text{ goes} \ 8/9 \text{ to clean the} \ 9/10 \text{ barn and stables. (NoCC Match)} \\
& b. \text{After} \ 1/2 \text{ each boy} \ 2/3 \text{ fetches a bucket} \ 3/4 \text{ of water} \ 4/5 \text{ from the well} \ 5/6 \text{ on foot, 6/7 she} \ 7/8 \text{ goes} \ 8/9 \text{ to clean the} \ 9/10 \text{ barn and stables. (NoCC Mismatch)}
\end{align*}
c. On foot, 1/2 each boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 before 6/7 he 7/8 goes 8/9 to clean the 9/10 barn and stables. (CC Match)
d. On foot, 1/2 each boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 before 6/7 she 7/8 goes 8/9 to clean the 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, 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 (12). 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 (20) 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, 10 participants’ data were removed, leaving a total of 65 participants for analysis.

Across all conditions and fillers, participants answered the comprehension questions with a mean accuracy of 0.93 (SE = 0.006). Table 7 shows the mean proportion of correct responses by condition, with no major effect of condition on comprehension.

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoCC</td>
<td>.92 (.017)</td>
<td>.90 (.017)</td>
</tr>
<tr>
<td>CC</td>
<td>.92 (.015)</td>
<td>.89 (.020)</td>
</tr>
</tbody>
</table>

Table 7: 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 8 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.
Table 8: Mean RTs (ms) by region in Experiment 3

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoCC</td>
<td>487</td>
<td>595</td>
<td>707</td>
<td>652</td>
<td>689</td>
<td>642</td>
<td>500</td>
<td>524</td>
<td>570</td>
<td>654</td>
</tr>
<tr>
<td></td>
<td>476</td>
<td>618</td>
<td>717</td>
<td>677</td>
<td>698</td>
<td>665</td>
<td>543</td>
<td>599</td>
<td>638</td>
<td>708</td>
</tr>
<tr>
<td>Mismatch</td>
<td>553</td>
<td>666</td>
<td>682</td>
<td>659</td>
<td>653</td>
<td>510</td>
<td>432</td>
<td>498</td>
<td>554</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>563</td>
<td>646</td>
<td>723</td>
<td>691</td>
<td>645</td>
<td>530</td>
<td>468</td>
<td>582</td>
<td>606</td>
<td>710</td>
</tr>
</tbody>
</table>

NoCC

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>CC</td>
<td>553</td>
<td>666</td>
<td>682</td>
<td>659</td>
<td>653</td>
<td>510</td>
<td>432</td>
<td>498</td>
<td>554</td>
<td>600</td>
</tr>
<tr>
<td></td>
<td>563</td>
<td>646</td>
<td>723</td>
<td>691</td>
<td>645</td>
<td>530</td>
<td>468</td>
<td>582</td>
<td>606</td>
<td>710</td>
</tr>
</tbody>
</table>

CC

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. Region 7 converged on the maximal model. Region 8 was fit to a model with the random slope for interaction removed for items. In region 9, for participants, the random effect structure only included the random slope for structure type.

Table 9: Summary of statistical analysis for Experiment 3

<table>
<thead>
<tr>
<th>Region 7 (pronoun)</th>
<th>Est.</th>
<th>SE</th>
<th>t</th>
<th>Est.</th>
<th>SE</th>
<th>t</th>
<th>Est.</th>
<th>SE</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure Type</td>
<td>-0.074</td>
<td>0.010</td>
<td>-7.734***</td>
<td>-0.022</td>
<td>0.011</td>
<td>-1.933 .</td>
<td>-0.018</td>
<td>0.010</td>
<td>-1.859 .</td>
</tr>
<tr>
<td>Gender Match</td>
<td>-0.035</td>
<td>0.009</td>
<td>-3.823 ***</td>
<td>-0.065</td>
<td>0.011</td>
<td>-6.038 ***</td>
<td>-0.044</td>
<td>0.009</td>
<td>-4.734 ***</td>
</tr>
<tr>
<td>Type × Match</td>
<td>-0.003</td>
<td>0.008</td>
<td>-0.342</td>
<td>-0.013</td>
<td>0.009</td>
<td>-1.451</td>
<td>-0.002</td>
<td>0.010</td>
<td>-0.211</td>
</tr>
</tbody>
</table>

Table 9: Summary of statistical analysis for Experiment 3

* p < 0.10; ** p < 0.05; *** p < 0.01

Region 7: The analysis revealed significant main effects of structure type, with longer reading times in the NoCC conditions, and of gender match (in the expected direction). Further analysis found significant simple effects of gender match in the NoCC condition (Est. = 0.064, SE = 0.024, t = 2.694, p < 0.01) and in the CC condition (Est. = 0.075, SE = 0.021, t = 3.557, p < 0.001).

Region 8: The analysis revealed a significant main effect of gender match. Further analysis found significant simple effects of gender match in the NoCC condition (0.104, SE = 0.034, t = 3.104, p < 0.01) and in the CC condition (Est. = 0.157, SE = 0.029, t = 5.454, p < 0.001).
Region 9: The analysis revealed a significant main effect of gender match. Further analysis revealed significant simple effects of gender match in the NoCC condition (Est. = 0.085, SE = 0.034, t = 2.525, p < 0.05) and in the CC condition (Est. = 0.095, SE = 0.025, t = 3.794, p < 0.001).

4.4 Discussion
We find GMMEs in both the NoCC and CC conditions across all three regions of interest. In contrast to Moulton and Han’s (2018) Experiment 2, with overall parallel syntactic structures, 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 the 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, there is a combination of no indication of a delay in GMMEs relative to controls (here the CC condition; DP antecedents in Experiment 2), and early significant GMMEs in the NoCC condition. Thus, the presence or absence of c-command between a potential 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 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, despite a lack of statistically significant interaction, there appears to be a numerical difference in the size of the GMMEs, slightly larger in the CC condition, with 84ms, than in the NoCC condition, with 75ms. But in the absence of positive evidence revealing such potential differences, we will proceed to our general discussion taking the present result patterns at face value.

5. General Discussion
Our main question was whether non-c-commanding QP antecedents in temporal adjunct clauses could give rise to GMMEs for pronouns in the main clause of the sentence, once certain features of these sentences made a covarying interpretation more easily accessible. The background for this was the absence of GMMEs in Moulton and Han (2018) for sentences with this general structure, as well as Kush and Eik’s (2019) finding that another class of sentences involving covarying interpretations of pronouns with a non-c-commanding antecedent, namely donkey sentences, did give rise to immediate GMMEs. The latter bore directly on a theoretical interpretation of Moulton and Han’s findings, fleshed out in more detail in Moulton (2017), that tied the presence of GMMEs to the personal pronoun standing in a standard c-command binding relation to its antecedent. On that analysis, the case of the non-c-commanding QP antecedents in their temporal adjunct clauses only involves binding of a situation variable by another operator, and covariation for the pronoun results from the situational dependency of its D-type interpretation in a way that does not (or not immediately) trigger GMMEs. Assuming donkey sentences also involve such situational binding and D-type pronoun interpretations (as their analysis is perhaps the most prominent application of this approach) predicts them to exhibit the same absence of GMMEs. The fact that Kush and Eik find
GMMEs in donkey sentences thus is problematic for the Moulton and Han analysis (at least under the various assumptions above). The new data reported here contribute to this debate by looking at variations of the temporal adjunct clauses from Moulton and Han (2018), which implement suggestions by Kush and Eik (2019) for making the covarying interpretation more easily accessible in this configuration. Extending the cue-based approach from Kush et al. (2015) to allow for an ACCESSIBLE feature whose setting can be affected by contextual considerations relevant to determining the scope of potential antecedent QPs, they predict that changing the Moulton and Han stimuli to have a more scripted interpretation will facilitate the relevant scope interpretation and thus lead to fuller and earlier consideration of the QP as a potential antecedent, which in turn should lead to GMMEs here as well. Our initial experiment replicates one of Moulton and Han’s to validate their original materials as a baseline in our setup. We largely replicate the overall patterns of their findings, with an initial interaction due to GMMEs for DP antecedents but not QP antecedents. However, in the second spillover region, we find an effect of gender for QPs as well. This pattern on its own is compatible with the Moulton and Han analysis, as long as one allows for later effects of gender in the case of covariation due to binding of a situation variable in D-type pronouns. However, Experiment 2, where the stimuli are adapted to yield a more scripted, non-accidental interpretation, yields no such interactions, and GMMEs arise equally quickly for both DP and QP antecedents. Finally, Experiment 3 confirms that the non-c-commanding QP antecedents from the first two studies lead to GMMEs as quickly as c-commanding ones (adapting the second experimental format from Moulton & Han).

Taken together, these new findings are incompatible with proposals that c-commanding QP antecedents are in general privileged in pronoun processing, i.e., exclusively considered upon encountering the pronoun prior to potential antecedents that don’t c-command it. This includes Moulton and Han’s (2018) proposal, spelled out in technical detail in Moulton (2017), where co-variation due to binding of situation variables is claimed to not (or at least not immediately) involve evaluation of gender features, thus not giving rise to GMMEs (or at least not as quickly as in control conditions with co-referential DPs or c-commanding QPs). The contextual manipulation we implemented for sentences with non-c-commanding QP antecedents in temporal adjunct clauses clearly was able to override whatever disadvantages for establishing or considering an anaphoric dependency there were in the original Moulton and Han stimuli. Thus, the structural and semantic properties of these constructions alone do not suffice to induce delay or absence of GMMEs. To an 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 Cunnings 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 findings 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. 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 Experiment 3. Alternative research paradigms (e.g., see Badecker & Straub, 2002)

11 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).
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.

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 their prediction for them, the emerging empirical picture is entirely consistent with the uniform antecedent retrieval mechanism proposed by Kush and Eik. This involves an ACCESSIBLE feature for possible antecedent noun phrases whose setting is sensitive to a variety of factors. Given Kush and Eik's findings for donkey sentences, the indefinite antecedents of donkey pronouns must be accessible right away, while our data support immediate accessibility of QP antecedents in temporal adjuncts. Since the latter, but not the former, are a case of the QP taking scope over the position of the pronoun, our results genuinely broaden our understanding of antecedent retrieval in processing across different cases of exceptional covariation. Future research should extend this approach to further types of exceptionally covarying constructions, such as the various others documented in Barker (2012).

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. Of particular importance in the present context is the question of why we find variation of accessibility of QP antecedents in the temporal adjunct configuration, but not in c-command configurations. At least at a global level, the CC and NoCC sentence variants in Moulton and Han's (2018) study (illustrated in (5) above) essentially convey the same propositional meaning (given the switch between before and after),

suggesting that whatever factors may make the covarying interpretation less accessible in the NoCC condition with the QP in a temporal adjunct are not at play when the QP is in the matrix clause and c-commands the pronoun. Unless some key differences in the way that the information unfolds over the course of the sentence up to the pronoun can be identified that would explain variation in the availability of the relevant scope (or whatever other property decreases accessibility in the NoCC condition), this contrast could be interpreted as support for the notion that c-commanding QP antecedents have a distinguished status of sorts in processing after all; while QP antecedents in temporal adjunct clauses require contextual support to make the immediate establishment of an anaphoric dependency possible in processing, c-commanding antecedents don’t seem to require any particular contextual setup of this sort. (Donkey pronouns will need to be folded into the picture, here, too—we’ll return to these shortly.)

While we aren’t in a position to present a fully fleshed out formal analysis, we offer a more detailed sketch of the changes we implemented, 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 that the quantificational force of the tense operator has 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 (13b) of the meaning of the sentence in (13a) from Moulton (2017), which roughly follows the detailed semantic analysis of temporal adjunct clauses in Artstein (2005):

(13)

---

12 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.
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.he}(s')(s') & \text{after.(s)(s')} & \text{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 situation $s'$ that temporally follows $s$ in which $x$ practices piano (‘M’ here represents a matching function in the sense of Rothstein, 1995). On a situation-based D-Type analysis, he$_s$ 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. (If one makes do without syntactic constraints for binding, this could, of course, also just be represented as ‘$x$’ in the formula; see (17c) below.) On such a one-event reading, this 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:

(14)

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.he}(s')(s') & \text{after.(s)(s')} & \text{M}(x)(s')]]]

With the two universals—quantifying over boys and situations—taking highest scope together, this now becomes a much more natural 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 more natural seeming generalization is 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 are in principle equivalently available for both configurations (changing up the examples slightly to side-step the pronoun binding issue):

(15)

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{serve}(y)(s') & \text{after.(s)(s')} & \text{M}(x)(s')]]]

(16)

a. After each boy comes home, a snack is served
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')} & \text{M}(x)(s')]]]

So, the key claim based on Kush and Eik’s (2019) suggestion is that relative to the alternative reading parallel to (16b), the reading in (14b) is more easily available, compared to the availability of (13b) relative to the alternative reading parallel to (15b). The speculation we offer as a potential

---

13 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 can ignore that for our purposes.
explanation is 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 is subject to further development and investigation, and we hope this is a useful first step towards that.

What can one say from this perspective about any potential role of c-command for facilitating the availability of covarying interpretations independent of contextual support, as we argue is suggested by Moulton and Han’s (2018) data? While there does not seem to be a purely mechanical constraint on considering only potential antecedent noun phrases that c-command the pronoun, c-command does, of course, have a fairly natural connection to scope, in that c-command does generally align with surface scope interpretations, which in turn have been argued to be preferred in processing (Anderson, 2004). Thus, while perhaps semantic scope does not imply c-command, c-command (as typically assessed in the surface structure) does quite generally imply (the easy availability of) corresponding semantic scope. The rapid GMMEs in Moulton and Han’s c-command conditions 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, quantificational noun phrases in temporal adjunct clauses require something extra, whether that’s covert LF-movement or higher-type interpretations of one sort or another, to take scope over the position of the pronouns in the matrix clause. This in and of itself seems to cause some delays in processing (if it happens at all), but such effects can be ameliorated if there are other aspects of the sentence’s interpretation that promote the availability of the relevant scope, as in our variant of the stimuli. The broader prediction that follows from this, which is worth investigating, is that the presence of GMMEs is going to correlate with the availability of relevant scopes (as independently measured), and testing this more generally seems like a formidable challenge to be taken up in future work.

Returning briefly to the question of how donkey sentences fit into this picture, one important point is that they differ from the cases at hand in that the indefinite antecedent noun phrase does not take scope over the position of the pronoun. The pronoun is either bound by a higher quantifier over situations (on D-Type accounts), or dynamically bound by a special dynamic quantifier that can bind beyond its scope. While it’s not completely clear whether or not binding in donkey sentences could ever run into issues of insufficient availability of covarying interpretations due to contextual factors (as Kush & Eik’s, 2019 Norwegian stimuli are hard to assess in this regard), it wouldn’t be surprising if they aren’t, on the perspective outlined above, since they are not dependent on scope configurations in the same way.

The final question—in many ways the biggest and most important, at least from a theoretical perspective—is what the current state of the empirical picture means for the grammatical mechanisms underlying covariation. The results and their interpretation from Moulton and Han (2018) owed their intrigue to the notion that there was processing evidence supporting the two mechanisms of covariation view, with a privileged role for c-commanding antecedents in processing. 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 a 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. This, of course, aligns well with Barker's (2012) proposal that a scope constraint is all that is needed in terms of 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:

\[
(17)
\]

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

There does not seem to be any principled reason forcing us to expect that one of these should be linked to a slower or fundamentally different cognitive process in comprehension. Thus, we see no general reason to favor either a one-mechanism or two-mechanism theory based on the reading time data considered here. What should be acknowledged, though, is the intricate link between underlying structures, and in particular, c-command relationships in the hierarchical structure of the surface string, and the availability of scope interpretations aligning with them. As we stressed above, the availability of covariation with a c-commanding QP does not seem to interact with subtle interpretive properties when it comes to GMMEs in the way it does for cases of non-c-commanding QPs in temporal adjunct clauses. Thus, there is an indirect role for c-command in processing covariation, as it facilitates certain scope interpretations. Beyond that, perhaps the 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. 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 GMMEs 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.

---

14 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.
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.

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

Authors' contributions
[Omitted for peer review.]

References


Appendix A: Experiment Demos

(18) Experiment 1: https://farm.pcbex.net/r/WRsDZO/
(19) Experiment 2: https://farm.pcbex.net/r/BooSRB/
(20) Experiment 3: https://farm.pcbex.net/r/ngjyma/

Appendix B: Experiment 1 Materials

1 QP MM 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 M 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 he – 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 she – 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

25
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

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

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

When the groomsman was on the quick 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

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

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

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

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

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

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

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

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

When the little girl was at the park playing with friends from school it appears that he practiced playing the piano for half an hour. Was the ice cream chocolate-flavored? N

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

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

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

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

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

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

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 house down the road? Y

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

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
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 nearly.

Were the shelves in the storage room? N

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 nearly.

Were the shelves in the storage room? N

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 nearly.

Were the shelves in the storage room? N

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 nearly.

Were the shelves in the storage room? N

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

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

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

Was the ice at the rink initially fresh? Y

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

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

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

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

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

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

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

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

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

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

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

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

Did the doctor say there was a tragedy? N


Did the water come from the kitchen? Y


Did the water come from the kitchen? Y


Did the water come from the kitchen? Y


Did the water come from the kitchen? Y

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

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


Was the chairman retiring next week? N


Did the waiter offer a margarita? Y

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
Appendix C: Experiment 2 Materials

| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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 |
| 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.

4. DP MM After the policeman - investigates - an incident and - interviews witnesses - she - files a - police report - detailing the event.

5. QP M Before - each woman - chooses a perfume - to wear to - the banquet - she samples several different scents - at the store.

5. QP MM Before - each woman - chooses a perfume - to wear to - the banquet - he samples several different scents - at the store.

5. DP M Before - each woman - chooses a perfume - to wear to - the banquet - he samples several different scents - at the store.

5. DP MM Before - the woman - chooses a perfume - to wear to - the banquet - he samples several different scents - at the store.

6. QP M After - each fireman - retrieves an axe - and gas mask - from the firetruck - he - searches the scene for - trapped victims.

6. QP MM After - each fireman - retrieves an axe - and gas mask - from the firetruck - she - searches the scene for - trapped victims.

6. DP M After - the fireman - retrieves an axe - and gas mask - from the firetruck - he - searches the scene for - trapped victims.

6. DP MM After - the fireman - retrieves an axe - and gas mask - from the firetruck - she - searches the scene for - trapped victims.

7. QP M When - each waitress - prepares a table - for customers - at the restaurant - she polishes the - nice silverware - meticulously.

7. QP MM When - each waitress - prepares a table - for customers - at the restaurant - he polishes the - nice silverware - meticulously.

7. DP M When - the waitress - prepares a table - for customers - at the restaurant - she polishes the - nice silverware - meticulously.

7. DP MM When - the waitress - prepares a table - for customers - at the restaurant - she polishes the - nice silverware - meticulously.

8. QP M When - each groomsman - stands up - to deliver a speech - on stage - he raises a toast to the bride and groom.

8. QP MM When - each groomsman - stands up - to deliver a speech - on stage - she raises a toast to the bride and groom.

8. DP M When - the groomsman - stands up - to deliver a speech - on stage - he raises a toast to the bride and groom.

8. DP MM When - the groomsman - stands up - to deliver a speech - on stage - she raises a toast to the bride and groom.

9. QP M After - each salesman - makes a sale - while traveling - on the job - he returns to - the office to - report the sale.

9. QP MM After - each salesman - makes a sale - while traveling - on the job - she returns to - the office to - report the sale.

9. DP M After - the salesman - makes a sale - while traveling - on the job - he returns to - the office to - report the sale.

9. DP MM After - the salesman - makes a sale - while traveling - on the job - she returns to - the office to - report the sale.

10. QP M When - each actress - recites a line - while preparing - for the new play - she holds the - script nearby - for reference.

10. QP MM When - each actress - recites a line - while preparing - for the new play - she holds the - script nearby - for reference.

10. DP M When - the actress - recites a line - while preparing - for the new play - she holds the - script nearby - for reference.

10. DP MM When - the actress - recites a line - while preparing - for the new play - she holds the - script nearby - for reference.

11. QP M After - each bridesmaid - picks out - a dress from - the bridal boutique - she - texts a picture to - the bride - for approval.

11. QP MM After - each bridesmaid - picks out - a dress from - the bridal boutique - he - texts a picture to - the bride - for approval.

11. DP M After - the bridesmaid - picks out - a dress from - the bridal boutique - she - texts a picture to - the bride - for approval.

11. DP MM After - the bridesmaid - picks out - a dress from - the bridal boutique - he - texts a picture to - the bride - for approval.

12. QP M When - each little girl - is given a word to spell at - the spelling bee - she is told the - definition - and etymology.

12. QP MM When - each little girl - is given a word to spell at - the spelling bee - he is told the - definition - and etymology.

12. DP M When - the little girl - is given a word to spell at - the spelling bee - she - is told the - definition - and etymology.

12. DP MM When - the little girl - is given a word to spell at - the spelling bee - he - is told the - definition - and etymology.

13. QP M After - each plumber - installs a pipe - in the bathroom - of the new house - he inspects the - joints with - a special tool.

Does the police report detail the incident? y

Does the police report detail the incident? y

Is the perfume for a business meeting? n

Is the perfume for a business meeting? n

Is the perfume for a business meeting? n

Is the perfume for a business meeting? n

Does the axe come from the firetruck? y

Does the axe come from the firetruck? y

Does the axe come from the firetruck? y

Does the axe come from the firetruck? y

Is the table for hotel patrons? n

Is the table for hotel patrons? n

Is the table for hotel patrons? n

Is the table for hotel patrons? n

Is the toast to the bride and groom? y

Is the toast to the bride and groom? y

Is the toast to the bride and groom? y

Is the toast to the bride and groom? y

Is the sale made while at the office? n

Is the sale made while at the office? n

Is the sale made while at the office? n

Is the sale made while at the office? n

Is the perfume for a business meeting? y

Is the perfume for a business meeting? y

Is the perfume for a business meeting? y

Is the perfume for a business meeting? y

Is the etymology of the word provided? y

Is the etymology of the word provided? y

Is the etymology of the word provided? y

Is the etymology of the word provided? y

Is the pipe installed in the kitchen? n
After each plumber - installs a pipe - in the bathroom - of the new house - she - inspects the - joints with - a special tool.

After each secretary - types up a memo - for the employees - to read - she - emails the boss - with relevant - important information.

After each secretary - types up a memo - for the employees - to read - he - emails the boss - with relevant - important information.

After the secretary - types up a memo - for the employees - to read - she - emails the boss - with relevant - important information.

After each hockey player - buys a pair - of skates - at the arena - he - spends time - on the rink - wearing them in.

After each hockey player - buys a pair - of skates - at the arena - she - spends time - on the rink - wearing them in.

Is the pipe installed in the kitchen? Is the pipe installed in the kitchen? Is the pipe installed in the kitchen? Is the pipe installed in the kitchen?

Is the memo for employees? Is the memo for employees? Is the memo for employees? Is the memo for employees?

Is the memo for employees? Is the memo for employees? Is the memo for employees? Is the memo for employees?

Is the memo for employees? Is the memo for employees? Is the memo for employees? Is the memo for employees?

Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet?

Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet?

Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet?

Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet?

Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet? Are the skates bought from the internet?

Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend?

Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend?

Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend?

Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend?

Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend?

Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend? Is the family trip on a weekend?

Are the family staff working hard? Were the family staff working hard? Were the family staff working hard? Were the family staff working hard?

Are the family staff working hard? Were the family staff working hard? Were the family staff working hard? Were the family staff working hard?

Are the family staff working hard? Were the family staff working hard? Were the family staff working hard? Were the family staff working hard?

Are the family staff working hard? Were the family staff working hard? Were the family staff working hard? Were the family staff working hard?
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 contain milk?

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 contain milk?

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 contain milk?

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 contain milk?

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?

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?

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?

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?

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?

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?

3. NoCC  M  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?

3. NoCC  MM  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?
Thoroughly, - each policeman - investigates - an incident and - interviews witnesses - before - he - files a - police report - detailing the event.

Thoroughly, - each policeman - investigates - an incident and - interviews witnesses - before - she - files a - police report - detailing the event.

After - each policeman - investigates - an incident and - interviews witnesses - thoroughly, - he - files a - police report - detailing the event.

After - each policeman - investigates - an incident and - interviews witnesses - thoroughly, - she - files a - police report - detailing the event.

At last, - each woman - chooses a perfume - to wear - to the banquet - after she - samples several - different scents - at the store.

At last, - each woman - chooses a perfume - to wear - to the banquet - after she - samples several - different scents - at the store.

Before - each woman - chooses a perfume - to wear - to the banquet - at last, - she - samples several - different scents - at the store.

Before - each woman - chooses a perfume - to wear - to the banquet - at last, - she - samples several - different scents - at the store.

With strength, - each fireman - retrieves an axe - and gas mask - from the firetruck - before - he - searches the - scene for - trapped victims.

With strength, - each fireman - retrieves an axe - and gas mask - from the firetruck - before - she - searches the - scene for - trapped victims.

After - each fireman - retrieves an axe - and gas mask - from the firetruck - with strength, - he - searches the - scene for - trapped victims.

After - each fireman - retrieves an axe - and gas mask - from the firetruck - with strength, - she - searches the - scene for - trapped victims.

At mealtime, - each waitress - prepares a table - for customers - at the restaurant - after - she - polishes the - nice silverware - meticulously.

At mealtime, - each waitress - prepares a table - for customers - at the restaurant - after - she - polishes the - nice silverware - meticulously.

Before - each waitress - prepares a table - for customers - at the restaurant - at mealtime, - she - polishes the - nice silverware - meticulously.

Before - each waitress - prepares a table - for customers - at the restaurant - at mealtime, - she - polishes the - nice silverware - meticulously.

One-by-one, - each groomsmen - stands up - to deliver a speech - on stage - before - he - raises a - toast to the bride and groom.

One-by-one, - each groomsmen - stands up - to deliver a speech - on stage - before - she - raises a - toast to the bride and groom.

After - each groomsmen - stands up - to deliver a speech - on stage - one-by-one, - he - raises a - toast to the bride and groom.

After - each groomsmen - stands up - to deliver a speech - on stage - one-by-one, - she - raises a - toast to the bride and groom.

Through effort, - each salesman - makes a sale - while traveling - on the job - before - he - returns to - the office to - report the sale.

Through effort, - each salesman - makes a sale - while traveling - on the job - before - she - returns to - the office to - report the sale.

After - each salesman - makes a sale - while traveling - on the job - through effort, - he - returns to - the office to - report the sale.

After - each salesman - makes a sale - while traveling - on the job - through effort, - she - returns to - the office to - report the sale.

With grace, - each actress - recites a line - to prepare for the new play - while - she - holds the - script neatly - for reference.

With grace, - each actress - recites a line - to prepare for the new play - while - he - holds the - script neatly - for reference.

While - each actress - recites a line - to prepare for the new play - with grace, - she - holds the - script neatly - for reference.

While - each actress - recites a line - to prepare for the new play - with grace, - he - holds the - script neatly - for reference.

With care, - each bridesmaid - picks out - a dress from - the bridal boutique - before - she - texts a picture - to the bride - for approval.

With care, - each bridesmaid - picks out - a dress from - the bridal boutique - before - he - texts a picture - to the bride - for approval.

After - each bridesmaid - picks out - a dress from - the bridal boutique - with care, - she - texts a picture - to the bride - for approval.

After - each bridesmaid - picks out - a dress from - the bridal boutique - with care, - he - texts a picture - to the bride - for approval.

At first, - each little girl - is given a - word to spell at - the spelling bee - before - she - is told the definition and etymology.

At first, - each little girl - is given a - word to spell at - the spelling bee - before - he - is told the definition and etymology.

Does the police report detail the incident?

Does the police report detail the incident?

Is the perfume for a business meeting?

Is the perfume for a business meeting?

Is the axe come from the firetruck?

Is the axe come from the firetruck?

Is the table for hotel patrons?

Is the table for hotel patrons?

Is the toast to the bride and groom?

Is the toast to the bride and groom?

Does the sale made while at the office?

Does the sale made while at the office?

Is the picture get emailed to the bride?

Is the picture get emailed to the bride?

Is the etymology of the word provided?

Is the etymology of the word provided?
After each little girl - is given a - word to spell at - the spelling bee - at first, - he - is told - the definition - and etymology.

With a wrench, - each plumber - installs a pipe - in the bathroom - of the new house - before - he - inspects the - joints with - a special tool.

With a wrench, - each plumber - installs a pipe - in the bathroom - of the new house - before - she - inspects the - joints with - a special tool.

After each plumber - installs a pipe - in the bathroom - of the new house - with a wrench, - he - installs the - joints with - a special tool.

After each plumber - installs a pipe - in the bathroom - of the new house - with a wrench, - she - inspects the - joints with - a special tool.

With speed, - each secretary - types up a memo - for the employees - to read - before - she - emails the boss - with relevant - important information.

With speed, - each secretary - types up a memo - for the employees - to read - before - he - emails the boss - with relevant - important information.

After each secretary - types up a memo - for the employees - to read - with speed, - she - emails the boss - with relevant - important information.

After each secretary - types up a memo - for the employees - to read - with speed, - he - emails the boss - with relevant - important information.

With cash, - each hockey player - buys a pair - of skates - at the arena - before - he - spends time - on the rink - wearing them in.

With cash, - each hockey player - buys a pair - of skates - at the arena - before - she - spends time - on the rink - wearing them in.

After each hockey player - buys a pair - of skates - at the arena - with cash, - he - spends time - on the rink - wearing them in.

After each hockey player - buys a pair - of skates - at the arena - with cash, - she - spends time - on the rink - wearing them in.

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.

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.

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.

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.

With joy, - each father - plans a - family trip - for a weekday - when - he - sets aside - vacation days - from work.

With joy, - each father - plans a - family trip - for a weekday - when - she - sets aside - vacation days - from work.

When each father - plans a - family trip - for a weekday - with joy, - he - sets aside - vacation days - from work.

When each father - plans a - family trip - for a weekday - with joy, - she - sets aside - vacation days - from work.

For pleasure, - each old woman - attends a bingo night - at the rec center - downtown - after - she - gets multiple boards - to increase - potential wins.

For pleasure, - each old woman - attends a bingo night - at the rec center - downtown - after - he - gets multiple boards - to increase - potential wins.

Before each old woman - attends a bingo night - at the rec center - downtown - for pleasure, - she - gets multiple boards - to increase - potential wins.

Before each old woman - attends a bingo night - at the rec center - downtown - for pleasure, - he - gets multiple boards - to increase - potential wins.

With care, - each businessman - files a tax return - at the end - of the quarter - before - he - faces review - and scrutiny from the IRS.

With care, - each businessman - files a tax return - at the end - of the quarter - before - she - faces review - and scrutiny from the IRS.

After each businessman - files a tax return - at the end - of the quarter - with care, - he - faces review - and scrutiny from the IRS.

After each businessman - files a tax return - at the end - of the quarter - with care, - she - faces review - and scrutiny from the IRS.

With joy, - each mother - sets up - a playdate - for the kids - after - she - checks to - make sure the date - is available.

With joy, - each mother - sets up - a playdate - for the kids - after - he - checks to - make sure the date - is available.

Is the etymology of the word provided?  y
Is the pipe installed in the kitchen?  n
Is the pipe installed in the kitchen?  n
Is the pipe installed in the kitchen?  n
Is the pipe installed in the kitchen?  n
Is the memo for employees?  y
Is the memo for employees?  y
Is the memo for employees?  y
Are the skates bought from the internet?  n
Are the skates bought from the internet?  n
Are the skates bought from the internet?  n
Are the skates bought from the internet?  n
Are the swings taken at the plate?  y
Are the swings taken at the plate?  y
Are the swings taken at the plate?  y
Is the family trip on a weekend?  n
Is the family trip on a weekend?  n
Is the family trip on a weekend?  n
Is the family trip on a weekend?  n
Is the bingo night downtown?  y
Is the bingo night downtown?  y
Is the bingo night downtown?  y
Is the bingo night downtown?  y
Are the tax returns filed at the beginning of the quarter?  n
Are the tax returns filed at the beginning of the quarter?  n
Are the tax returns filed at the beginning of the quarter?  n
Are the tax returns filed at the beginning of the quarter?  n
Is the playdate for the kids?  y
Is the playdate for the kids?  y
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 (11a) and (11b) above, are left unchanged in the stimuli for this experiment variant, labeled NoCC in (21). 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.

(21)
a. After 1/2 each boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 he 6/7 goes 7/8 to clean the 8/9 barn and stables. *(NoCC Match)*

b. After 1/2 each boy 2/3 fetches a bucket 3/4 of water 4/5 from the well 5/6 she 6/7 goes 7/8 to clean the 8/9 barn and stables. *(NoCC Mismatch)*

c. Each boy 1/2 fetches a bucket 2/3 of water 3/4 from the well 4/5 before 5/6 he 6/7 goes 7/8 to clean the 8/9 barn and stables. *(CC Match)*

d. Each boy 1/2 fetches a bucket 2/3 of water 3/4 from the well 4/5 before 5/6 she 6/7 goes 7/8 to clean the 8/9 barn and stables. *(CC Mismatch)*

Participants ran: 76
Participants analyzed after removal: 70

Overall comprehension accuracy: 0.92 (SE = .007)

<table>
<thead>
<tr>
<th></th>
<th>Match</th>
<th>Mismatch</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoCC Match</td>
<td>0.92</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Mismatch</td>
<td></td>
<td>0.89</td>
</tr>
<tr>
<td>CC Match</td>
<td>0.91</td>
<td>(0.015)</td>
</tr>
<tr>
<td>Mismatch</td>
<td></td>
<td>0.89</td>
</tr>
</tbody>
</table>

*Table 10: Mean accuracy rates of comprehension question responses (SE)*

Experimental trials removed: 36 (3%)

<table>
<thead>
<tr>
<th>Region</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
</tr>
</thead>
<tbody>
<tr>
<td>NoCC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>468</td>
<td>586</td>
<td>710</td>
<td>673</td>
<td>684</td>
<td>509</td>
<td>535</td>
<td>592</td>
<td>675</td>
</tr>
<tr>
<td>Mismatch</td>
<td>472</td>
<td>564</td>
<td>716</td>
<td>661</td>
<td>673</td>
<td>563</td>
<td>605</td>
<td>642</td>
<td>742</td>
</tr>
<tr>
<td>CC</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Match</td>
<td>563</td>
<td>710</td>
<td>690</td>
<td>677</td>
<td>525</td>
<td>444</td>
<td>517</td>
<td>570</td>
<td>668</td>
</tr>
<tr>
<td>Mismatch</td>
<td>596</td>
<td>695</td>
<td>677</td>
<td>694</td>
<td>541</td>
<td>477</td>
<td>571</td>
<td>626</td>
<td>719</td>
</tr>
</tbody>
</table>

*Table 11: Mean RTs (ms) by region*

*Figure 4: Log-transformed mean RTs by region*

Regions 6 and 8 converged on the maximal model.
Region 7 was fit to a model with the random slope for interaction removed for items.
Region 6: The analysis revealed significant main effects of structure type, with longer reading times in the NoCC conditions, and of gender match (in the expected direction). Further analysis found a marginally significant simple effect of gender match in the NoCC condition (Est. = 0.056, SE = 0.032, t = 1.746, p < 0.10) and a significant simple effect of gender match in the CC condition (Est. = 0.057, SE = 0.021, t = 2.716, p < 0.01).

Region 7: The analysis revealed significant main effects of structure type and gender match. Further analysis found significant simple effects of gender match in the NoCC condition (Est. = 0.092, SE = 0.029, t = 3.240, p < 0.01) and in the CC condition (Est. = 0.082, SE = 0.027, t = 3.040, p < 0.01).

Region 8: The analysis revealed a significant main effect of gender match. Further analysis revealed significant simple effects of gender match in the NoCC condition (Est. = 0.061, SE = 0.028, t = 2.165, p < 0.05) and in the CC condition (Est. = 0.097, SE = 0.034, t = 2.850, p < 0.01).

<table>
<thead>
<tr>
<th>Region 6 (pronom)</th>
<th>Region 7 (spillover)</th>
<th>Region 8 (spillover)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Structure Type</strong></td>
<td><strong>Gender Match</strong></td>
<td><strong>Type × Match</strong></td>
</tr>
<tr>
<td><strong>Est.</strong></td>
<td><strong>SE</strong></td>
<td><strong>t</strong></td>
</tr>
<tr>
<td>-0.068</td>
<td>0.010</td>
<td>-6.598</td>
</tr>
<tr>
<td>-0.028</td>
<td>0.010</td>
<td>-2.897</td>
</tr>
<tr>
<td>-0.000</td>
<td>0.010</td>
<td>-0.017</td>
</tr>
</tbody>
</table>

Table 12: Summary of statistical analysis

. p < 0.10; * p < 0.05; ** p < 0.01; *** p < 0.001