C-Command vs. Scope: An Experimental Assessment of Bound Variable Pronouns

Keir Moulton & Chung-hye Han
Simon Fraser University

To appear in *Language*

Abstract While there are very clearly some structural constraints on pronoun interpretation, debate remains as to their extent and proper formulation (Bruening 2014). Since Reinhart 1976 it is commonly reported that bound variable pronouns are subject to a C-Command requirement. This claim is not universally agreed on and has been recently challenged by Barker 2012, who argues that bound pronouns must merely fall in the semantic scope of a binding quantifier. In the processing literature, recent results have been advanced in support of C-Command (Kush et al. 2015, Cunnings et al. 2015). However, none of these studies separates semantic scope from structural C-Command. In this article, we present two self-paced reading studies and one off-line judgment task which show that when we put both C-Commanding and non-Commanding quantifiers on equal footing in their ability to scope over a pronoun, we nonetheless find a processing difference between the two. Semantically legitimate, but non-C-Commanded bound variables do not behave like C-Commanded bound variables in their search for an antecedent. The results establish that C-Command, not scope alone, is relevant for the processing of bound variables. We then explore how these results, combined with other experimental findings, support a view in which the grammar distinguishes between C-Commanded and non-C-Commanded variable pronouns, the latter perhaps being disguised definite descriptions (Cooper 1979, Evan 1980, Heim 1990, Elbourne 2005). *

Keywords: C-Command, Scope, Bound variable/Co-varying Pronouns, Gender Mismatch Effects, Self-paced Reading

*We are extremely indebted to the two anonymous reviewers and Grant Goodall, an Associate Editor of *Language*, for their insightful comments that were crucial in improving this paper. We also thank the members of the Experimental Syntax Lab at Simon Fraser University for their assistance in running the studies reported here. This work was partially supported by SSHRC 430-2014-01034 to Moulton and SSHRC 435-2014-0161 to Han.

1. Does C-Command matter? Pronouns in natural language can have a referential interpretation, as in (1a), or an interpretation that *co-varies* as a function of an antecedent quantifier, as in (1b).[1]

(1) a. *The dog* ate *his* bone. referential pronoun
b. *No/Each dog* ate *his* bone. co-varying pronoun

In one tradition, it is claimed that there is a structural constraint on co-varying pronouns. Attributed to Reinhart 1976, the claim is that co-varying pronouns must be C-Commanded by their quantificational binders.[2] Contrasts like (2) are often cited in support of this: only the C-Commanded pronoun in (2a) gets a co-varying reading.

(2) a. [*No dog* was fed [before Sal fed *it*]]. ✓co-varying
b. [[Bob fed *no dog*] but [Sal fed *it*]]. ✗co-varying

Recent experimental work has explored whether C-Command plays a role in on-line processing (Kush 2013, Kush, Lidz, and Phillips 2015, and Cunnings, Patterson, and Felser 2015). Kush et al. tested sentences in which the quantifier *any janitor* C-Commands the pronoun in an adjunct *when*-clause (3a) and those in which the antecedent does not C-Command the pronoun in a co-ordinated *but* clause (3b). These were compared to non-quantificational antecedents, because referential pronouns do not require C-command by their antecedents (3c–d).

(3) Kush et al. stimuli (Exp 1a,1b,1c)

a. Kathi didn't think *any janitor* liked performing his custodial duties when *he* had to clean up messes left after prom.
b. Kathi didn't think *any janitor* liked performing his custodial duties, but *he* had to clean up messes left after prom anyway.
c. Kathi didn't think *the janitor* liked performing his custodial duties when *he* had to clean up messes left after prom.
d. Kathi didn't think *the janitor* liked performing his custodial duties, but *he* had to clean up messes left after prom anyway.

In eye-movement data, Kush et al. 2015 found that at the post-pronoun region, the C-Command manipulation had no effect on the referential pronouns (3c–d), but did affect the variable pronouns (3a–b). Sentences with a C-Commanding quantified antecedent were read more quickly than those with a non-C-Commanding quantified antecedent, suggesting that participants easily accessed the C-Commanding quantificational antecedent, but not the non-C-Commanding one. Further experiments in Kush et al. 2015 and Cunnings et al. 2015 showed that the processor accesses structurally appropriate (i.e. C-Commanding) quantificational antecedents at very early points in processing but not structurally inappropriate antecedents. Building on Kush 2013, Kush et al. 2015 propose an antecedent retrieval mechanism in the processor that deactivates structurally inappropriate antecedents, rendering them inaccessible in working memory.

The structures employed by Kush et al. and Cunnings et al., however, do not allow the pronouns to fall in the scope of the antecedent quantifier, a necessary condition for co-varying pronouns. Scope is a semantic notion: one quantifier *out-scopes* or *scopes over* a second quantifier if the second quantifier is interpreted as part of the proposition to which the out-scoping quantifier is attached. Sentences containing two quantifiers often allow either quantifier to out-scope the other. In (4) the universal quantifier may scope over the indefinite numeral *two toys*, which is compatible with a scenario in which each dog played with a distinct pair of toys. Or (4) can covey that there are the same two toys that every dog plays with. In that case the numeral out-scopes the universal subject.

(4) Every dog played with two toys.
  i. *every* > 2: For every dog x [there are 2 toys that x ate]
  ii 2 > *every*: There are 2 toys Y [for every dog x, x played with Y]

What's relevant here is that bound variable pronouns must fall in the scope of their binding quantifiers. We corroborate this by showing that quantifier scope and bound variable interpretation travel together: the quantifier in (5), *every dog*, must put the quantified phrase that contains the pronoun in its scope if there is to be a bound variable interpretation for the pronoun.

(5) Every dog$_i$ played with two of his$_i$ toys.
  every > 2
  ∗2 > *every his*

The DP *two of his toys* in (5), when construed with a bound variable interpretation, must take narrow scope with respect to the universally quantified subject. Scope, then, is a necessary condition for bound variable interpretations.

The scopal properties of natural language quantifiers is constrained, however. For instance, the negative polarity item (NPI) *any janitor* that is used in Kush et al.'s materials (3b) cannot scope beyond its licencing negation, and therefore cannot scope into the *but*-clause. The NPI does scope into the *when*-clause (5a), which it also C-Commands. As in many cases, here C-Command and quantifier scope coincide. But this is not always the case. Barker 2012 documents many cases in which a quantifier merely scopes over a co-varying pronoun but does not C-Command it. None of the quantified antecedents in (6), a selection of examples gathered by Barker 2012, C-Command the pronoun, but all take scope over the pronoun, and that is sufficient for co-variation.

(6) a. One page in *every book* had something written on *it*.
  b. We [will sell *no wine*] before *its* time.
  c. [After unthreading *each screw*], but before removing *it*, make sure to hold the screw in place while seperating [sic] the screw from the driver.
  d. . . . [after seeing *each animal*] but before categorizing *it* on the computer or recording it on their response sheet.
  e. . . . [after fetching *each pointer*], but before dereferencing *it* . . .

f. The grade [that *each student* receives] is recorded in *his* file.
(Barker 2012)

Barker argues that a C-Command requirement should be entirely abandoned since there are so many cases where the C-Command requirement does not hold for pronominal co-variation. C-Command is at best a redundant requirement in a subset of cases; all that is required for co-variation is that the pronoun fall in the scope of the quantifier.

The results in Kush et al. and Cunnings et al. are equally compatible with Barker's claim: scopally illicit antecedents are likewise rendered inaccessible to the processor. This raises the question, then, about the processing of scope-compliant pronouns: those that fall in the scope, but not the C-Command domain, of a quantifier.

In this article, we test such configurations with stimuli modelled after sentences Barker collected in which the non-C-Commanding quantifier is housed in a clause-initial temporal adjunct, as in (6c–e). We establish through off-line measures in a forced-choice comprehension task study that these pronouns indeed are readily interpreted in the scope of the quantifier and therefore as co-varying. We then show in two self-paced reading studies, however, that the on-line processing profile of such pronouns is different from those that meet both the scope *and* C-Command requirement. In particular, they fail to exhibit the same kind of effects that diagnose the search for an appropriate antecedent that c-commanded pronouns do. In this respect their processing resembles that of the non-C-Commanded pronouns reported in Kush et al. 2015 and Cunnings et al. 2015. This confirms that C-Command, not merely semantic scope, is relevant to the processing of co-varying pronouns.

We then discuss how a processing-based explanation, along the lines of the retrieval mechanisms proposed by Kush 2013 and Kush et al. 2015, can account for the results. We contrast this interpretation with one that posits, contra Barker 2012, that pronouns are ambiguous between standard individual-denoting variables which are subject to a C-Command requirement and pronouns that can co-vary without C-Command in virtue of a different mechanism, perhaps as disguised definite descriptions (Cooper 1979, Evans 1980, Heim 1990, Elbourne 2005). We suggest that our results, taken together with previous experimental results (Carminati, Frazier, and Rayner 2002, Anderssen 2011) support a *grammatical* distinction among pronouns.

In the next section, we review in more detail the previous experimental work on the processing of co-varying pronouns.

2. Scope vs. C-Command: Previous Experimental Work

Carminati et al. 2002 tested intra- (=C-Commanded, CC) vs. cross-clausal (=non-C-Commanded, NoCC) pronominal interpretation with stimuli such as (7).

(7) Carminati et al. 2002 stimuli

a. *Every British soldier* thought that *he* killed an enemy soldier. *CC QP*
b. *The British soldier* thought that *he* killed an enemy soldier. *CC DP*
c. *Every British soldier* aimed and then *he* killed an enemy soldier. *NoCC QP*
d. *The British soldier* aimed and then *he* killed an enemy soldier. *NoCC DP*

The eye-movement data showed no difference between CC and NoCC conditions. There were only re-reading differences between quantified and referential conditions, with the former taking longer. The lack of any penalty for NoCC co-varying pronouns suggested to Carminati et al. that such interpretations are readily available and that there is no special status for quantifier-variable configurations that obey C-Command.

Kush 2013 points out, however, that the NoCC configurations that Carminati et al. employed are instances of *telescoping* (Roberts 1987, Poesio and Zucchi 1992). Telescoping is the name for phenomena that involve semantic binding across a sentence boundary. Such cases clearly do not obey C-Command, and they have a special semantic status. For instance, not all quantifiers are able to telescope. Negative quantifiers, for instance, resist telescoping (8b).

(8) a. *No British soldier* thought that *he* killed an enemy. *CC*
b. **No British soldier* picked up his rifle. Then *he* killed a soldier. *NoCC (telescoping)*

Kush 2013 and Kush et al. 2015 argue that telescoping does not give rise to truly bound variable interpretations and so they block the possibility of telescoping by using NPIs as quantified antecedents in their stimuli (3b). As discussed above, they found that at the post-pronoun region, the C-Command manipulation had no effect on the referential conditions, but did affect the quantificational cases, such that C-Commanded pronouns showed greater ease of processing. The CC QP sentences were read more quickly than the NoCC QP sentences, suggesting that participants easily accessed the C-Commanding quantificational antecedent, but not the non-C-Commanding one.

Cunnings et al. 2015 report similar findings, using a gender mismatch paradigm (Garnham 2001, van Gompel and Liversedge 2003, Sturt 2003). Their materials placed the quantifier either in a position that C-Commands the pronoun, as in (9a,b), or inside a relative clause modifying the subject, a position that does not C-Command the pronoun in the matrix clause (NoCC), as in (9c,d). They crossed this manipulation with whether the pronouns matched the QP in gender (*Match*) or not (*Mismatch*).

(9) Cunnings et al. 2015 (Exp. 2)

a. Being in hospital can be quite difficult at times.
The surgeon saw that *every old man* on the emergency ward silently wished that *he* could go a little bit faster. *CC Match*

b. Being in hospital can be quite difficult at times.
The surgeon saw that *every old woman* on the emergency ward silently wished that *he* could go a little bit faster. *CC Mismatch*

c. Being in hospital can be quite difficult at times.
The surgeon who *every old man* on the emergency ward saw silently wished that *he* could go a little bit faster. *NoCC Match*

d. Being in hospital can be quite difficult at times.
The surgeon who *every old woman* on the emergency ward saw silently wished that *he* could go a little bit faster. *NoCC Mismatch*

Processing difficulty at gender mismatching pronouns reflects an attempt to establish a

dependency (van Gompel and Liversedge 2003, Sturt 2003, Garnham 2001, Kazanina, Lieberman, Yoshida, and Phillips 2007). In their eye-movement data, Cunnings et al. found an interaction such that there was a gender mismatch effect (GMME) in the C-Command conditions but not in the non-C-Commanding conditions. They conclude that the processor attempts to retrieve a C-Commanding QP upon encountering the pronoun, but no such attempt is made if the QP does not C-Command the pronoun. (In a further experiment, Cunnings et al. controlled for the fact that the QP is in a less salient position in NoCC than CC, something we also do in Experiment Three.)

The results of Kush et al. 2015 and Cunnings et al. 2015 are important, chiefly because they show on-line sensitivity to structural factors that regulate pronominal co-variation. From a processing perspective, the results are important for understanding how antecedents are retrieved in memory. One popular model of memory retrieval in sentence processing involves a cue-based, content addressable memory (McElree 2000, McElree, Foraker, and Dyer 2003, Lewis, Vasishth, and Van Dyke 2006). Items in a content addressable memory are tagged with cues, like gender or number. Retrieval involves matching a pronoun's features with a set of retrieval cues: the best match activates that item in memory, allowing retrieval. Features like gender and number are natural cues, easily modeled in a content addressable memory as they are, in a sense, intrinsic properties of a noun. In contrast, a relational constraint like C-Command is only calculated with respect to some (not-yet-encountered) pronoun and so cannot be encoded easily as a cue in memory. Kush 2013 and Kush et al. 2015 propose a mechanism that resolves this tension, which we return to in Section 6.

The results of Kush et al. 2015 and Cunnings et al. 2015 do not, however, actually determine whether the operative relational constraint is C-Command or quantifier scope. This is because, as noted, the position of the quantifier in the NoCC conditions in both sets of stimuli is not one from which the quantifier can semantically *scope* over the pronoun. In Kush et al.'s stimuli, the QP is an NPI, which cannot scope wide enough. In Cunnings et al.'s stimuli, the QP is inside a relative clause, which constitutes a scope island in most cases. In an off-line sentence judgement task experiment, Kush et al. found that in sentences such as (3b), participants clearly preferred the single-individual interpretation of the pronoun *he* to the co-varying interpretation, and Cunnings et al. found that in sentences such as (9c), participants significantly preferred the DP *the surgeon* to the non-C-Commanding QP *every old man* as an antecedent of *he*.

To untangle scope from C-Command, we will compare cases like those Barker 2012 found where a QP that does not C-Command can easily scope over a pronoun and thus allow it co-vary with the QP. To that end, we have chosen to test QPs in left-adjoined temporal clauses, as in (10).[3]

(10) After each boy brought fresh water from the kitchen, it seems he went on an early break.

Like Barker, and unlike Kush et al. and Cunnings et al., we have chosen to use the wide-scoping *each*. *Each* is notorious for taking wide scope (Vendler 1962, Kroch 1974, Ioup 1975, Beghelli and Stowell 1997, Tunstall 1998, Brasoveanu and Dotlačil 2015 among many others), even from temporal adjunct clauses (see Artstein 2005 and

references therein). Universal QPs like *every NP* or *all the NP* do not so readily take wide scope. This is what drives the felicity contrast between (11a) and (11b). In the first case, where the universal cannot scope widest, the resulting anomalous meaning is detected very quickly (i.e. the odd situation in which parents come up with a name after all children are born). With *each*, however, the wide scope reading is easy and immediate: there are separate naming-events for each child.

(11) a. #After all the children were born, the parents came up with a good name.
b. After each child was born, the parents came up with a good name.

The next set shows a contrast, but coming from the opposite direction.

(12) a. After all the disciples spoke to Jesus, he died.
b. #After each disciple spoke to Jesus, he died.

To our ears, (12b) suggests multiple deaths, unlike (12a). So *each*, unlike *all*, prefers to scope over the entire root clause, and it appears that this interpretation is pursued readily. We call these stimuli *scope compliant, No C-Command*.

3. Experiment One We begin by establishing that our No C-command (NoCC) stimuli—unlike the NoCC stimuli of Kush et al. 2015 and Cunnings et al. 2015—readily allow for co-varying interpretations for non-C-Commanded pronouns. We conducted a forced-choice comprehension study to establish the availability of co-varying interpretations and thus the ease with which the QP can semantically scope over the pronoun.

3.1. Materials Twenty sentence pairs as in (13) were constructed in the following way (all materials appear appended to this document). The NoCC sentence (13b) began with a sentence-initial temporal adjunct clause (headed either by *when, after, before*) of which the subject was the antecedent. The matrix clause began with an expletive construction (*it seems, it appears*) which embedded a finite clause whose subject was the target pronoun. The CC sentence (13a) began with the expletive construction selecting a clause containing a QP; the target pronoun was the subject of a temporal adjunct clause (*when, after, before*) adjoined below the QP subject. Sentences were followed by a forced-choice comprehension question to determine the interpretation of the pronoun by asking whether one individual performed the action or whether all the individuals quantified over did so. A choice of the second indicates a co-varying interpretation of the pronoun.

(13) Scope Compliant Stimuli
a. It seems *each boy* brought fresh water from the kitchen quickly right before *he* went on an early break. *CC*
Who went on an early break?
One person.
Each of the boys.

b. After *each boy* brought fresh water from the kitchen quickly it seems that *he* went on an early break. *NoCC*
Who went on an early break?
One person.
Each of the boys.

We compared these with the quantified sentences from Kush et al. 2015—in which the QP does not semantically scope over the pronoun in the NoCC condition—and asked the same forced-choice question, as in (14).[4]

(14) Kush et al. 2015 stimuli

a. Kathi didn't think any janitor enjoyed performing his custodial duties when he had to clean up the messes left after prom.
Who had to clean up messes? *CC*
One person.
Janitors.

b. Kathi didn't think any janitor enjoyed performing his custodial duties, but he had to clean up the messes left after prom anyway. *NoCC*
Who cleaned up messes?
One person.
Janitors.

We predict that the scope compliant stimuli—regardless of structural manipulation—will allow co-varying interpretations. We also predict that we will replicate the off-line results of Kush et al. 2015 for their materials. Overall, then, we predict an interaction, such that there is an effect of C-Command in Kush et al.'s stimuli, but not in the scope compliant stimuli.

Forty-two filler sentences were included to gauge participants' general sensitivity to co-varying pronouns in cases established already in the experimental and theoretical literature. There were five types of filler sentences: (15a): standard forward QP-binding into complement clauses, which we call Good Co-variation;[5] (15b): strong crossover violations where binding is straightforwardly unacceptable; (15c): ungrammatical telescoping based on the results and stimuli of Anderssen 2011; (15d): cases of Binding out of DP and successful telescoping which are thought to be instances of D-type pronouns (Büring 2004, Elbourne 2005, Anderssen 2011); (15e): cases in which the pronoun is biased to a referential interpretation.

(15) a. Good Co-variation
Each doctor said that he was going to leave the clinic.

b. Strong Crossover (*co-variation)
Sally knows that he says nice things about each fireman.

c. Bad Telescoping (*co-variation)
Each student in the syntax class was accused of cheating on the exam and he has a Ph.D. in astrophysics.

d. Good D-type interpretation/Telescoping

The wife of each guy admitted he watched a lot of sports on TV and read no books.

e. Biased to Referential Pronoun
I met the surgeon who helped every patient in town and he was nice.

3.2. Participants and Procedure Two lists were created so that no participant saw any one item set in both its CC and NoCC instantiation, but all participants read all filler items. Each list was uniquely randomized per participant. The materials were prepared using TurkTools (Erlewine and Kotek 2016). Twenty-four participants completed the questionnaire via AmazonTurk. They received $1.25 USD. All 24 participants identified themselves as native English speakers.

3.3. Results The mean proportion of co-varying responses is given in Figure 1.

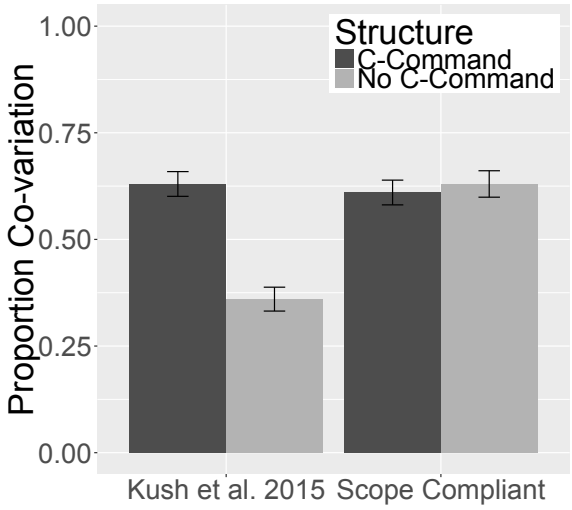

Figure 1: Mean proportion of co-varying responses and standard error

The response data were analyzed by means of a logistic mixed effects model in R (R Development Core Team 2012). Models were fit using the lme4 package (Bates 2005) and the lmerTest package was used to obtain *p*-values (Kuznetsova, Brockho, and Christensen 2014). In analyses of data obtained from all experiments reported in this paper, we first attempted to fit a maximal random effects structure with random intercepts and random slopes for participants and items (Barr, Levy, Scheepers, and Tily 2013). If that model did not converge, we fit a model just like the maximal model, but with the random correlation parameter for the interaction term removed for both participants and items. Moreover, the predictors in all analyses reported in this paper were sum coded, with one of the levels coded as 1, and the other coded as -1.

We fit a mixed model to the response data with a fixed factor of Structure (CC vs. NoCC) and a between-items fixed factor of Item Type, with the levels Kush et al. 2015 and. Scope Compliant. We found a main effect of Item Type (Est.=-0.45,

SE=0.20, z=-2.27, p<.05), such that the scope compliant conditions had a higher rate of co-variation response than the Kush et al. 2015 condition, regardless of Structure, and a main effect of Structure (Est.=0.38, SE=0.12, z= 3.09, p<.01), such that the C-Command condition had a higher rate of co-variation response than the No C-Command condition, regardless of Item Type. And crucially, we found an interaction between the two factors (Est.=0.41, SE=0.13, z=3.27, p<.01). According to planned comparisons, this interaction was due to the fact that in the Kush et al. 2015 stimuli, the C-Command condition had significantly higher rate of co-variation response than the No C-Command condition (Est.=0.78, SE=0.16, z= 4.80, p<.001), but the two scope compliant conditions were equally likely to be given a co-varying response (Est.=-0.08, SE=0.17, z= -0.50, p=.62).

3.4. Discussion The scope compliant NoCC sentences allow co-variation to the same extent as both CC conditions, and significantly more than Kush et al.'s NoCC stimuli. This confirms the intuitions offered in Barker 2012—that quantifier scope, not C-Command, is all that is required for a pronoun to have a co-varying interpretation. It also fits with the results of Carminati et al. 2002, where the lack of C-Command does not attenuate the availability of a co-varying pronoun. The validity of these results is supported by those of the filler sentences given in Table 1. The proportions of co-varying responses are as expected with high proportions for Good Co-variation and Good D-type sentences and low proportions for Strong Crossover, Bad Telescoping and Biased Referential sentences. It is true that the proportion of co-varying responses in the experimental items (61–63%) is somewhat lower than the co-varying responses for the filler items that allow binding (73–75%). We think that this is due to the complexity of the experimental items as compared to the filler items.

Good Co-variation	Strong Crossover	Bad Telescoping	Good D-type	Biased Referential
.73 (.026)	.09 (.02)	.28 (.03)	.75 (.04)	.06 (.02)

Table 1: Proportion of Co-varying Responses (SE) for Fillers

We now turn to the self-paced reading studies, which in contrast to the off-line judgments reveal an on-line effect of C-Command.

4. Experiment Two Following the GMME design of Cunnings et al. 2015, we crossed two two-level factors using the scope compliant stimuli tested in Experiment One: Structure (C-Command (CC) vs. No-command (NoCC)) and Gender (Match vs. Mismatch). The sentences were identical to those tested in Experiment One, modulo the Gender manipulation of the pronoun.

(16) a. It seems $_{1}/_{2}$ *each boy* $_{2}/_{3}$ brought fresh water $_{3}/_{4}$ from the kitchen $_{4}/_{5}$ quickly $_{5}/_{6}$ right $_{6}/_{7}$ before *he* $_{7}/_{8}$ went $_{8}/_{9}$ on an early $_{9}/_{10}$ break. *CC Match*

b. It seems $_{1}/_{2}$ *each boy* $_{2}/_{3}$ brought fresh water $_{3}/_{4}$ from the kitchen $_{4}/_{5}$ quickly $_{5}/_{6}$ right $_{6}/_{7}$ before *she* $_{7}/_{8}$ went $_{8}/_{9}$ on an early $_{9}/_{10}$ break. *CC Mismatch*

c. 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. *NoCC Match*
d. 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. *NoCC Mismatch*

The region of interest, Region 7, invariably contained a subordinator *before/after/when/that* followed by the critical pronoun. If scope, not C-Command, is the relevant constraint, and since all cases are scope compliant, we do not expect an interaction between structure and gender. That is, we expect gender-mismatch pronouns to incur a processing cost in all cases. On the other hand, a difference between NoCC and CC in regards to gender-mismatch effects would support a role for C-Command.

4.1. Materials, Participants and Procedure Twenty item sets like (16) were created, where the antecedent quantifier phrase contains a noun phrase with either definitional gender (*man, woman, boy*) in 17 items or a stereo-typical gender (*hockey player, secretary*) in three items.[6]

The items were distributed over four lists in a Latin Square design so that no participant saw any one item in more than one condition. In addition, each list contained a set of 36 fillers. The sentences were presented in PsychoPy (Peirce 2007) in a uniquely generated random order for each participant. We used the moving-window paradigm in our SPR experiment (Just, Carpenter, and Woolley 1982). At the beginning of each trial, participants see, on a computer screen, several groupings of dashes, each representing a word to be read, separated by spaces. When the participant presses on the space bar, the first grouping of dashes is replaced with the words they represent. The next space bar press reverts these words back to dashes, and replaces the next grouping of dashes with the words they represent. Each experimental sentence was divided into ten regions, and participants could advance to the next region only by a space bar press. No region could be displayed more than once. After each experimental sentence was read, a comprehension question was presented, which could be answered by pressing a key for 'yes' and another key for 'no.' The comprehension questions tested participants' understanding of the sentence but not for their interpretation of the pronoun.

61 native-English participants from the university community completed the experiment, receiving either $10 or course credit.

4.2. Results Two participants were eliminated whose average reading speeds per region was below 400ms, leaving 59 participants. These eliminated participants also had the two lowest comprehension question scores overall, averaging 62%. The grand mean comprehension question scores, including the scores on fillers, was 80%.

Comprehension question response accuracy The mean correct responses for the comprehension questions are reported in Table 2. Although the comprehension questions did not themselves test for co-varying interpretations (they merely tested participants' attention to the sentence content), the results show no impact of the

manipulated factors on comprehension generally. No participant fell below 75% accuracy.

	Match	Mismatch
CC	.87 (.006)	.82 (.007)
NoCC	.85 (.006)	.82 (.007)

Table 2: Proportion of Correct Responses (SE)

Reading times Raw reading times for each region are reported in Table 3. These represent reading times for all data, whether the comprehension question was answered correctly or not. The graphs in Figure 2 represent log reading times (log(RTs)) for the CC conditions (left panel) and the NoCC conditions (right panel) for all data.

		Region									
		1	2	3	4	5	6	7	8	9	10
CC	Match	708	908	1008	878	805	586	582	749	763	912
	Mismatch	704	882	1033	915	799	598	705	841	845	919
NoCC	Match	590	869	1018	976	830	715	607	742	804	857
	Mismatch	648	892	1056	940	823	707	646	760	828	836

Table 3: Average Raw Reading Times (ms)

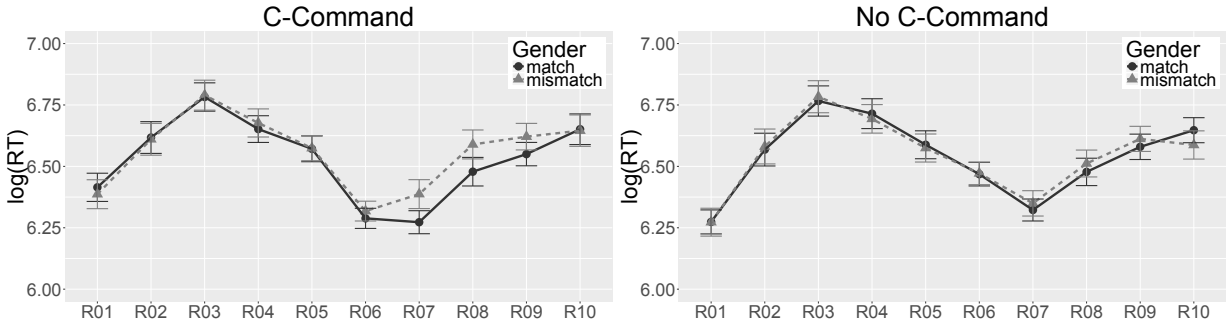

Figure 2: Log Reading Times

We analyzed each region's log reading times (log(RTs)) with a mixed model, with a random effects structure as described for Experiment One.

Region 7 (critical region) In analyzing the reading times in Region 7, there was a main effect of Gender (Est.=-0.04, SE=0.01, df=84.53, t=-3.16, p<.01) and an interaction of Gender with Structure (Est.=-0.02, SE=0.01, df=37.77, t=-2.01, p=.05). According to planned comparisons, this interaction was driven by the fact that while the mismatch condition had significantly longer reading time than the match condition in the C-Command (Est.=0.11, SE=0.03, df=297.06, t=3.86, p<.001), the two conditions had similar reading times in the No C-Command (Est.=-0.01, SE=0.01, df=57.62, t=-0.95, p=.35).

Region 8 (spillover region) The analysis showed a main effect of Gender (Est.=-0.04, SE=0.01, df=28.93, t=-3.22, p<.01), such that overall, the mismatch condition showed significantly slower reading time than the match condition. However, planned comparisons revealed that the main effect of gender was mainly driven by the two C-Command conditions: while the mismatch condition took significantly longer to read than the match condition in C-Command (Est.=-0.06, SE=0.02, df=55.68, t=-3.52, p<.001), the two No C-Command conditions showed no difference in reading time (Est.=-0.02, SE=0.02, df=14.80, t=-1.01, p=.33).

Region 9 In the analysis of data in Region 9, we again found a main effect of Gender (Est.=-0.03, SE=0.01, df=39.80, t=-2.26, p<.05) (Est.=-0.03, SE=0.01, df=39.80, t=-2.26, p<.05). However, as in Region 8, planned comparisons revealed that the main effect of gender was mainly driven by the two C-Command conditions: while the mismatch condition had significantly longer reading time than the match condition in C-Command (Est.=-0.04, SE=0.01, df=50.91, t=-2.39, p<.05), the two No C-Command conditions showed no difference in reading time (Est.=-0.02, SE=0.01, df=450.10, t=-1.11, p=.27).

4.3. Discussion The results show a GMME in the CC conditions but not in the NoCC conditions in the critical region. Moreover, the main effect of Gender evidenced in subsequent regions is driven by the CC conditions, and not the NoCC conditions. The interaction—and lack of GMME in NoCC cases—indicates that even when quantifiers can easily take scope over a pronoun, C-Command is a significant factor in how an antecedent is accessed in on-line processing.

We can imagine one interpretation that is consistent with scope being relevant: the lack of GMME at the pronoun in the NoCC conditions could be due to the fact that participants have not computed wide scope for the quantifier by the time they encounter the pronoun. If that were the case, then they may not attempt to link the pronoun to the quantified antecedent, and no GMME would be expected. While this interpretation certainly deserves further investigation, there are a number of reasons to reject it. Our off-line data in Experiment One reported in Section 3 suggested no difficulty in interpreting the QP with wide scope. We might have expected that if a wide scope interpretation of the quantifier is somehow 'delayed' from a clause-initial adjunct there would be effects even in off-line acceptability and interpretation. It has been argued that wide scoping is a costly processing option, and that this is evidenced in off-line responses (Kurtzman and MacDonald 1993, Tunstall 1998, Anderson 2004 among many others, but cf. Brasoveanu and Dotlačil 2015,b). However, the high proportion of bound variable responses in our off-line data (Experiment One) suggests no such difficulty when *each* takes scope from a clause-initial adjunct. Furthermore, Dotlačil and Brasoveanu 2015 found that re-analysis to give quantified temporal adjuncts wide scope is quite rapid: they found facilitation effects for the plural 'the children' in the second sentence of (17), suggesting that by this point readers had successfully computed a scope which put the quantifier in the first sentence above the object *a child*.

(17) A caregiver comforted a child every night. The caregivers wanted the children to get some rest. (Dotlačil and Brasoveanu 2015)

While the time-course of processing quantifiers remains an active question, we would like to point out that in our stimuli the quantifier linearly *precedes* its bindee, which puts it on better footing to take wide scope than the cases tested in (17).

While we found no GMME in the NoCC conditions, the experiment also found no particular processing difficulty for NoCC pronouns. Post-hoc pairwise comparisons using Tukey between CC-Match and NoCC-Match did not show any difference in Region 7 ($p>.5$). If the NoCC pronoun initiated a search for an antecedent but failed to find one, we might have expected to find an unheralded pronoun effect (Filik, Sandford, and Leuthold 2008, Nieuwland 2014), but we did not.[7] In comparison, in Kush et al.'s Experiment 1c, non-C-commanded pronouns showed a processing difficulty (in post pronoun region). In the experiments that used a gender mismatch paradigm in Kush et al. 2015 a matching C-Commanding antecedent *facilitated* reading times and the raw reading times of the non-C-Commanded pronouns were elevated in comparison.[8] Those were eye-tracking studies, however, and the self-paced reading methodology may obscure finer processing differences among NoCC pronouns. Nonetheless, the lack of any apparent penalty for the (matched) NoCC pronouns squares with the off-line results, which showed that readers readily accepted the bound variable interpretation. This does raise the question of how to reconcile these results with the fact that the lack of GMMEs suggest that the processor did *not* seek NoCC antecedents. We take this up in the General Discussion.

5. Experiment Three Experiment Three is designed to eliminate the possibility that the results of Experiment Two are tied to the fact that the antecedent in the NoCC cases is within an adjunct clause, while it is part of a matrix clause in the CC case. It is possible that this renders the antecedent in NoCC conditions less salient or prominent, and therefore not a target for anaphora resolution. In that case, the absence of a GMME would not be related to C-Command. This same question arises in the experiments conducted by Cunnings, Patterson, and Felser 2014 and Cunnings et al. 2015, where potential antecedents were part of a relative clause and therefore possibly less salient or prominent (Gordon and Hendrick 1998, Foraker and McElree 2007). We follow Cunnings et al.'s method for ruling out this possible interpretation of our results. We compared our NoCC stimuli containing quantificational (QP) antecedents with Referential (DP) antecedents in the same position. There is no constraint that requires referential DPs to C-Command co-referring pronouns (e.g. *This picture of the count makes him look weird* vs. **This picture of each count makes him look weird*). Consequently, if all we found in Experiment Two were an effect of non-prominence, this should carry over to DPs. The stimuli are the result of crossing Antecedent Type with Gender of the pronoun (Match vs. Mismatch) to probe GMME. This resulted in the four conditions in (18).

(18) 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*

If the lack of GMMEs in the scope compliant NoCC sentences in Experiment Two were due to the antecedent not being in a position salient enough to be a candidate for anaphora resolution, then we expect that even referential (DP) antecedents as in (18c,d) will fail to exhibit GMMEs. In contrast, if the lack of GMME in scope compliant cases is unique to quantification and co-variation, then we expect an interaction, such that DP conditions will be sensitive to gender mismatch, but QP conditions will not.

5.1. Materials, Participants and Procedure Twenty item sets like (18) were created and distributed over four lists in a Latin Square design. In addition, each list contained a set of 36 fillers. The sentences were presented in PsychoPy, following the same procedure as Experiment Two. Sixty native-English participants from the university community completed the experiment, receiving either $10 or course credit. None of them had participated in Experiment Two.

5.2. Results

Comprehension question response accuracy The comprehension question accuracy is given in Table 4. The range of participant means was 70%–100%. We excluded no participants in the analysis.

	Match	Mismatch
QP	.85 (.006)	.76 (.005)
DP	.93 (.005)	.78 (.007)

Table 4: Proportion of Correct Responses (SE)

Reading times Raw reading times for each region are reported in Table 5. These represent reading times for all data, whether the comprehension question was answered correctly or not.

The graphs in Figure 3 represent log(RTs) for the DP conditions (left panel) and the QP conditions (right panel) for all data.

As in the analyses performed in Experiment Two, here we analyzed each region's log(RTs) with a mixed model, with a random effects structure as described for Experiment One.

Region 7 (critical region) The analysis did not reveal any effect.

Region 8 (spillover region) The analysis of all data (correct and incorrect on the comprehension question) revealed a significant main effect of Gender (Est.=-0.06,

		Region									
		1	2	3	4	5	6	7	8	9	10
QP	Match	539	807	1091	980	872	707	600	762	802	883
	Mismatch	526	894	1041	897	858	677	618	818	843	929
DP	Match	554	732	1032	1114	888	724	595	695	779	870
	Mismatch	587	793	1035	911	855	710	646	885	840	938

Table 5: Average Raw Reading Times (ms)

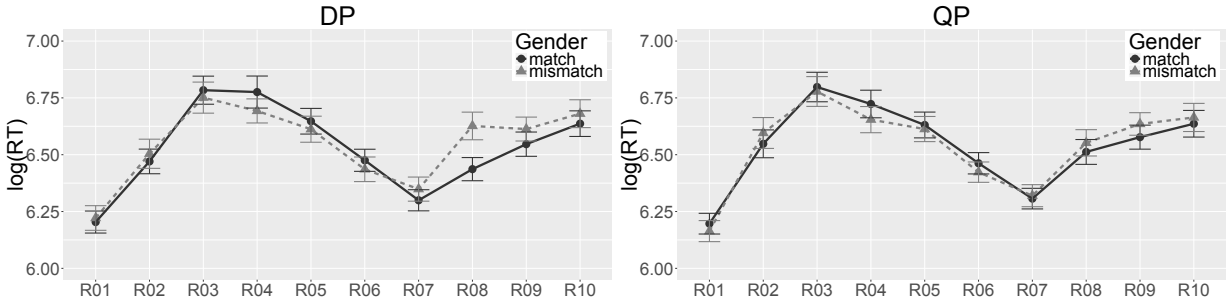

Figure 3: Log Reading Times

SE=0.02, df=19.00, t=-3.01, p<.01) and a significant interaction between Gender and Antecedent Type (Est.=.04, SE=0.01, df=39.42, t=3.09, p<.01). Planned comparisons found that the interaction is due to the mismatch condition having significantly longer reading time than the match condition with a DP antecedent (Est.=-0.10, SE=0.02, df=21.25, t=-4.34, p<.001). In contrast, the two QP conditions exhibited similar reading times (Est.=-0.02, SE=0.02, df=17.34, t=-1.03, p=.32).

Region 9 The only significant effect in Region 9 was a main effect of Gender (Est.=-0.03, SE=0.01, df=24.34, t=-2.50, p<.05), such that overall the mismatch condition had significantly longer reading time than the match condition. However, planned comparisons revealed that this main effect was mainly due to the two DP conditions, which showed a significant difference in reading times (Est.=-0.03, SE=0.01, df=13.63, t=-2.24, p<.05). The two QP conditions did not exhibit a significant difference in reading times (Est.=-0.03, SE=0.02, df=43.13, t=-1.76, p=.09).

5.3. Discussion We found a GMME when a referential antecedent does not C-Command a pronoun, but not with quantificational antecedents in the same position. The results thus confirm that GMMEs are not triggered by scope compliant NoCC pronouns, and that this lack of effect is not due to the fact that the quantificational antecedent is in a less-salient position in a preposed adjunct clause: referential antecedents in the same position trigger GMMEs. We can maintain our conclusion from Experiment One that C-Command, not scope, regulates the processing availability of quantified antecedents.

The GMME for the referential cases was found in the post-pronoun region, unlike in Experiment Two where the GMME effect of a C-Commanding quantifier registered

at the gender mismatching pronoun itself. We might speculate that this difference is due to the fact that with DPs, but not QPs, the anaphoric dependency can be one of co-reference, and not binding. While the interpretation of a bound pronoun comes for "free", so to speak, in some frameworks (Reuland 2011), a co-referent interpretation requires a possibly more costly operation of finding a referent in a discourse model, which could delay gender-mismatch effects. However, we are cautious in putting too much stock in the location of the effect in the spillover region for two reasons. First, Cunnings et al. 2015 found GMMEs on the pronoun region itself even when the antecedent was a DP and not a QP. Theirs was an eye-tracking study, unlike ours, which may account for the difference. Additionally, the DP antecedent in their sentences was within a relative clause while ours was in a preposed temporal adjunct: there may be differences in the retrieval of antecedents from these types of clauses. Second, Cunnings et al. 2014 did not find any evidence that variable binding is computed before coreference assignment, so it is not obvious that such an interpretation is generally available.

6. General Discussion We began by showing in Experiment One that our experimental stimuli, modelled after Barker 2012, allowed a quantifier to easily scope over a pronoun it did not C-Command. These stimuli were shown to be different from from the ones in previous experimental work (Kush et al. 2015, Cunnings et al. 2015), which did not allow the quantifier to semantically scope over the pronoun. Our stimuli, therefore, allowed us to tease apart the role of C-Command from semantic scope. In two self-paced reading experiments (Experiments Two and Three) we found gender mismatch effects only in C-Command configurations and not in the non-C-Command configurations, suggesting that even when a quantifier scopes over a pronoun and the co-varying interpretation is available C-Command is a factor. The results confirm that C-Command, independently of semantic scope, plays a role in the on-line processing of co-varying pronouns. As we outlined above, the results in Kush 2013, Kush et al. 2015, and Cunnings et al. 2013 did not distinguish between scope and C-Command as their materials confounded the two. We turn next to the implications of our results for both grammar and processing, and how they serve to refine the processing mechanisms proposed by Kush et al. 2015.

6.1. Grammar or Processing The question remains, however, whether our results show that C-Command is relevant for processing alone. An alternative option, discussed in the semantics literature, is that English has two type of pronouns: standard bound variable pronouns that must meet the C-Command requirement and the scope requirement, and a set of homophonous pronouns, variously called E-type, D-type or donkey pronouns, that must merely meet the scope requirement. Barker 2012 has argued against this division among pronouns and claims that co-varying pronouns are uniform in their syntax and semantics and must meet only the scope requirement. It has been suggested to us that the apparent discrepancy between the off-line and on-line results show that only the processing mechanism responsible for antecedent retrieval is sensitive to C-Command, but the grammar of the pronouns is uniform.[9]

That is, participants reach a co-varying interpretation for the NoCC pronoun, but at a point that is later than the point at which a GMME could register. An alternative is that the grammar of English distinguishes between co-varying pronouns that require C-Command and those that do not, and the processing profiles we found reflect this. We discuss these two possibilities in turn below.

6.2. A PROCESSING (ONLY) EXPLANATION Kush 2013 and Kush et al. 2015 propose a retrieval mechanism that implements the effects of C-Command into a processing architecture that allows only content-addressable memory (Lewis et al. 2006). As noted above, a content-addressable memory can tag a noun phrase in memory with inherent features like number, gender or animacy. And those can be used in retrieval. C-Command, however, is not a static, inherent property of a noun phrase; it is a relation that differs depending on the position of one noun phrase with respect to another. Because of this, C-Command cannot be encoded as a fixed feature of a noun phrase. To address this, Kush proposes that all antecedents initially have a feature, ACCESSIBLE, which is needed to be a visible antecedent for retrieval by a pronoun. This feature can be deactivated at points in processing. Kush et al. 2015 propose that a QP's ACCESSIBLE feature is deleted when the parser moves to a higher level of embedding, which defines a C-Command domain. In Kush et al., it is proposed that the parser deletes the QP's feature at the edge of the QP's "scope domain" (p. 36). Since, in our stimuli, the sentence-initial adjunct clause does *not* delimit the scope domain of the quantifier (Artstein 2015), this would not deactivate ACCESSIBLE, which should lead to GMMEs, contrary to what we found. It is for this reason that we interpret the conditions that deactivate ACCESSIBLE as structural, not semantic. C-command will not deactivate a referential DP's ACCESSIBLE feature because C-command does not limit referential antecedents to C-Command configurations. This version of the retrieval mechanism explains the results of Kush et al. 2015 and Cunnings et al. 2015 where pronouns do not retrieve a non-C-Commanding QP antecedent. It could be that non-C-Commanding antecedents can eventually be retrieved, but perhaps later than C-Commanding antecedents or less often. For instance, Kush et al. (Experiment 2b) found that while initial retrieval did not target non-C-Commanding antecedents, they found suggestive evidence that later processing may be affected by a structurally inappropriate antecedent. This also makes plausible the idea that scope compliant NoCC antecedents are accessed later in processing, and hence the lack of a GMME is only telling of the time-course or processing, not the grammar, of variable binding. The fact that the comprehension results in Experiment 1 showed that participants arrived at a co-varying interpretation for NoCC pronouns show that the NoCC antecedents are successfully accessed, but possibly later in the time course of processing.

If retrieval were delayed for scope compliant NoCC pronouns, however, we might have expected processing difficulty at the NoCC pronoun or soon after, which we did not find. In none of our on-line studies did we find evidence of an unheralded pronoun effect for matched NoCC pronouns (Filik et al. 2008, Nieuwland 2014). Post-hoc analyses of region 7 and beyond in Experiment Two showed that gender matched pronouns in the NoCC condition had similar reading time as the gender matched

pronouns in the CC condition ($p>.5$). In comparison, in Kush et al.'s Experiment 1c, non-C-Commanded pronouns showed a processing difficulty (in post pronoun region). In the experiments that used a gender mismatch paradigm in Kush et al. and Cunnings et al. 2015, a matched C-Commanding antecedent *facilitated* reading times; the raw reading times of the non-C-Commanded pronouns were elevated in comparison. It is of course possible that self-paced reading is not sensitive enough to reveal a processing difference between NoCC and CC match conditions. At the same time, the numerical trend in our results shows the inverted pattern of Kush et al.: in our Experiment Two, the NoCC conditions are numerically on par with matched C-Command cases, with the mismatched C-Commanded pronouns being elevated. The lack of processing difficulty for NoCC pronouns that we found is consistent with Carminati et al. 2002's results for NoCC pronouns: even with eye-tracking, they found no reading time penalty for non-C-Commanded co-varying pronouns.

Moreover, many on-line studies of the time course of anaphora processing report that potential antecedents of anaphors that meet structural constraints such as the Binding Principles are accessed quickly (Xiang, Dillon, and Phillips 2009, Dillon, Mishler, Sloggett, and Phillips 2013), although structurally inappropriate but pragmatically prominent ones may be accessed in the later stage (Nicol and Swinney 1989, 2003, Sturt 2003). Our results again do not suggest that NoCC pronouns display this lag, as evidenced by the off-line and on-line data. In this respect, we might have expected the off-line comprehension data to show some penalty for NoCC pronouns. However, this is not what we found.[10]

If scope compliant (matched) NoCC pronouns are quickly and readily interpreted as co-varying, as our initial results suggest, then how can we reconcile this with the fact that mismatched NoCC pronouns did not trigger a GMME, which is typically taken as evidence that the processor is not initiating a search for an antecedent? We discuss how supplementing Kush's retrieval mechanisms with a grammatical distinction among pronouns provides such a reconciliation.

6.3. Grammar plus processing An alternative to the processing-only explanation is one which posits an ambiguity in third person pronouns in English.[11] In addition to the cases Barker 2012 identified and on which we modeled our examples, there are many well-known cases where a pronoun co-varies with a non-C-Commanding quantifier (Bach and Partee 1980). The most famous of these are donkey anaphora (Geach 1962, Evans 1980), in which a co-varying pronoun takes a non-C-Commanding indefinite antecedent that is within a relative clause or a conditional (19). There are also cases of apparent binding across sentence boundaries, as in the phenomena known as modal subordination (20) (Roberts 1987) and telescoping (21), discussed above (Roberts 1987, Poesio and Zucchi 1992).

(19) a. If a farmer owns *a donkey*, he beats *it*.
b. Every farmer who owns *a donkey* beats *it*.

(20) If John bought a book, he'll be home reading it by now. *It* will be a mystery novel.

(21) *Each degree* candidate walked to the stage. *He* took his diploma from the dean and returned to his seat. (Roberts 1987)

In one wide-spread account, donkey pronouns and their relatives are treated as disguised definite descriptions (Cooper 1979, Evans 1980). In the E-type account, this definite contains silent pronominal elements that are bound; (22) provides an illustrative paraphrase for (19b) (see Heim and Kratzer 1998).

(22) Every farmer who owns *a donkey* beats *the donkey he owns*.

The D-type approach developed by Elbourne 2005 treats the pronoun as the spell-out of a definite determiner whose NP complement has been elided (Postal 1969).

(23) Every farmer who own *a donkey* beats *it ~~donkey~~*. it $\leadsto$ the

The definite co-varies because it is interpreted with respect to a situation variable which co-varies as it is bound by *every*, which quantifies over individual-situation pairs (Berman 1987). Restrictions on the size of the situation ensures that there is only one donkey per bound situation, satisfying the uniqueness requirements of the definite (Heim 1990).[12]

The D-type analysis has been extended to other cases of non-C-Commanded co-varying pronouns, such as binding out of DP (Büring 2004), modal subordination (Elbourne 2005) and telescoping (Anderssen 2011).[13] The virtue of the situation-based D-type approach is that it naturally extends to the kinds of stimuli we tested with temporal adjunct clauses. Since situations have a temporal ordering, this makes situation quantification a straightforward approach. A simplified version of one of the NoCC stimuli is in (24), and a sketch of its logical form in (25).

(24) After each boy came home, he practiced piano.

(25) For each x, s_b such that s_b is a minimal situation in which a boy x comes home, there is an extended situation s_e, $s_b \leq s_e$ such that **the boy in s**$_b$ practices piano in s_e and s_b temporally precedes s_e.[14]

In addition to these disguised definites, pronouns can also simply denote individual variables whose referent is determined by an assignment function, with gender features acting as a presupposition (Heim 1982, 1998). Such a standard pronoun can only co-vary if it falls in the C-Command domain of a quantifier, while a D-type pronoun merely needs to be evaluated with respect to a co-varying situation.

In terms of processing, then, we make the following observations about the experimental results. When a D-type analysis is available, as it is in scope compliant NoCC sentences, participants may quickly and successfully pursue it—despite, apparently, the fact that its antecedent is not marked Accessible. We do not expect a processing difficulty (unherladed pronoun effect) in this case.

(26) After each boy brought fresh water from the kitchen quickly, it seems that he ($\leadsto$

the boy) went on an early break.

This would also explain why in the off-line study participants are so accepting of the NoCC pronoun.

When the parser meets a gender mismatched NoCC pronoun, it cannot be interpreted as a co-varying D-type pronoun. In that case, it will be a standard pronoun. It will not introduce a GMME because the antecedent QP is not Accessible in Kush's system. We might expect it to show an unheralded pronoun effect, a much weaker effect distinct from a GMME. We found no evidence of an unheralded pronoun effect in Experiment Two. However, in post-hoc analyses of region 8 in Experiment Three, we found that while reading times for the matched pronoun with a QP antecedent (the hypothesized D-type pronoun) were no different than the matched pronouns with a DP antecedent ($p>.2$), mismatched pronouns in the QP condition had significantly longer reading time than the matched pronouns with a DP antecedent ($p=.02$). So while non-C-Commanded mismatched pronouns do not induce GMMEs, as indicated by the fact that they were read as fast as matched NoCC pronouns, they may still register a (much weaker) unheralded pronoun effect, when compared to the best case situation where pronouns occur with appropriate antecedent DPs.[15]

In C-Command configurations, a standard pronoun is chosen.[16] This leads to the automatic search for an antecedent, sensitive to C-Command along the lines proposed with the feature Accessible. A matching antecedent is quickly found; a mismatching antecedent produces a GMME.

To summarize, invoking a grammatical distinction between standard and D-type pronouns offers a way to reconcile why scope compliant pronouns appear to be efficiently processed but do not show evidence of triggering GMMEs. GMMEs are triggered just in cases where an antecedent is Accessible, as Kush has proposed. On this hypothesis, a D-type pronoun must be able to successfully, and possibly very quickly, retrieve its antecedent by a mechanism that does not rely on Accessible features. Of course, the mechanisms by which scope compliant, non-C-Commanded pronouns are processed requires investigation.[17] There is, however, already some independent experimental support pointing in the proposed direction. Invoking a D-type analysis for our stimuli fits closely with the results of Carminati et al. 2002 and Anderssen 2011 on telescoping. Recall that Carminati et al. 2002 found no evidence for processing difficulty for telescoped pronouns. Anderssen 2011 has argued that telescoped pronouns are D-type pronouns. In an acceptability study, Anderssen 2011 showed that telescoping is sensitive to whether the sentence containing the quantifier and sentence containing the co-varying pronoun are in a relation that expresses a "non-accidental" generalization. This is the case in (27a) but not (27b).

(27) Jeder Hausmeister trägt einen grossen Schlüsselbund mit sich herum.
Every janitor carries a large key chain with self around
'Every janitor carries around a large key chain.'

a. Er hat damit Zugang zu allen Räumen.
he has with.that access to all rooms
'With that he has access to all rooms.'

b. Er hat auch schon seit einigen Jahren graue Haare.
he has also since several years gray hairs
'He has also had gray hair for several years.'
(Anderssen 2011: 41 (51a–c))

The accidental generalization sentences were rated significantly worse than non-accidental generalization sentences. Crucially, the telescoping sentences that meet the non-accidental generalization condition were rated as acceptable as the fully acceptable referential cases. As with Carminati et al. 2002, telescoping—when semantically and pragmatically licensed—shows no deleterious effects on acceptability. Anderssen goes on to show that treating telescoping pronouns as D-type pronouns captures their sensitivity to the types of situations involved using a generic quantification over situations (Kratzer 1989). All this is consonant with the pattern of results we found: if the NoCC pronouns we tested are D-type pronouns as are telescoping pronouns, then we do not expect any comprehension or severe processing difficulties. However, the process of searching for an antecedent for such anaphoric forms will be different from those anaphoric forms that can be interpreted as standard individual-denoting pronouns, i.e. C-Commanded pronouns. And this is what we found: standard, C-Commanded pronouns exhibited gender mismatch effects but our hypothesized D-type pronouns did not.

In summary, there are good reasons to think that D-type pronouns are needed independently, for phenomena like telescoping and others. Previous experimental work has shown that such D-type pronouns do not exhibit processing difficulty. The fact that the non-C-Commanded pronouns in our studies showed no evidence of comprehension or processing difficulty *per se*, and yet did not trigger the kinds of gender-mismatch effects that standard pronouns do, suggests that treating these differently in the grammar is a viable hypothesis.

The results of our Experiments Two and Three taken together thus raise serious questions for Barker's approach that all co-varying pronouns are uniform, and calls for the necessity of testing the on-line processing of other D-type pronouns, as begun by Grosz, Patel-Grosz, Fedorenko, and Gibson 2014, Foppolo 2009, to determine whether the present results reflect a grammatical implementation that distinguishes individual denoting co-varying pronouns from D-type ones.

7. Conclusion We have reported that structural C-Command governs the processing of co-varying pronouns. Unlike the previous studies, however, we ruled out semantic scope as the relevant constraint (cf. Barker 2012). We found that pronouns that fall in scope but not the C-Command domain of the quantified antecedent do not show gender-mismatch effects, but C-Commanded pronouns do. We also showed that referential antecedents in the same non-C-commanding position *did* induce gender mismatch effects, confirming that co-varying pronouns that are semantically bound by non-C-Commanding quantified phrases do not activate the same kind of antecedent retrieval search that C-Commanded co-varying pronouns (and referential pronouns in general) do. Crucially, then, structural relations, not merely semantic relations, are implemented by the on-line processor, as argued for a range of other grammatical

relations that are syntactically constrained, such as reflexive binding (Nicol and Swinney 1989, Xiang et al. 2009, Dillon et al. 2013) and islands for movement dependencies (Stowe 1986, Phillips 2006).

We then addressed whether C-Command plays a role only in the on-line processing of co-varying pronouns (i.e. made them more readily or quickly accessible) or whether the grammar itself distinguishes between C-Commanded and non-C-Commanded pronouns. If C-Command merely regulated how readily or how often the processor considered quantificational antecedents, we may have expected scope compliant co-varying pronouns to exhibit some processing difficulty or a hint of reduced comprehension. Here the results were more equivocal, but the numerical trends suggest that scope compliant, non-C-Commanded pronouns are easily processed, a finding consistent with Carminati et al. 2002. We entertained augmenting Kush's processing mechanisms with a grammatical distinction, in which pronouns are ambiguous between standard variables (for C-Command configurations) and disguised definite descriptions (for scope compliant, non-C-Command configurations). Our hypothesis comes with the prediction that we will find this processing profile for other types of anaphora, such as other D-type pronoun configurations and dependent definites generally. If that is the case, then the mechanisms of antecedent retrieval for these varieties of anaphora need to be explicated.

A Experimental Materials: Experiment One

Fillers: Good Co-variation

(1) Each doctor said that he was going to leave the clinic.
Who was going to leave the clinic?
One person.
Each of the doctors.

(2) Every boy at the birthday party wants to tell his mother that he was upset.
Who was upset?
One person.
Each of the boys.

(3) John didn't want any girl to know she was getting a new sports car.
Who was getting a new sports car?
One person.
Girls.

(4) It seems that each bridesmaid regrets that she did not opt for the more expensive dress.
Who did not opt for the more expensive dress?
One person.
Each of the bridesmaids.

(5) The committee members said that each applicant should have sent an updated resume before he came to the meeting.
Who came to the meeting?
One man.
Each of the candidates.

(6) It appears that each teacher in the meeting regretted that she made mean comments to the principal.
Who made mean comments?
One person.
Each of the teachers.

(7) Each football player on the team admitted that he had taken lots of drugs over the weekend.
Who took drugs over the weekend?
One person.
Each of the football players.

(8) Hugh doesn't think that any ballerina will regret that she was given a small role.
Who was given a small role?
One person.

Ballerinas.

(9) Sally does not believe that any policeman will be given a medal because he saved a dog.
Who saved a dog?
One person.
Policemen.

(10) Each cello player came in early and he tuned his instrument.
Who tuned his instrument?
One person.
Each of the cello players.

(11) Each old man at the coffee shop remembered that he had forgotten his pill.
Who forgot his pill?
Some guy.
Each of the old men.

(12) The woman at the counter didn't think any waiter should admit that he is unhappy.
Who is unhappy?
One person.
Waiters.

Fillers: Strong Crossover

(13) Sally knows that he says nice things about each fireman.
Who says nice things?
Some guy.
Each of the firemen.

(14) She admitted that each princess was stealing jewelry from the museum.
Who admitted it?
One person.
Each of the princesses.

(15) The newspaper claimed that he sent a bad check to each businessman.
Who sent a bad check?
One person.
Each of the businessmen.

(16) Roger didn't know that she helped out any of the little girls at the dance.
Who helped out at the dance?
One person.
Each of the little girls.

(17) He offered a home to each lawyer's client.
Who offered a home?
One person.
Each of the lawyers.

(18) Many people said the she would send each kindergarten teacher away for the fall.
Who would send people away?
One person.
Each of the kindergarten teachers.

(19) He prevented the nurse from talking to every doctor.
Who prevented the nurse from talking?
One person.
Each of the doctors.

(20) I think she was going to place an order and give each little girl a new dress.
Who was going to place an order?
One person.
Each of the little girls.

(21) It seems to me that he was not aware that every janitor in the building was about to be fired.
Who was not aware of it?
One person.
Every janitor.

(22) She made a really impressive effort to provide each bride with an itinerary for the day.
Who provided an itinerary?
One person.
Each of the brides.

Fillers: Bad Telescoping

(23) Each student in the syntax class was accused of cheating on the exam and he has a Ph.D. in astrophysics.
Who has a Ph.D. in astrophysics?
Some guy.
Each of the students.

(24) Every rice-grower in Korea owns a wooden cart. He also has black hair.
Who has black hair?
One person.
Every rice-grower.

(25) Mary doesn't think that any barber has been here for a while. We learned he was outside smoking.
Who was outside smoking?
Some guy.
Barbers.

(26) Each nurse was given a form to fill out concerning sanitation. She sent for the doctor.
Who sent for the doctor?
One person.
Each of the nurses.

(27) Every janitor carries around a large keychain. He has also had grey hair for years.
Who has grey hair?
One person.
Every janitor.

(28) Mary wrote about every basketball player from the school. The editors revealed he was taking steroids.
Who was taking steroids?
Some guy.
Every basketball player.

(29) Each lawyer for the defendant wrote a long description of the crime scene. He graduated from a really good law school.
Who graduated from a really good law school?
Some guy.
Each of the lawyers.

(30) Each baby-sitter was paid a lot of money for taking care of the terrible children. Then she went and finished her homework.
Who finished her homework?
One person.
Each of the baby-sitters.

(31) Sally doesn't like any policeman on the force. He doesn't like Sally either. Who doesn't like Sally?
One person.
Policemen.

(32) If the chef should not speak to each waitress, we think she should resign.
Who should resign?
One person.
Each of the waitresses.

Fillers: Good D-type interpretation/Telescoping

(33) The wife of each guy admitted he watched a lot of sports on TV and read no books.
Who watched a lot of sports?
One guy.
Each of the guys.

(34) The detective spoke to each teacher and asked her to stick around.
Who was asked to stick around?
One person.
Each of the teachers.

(35) Reports submitted by each policeman had to have his signature on it.
Whose signature had to be on the report?
Some guy's.
Each policeman's.

(36) Right after meeting each baseball player, the manager gave him a raise.
Who got a raise?
Some guy.
Each baseball player.

Fillers: Biased to referential Pronoun

(37) The manager put each waitress on the night shift at some point in the week and he regretted it.
Who regretted it?
One person.
Each manager.

(38) Every woman wanted to meet Rachel but she had already left town.
Who had already left town?
One person.
Every woman.

(39) Although he is popular, Patrick didn't think any other little boy would come to his party.
Who is popular?
One person.
Little boys.

(40) After she met with the employees, Rita chose to single out each client for a special deal.
Who met with the employees?
One person.

Each client.

(41) I met the surgeon who helped every patient in town and he was nice.
Who was nice?
One person.
Every patient.

(42) Sally gave a medal to the lady who taught every boy the piano and she liked it.
Who liked it?
One person.
Boys.

Scope Compliant (a) = CC, (b) = NoCC

(43) a. It seems each boy brought fresh water from the kitchen quickly right before he went on an early break.
Who went on an early break?
One person.
Each of the boys.
b. After each boy brought fresh water from the kitchen quickly it seems that he went on an early break.
Who went on an early break?
One person.
Each of the boys.

(44) a. It appears each ballerina fainted on stage dramatically during the recital right after she practiced the dance steps carefully.
Who practiced the dance steps carefully?
One person.
Each of the ballerinas.
b. Before each ballerina fainted on stage dramatically during the recital it appears that she practiced the dance steps carefully.
Who practiced the dance steps carefully?
One person.
Each of the ballerinas.

(45) a. It seems each prince went hunting in the woods for a stag just after he poisoned the food with cyanide.
Who poisoned the food?
One person.
Each of the princes.
b. Before each prince went hunting in the woods for a stag it seems that he poisoned the food with cyanide.
Who poisoned the food?
One person.

Each of the princes.

(46) a. It appears each policeman went on duty at the event downtown right before he investigated the crime scene in the dark.
Who investigated the crime scene?
One person.
Each of the policemen.
b. After each policeman went on duty at the event downtown it appears that he investigated the crime scene in the dark.
Who investigated the crime scene?
One person.
Each of the policemen.

(47) a. It seems each woman got worried about proper safety procedures only after she checked on the operating room several times.
Who checked on the operating room?
One person.
Each of the women.
b. Before each woman got worried about proper safety procedures it seems that she checked on the operating room several times.
Who checked on the operating room?
One person.
Each of the women.

(48) a. It appears each fireman carried the hose from the firetruck to the house right before he called for backup to help quickly.
Who called for backup?
One person.
Each of the firemen.
b. After each fireman carried the hose from the firetruck to the house it appears that he called for backup to help quickly.
Who called for backup?
One person.
Each of the firemen.

(49) a. It seems each waitress spilled the drinks all over the floor of the restaurant just when she went quickly back to the kitchen.
Who went back to the kitchen?
One person.
Each of the waitresses.
b. When each waitress spilled the drinks all over the floor of the restaurant it seems that she went quickly back to the kitchen.
Who went back to the kitchen?
One person.
Each of the waitresses.

(50) a. It appears each groomsman was on the dance floor having fun with friends right when he asked one of the little flower girls to dance.
Who asked one of the flower girls to dance?
One person.
Each of the groomsmen.
b. 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.
Who asked one of the flower girls to dance?
One person.
Each of the groomsmen.

(51) a. It seems each actress was practicing lines for the play at home right when she learned that the play would appear on Broadway.
Who learned the play would appear on Broadway?
One person.
Each of the actresses.
b. When each actress was practicing lines for the play at home it seems that she learned that the play would appear on Broadway.
Who learned the play would appear on Broadway?
One person.
Each of the actresses.

(52) a. It appears each salesman sold a brand new Mercedes to the rich man for a big profit well before he got a big bonus this year from the boss.
Who got a big bonus this year?
One person.
Each of the salesmen.
b. 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.
Who got a big bonus this year?
One person.
Each of the salesmen.

(53) a. It appears each little girl was at the park playing with friends from school right when she got a strawberry ice-cream cone for free.
Who got a strawberry ice-cream cone for free?
One girl.
Each of the girls.
b. 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.
Who got a strawberry ice-cream cone for free?
One girl.
Each of the girls.

(54) a. It seems each little boy came home late from a long day at school just before

he practiced playing the piano for half an hour.
Who practiced piano for half an hour?
One person.
Each of the boys.

b. 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.
Who practiced piano for half an hour?
One person.
Each of the boys.

(55) a. It seems each mother arrived at school for a meeting early in the morning straight after she phoned the principal to complain about bullies.
Who phoned the principal?
One person.
Each of the mothers.

b. Before each mother arrived at school for a meeting early in the morning it seems that she phoned the principal to complain about bullies.
Who phoned the principal?
One person.
Each of the mothers.

(56) a. It appears each businessman spoke at the city council meeting last night only when he was angry about high taxes and new bylaws.
Who was angry?
One person.
Each of the businessmen.

b. When each businessman spoke at the city council meeting last night it appears that he was angry about high taxes and new bylaws.
Who was angry?
One person.
Each of the businessmen.

(57) a. It seems each old woman came quickly out of the house down the road right when she yelled mean things loudly for hours.
Who yelled mean things?
One person.
Each of the old women.

b. When each old woman came quickly out of the house down the road it seems that she yelled mean things loudly for hours.
Who yelled mean things?
One person.
Each of the old women.

(58) a. It appears each hockey player changed into skates carefully at the rink right before he skated on the fresh ice for several hours.

Who skated on the fresh ice?
One person.
Each of the hockey players.

b. After each hockey player changed into skates carefully at the rink it appears that he skated on the fresh ice for several hours.
Who skated on the fresh ice?
One person.
Each of the hockey players.

(59) a. It seems each secretary typed up the minutes in the office from the meeting yesterday well after she organized the shelves in the photocopying room very neatly.
Who organized the shelves?
One person.
Each of the secretaries.

b. After each secretary typed up the minutes in the office from the meeting yesterday it seems that she organized the shelves in the photocopying room very neatly.
Who organized the shelves?
One person.
Each of the secretaries.

(60) a. It appears each plumber fixed a leaky sink in the kitchen in the morning immediately before he picked out a new faucet for the bathroom in the afternoon.
Who picked out a new faucet?
One person.
Each of the plumbers.

b. 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.
Who picked out a new faucet?
One person.
Each of the plumbers.

(61) a. It appears each bridesmaid had her hair done up fancy for the wedding just before she helped the bride get to the church on time.
Who helped the bride get to the church?
One person.
Each of the bridesmaids.

b. 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.
Who helped the bride get to the church?
One person.
Each of the bridesmaids.

(62) a. It seems each new father arrived at the hospital in a rush from work right when he was told by the doctor that everything would be fine.
Who was told that everything would be fine?
One person.
Each of the fathers.
b. 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.
Who was told that everything would be fine?
One person.
Each of the fathers.

B Experimental Materials: Experiment Two (a) = NoCC, (b) = CC;

(1) a. After - each boy - brought fresh water - from the kitchen - quickly - it seems - that he/she - went - on an early - break.
b. It seems - each boy - brought fresh water - from the kitchen - quickly - right - before he/she - went - on an early - break.

(2) a. Before - each ballerina - fainted on stage - dramatically - during the recital - it appears - that he/she - practiced - the dance steps - carefully.
b. It appears - each ballerina - fainted on stage - dramatically - during the recital - right - after he/she - practiced - the dance steps - carefully.

(3) a. Before - each prince - went hunting - in the woods - for a stag - it seems - that he/she - poisoned - the food - with cyanide.
b. It seems - each prince - went hunting - in the woods - for a stag - just - after he/she - poisoned - the food - with cyanide.

(4) a. After - each policeman - went on duty - at the event - downtown - it appears - that he/she - investigated the - crime scene - in the dark.
b. It appears - each policeman - went on duty - at the event - downtown - right - before he/she - investigated the - crime scene - in the dark.

(5) a. Before - each woman - got worried - about proper - safety procedures - it seems - that he/she - checked on - the operating room - several times.
b. It seems - each woman - got worried - about proper - safety procedures - only - after he/she - checked on - the operating room - several times.

(6) a. After - each fireman - carried the hose - from the firetruck - to the house - it appears - that he/she - called for - backup to help - quickly.
b. It appears - each fireman - carried the hose - from the firetruck - to the house - right - before he/she - called for - backup to help - quickly.

(7) a. When - each waitress - spilled the drinks - all over the floor - of the restaurant - it seems - that he/she - went quickly - back to - the kitchen.
b. It seems - each waitress - spilled the drinks - all over the floor - of the

restaurant - just - when he/she - went quickly - back to - the kitchen.

(8) a. When - each groomsman - was on the dance - floor having fun - with friends - it appears - that he/she - asked one of - the little flower girls - to dance.
b. It appears - each groomsman - was on the dance - floor having fun - with friends - right - when he/she - asked one of - the little flower girls - to dance.

(9) a. When - each actress - was practicing - lines for the play - at home - it seems - that he/she - learned that the - play would appear - on Broadway.
b. It seems - each actress - was practicing - lines for the play - at home - right - when he/she - learned that the - play would appear - on Broadway.

(10) a. After - each salesman - sold a brand new Mercedes - to the rich man - for a big profit - it appears - that he/she - got a big bonus - this year - from the boss.
b. It appears - each salesman - sold a brand new Mercedes - to the rich man - for a big profit - well - before he/she - got a big bonus - this year - from the boss.

(11) a. When - each little girl - was at the park - playing with friends - from school - it appears - that he/she - got a strawberry - ice-cream cone - for free.
b. It appears - each little girl - was at the park - playing with friends - from school - right - when he/she - got a strawberry - ice-cream cone - for free.

(12) a. After - each little boy - came home - late - from a long day at school - it seems - that he/she - practiced playing - the piano - for half an hour.
b. It seems - each little boy - came home - late - from a long day at school - just - before he/she - practiced playing - the piano - for half an hour.

(13) a. Before - each mother - arrived at school - for a meeting - early in the morning - it seems - that he/she - phoned the principal - to complain - about bullies.
b. It seems - each mother - arrived at school - for a meeting - early in the morning - straight - after he/she - phoned the principal - to complain - about bullies.

(14) a. When - each businessman - spoke at the - city council meeting - last night - it appears - that he/she - was angry about - high taxes - and new bylaws.
b. It appears - each businessman - spoke at the - city council meeting - last night - only - when he/she - was angry about - high taxes - and new bylaws.

(15) a. When - each old woman - came quickly - out of the house - down the road - it seems - that he/she - yelled mean - things loudly - for hours.
b. It seems - each old woman - came quickly - out of the house - down the road - right - when he/she - yelled mean - things loudly - for hours.

(16) a. After - each hockey player - changed into skates - carefully - at the rink - it appears - that he/she - skated on the - fresh ice - for several hours.

b. It appears - each hockey player - changed into skates - carefully - at the rink - right - before he/she - skated on the - fresh ice - for several hours.

(17) a. After - each secretary - typed up the minutes - in the office - from the meeting yesterday - it seems - that he/she - organized the shelves - in the photocopying room - very neatly.
b. It seems - each secretary - typed up the minutes - in the office - from the meeting yesterday - well - after he/she - organized the shelves - in the photocopying room - very neatly.

(18) a. After - each plumber - fixed a leaky sink - in the kitchen - in the morning - it appears - that he/she - picked out a new - faucet for the bathroom - in the afternoon.
b. It appears - each plumber - fixed a leaky sink - in the kitchen - in the morning - immediately - before he/she - picked out a new - faucet for the bathroom - in the afternoon.

(19) a. After - each bridesmaid - had her hair - done up fancy - for the wedding - it appears - that he/she - helped the bride - get to the church - on time.
b. It appears - each bridesmaid - had her hair - done up fancy - for the wedding - just - before he/she - helped the bride - get to the church - on time.

(20) a. When - each new father - arrived at the - hospital in a rush - from work - it seems - that he/she - was told by - the doctor that - everything would be fine.
b. It seems - each new father - arrived at the - hospital in a rush - from work - right - when he/she - was told by - the doctor that - everything would be fine.

C Experimental Materials: Experiment Three Test materials of Experiment Three were the same as the ones used in the NoCC condition of Experiment Two, with the addition of *each/the* manipulation.

References

Anderson, Catherine. 2004. The structure and Real-Time Comprehension of Quantifier Scope Ambiguity. Doctoral Dissertation, Northwestern University.

Anderssen, Jan. 2011. Qauntification, Misc. Doctoral Dissertation, University of Massachusetts at Amherst.

Artstein, Ron. 2005. Quantificational arguments in temporal adjunct clauses. *Linguistics and Philosophy* 28:541–597.

Bach, Emmon, and Barbara H. Partee. 1980. Anaphora and semantic structure. In *Papers from the Parasesion on Pronouns and Anaphora*, ed. Jody Kreiman and Almerindo E. Ojeda, 1–28. Chicago Illinois: University of Chicago.

Barker, Chris. 2012. Quantificational binding does not require c-command. *Linguistic Inquiry* 4:614–633.

Barker, Chris, and Geoffrey K. Pullum. 1990. A theory of command relations. *Linguistics and Philosophy* 13:1–34.

Barker, Chris, and Chung-Chieh Shan. 2008. Donkey anaphora as in-scope binding. *Semantics and Pragmatics* 1:1–34.

Barr, Dale J.; Roger Levy; Christoph Scheepers; and Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68:255–278.

Bates, Douglas M. 2005. Fitting linear mixed models in R: Using the lme4 package. *R News: The Newsletter of the R Project* 5:27–30.

Beghelli, Filippo. 1993. A minimalist approach to quantifier scope. In *Proceedings of the North East Linguistic Society 23*, ed. Amy J. Schafer, 65–80. University of Ottawa: Graduate Linguistic Student Association.

Beghelli, Filippo, and Tim Stowell. 1997. Distributivity and negation: The syntax of *each* and *every*. In *Ways of Scope Taking*, ed. Anna Szabolcsi, 71–108. Kluwer Academic Publishers.

Berman, Stephen. 1987. Situation-based semantics for adverbs of quantification. In *University of Massachusetts Occasional Papers Vol. 12*, ed. Jim Blevins and Ann Vainikka, 8–23. University of Massachusetts at Amherst.

Brasoveanu, Adrian. 2008. Donkey pluralities: plural information states versus non-atomic individuals. *Linguistics and Philosophy* 31:129–209.

Brasoveanu, Adrian, and Jakub Dotlačil. 2015. Sentence-internal "same" and its quantificational licensors: A new window into the processing of inverse scope. *Semantics and Pragmatics* 8:1–52.

Brasoveanu, Adrian, and Jakub Dotlačil. 2015. Strategies for scope taking. *Natural Language Semantics* 23:1–19.

Bruening, Benjamin. 2014. Precede-and-command revisited. *Language* 90:342–388.

Büring, Daniel. 2004. Crossover situations. *Natural Language Semantics* 12:23–62.

Büring, Daniel. 2005. Bound to bind. *Linguistic Inquiry* 36:259–274.

Carminati, Maria Nella; Lyn Frazier; and Keith Rayner. 2002. Bound Variables and C-Command. *Journal of Semantics* 19:1–34.

Chierchia, Gennaro. 1995. *Dynamics of Meaning*. Chicago: University of Chicago Press.

Cooper, Robin. 1979. The interpretation of pronouns. In *Syntax and Semantics 10: Selections from the Third Groningen Round Table*, ed. Frank Heny and H. Schnelle, 61–92. New York: Academic Press.

Cunnings, Ian; Clare Patterson; and Claudia Felser. 2014. Variable binding and coreference in sentence comprehension: Evidence from eye movements. *Journal of Memory and Language* 71:39–56.

Cunnings, Ian; Clare Patterson; and Claudia Felser. 2015. Structural constraints on pronoun binding and coreference: evidence from eye movements during reading. *Frontiers in Psychology* 6.

Dillon, Brian; Alan Mishler; Shayne Sloggett; and Colin Phillips. 2013. Contrasting intrusion profiles for agreement and anaphora: experimental and modeling evidence. *Journal of Memory and Language* 69:85–103.

Dotlačil, Jakub, and Adrian Brasoveanu. 2015. The manner and time course of updating quantifier scope representations in discourse. *Language, Cognition and Neuroscience* 30:305–323.

Elbourne, Paul D.. 2001. E-type anaphora as NP deletion. *Natural Language Semantics* 9:241–288.

Elbourne, Paul D. 2005. *Situations and individuals*. MIT Press.

Erlewine, Michael Yoshitaka, and Hadas Kotek. 2016. A streamlined approach to online linguistic surveys. *Natural Language and Linguistic Theory* 34:481–495.

Evans, Gareth. 1980. Pronouns. *Linguistic Inquiry* 11:337–362.

Filik, Ruth; Anthony J. Sandford; and H. Leuthold. 2008. Processing pronouns without antecedents. *Journal of Cognitive Neuroscience* 20:1315–1326.

Foppolo, Francesca. 2009. The puzzle of donkey anaphora resolution. In *Proceedings of the 38th Annual Meeting of the North East Linguistic Society*. GLSA. Amherst, MA. 297–310.

Foraker, Stephani, and Brian McElree. 2007. The role of prominence in pronoun resolution: Availability versus accessibility. *Journal of Memory and Language* 56:357–383.

Frazier, Lyn, and Charles Jr. Clifton. 2000. On bound variable interpretations: the LF-only hypotehsis. *Journal of Psycholinguistic Research* 29:125–139.

Garnham, Alan. 2001. *Mental models and the interpretation of anaphora*. Psychology Press.

Geach, Peter. 1962. *Reference and generality*. Ithaca, NY: Cornell University Press.

Geurts, Bart. 2002. Donkey business. *Linguistics and Philosophy* 25:129–156.

Geurts, Bart. 2011. Accessibility and anaphora. In *Semanitcs: An International Handbook of Natural Language Meaning*, ed. Klaus von Heusinger, Claudia Maienborn, and Paul Portner, volume 2, 1988–2011. Gruyter.

Gordon, Peter C., and Randall Hendrick. 1998. The representation and processing of coreference in discourse. *Cognitive Science* 22:389–424.

Grodzinsky, Yosef, and Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24:69–102.

Groenendijk, J., and M. Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14:38–100.

Grosz, Patrick G.; Pritty Patel-Grosz; Evelina Fedorenko; and Edward Gibson. 2014. Constraints on donkey pronouns. *Journal of Semantics* 32: 619–648.

Heim, Irene. 1982. The semantics of definite and indefinite Noun Phrases. Doctoral Dissertation, University of Massachusetts at Amherst.

Heim, Irene. 1990. E-type pronouns and donkey anaphora. *Linguistics and Philosophy* 13:137–178.

Heim, Irene. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart's approach. In *MIT Working Papers in Linguistics 25: The Interpretive Tract*, ed. Orin Percus and Uli Sauerland, 205–246. Cambridge, Massachusetts: MIT, Department of Linguistics.

Ioup, Georgia. 1975. Some universals for quantifier scope. In *Syntax and Semantics*, ed. John Kimball, volume 4, 37–58. Academic Press.

Just, Marcel A.; Patricia A. Carpenter; and Jacqueline D. Woolley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111:228–238.

Kamp, Hans. 1981. A theory of truth and semantic representation. In *Formal methods in the study of language*, ed. Jeroen A. G. Groenendijk, T. M. V. Janssen, and Martin B. J. Stokhof, 277–322. Amsterdam: Mathematical Center.

Kamp, Hans, and Uwe Reyle. 1993. *From discourse to logic*. Dordrecht: Kluwer Academic Publishers.

Kazanina, Nina; Ellen F. Lau; Moti Lieberman; Masaya Yoshida; and Colin Phillips. 2007. The effect of Syntactic Constraints on the Processing of Backwards Anaphora. *Journal of Memory and Language* 56:384–409.

Kratzer, Angelika. 1989. An investigation of the lumps of thought. *Linguistics and Philosophy* 12:607–653.

Kroch, Anthony. 1974. The semantics of scope in English. Doctoral Dissertation, MIT.

Kurtzman, Howard S., and Maryellen C. MacDonald. 1993. Resolution of quantifier scope ambiguities. *Cognition* 48:243–279.

Kush, Dave. 2013. Respecting relations: Memory access and antecedent retrieval in incremental sentence processing. Doctoral Dissertation, University of Maryland.

Kush, Dave; Jeffrey Lidz; and Colin Phillips. 2015. Relation-sensitive retrieval: evidence from bound variable pronouns. *Journal of Memory and Language* 82:18–40.

Kuznetsova, Alexandra; Per Bruun Brockho; and Rune Haubo Bojesen Christensen. 2014. lmerTest: Tests for Random and Fixed Eects for Linear Mixed Eect Models (lmer Objects of lme4 Package). R package version 2.0-11.

Leu, Thomas. 2005. Donkey pronouns: void descriptions? In *Proceedings of Thirty-Fifth Annual meeting of the North East Linguistic Society*, volume 2. GLSA. Amherst, MA. 379–390.

Lewis, David; Shravan Vasishth; and J. Van Dyke. 2006. Computation principles of working memory in sentence comprehension. *Trends in Cognitive Science* 10:447–454.

Lewis, Shevaun, and Colin Phillips. 2015. Aligning Grammatical Theories and Language Processing Models. *Journal of Psycholinguistic Research* 44:27–46.

McElree, Brian. 2000. Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research* 29:111–123.

McElree, Brian; Foraker, Stephani; Dyer, Lisbeth. 2003. Memory structures that subserve sentence comprehension. *Journal of Memory and Language* 48:67–91.

Nicol, Janet, and David Swinney. 1989. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18:5–19.

Nicol, Janet, and David Swinney. 2003. The psycholinguistics of anaphora. In *Anaphora: A reference guide*, ed. Andrew Barss, 72–104. Oxford University Press.

Nieuwland, M. S. 2014. Who's he: Event-related brain potentials and unbound pronuns. *Journal of Memory and Language* 76:1–28.

Patel-Grosz, Pritty, and Patrick Grosz. to appear. Revisiting pronominal typology. *Linguistic Inquiry*.

Peirce, Jonathan W. 2007. Psychopy—psychophysics software in python. *Journal of Neuroscience Methods* 162:8–13.

Phillips, Colin. 2006. The Real-Time Status of Island Phenomena *Language* 82.2:795–823.

Poesio, Massimo, and Alessandro Zucchi. 1992. On telescoping. In *Semantics and Linguistic Theory*, volume 2, ed. Chris Barker and David Dowty, 347–366. Columbus, Ohio: Ohio State University

Postal, Paul M. 1969. On so-called 'pronouns' in English. In *Modern studies in English*, ed. D. Reibel and Sandford Schane, 201–244. Englewood Cliffs, New Jersey: Prentice-Hall.

R Development Core Team. 2012. R: A language and environment for statistical computig. http://www.R-project.org.

Reinhart, Tanya. 1976. The syntactic domain of anaphora. Doctoral Dissertation, Massachusetts Institute of Technology.

Reinhart, Tanya. 1983. *Anaphora and semantic interpretation*. Chicago, Illinois: University of Chicago Press.

Reuland, Eric. 2011. *Anaphora and Language Design*. Cambridge, Mass: MIT Press.

Roberts, Craige. 1987. Modal subordination, anaphora and distributivity. Doctoral Dissertation, University of Massachusetts at Amherst.

Rothstein, Susan. 1995. Adverbial quantification over events. *Natural Language Semantics* 3:1-32.

Schwarz, Florian. 2009. Two types of definites in natural language. Doctoral Dissertation, University of Massachusetts at Amherst.

Stowe, Leslie. 1986. Parsing WH-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1.3:227–245.

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

Tunstall, Susanne Lynn. 1998. The interpretation of quantifiers: semantics & processing. Doctoral Dissertation, University of Massachusetts at Amherst.

van Gompel, R.P.G, and Simon P. Liversedge. 2003. The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory and Cognition* 1:128–139.

Vendler, Zeno. 1962. Each and every, any and all. *Mind* 71:145–160.

Xiang, Ming; Brian Dillon; and Colin Phillips. 2009. Illusory licensing effects across dependency types: ERP evidence. *Brain and Language* 108:40–55.

Notes

[1]From here on, we avoid the traditional term *bound variable* because the word *bound* is also used in Chomsky's 1981 Binding Principles. The Binding Principles, however, take C-Command as a requirement for binding. We use the more neutral term *co-varying pronoun*, which just indicates that the value of the pronoun co-varies with the quantified noun.
[2]In its most cited form, C-Command is defined as:

(1) Node A C-Commands Node B iff every branching node that dominates A also dominates B (and neither A nor B dominate the other). (after Reinhart 1983)

There are variations, with respect to the status of branching and dominating nodes, that do not play a role here. See Barker and Pullum 1990 for discussion.
[3]These are not instances of telescoping because the quantifier takes semantic scope over other quantifiers in the position of the pronoun. The QP in the temporal adjunct clause in (2a) can scope over the indefinite in the main clause, such that for each boy there is a possibly different cup that went missing. This is not possible in the telescoping case in (2b), although pronouns can certainly be telescoped in the same configuration: *Each boy left. Then he went to his mother's*.

(2) a. After each boy left, a (different) cup went missing.
b. Each boy left. Then a (#different) cup went missing.

It is true, though, that as in telescoping, negative quantifiers do not bind out of temporal adjunct clauses, but this is because the scope-taking properties of such quantifiers is independently limited (Beghelli and Stowell 1997, among others).
[4]Since materials in Kush et al. 2015 used NPIs as the QP, we used bare plurals as representative of the co-varying interpretation.
[5]Some of these used NPI quantifiers (like Kush et al.) and some used *each*.
[6]For Experiment Two and Experiment Three, we ran a model where the difference in gendered antecedent types was coded as a variable to see if that had an effect on the results. In both experiments, the gendered antecedent type showed no main effect nor any reliable interaction with other factors. We thus removed gendered antecedent type variable from the analyses reported here.
[7]A follow-up study to confirm that NoCC pronouns with quantificational antecedents do not give rise to unheralded pronoun effects would compare the following two sentences.

(3) a. After each boy delivered the fresh water, it seems that he went back to the kitchen.
b. After the fresh water was delivered, it seems that he went back to the kitchen.

Our hypothesis is that (3a) will not show an unheralded pronoun effect, but (3b) will.
[8]Cunnings et al. 2015 found a similar pattern. Their stimuli, however, always introduced an early, viable antecedent for the NoCC mismatching pronoun, suggesting that elevated

reading times can be associated with effects distinct from unherladed pronoun effects or gender mismatch.

[9]We thank a reviewer for making us take this possibility seriously.

[10]Further investigation is required to determine the time-course of the processing of scope compliant NoCC pronouns. For instance, in (4) the negative quantifier *no boy* cannot scope out of the *after*-clause, while *each* can, as we saw.

(4) After {each boy/no boy} brought fresh water from the kitchen quickly, it seems that he/she...

Given our results and those of previous studies, we expect neither quantifier will induce a gender mismatch effect at the mismatched pronoun. The question is whether, particularly in eye-tracking, the match pronouns, or regions following, would be processed differently depending on the quantifier.

[11]We do not intend here to suggest that the processor is fundamentally distinct from the grammar. The system that implements the grammar in real-time may be the same system we call the grammar (Lewis and Phillips 2015). Here we simply use *processing* to refer to influences of the memory retrieval mechanisms.

[12]More specifically, interpreting the pronoun as definite involves 'extended' situations in the nuclear scope of the quantifier

(5) For every x, s_b such that s_b is a minimal situation in which a farmer x owns a donkey, there is an extended situation s_e, $s_b \leq s_e$ such that x beats in s_e **the donkey in s_b**.
(adapted from Büring 2004, p. 29: (39))

The preceding formula, from Büring 2005, differs from formulations given in Elbourne 2005 in terms of which situation (s_e or s_b) the definite is evaluated in.

[13]In contrast, other approaches keep the basic syntax and semantics of pronouns constant across C-Command and non-C-Command configurations, but alter how quantifiers can take scope. This is the kind of system advocated by Barker 2012 (see Barker and Shan 2008) and it is assumed in dynamic semantic approaches (Kamp 1981, Heim 1982, Groenendijk and Stokhof 1991, Kamp and Reyle 1993, Chierchia 1995, Geurts 2002, 2011, Brasoveanu 2008).

[14]This is not meant to be a thorough analysis of temporal adjunct clauses in situation semantics. For instance, we would need to add matching functions in the sense of Rothstein 1995. Additionally, the situation variable in which the definite pronoun is resolved could be the extended situation not the 'base' situation (s_b), although this depends on how situation extension is formulated (see Elbourne 2001, Schwarz 2009 for discussion).

[15]A reviewer asks how the parser chooses between a D-type and standard bound variable pronoun. If the parser pursued a D-type analysis here for the mismatched NoCC pronoun, instead of a standard variable analysis, this would still involve retrieving an antecedent, albeit restricted by scope not c-command. The reviewer asks why such a retrieval process wouldn't be susceptible to gender mismatch effects. Although we are not in a position to propose a retrieval mechanism for D-type pronouns at this time, we

note that our results suggest an implementation distinct from Kush's ACCESSIBLE-based system so that GMMEs are not predicted. We must leave it to future research to address under what range of conditions the parser pursues a D-type analysis instead of a standard variable pronoun analysis.

[16]A D-type pronoun would induce a Condition C violation in the C-Command/Match condition. Some researchers have argued that all personal pronouns have a D-type analysis (Elbourne 2005, Patel-Grosz and Grosz to appear). We leave this option for further research.

[17]For instance, if these pronouns are interpreted as definite descriptions whose NP complement is elided under identity with an overt noun, then the search for that noun must be understood. The present hypothesis is that such a search does not lead to gender mismatch effects. However, Grosz et al. 2014 provide processing evidence against the NP-deletion account and in support of a dynamic binding account. Their focus was on donkey configurations, where the antecedent is an indefinite and the quantification is provided by a different element (e.g. an *if*-clause). In our configurations, the antecedent NP is part of the phrase that carries the quantificational force.