

## RESEARCH

# Processing adjunct control: Evidence on the use of structural information and prediction in reference resolution

Jeffrey J. Green<sup>1</sup>, Michael McCourt<sup>2</sup>, Ellen Lau<sup>2</sup> and Alexander Williams<sup>2</sup><sup>1</sup> University of Illinois at Urbana-Champaign, US<sup>2</sup> University of Maryland, USCorresponding author: Jeffrey J. Green ([jgreen88@terpmail.umd.edu](mailto:jgreen88@terpmail.umd.edu))

The comprehension of anaphoric relations may be guided not only by discourse, but also syntactic information. In the literature on online processing, however, the focus has been on audible pronouns and descriptions whose reference is resolved mainly on the former. This paper examines one relation that both lacks overt exponence, and relies almost exclusively on syntax for its resolution: adjunct control, or the dependency between the null subject of a non-finite adjunct and its antecedent in sentences such as *Mickey talked to Minnie before \_\_\_ eating*. Using visual-world eyetracking, we compare the timecourse of interpreting this null subject and overt pronouns (*Mickey talked to Minnie before he ate*). We show that when control structures are highly frequent, listeners are just as quick to resolve reference in either case. When control structures are less frequent, reference resolution based on structural information still occurs upon hearing the non-finite verb, but more slowly, especially when unaided by structural and referential predictions. This may be due to increased difficulty in recognizing that a referential dependency is necessary. These results indicate that in at least some contexts, referential expressions whose resolution depends on very different sources of information can be resolved approximately equally rapidly, and that the speed of interpretation is largely independent of whether or not the dependency is cued by an overt referring expression.

**Keywords:** reference resolution; adjunct control; pronouns; prediction; structural information

## 1 Introduction

A speaker of (1) can use both *Mickey Mouse* in the main clause and *he* in the embedded clause to refer to Mickey Mouse.

(1) Mickey Mouse talked to Minnie Mouse before he ate.

To understand this use of the name *Mickey Mouse*, it might be enough for a listener to access its representation in his or her mental lexicon, as this might be linked to the singular concept MICKEY MOUSE. But understanding *he* requires more than lexical access, since the pronoun itself does not lexically encode concepts like MICKEY MOUSE. Instead, interpreting *he* also requires consultation of other representations in memory, such as a representation of the discourse in which the pronoun occurs, of the syntax of the local sentence, of the speaker's interests and intent, and so on.

Much work in psycholinguistics has therefore been directed at the online processing of pronouns and other anaphoric expressions, since its study promises to illuminate the mechanisms by which disparate information sources are integrated in language comprehension

(Ehrlich & Rayner 1983; Blanchard 1987; Arnold et al. 2000; Stewart, Pickering & Sanford 2000; Kehler & Rohde 2013). Given the potential complexity of this task and the difficulties of measuring it experimentally, our initial questions only concern timing: How long does it take to resolve reference? And in general, is resolution initiated immediately upon perception of the anaphoric expression? Or do we, often enough, resolve reference only when our practical goals demand it? There is active pursuit of these questions in the literature (see, e.g. Stewart Holler & Kidd 2007; Karimi & Ferreira 2016; for a discussion on other instances of shallow processing, see Ferreira, Bailey & Ferraro 2002).

Seminal work by Arnold et al. (2000) using the visual-world paradigm has shown that pronouns can be rapidly interpreted to infer the referent intended by the speaker, especially when gender and number features on the pronoun agree with a single entity in the discourse. In their experiments, participants viewed pictures on a screen and listened to descriptions in which a pronoun referred to one of the two characters in the picture. Results showed that when the two characters mismatched in gender, participants looked to the correct character within approximately 200 ms after the offset of the pronoun, suggesting that they were able to use the gender features of the pronoun to successfully resolve reference by that time.

However, such studies target only one specific type of anaphoric dependency, in which the cue to anaphora is an audible noun phrase, and its resolution is guided mainly by properties of the discourse—like what its topic is (Crawley, Stevenson & Kleinman 1990; Grosz, Weinstein & Joshi 1995; Kehler & Rohde 2013) or how it could continue most coherently (Hobbs 1979; Kehler et al. 2008; Kehler & Rohde 2013)—or perhaps also by the attentional state of the comprehender (Arnold & Lao 2015). But there are other anaphoric dependencies that share neither property. The referential expression may be inaudible (such as the null *pro* that is said to be present as the subject of sentences like the Spanish *Te amo*), or reference may be resolved mainly by the syntax (as with English *himself*). In this paper, we investigate a case of both, where no referential expression is audible and the anaphoric dependency is resolved by the syntax. We ask if it matters to the timing of anaphora resolution whether the surface cue to anaphora is an audible pronoun, like *he* in (1), or instead a non-finite participial verb in a control structure, like *eating* in (2).

(2) Mickey Mouse talked to Minnie Mouse before eating.

A speaker can use (2) just like (1), to say that Mickey talked to Minnie before he, Mickey, ate. But now the process of understanding the speaker is different in two important ways. First, understanding that it was Mickey who ate depends on different sorts of information in the two cases. For (2) but not (1) the sentence itself determines the interpretation, due to its structure and meaning.<sup>1</sup> Only (1) can be used to say something else—for example, that Mickey talked to Minnie before Donald ate—since *he* may be used to refer to any male salient in the discourse. Second, only (1) contains an *overt* and *unambiguous* sign of reference to a particular eater, namely *he*. For (2), in contrast, the signal of such reference is both implicit and temporarily ambiguous. It is implicit, because it consists not in the lexical semantic properties of any one word, but in the fact that *eating* is here the predicate of a non-finite clause; and such clauses, when lacking an audible subject, generally have their entailed subject role filled anaphorically. And it is temporarily ambiguous, because at the point in which it is encountered, it is in principle uncertain whether *eating*

<sup>1</sup> This is only true when “Obligatory Control” is involved. See Landau (2017; forthcoming) and Green (2019a; b) for discussions on when this is not the case.

is a signal to anaphora, as (3) makes plain. Here *eating* is not the predicate of any clause, but rather the gerundival subject of one, and on this use it does not signal anaphoric reference to any particular eater.

(3) Mickey Mouse talked to Minnie Mouse before eating was forbidden.

Determining whether these differences in understanding synonymous uses of (1) versus (2) impact the timing of reference resolution is an important step toward understanding which aspects of processing anaphora are fully general, and which are instead specific to the kind of anaphoric dependency involved—for example, whether the referential expression is audible or inaudible, and whether its resolution is guided mainly by discourse or syntax.

In contrast to pronouns, there has been relatively little work on the timing of the resolution of control structure anaphora. And the majority of what we do know comes from studies of complement control as in (4), rather than adjunct control (5).

- (4)
- a. Mickey<sub>i</sub> managed [<sub>i</sub> to eat].
  - b. Mickey<sub>i</sub> promised Minnie [<sub>i</sub> to eat].
  - c. Mickey persuaded Minnie<sub>i</sub> [<sub>i</sub> to eat].

(5) Mickey<sub>i</sub> talked to Minnie [before <sub>i</sub> eating].

Studies on complement control (e.g. Boland, Tanenhaus & Garnsey 1990; Demestre et al. 1999) have generally found that the anaphor can be processed quickly. But complement control differs from both adjunct control and pronominal anaphora in an important way: the matrix verb gives early cues that a control structure is coming as well as about what the referent of the null subject will be. After the verb *persuade* in (4c), for example, it is likely that an infinitival complement is to follow, and that its null subject will be controlled by the direct object. It is therefore plausible that the rapid processing seen in complement control structures is due at least in part to information conveyed by the main-clause verb. With adjunct control, the participial clause is not selected by any element of the matrix clause, and accordingly its processing may differ from complement control. Betancort, Carreiras & Acuña-Fariña (2006) found slowdowns in the processing of one type of adjunct control when compared to complement control in an eyetracking while reading study, suggesting that verb information may have led to prediction of control and/or the referent of the null subject. Because adjunct control is less predictable than complement control, it may make for a more minimal comparison with audible pronouns, since in general the occurrence of a pronoun is also not strongly predicted by a prior verb.

This paper compares the processing of audible pronouns and the implicit anaphora in adjunct control structures to examine the speed with which structural information can be used in reference resolution. In two experiments, we use visual-word eyetracking to measure the timecourse of reference resolution in sentences such as (6).

- (6) Mickey<sub>i</sub> ran into Daisy<sub>j</sub> in front of the school...
- a. ...[before <sub>i</sub> picking up a blue ball].
  - b. ...[before he<sub>i</sub> picked up a blue ball].

There are several differences between anaphora involving overt pronouns and adjunct control that might cause the processing and interpretation of the null subject in adjunct

control structures to proceed relatively slowly, but there are also reasons to expect it to happen quickly. In what follows, we outline the reasons for both expectations.

As was already mentioned, one might expect the processing of adjunct control to proceed slowly because control structures have no overt morpheme dedicated to reference in their subject position, in contrast to when there is an overt pronoun. Instead, the first indication that a referent is needed is given indirectly by the non-finite morphology (*-ing*) on the embedded verb. Therefore, if processing is taking place incrementally, then at the embedded verb, comprehenders must interpret the verb to identify the event it corresponds to, notice that the verb is missing a subject,<sup>2</sup> and (if indeed this prompts immediate resolution of its reference) determine which character the speaker intended to refer to with that missing subject.<sup>3</sup> Furthermore, as mentioned above, the non-finite verb does not unambiguously indicate that a control relation is necessary, and this may not become clear until several words after the verb. If (7a) is spoken on its own, then an anaphoric dependency for the null subject is necessary; it must be understood that Mickey talked to Minnie before he, Mickey, ate pizza at the park. But if (7a) is continued with (7b), then the non-finite clause becomes the gerundival subject of the adjunct, and no anaphoric dependency is necessary.<sup>4</sup> Because the structure of the adjunct will not be clear to the incremental processor at the point of the verb, comprehenders may wait to interpret the null subject until the structure is disambiguated.

- (7) a. Mickey talked to Minnie before eating pizza at the park...  
 b. ... was forbidden.

It has been demonstrated, though, that listeners often do not wait for disambiguating information before building a preferred parse (e.g. Marslen-Wilson 1975; Kutas, DeLong & Smith 2011). Therefore, hearing the verb may be enough to cause people to immediately attempt to resolve the structural ambiguity in favor of a control structure, and to quickly identify the arguments of the verb. In other words, hearing a verb makes it likely that a subject is needed, and people may therefore look for a potential subject and form an anaphoric dependency at the earliest possible point, despite the fact that that dependency may not end up being necessary.

But even if at the non-finite verb listeners immediately assume a control dependency is present, using structural information to retrieve a referent from memory may still be more difficult than using the morphological and discourse information relevant for overt pronouns. In cue-based retrieval models of sentence processing (e.g. Lewis & Vasishth 2005), structural dependencies between elements of a sentence may be difficult to use as a cue to anaphora resolution (Kush, Lidz & Phillips 2015). If this is the case, then reference resolution in (6a) may be slower than in (6b), due to the difficulty in retrieving the antecedent based on structural information. On the other hand, Parker & Phillips (2017)

<sup>2</sup> This is not to say, necessarily, that that missing subject is present syntactically. A major question in the study of adjunct control concerns exactly what the adjunct control dependency is. For simplicity, we will assume the presence of a null subject, and use the “PRO” label for clarity. However, our results and conclusions are not dependent on a PRO analysis of the null subject of control structures, or on the presence of a syntactically-represented subject. For a discussion on the debate over the representation of PRO, see Hornstein (2003).

<sup>3</sup> A reviewer suggests the fact that the verb provides the theta role of the null subject may counteract this potential slowdown, since it provides conceptual information about the antecedent, and the plausibility of potential arguments has been shown to affect resolution of other, similar dependencies such as in filler-gap constructions (Garnsey, Tanenhaus & Chapman 1989).

<sup>4</sup> A reviewer notes that fully interpreting the sentence still requires someone to do the eating, even if it is an arbitrary “anyone”. Even if that were to be true, though, this is crucially different from the anaphora we are describing, because it does not involve coreference with a previous expression.

argued that structural information not only *can* be used as a cue to antecedent retrieval, but that it is weighed more heavily in the retrieval process for reflexives such as *himself*, which similarly require a structurally accessible antecedent. If this is the case, then using structural information may not cause a slowdown in reference resolution.

Another reason anaphora resolution in adjunct control may be fast is that words like *before* are often followed by control structures.<sup>5</sup> Listeners may therefore predict a control structure before even encountering the verb, making it easier to identify the control dependency and begin the retrieval process.

Our results show that when adjunct control structures are highly frequent within the experiment, listeners are just as quick to resolve reference when they hear the non-finite verb in a sentence such as (6a) as they are when they hear the overt pronoun in a sentence like (6b). This suggests that the large differences between the two forms of anaphora—that one is implicit and relies mainly on structural information and the other is explicit and relies on mostly morphological and discourse information—are in at least some contexts irrelevant to the timing of anaphora resolution. When adjunct control structures are less frequent within the experiment, however, anaphora resolution in adjunct control slows in comparison to overt pronouns. We argue that this is not due to greater difficulty in using structural information, but rather to difficulty in identifying the presence of a control dependency.

## 2 Experiment 1

In a design similar to that of Arnold et al. (2000), our first experiment uses visual-world eyetracking to examine the timecourse of interpretation of the null subject in adjunct control. Several experiments have shown that upon encountering a pronoun, comprehenders often look to the image of the character in a visual-world scene that they believe the speaker is referring to with the pronoun (e.g. Arnold 1998; Arnold et al. 2000; Arnold & Lao 2015). This is taken as evidence for the timing of resolution of the pronoun, especially when the task encourages participants to attend to what the sentence is talking about by having them verify the sentence they hear against the picture it describes. We similarly expect that upon encountering the non-finite verb in a sentence like (6a), comprehenders will look to the referent evoked by the null subject if it is being interpreted immediately. Looks to the correct referent will therefore be taken as evidence that reference resolution has been successfully completed. We measure the timecourse of looks to that referent, comparing it to the timecourse of the interpretation of an overt pronoun used to refer to the same character, as in (6b), as a baseline.

Both (6a) and (6b) involve coreference with the main clause subject. Pronouns are often biased to corefer with a prior subject rather than with an object (e.g. Arnold et al. 2000). In this way, English subjects seem to have a special role in discourse (Grosz, Weinstein & Joshi 1995). As a measure of whether a subject bias was in effect, we included a control condition with a pronoun coreferring with the main clause object. If, for example, participants are faster at resolving both subject-oriented pronouns and the null subject than object-oriented pronouns, this may be due simply to a bias to look at the character corresponding to the subject. If, on the other hand, participants are just as quick to resolve both types of pronouns, then it is unlikely that reference is being strongly influenced by such biases.

<sup>5</sup> In the Corpus of Contemporary American English (COCA) (Davies 2008–), *before*, *after*, and *while* are followed by a gerundival verb 11.3%, 13.2%, and 16.6% of the time, respectively.



## 2.1 Materials and methods

### 2.1.1 Participants

Thirty participants (20 female, 9 male, 1 declined response) were recruited at the University of Maryland. Participants were all native speakers of American English, and were at least 18 years of age (mean age = 22.5). Each received \$10 for their time. Three additional participants were excluded, two for low accuracy on comprehension questions (less than 80%), and one for equipment failure.

### 2.1.2 Materials

Participants listened to auditory descriptions of illustrated scenes, and were asked to indicate whether the description matched the scene. The scenes involved two main participants selected out of a set of four well-known characters: Mickey and Minnie Mouse, and Donald and Daisy Duck. The auditory stimuli were composed of two sentences: an initial sentence to focus attention on one of the two characters, and a second sentence describing the content of the picture. The second sentence began with a main clause predicate, with the two characters as arguments (the subject and the object/indirect object). This was followed by a prepositional phrase describing the third element of the image in order to draw attention away from the two characters immediately before the critical region. The description concluded with a temporal adjunct headed by *before*, *after*, or *while*. A sample set of stimuli is given in Table 1. Each set was paired with a single image, as will be described below.

In a  $2 \times 3$  design, stimuli varied with respect to which character was the attentional focus of the first sentence of the preamble (the subject or the object of the main clause), and with respect to the referential expression used as the subject of the temporal adjunct. Because the preference for pronouns to co-refer with the subject of a preceding clause can be reduced if attention is placed on another character at the beginning of the discourse (Arnold & Lao 2015), the focus manipulation was included as an attempt to account for the possibility that participants might be strongly biased to expect subsequent reference to the character named by the subject of the main clause. In the “Focus<sub>subj</sub>” condition, attention was drawn to the referent of the preamble’s subject, and in the “Focus<sub>obj</sub>” condition, the referent of its object. As for the subject of the temporal adjunct, in what we call the “PRO” condition, the temporal adjunct was non-finite, including a null “PRO” subject coreferential with the subject of the main clause. In the other conditions, the temporal adjunct was finite and had an overt pronominal subject coreferential with either the main clause subject (“pron<sub>subj</sub>” conditions) or the main clause object (“pron<sub>obj</sub>” conditions). In the critical stimuli the two characters were of opposite gender, and the pronominal subject (*he/she*) was therefore unambiguous. For the PRO condition, because the subject was implicit, the first indication that a referential dependency was required was the non-finite

**Table 1:** Sample stimuli.

Condition	Preamble
Focus <sub>subj</sub>	Look there’s Minnie! She was talking to Mickey in front of a huge tree after
Focus <sub>obj</sub>	Look there’s Mickey! Minnie was talking to him in front of a huge tree after
	<b>Critical region</b>
PRO	<b>putting</b> on a nice new bow, and they seem to be having a good time.
pron <sub>subj</sub>	<b>she</b> put on a nice new bow, and they seem to be having a good time.
pron <sub>obj</sub>	<b>he</b> put on a nice new hat, and they seem to be having a good time.

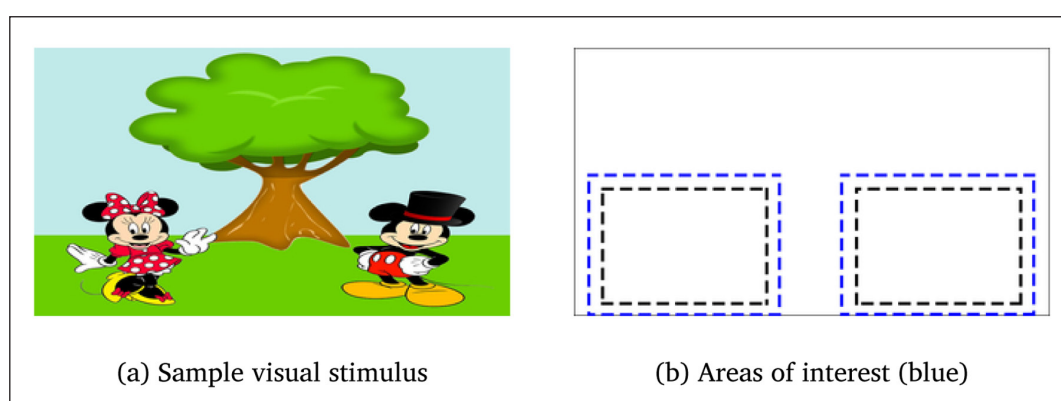
verb. Within the experiment this cue was also unambiguous, identifying the main clause subject as the antecedent of the null subject of the temporal adjunct.<sup>6</sup>

Auditory stimuli were recorded in a sound-attenuated room by a male speaker of American English. Each stimulus set was recorded in parts, with each part in a carrier phrase, as in (8). After normalizing for intensity, the two recordings were spliced together at the point indicated by the vertical bar. This was done in order to minimize potentially confounding differences in the acoustics of the stimuli before the critical word. The splice point immediately preceded the non-finite verb in the PRO conditions, and the pronoun otherwise, so that there would be no coarticulatory clues about the content of the upcoming critical word. After splicing, the stimuli were filtered to remove noise. All recording and processing of the auditory stimuli was done in Praat (Boersma 2001; Boersma & Weenink 2017). The average duration of the critical word in the PRO conditions (the non-finite verb) was 314 ms. In the  $\text{pron}_{\text{subj}}$  and  $\text{pron}_{\text{obj}}$  conditions, the average critical word duration was 153 and 152 ms, respectively.

- (8) a. Look there's Mickey! Minnie was talking to him in front of a huge tree after  
| Tom left.  
b. Tom left after | putting on a nice new bow, and they seem to be having a  
good time.

Visual stimuli consisted of scenes containing two characters located in the bottom right and bottom left corners of the picture, equally distant from the center and of roughly equal size. The stimuli were counterbalanced with respect to which two characters were seen together and which character appeared on the left or right side of the screen. The pictures also contained a third prominent, inanimate element that was placed at the top center. A sample image is given in Figure 1(a). The two main regions of interest in the eyetracking analysis were fixed across trials, and included the entire image of the character and a small surrounding area to allow for gaze drift. This is illustrated in Figure 1(b). Any gazes not in these two areas of interest were classified as looks to “other.”

Thirty sets of stimuli were distributed across six lists using a Latin square design. Each participant saw 30 critical trials—5 from each condition. For all of the critical items, the



**Figure 1:** Sample visual stimulus and regions of interest. **(a)** is the image displayed for the sample stimuli in Table 1. In **(b)**, dotted black rectangles give the regions in which the characters were contained, and dotted blue rectangles show the regions of interest used in analysis.

<sup>6</sup> As discussed in §1, a non-finite verb in general is ambiguous in that it can be followed by continuations that do not require a control dependency, but such continuations were absent in these experiments.

auditory description matched the scene depicted. Twenty fillers were also included in each list, for a total of 50 items. The fillers were similar in form to the critical items, and contained an equal number of PRO, pron<sub>subj</sub>, and pron<sub>obj</sub> sentences. The only difference was that the filler stimuli contained discrepancies between the picture and the auditory description of it; the location of the discrepancy varied across filler items to include all parts of the auditory description (in both the preamble and the adjunct), and could involve any of the mentioned objects or characters; for a small number of fillers, detection of the discrepancy depended on successful reference resolution in the adjunct.<sup>7</sup> Varying the location of the discrepancy across items encouraged participants to attend to each part of the visual stimulus as it was mentioned in the description and to resolve all referential expressions. Two practice items were also included: one where the description matched the image, and one where it did not.

### 2.1.3 Procedure

Eye movements were recorded using an EyeLink 1000 tower-mounted eyetracker (S.R. Research, Mississauga, Ontario, Canada), interfaced with a PC, with a sampling rate of 1000 Hz. Visual stimuli were displayed on a 23-in. LCD monitor, approximately 104 cm away from where participants were seated. Viewing was binocular, but only the right eye was recorded. Auditory stimuli were presented at a comfortable volume using speakers situated next to the monitor.

The experiment was implemented using the Experiment Builder software (S.R. Research, <http://www.sr-research.com/eb.html>). At the beginning of the experiment, participants were introduced to the four characters to ensure familiarity. This was followed by a nine-point calibration procedure and presentation of the two practice items. Stimuli were then presented in five blocks of 10 items each, with recalibration at the beginning of each block and otherwise as needed, and with drift-correction prior to each trial. Visual and auditory stimuli were presented simultaneously, with the image remaining on the screen for the duration of the auditory description. At the conclusion of the description, the image disappeared, and participants were asked to indicate by pressing a button on a remote control whether the description matched the image.

Each participant was presented with the items from one of the six lists. The order of experimental and filler items within that list was randomized for each participant.

## 2.2 Main results

Eye movements were analyzed using the `eyetrackingR` package (Dink & Ferguson 2015) in R (R Core Team 2019). Because we were interested in the effect of the critical word (either the pronoun or the non-finite verb) on looks, and because it takes roughly 200 ms to plan and execute a saccade (Tanenhaus et al. 1995), our window of analysis began 200 ms after the onset of the critical word (the non finite-verb or the pronoun) and extended until 200 ms after the onset of the disambiguating word (the first word in the sentence after the critical word that unambiguously identifies the target referent, e.g. *bow/hat* in the items from Table 1). The average window length across trials was 1263 ms. Fifteen trials were excluded due to high trackloss (greater than 25%) in this window, resulting in a loss of 1.7% of the data. The mean trackloss in the remaining trials was 2.6%.

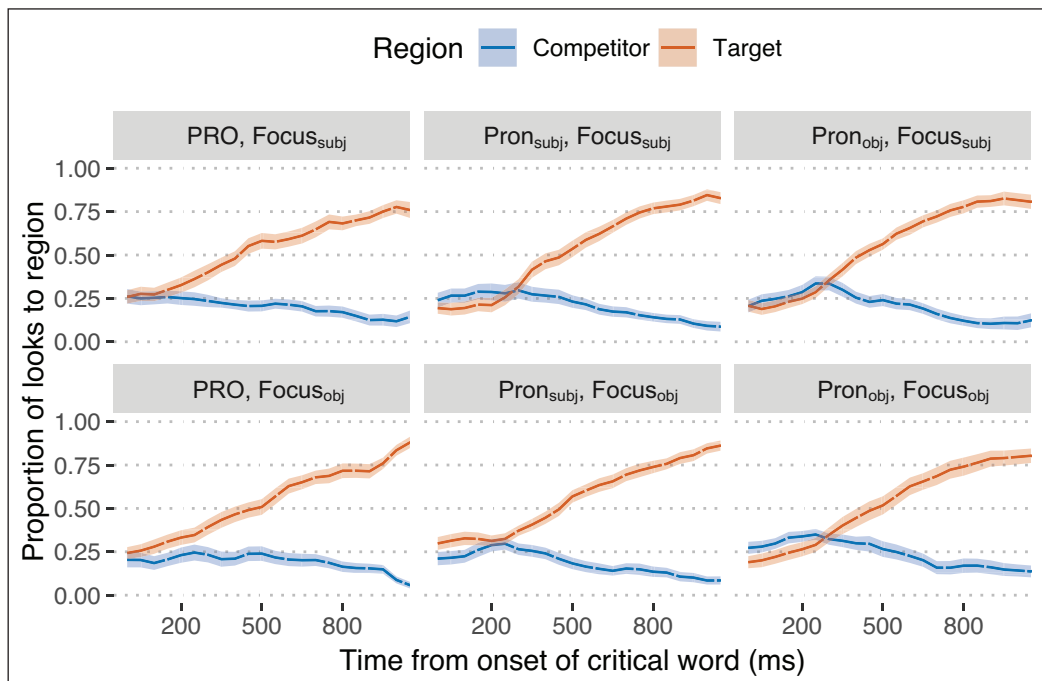
<sup>7</sup> A reviewer notes that since these fillers also included PRO or pronouns, they could in principle be analyzed with the critical items. The main reason they were not included was that in roughly half of the fillers, the discrepancy between the auditory stimulus and the visual scene occurred before the adjunct; for those items, participants could correctly answer that the scene and description did not match without having to resolve reference in the adjunct.



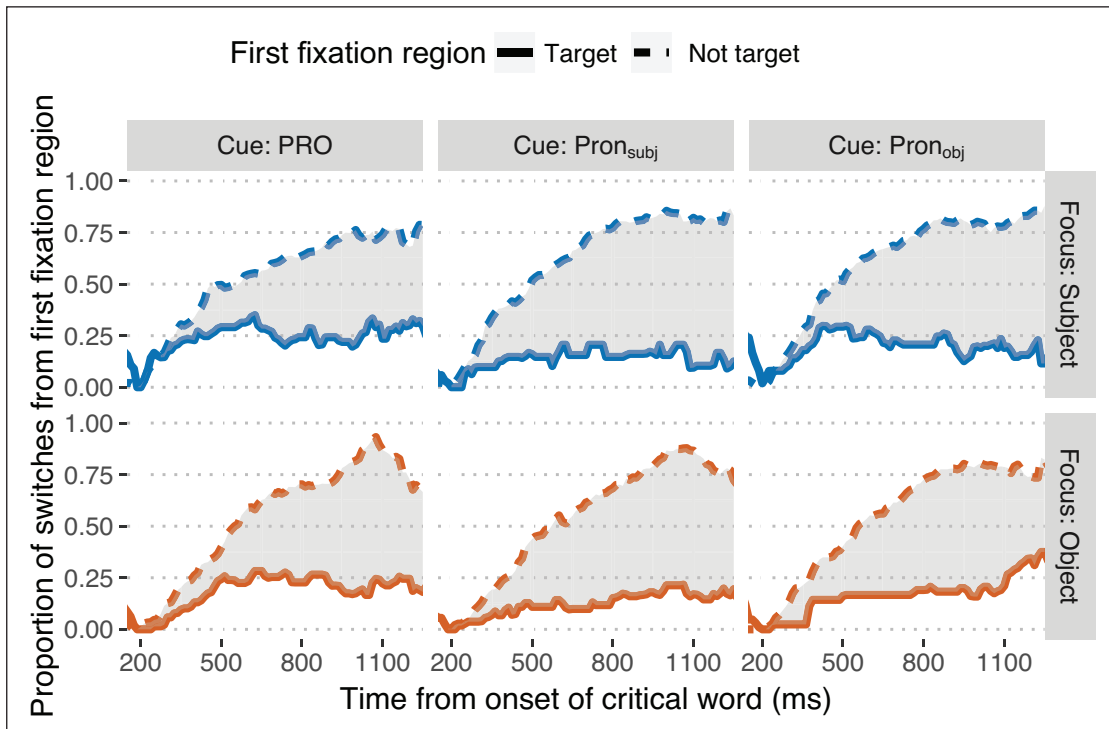
To determine whether participants looked to the target character after the onset of the critical word, we compared looks to the target or competitor in the critical window for each condition. The proportion of looks to each character at each time point is plotted in Figure 2. In each condition, looks to the target appear to diverge from looks to the competitor by around 300–400 ms, suggesting that successful resolution of reference was achieved by that point. This is consistent with the results in Arnold et al. (2000) for the interpretation of overt pronouns whose gender features unambiguously identify a referent.

Although traditional analyses comparing the total proportion of looks or time spent looking at the target or competitor (e.g. Arnold et al. 2000; Arnold & Lao 2015; Mirman, Dixon & Magnuson 2008) provide information about interpretation preferences, they do not directly measure differences in the *timing* of successful reference resolution across conditions. In order to determine how quickly reference was resolved in the PRO conditions compared to when there was an overt pronoun, we used an onset-contingent analysis (Fernald et al. 2008; Yurovsky & Frank 2015) to examine how quickly participants switched their gaze away from where they were fixated at the beginning of the critical window. Fixation regions were determined by taking the average gaze location over a rolling 50 ms window at each time point for each item. The initial fixation region was measured at 200 ms after the onset of the critical word in order to capture only those eye movements that could have been caused by hearing the critical word, rather than movements that had already been planned prior to the critical word. When the initial fixation point could not be determined due to trackloss, the trial was excluded from further analysis. The proportion of trials in which participants switched from their initial fixation region (either the target or any non-target region) at each time point is given in Figure 3.

If participants have successfully resolved reference, then they should be quicker to switch *toward* the target (moving away from non-target regions) than *away* from the target. We therefore calculated the first time point at which looks switched away from the initial fixation region. If a participant's fixation did not switch regions throughout the entire



**Figure 2:** Experiment 1: Proportion of looks to the character corresponding to the target (orange) or competitor (blue) for the PRO (left), pron<sub>subj</sub> (middle), and pron<sub>obj</sub> (right) conditions when either the subject character (top) or object character (bottom) was given focus.

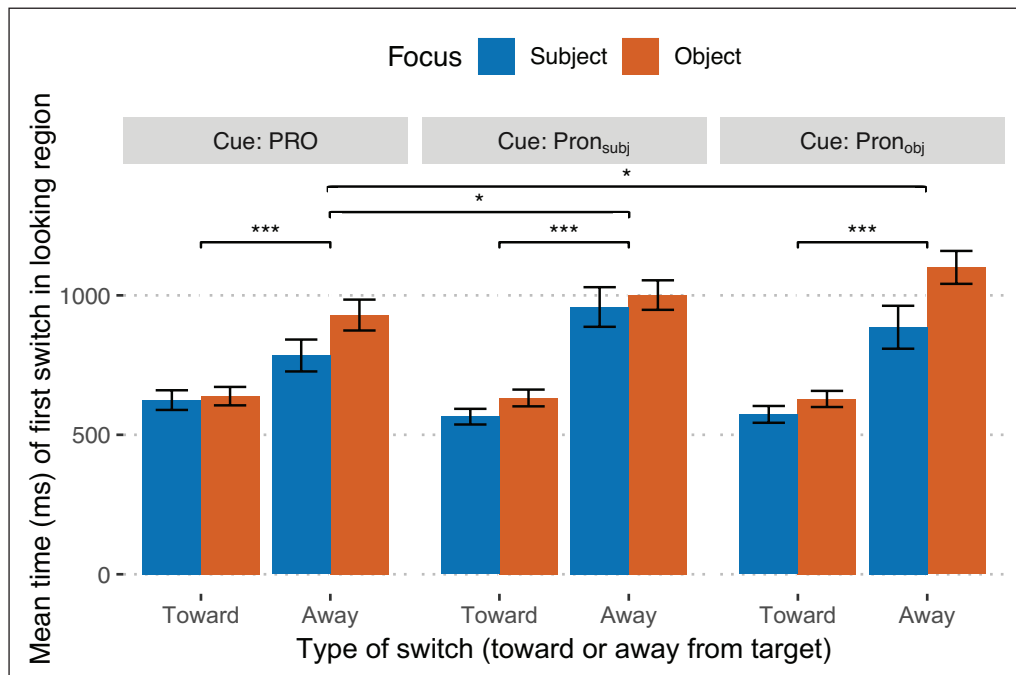


**Figure 3:** Experiment 1: Proportion of trials where participants switched from their initial fixation region across conditions, either toward the target (dotted line) or away from it (solid line).

critical window in a trial, then the first switch time was coded as the final time point in the analysis window for that item, indicating that there had been no change prior to that point. Mean switch times for each condition based on initial fixation region are plotted in Figure 4. The effects of the experimental manipulations on first switch times were tested using linear mixed-effects models (Bates et al. 2015). P-values were obtained by performing type-III ANOVAs on the models using the `lmerTest` package (Kuznetsova, Brockhoff & Christensen 2017), which uses Satterwaithe approximations to calculate degrees of freedom. Fixed effects in the model were the factors *switch type* (toward vs. away from the target character), *cue* (“PRO”,  $\text{pron}_{\text{subj}}$ ,  $\text{pron}_{\text{obj}}$ ), and *focus* (subject vs. object), as well as interactions between switch type and the other factors. Random intercepts were included for participants and items (including random slopes led the model to not converge). The results are given in Table 2.

The main effect of switch type was due to faster switches toward the target than away from it (estimated difference = 342 ms, SE = 25.8). The main effect of focus was the result of faster switches generally when focus was on the subject compared to when it was on the object (estimated difference = 87.2 ms, SE = 25.2), but since focus did not interact with whether switches were toward or away from the target, this effect is taken to be orthogonal to reference resolution and will not be discussed further.

Table 3 gives the results of pairwise Tukey-adjusted follow-up comparisons on the significant interaction, which were completed using the `emmeans` package (Lenth 2020) in R. The support of the null or alternative hypothesis was further tested for each comparison by computing Bayes factors using the `ttestBF` function of the `BayesFactor` package (Morey & Rouder 2018). These factors represent the ratio between the likelihood of the alternative hypothesis compared to the null hypothesis; a factor of  $K$  indicates that the alternative hypothesis is  $K$  times as likely as the null hypothesis. A value of  $K < 0.1$  should be taken as strong evidence for the null hypothesis, and a value of  $K$  between 0.33 and 0.1, as moderate evidence. On the other hand, a value of  $K > 10$  indicates strong



**Figure 4:** Experiment 1: Average time of first switch of looking region during critical window. Error bars represent standard error of the mean.

**Table 2:** Experiment 1: Main results.

Factor	F-value	p-value
Switch type	92.73	<0.001
Cue	1.87	0.15
Focus	12.15	<0.001
Switch type × Cue	3.94	0.02
Switch type × Focus	2.34	0.13
Switch type × Cue × Focus	0.99	0.41

**Table 3:** Experiment 1: Pairwise follow-up comparisons.

Comparison	Estimate	SE	t-ratio	p-value	K-value
Switch toward target – away from target by Cue					
PRO	-235	42.4	-5.55	<0.001	>100
Pron <sub>subj</sub>	-393	45.6	-8.63	<0.001	>100
Pron <sub>obj</sub>	-396	45.1	-8.78	<0.001	>100
PRO – Pron <sub>subj</sub> by switch type					
Toward target	32.16	32.8	0.98	0.59	0.20
Away from target	-125.91	52.0	-2.42	0.04	1.36
PRO – Pron <sub>obj</sub> by switch type					
Toward target	27.67	32.9	0.84	0.68	0.17
Away from target	-133.13	51.7	-2.58	0.03	1.43
Pron <sub>subject</sub> – Pron <sub>obj</sub> by switch type					
Toward target	-4.48	32.0	-0.14	0.99	0.11
Away from target	-7.22	55.1	-0.13	0.99	0.18

evidence for the alternative hypothesis, and a value between 3 and 10 as moderate support for it (Jeffreys 1961; Lee & Wagenmakers 2013).

Regardless of cue, switch times were faster toward the target than away from it; there was no difference between cues in switch times toward the target, but looks away from the target appear to have happened earlier in the PRO condition than for the pronoun conditions. However, although the differences pass a significance threshold of  $\alpha = 0.05$ , the Bayes factor comparisons indicate only very weak support for a real difference.

## 2.3 Discussion and exploratory analyses

### 2.3.1 Discussion of main results

Experiment 1 investigated the process of reference resolution for the null (PRO) subject of non-finite temporal adjuncts in comparison to overt pronouns. For PRO in the adjuncts used here, reference is determined based on structural features of the sentence. For overt pronouns, on the other hand, reference resolution is influenced by discourse factors, guided by the morphological features of the pronoun; structural features are only important insofar as they affect information structure.

Our results indicate that upon encountering a non-finite verb in a temporal adjunct, listeners are just as quick to look toward PRO's referent as they do an overt pronoun's. Participants may have looked away from the target sooner when interpreting PRO than they did with pronouns, but the evidence only weakly supports such a difference. If we infer that consistent looks to the correct character indicate successful reference resolution, then these results suggest that the interpretation of PRO in adjunct control is just as fast as the interpretation of a pronoun, at least in some instances.

The interpretation of PRO in our items required participants to recognize and interpret the non-finite verb, realize that it required a subject, and find a referent for that subject based on their knowledge of the adjunct control dependency. Where present, this dependency requires the use of structural information about the adjunct and its host sentence; PRO's referent is tied to that of the closest structurally-accessible antecedent (Hornstein 1999; Parker, Lago & Phillips 2015; Gerard 2016). To interpret an overt pronoun, participants needed to recognize the pronoun and find a salient referent in the discourse that matched the pronoun's gender features. Because there was no difference in how quickly participants looked toward the correct character when the cue to reference was a non-finite verb compared to when it was a pronoun, listeners appear to be able to use structural information in searching for a referent for PRO just as quickly as they can use the morphological features on the pronoun. That listeners would be able to use either kind of feature, structural or morphological, to resolve reference without delays in timing is not clearly consistent with cue-based retrieval accounts (e.g. Lewis & Vasishth 2005), in which, it has been claimed, syntactic relations such as the one involved in adjunct control should be difficult to use as retrieval cues in reference resolution (Kush, Lidz & Phillips 2015).

This fast interpretation of the null subject is also somewhat surprising considering how different a cue to reference resolution a non-finite verb is in comparison to an overt pronoun in terms of semantics and morphology. On top of those differences, the duration of the non-finite verb in the PRO condition was twice as long as the duration of the overt pronouns in the other conditions, with an average of 314 ms vs. 153 ms, meaning that even if participants were able to use either structural or morphological information just as quickly to search for a referent, the bottom-up cue to anaphora simply was not heard as quickly in the PRO condition as in the pronoun conditions, especially considering that the part of the verb that indicates a control dependency, the *-ing*, is at the end of the word.

Why, then, was the interpretation of PRO so fast? We investigate possible reasons in the exploratory analyses below.

### 2.3.2 Exploratory analyses

There are at least two additional factors beyond the bottom-up input of the non-finite verb that could have contributed to early resolution of PRO. First, participants may have been predicting upcoming reference to the main clause subject, making it easier to resolve control dependencies satisfying such predictions. Second, in addition to or instead of predicting upcoming reference to the subject character, participants may have been making structural predictions such that they assumed a control dependency would be needed. Either of these types of prediction could in turn have two different sources: individual item contexts, or the high frequency of coreference with the subject and of control structures within the experiment. If participants were actively predicting a control structure or reference to the subject character, either based on individual items or on frequency-based expectations, then the speed of looks to the correct character in the PRO condition would not necessarily reflect the speed of bottom-up interpretation of PRO. Instead, interpretation of PRO upon encountering a non-finite verb in an adjunct clause may actually take longer than resolution of an overt pronoun, but the difference may not have been seen if participants began that process early in the PRO condition based on prediction.

The effect on reference resolution of prediction about who is likely to be mentioned next is well known (e.g. Kehler & Rohde 2013). In order to measure whether individual item contexts led to prediction that a particular character would be mentioned next in the adjunct clause, or to structural predictions about an upcoming control dependency, we collected sentence completion (cloze) data in a separate experiment on Amazon MechanicalTurk ( $n = 60$ ). Participants were presented with the images for critical items from Experiment 1 and the preamble of the corresponding sentences (e.g. “Look there’s Mickey! Minnie was talking to him in front of a huge tree after \_\_\_\_.”), and asked to complete the sentence with the first thing that came to mind. We analyzed responses for whether they began with unambiguous reference to either character, and whether that reference was via a control dependency, an overt pronoun, or some other referential expression. Responses began with clear reference to the character named by the subject in 61% of responses to Focus<sub>subj</sub> preambles and 56% of responses to Focus<sub>obj</sub> preambles, with ranges of 13–97% and 7–93%, respectively. As for the likelihood of adjunct control, responses began with a control structure in 56% of responses to Focus<sub>subj</sub> preambles and 48% of responses to Focus<sub>obj</sub> preambles, with ranges of 27–87% and 23–80%, respectively. These results indicate that in the critical items of Experiment 1, reference to the subject character and the use of a control structure were highly predictable, but that this predictability varied across items. Therefore, it is possible that participants were using referential or structural predictions to get a head start on the resolution of PRO in some of the items in Experiment 1, leading to faster looks toward and/or slower switches away from the target than when control structures were not predicted.

In addition to item-based predictions, it may have been the case that the experimental context had a large enough proportion of either control structures or of initial reference to the subject character in the adjunct that participants began making relevant structural or referential predictions within the experiment. Two-thirds of our items required participants to look at the subject at the critical word, since that was the case for both the PRO and the pron<sub>subj</sub> conditions. However, if a subject bias were strongly contributing to the speed of interpretation of the null subject, then we would expect the PRO and pron<sub>subj</sub> conditions to both be faster than the pron<sub>obj</sub> condition. This was not the case; participants



were just as fast to look at the correct character in both pronoun conditions. In the  $\text{pron}_{\text{obj}}$  condition, this could not have been due to a subject preference. However, it could have been the case that participants were using different strategies for pronouns than for PRO. Even if a preference to look at the subject character did not influence looks in the pronoun conditions (perhaps because of the unambiguous gender information), that does not rule out the possibility that participants relied on such a preference in the PRO condition without actually resolving reference. And although control structures were found in only a third of the items participants saw, this is still more frequent than these structures are seen in real life.<sup>8</sup> It is therefore possible that the rapid resolution of reference in the PRO condition was due to participants learning to expect an upcoming control structure within the experiment, or to assume that any non-pronoun encountered in the temporal adjunct was a non-finite verb.

In order to test whether looks in PRO items were influenced by high predictability of reference to the subject character or of control structures, or by a learned strategy within the experiment, we fit a linear mixed effects regression on the PRO items measuring the influence on first switch times of switch type, focus, order within the experiment, cloze probability of reference to the subject character, and cloze probability of PRO (i.e. of a control structure), as well as interactions between these factors that included switch type, with random intercepts for participants. Order within the experiment and the two cloze measures were coded as a scaled, centered, continuous variables over items, based on when each item was seen by each participant and on responses to corresponding items in the offline cloze task. Significant results of this model are given in Table 4; all other effects in the model were non-significant ( $p > 0.1$  in each case).

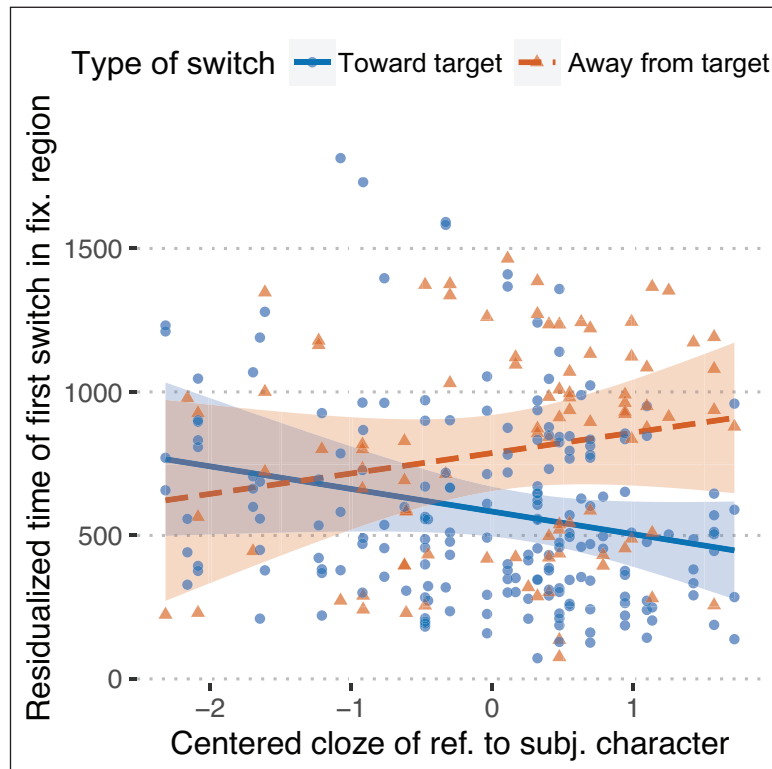
Besides the main effects of switch type and focus (which were also seen in the main analysis above), these results do not indicate that switch times were affected by experiment order or by cloze probability of a control structure. The only possible new effect revealed is a marginal interaction between switch type and cloze probability of reference to the subject character, illustrated in Figure 5. Using the `interactions` package in R (Long 2019), which estimates the slope of continuous effects in interactions, we determined that this marginal interaction was driven by looks away from the target; items with higher cloze probability of reference to the subject character saw numerically longer fixations on the target before looking away, but this slope of this effect did not reach significance ( $p = 0.10$ ).

If the interpretation of PRO were affected by the likelihood of referring to the subject character in the adjunct, then a similar effect would be expected to be evident in the  $\text{pron}_{\text{subj}}$  items, since the target character is the same for both conditions. As a comparison to the PRO items, we performed the same analysis as above for the  $\text{pron}_{\text{subj}}$  condition, testing the effects of order, cloze of reference to the subject character, and cloze of a control structure on switch times toward or away from the target character. Significant effects in this regression are given in Table 5 (for all other effects,  $p > 0.05$ ).

**Table 4:** Experiment 1, PRO analysis: Significant results.

Factor	F-value	p-value
Switch type	24.45	<0.001
Focus	5.80	0.02
Switch × Cloze <sub>subj.ref.</sub>	3.06	0.08

<sup>8</sup> See Footnote 5.



**Figure 5:** Experiment 1, PRO items: Residualized time of first switch of looking region during critical window based on cloze probability of reference to the subject character and switch type. Each point represents a single observation. Regression lines represent the estimated effect of cloze of subject reference. Shading indicates the standard error of the estimated slope.

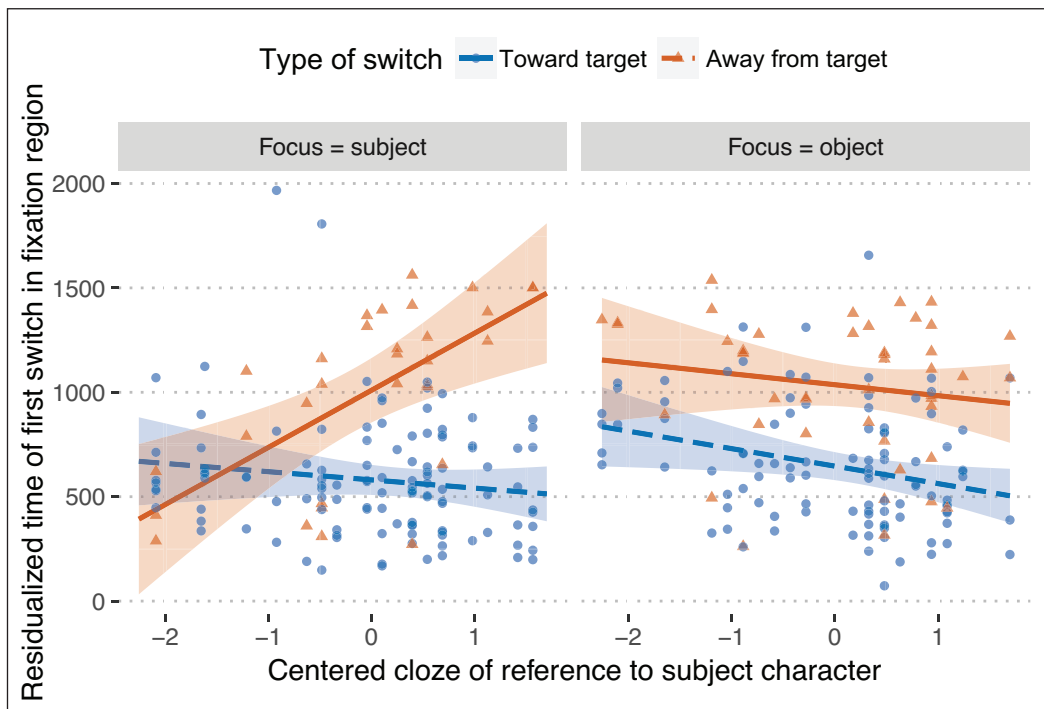
**Table 5:** Experiment 1,  $\text{pron}_{\text{subj}}$  analysis: Significant results.

Factor	F-value	p-value
Switch type	60.24	<0.001
Switch × Cloze <sub>subj.ref.</sub>	12.20	<0.001
Switch × Focus × Cloze <sub>subj.ref.</sub>	5.84	0.003
Switch × Order × Cloze <sub>subj.ref.</sub> × Cloze <sub>PRO</sub>	4.51	0.01

Follow-ups to the interactions between switch type and cloze of subject reference, and between those factors and focus indicate that higher cloze of reference to the subject character led to longer fixations on the target when focus was on the subject ( $p < 0.01$ ), and faster switches toward the target when focus was on the object ( $p = 0.03$ ). This is illustrated in Figure 6. Analysis of the four-way interaction between switch type, order, and the two cloze measures revealed that early in the experiment, cloze of subject reference only had an effect on switches toward the target when cloze of PRO was also high ( $p < 0.05$  when cloze of PRO was high,  $p > 0.05$  otherwise). This effect disappears, however, over the course of the experiment. There was no significant interaction between order and the cloze measures in switches away from the target ( $p > 0.1$ ).

### 2.3.3 Discussion of exploratory analyses

The interpretation of PRO in this experiment was surprisingly fast. If rather than just relying on bottom-up information to identify and resolve the necessary dependency, participants were predicting an upcoming control structure or reference to the subject character, then this fast interpretation could be explained. However, the results of the exploratory



**Figure 6:** Experiment 1,  $\text{pron}_{\text{subj}}$  items: Residualized time of first switch of looking region during critical window based on focus, cloze probability of reference to the subject character, and switch type. Each point represents a single observation. Regression lines represent the estimated effect of cloze of subject reference. Shading indicates the standard error of the estimated slope.

analysis show little effect of such prediction. There was no effect of item-wise predictability of a control structure in the PRO condition, suggesting that participants' looks were not being driven by structural predictions. There was also no effect of order within the experiment, suggesting that participants did not develop strategies over the course of the experiment to aid in their resolution of PRO. Higher predictability of reference to the subject character did lead to numerically longer looks at the target in the PRO condition, but this effect was small and statistically non-significant.

In contrast to PRO, reference resolution in the  $\text{pron}_{\text{subj}}$  condition was strongly affected by prediction. When focus was on the subject character, participants were much more likely to look longer at the target the more predictable reference to it was. This suggests that interpretation of an overt pronoun in this experiment was highly sensitive to participants' referential predictions, at least when initial focus was placed on that character. This strong effect, in contrast to the lack of an effect in the PRO condition, further suggests that the fast interpretation of PRO was not due to referential predictions. When focus was on the object character, higher predictability of subject reference was associated with faster looks toward the target in the  $\text{pron}_{\text{subj}}$  condition. This again suggests an effect of referential prediction, but since participants were also numerically faster to look away from the target, this may have just been a trend toward faster switches overall. We take this result to mean that focusing on the object made participants less likely to rely on strong predictions about reference to the subject character. The fact that they were still more likely to look toward the target than away from it meant that they were successfully using the morphological information on the pronoun to resolve reference, regardless of predictions made.

A surprising result in the  $\text{pron}_{\text{subj}}$  analyses was that the cloze probability of a control structure played a role, interacting with predictability of subject reference and experiment

order. The interaction indicated that early on in the experiment, participants were faster to look toward the target when both subject reference and a control structure were predictable. Why would prediction of a control structure affect interpretation of overt pronouns, but not the interpretation of PRO? Perhaps the speed-up in looks toward the target was due to multiple different factors all indicating reference to the subject character, including item-based prediction of subject reference, prediction of a control structure (which also generally entails reference to the subject character), the easily-accessible morphological information from the pronoun, in addition to the high proportion of items with subject reference in the adjunct. Perhaps for PRO, structural information from the control dependency was not as easy to access, and so early looks toward the target were not affected, despite possible structural or referential predictions. As to why this effect was only seen early in the experiment in the  $\text{pron}_{\text{subj}}$  items, participants may have stopped using item-based structural predictions, if they learned due to the high proportion of items with a control dependency within the experiment that item-based structural predictions were unnecessary.

From these results, it appears that the resolution of PRO based on structural information provided by the control dependency was as fast or nearly as fast as the resolution of an overt pronoun. But if looks in the  $\text{pron}_{\text{subj}}$  condition were influenced by referential and/or structural predictions, why weren't the PRO items? If we are correct that looks toward the target were not influenced by predictability in the PRO condition because structural information was more difficult to access, then why weren't switch times toward the target slower for PRO versus overt pronouns? And why were looks away from the target not affected by referential predictions for PRO the way they were for overt pronouns? One possibility is that the high proportion of items with control structures in Experiment 1 led participants to expect control structures generally, independent of individual item contexts. We attempted to test this by including experiment order in our exploratory linear regression and found no effect in the PRO condition. However, a linear order effect would only be seen if the strategy participants adopted affected looks linearly over time. It is possible that participants were affected by the within-experiment frequency of PRO so quickly that it did not show up as a linear effect, but rather as an effect over the whole experiment. We test this possibility in Experiment 2.

### 3 Experiment 2

Experiment 2 included all the items from Experiment 1, in addition to extra fillers designed to reduce the overall proportion of control structures and the within-experiment bias to refer to the subject character. Although these changes do not affect the predictability of reference to a single character or of control structures for any individual item, since the same critical items were used in both experiments, they do change the overall biases within the experiment, which in turn may affect participants' interpretive strategies.

#### 3.1 Materials and methods

##### 3.1.1 Participants

Thirty participants (19 female, 11 male) were recruited at the University of Maryland. Participants were all native speakers of English, and were at least 18 years of age (mean age = 21). Twenty-nine of the participants were compensated with course credit, and one with \$5. Two additional participants were excluded for high trackloss (> 33%).

##### 3.1.2 Materials

All stimuli from Experiment 1 were included. In addition to the 20 fillers from Experiment 1, 40 new filler items were added. These consisted of 10 items similar to the original fill-

ers, but all with a pronoun coreferring with the main clause object as the subject of the temporal adjunct, in order to reduce the overall proportion of items with subject reference in the adjunct. The remaining 30 new fillers had a subject in the adjunct referring to something or someone other than the two main characters (e.g. *Look there's Donald! Minnie found him outside of Daisy's house after **Daisy** kicked him out for being rude.*); often this was the third prominent element in the image. Half of these 30 fillers were true (i.e. the image matched the description), and half were false, as were the other 10 new fillers. This brought the total number of items each participant saw to 90 (the 30 critical items from Experiment 1 and 60 fillers), 45 of which should have been judged true. These new fillers reduced the within-experiment proportion of items containing control structures from 32% to 17.8%. The recording of the auditory stimuli for these fillers followed the same procedure as in Experiment 1, including splicing the temporal adjunct into the stimulus after the preposition introducing it.

### 3.1.3 Procedure

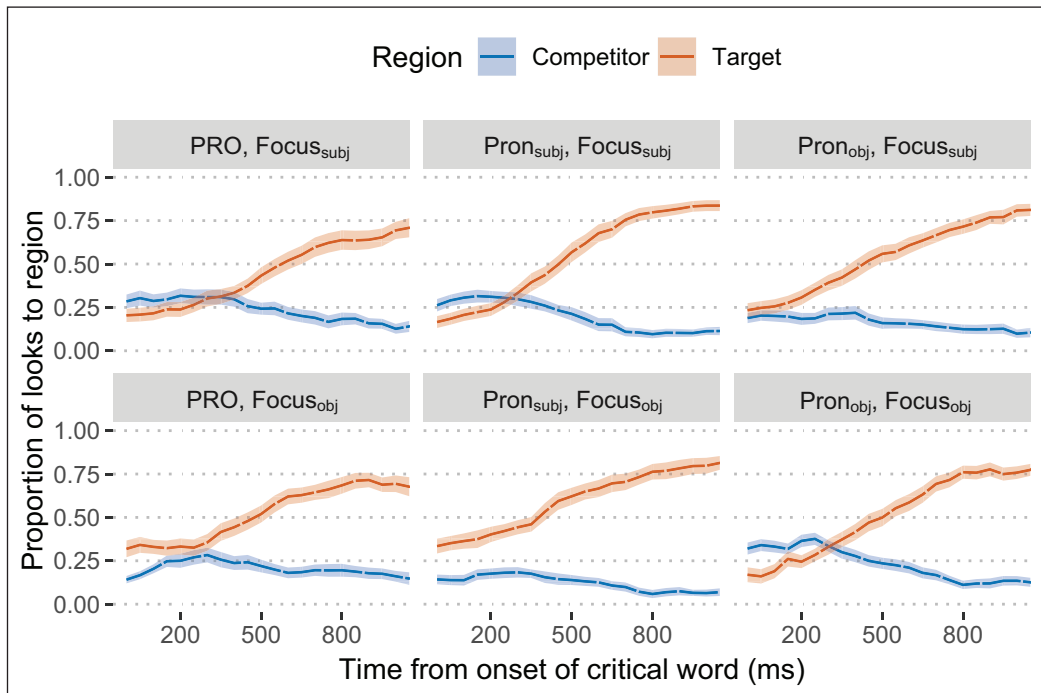
Stimuli were presented in nine blocks of 10 items each. The procedure followed that of Experiment 1 in all other respects.

## 3.2 Results

### 3.2.1 Main results

Twenty-seven trials were excluded for high trackloss (greater than 25%), resulting in a loss of 3% of the data. The mean trackloss per trial in the remaining data was 3.5%.

Looks to the target or competitor beginning at the onset of the critical word in each condition are plotted in Figure 7. As with Experiment 1, looks to the target diverge from the competitor in each case by around 400 ms, suggesting that reference resolution was successful by that point. However, unlike with Experiment 1, focus appears to have had a strong early effect. At least in the pron<sub>subj</sub> condition, when focus was on the character

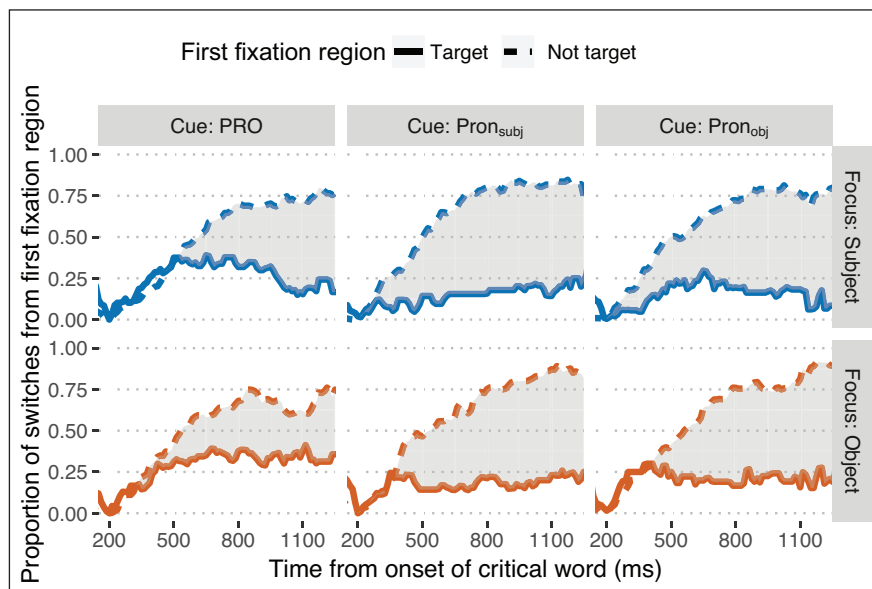


**Figure 7:** Experiment 2: Proportion of looks to the character corresponding to the target (orange) or competitor (blue) for the PRO (left), pron<sub>subj</sub> (middle), and pron<sub>obj</sub> (right) conditions when either the subject character (top) or object character (bottom) was given focus.

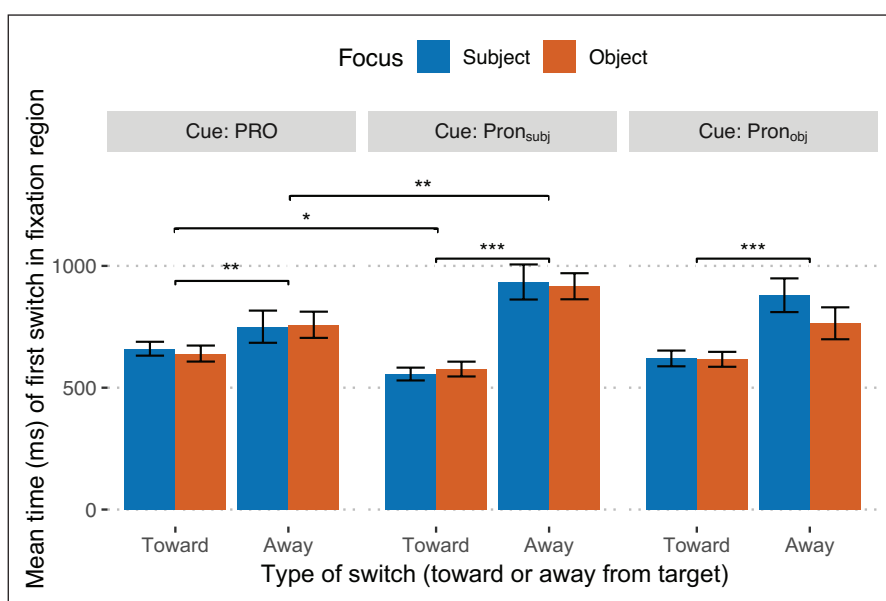


corresponding to the main clause object, there were more fixations on the target already at the onset of the critical word. Because this divergence appears so early, it could not be due to interpretation of the pronoun. This provides even greater justification for the use of an onset-based analysis, as it takes into account looking region at the onset of the critical region, and successful interpretation is determined based on whether participants are more likely to switch their gaze toward the target than away from it, rather than the overall proportion of fixations. Any biases in looking prior to the critical region are therefore controlled for. The proportion of trials where participants switched their fixation are given in Figure 8.

Mean switch times by condition based on whether that switch was toward or away from the target are given in Figure 9. The results of a linear mixed-effects model testing



**Figure 8:** Experiment 2: Proportion of trials where participants switched from their initial fixation region across conditions, either toward the target (dotted line) or away from it (solid line).



**Figure 9:** Experiment 2: Average time of first switch of looking region during critical window. Error bars represent standard error of the mean.

the effects of switch type, cue, and focus on first switch times (with random intercepts for participants and items) are given in Table 6. As with Experiment 1, the main effect of initial fixation region was due to participants being significantly faster to look toward the target than away from it. Pairwise comparisons on the interaction between switch type and cue are given in Table 7. These comparisons reveal that although participants were more likely to look toward the target than away from it for all three cue types, there was only weak evidence for such a difference in the PRO condition. In addition, participants were faster to look away from the target and slower to look toward it in the PRO condition compared to the  $\text{pron}_{\text{subj}}$  condition, as supported by both the frequentist and Bayesian comparisons.

### 3.2.2 Exploratory analyses

As with Experiment 1, two other linear mixed-effects models tested whether looks in the PRO and  $\text{pron}_{\text{subj}}$  conditions were affected by the cloze probability of reference to the subject character or of a control structure, or by trial order within the experiment. Including random effects in the  $\text{pron}_{\text{subj}}$  model led to non-convergence due to singular fit; therefore, for this analysis only, a general linear regression without random effects was used. Significant results are given in Table 8 for the PRO condition, and in Table 9 for the  $\text{pron}_{\text{subj}}$  condition ( $p > 0.05$  for all other effects).

**Table 6:** Experiment 2: Main results.

Factor	F-value	p-value
Switch type	50.93	<0.001
Cue	0.91	0.40
Focus	0.90	0.34
Switch type × Cue	4.47	0.01
Switch type × Focus	2.34	0.13
Switch type × Cue × Focus	0.42	0.79

**Table 7:** Experiment 2: Pairwise follow-up comparisons.

Comparison	Estimate	SE	t-ratio	p-value	K-value
Switch toward target – away from target by Cue					
PRO	-119	44.4	-2.69	0.007	2.29
$\text{Pron}_{\text{subj}}$	-364	43.6	-8.35	<0.001	>100
$\text{Pron}_{\text{obj}}$	-214	44.6	-4.79	<0.001	>100
PRO – $\text{Pron}_{\text{subj}}$ by switch type					
Toward target	81.8	33.4	2.45	0.04	6.69
Away from target	-162.6	51.3	-3.17	0.004	6.08
PRO – $\text{Pron}_{\text{obj}}$ by switch type					
Toward target	31.5	32.8	0.96	0.60	0.19
Away from target	-63.0	52.7	-1.20	0.46	0.31
$\text{Pron}_{\text{subject}}$ – $\text{Pron}_{\text{obj}}$ by switch type					
Toward target	-50.3	33.2	-1.52	0.28	0.50
Away from target	99.6	52.0	1.92	0.13	0.45

**Table 8:** Experiment 2, PRO analysis: Significant results.

Factor	F-value	p-value
Switch type	4.34	<b>0.04</b>
Switch × Focus × Cloze <sub>subj.ref.</sub> × Cloze <sub>PRO</sub>	3.30	<b>0.04</b>

**Table 9:** Experiment 2, pron<sub>subj</sub> analysis: Significant results.

Factor	F-value	p-value
Switch type	71.07	<b>&lt;0.001</b>
Switch × Order × Cloze <sub>subj.ref.</sub>	3.10	<b>0.047</b>

In PRO items, there was a significant four-way interaction between switch type, focus, and the two cloze measures. This interaction was driven by the object focus condition, in which the effect of cloze probability of subject reference on switch times away from the target was significant when the cloze of PRO was high ( $p = 0.01$ ), with higher cloze of subject reference leading to slower switches away from the target, but not with mid to low cloze of PRO ( $p > 0.05$ ), as illustrated in Figure 10. Switches toward the target were not affected ( $p > 0.1$ ). The interaction between switch type and the cloze measures did not persist when initial focus was on the subject.

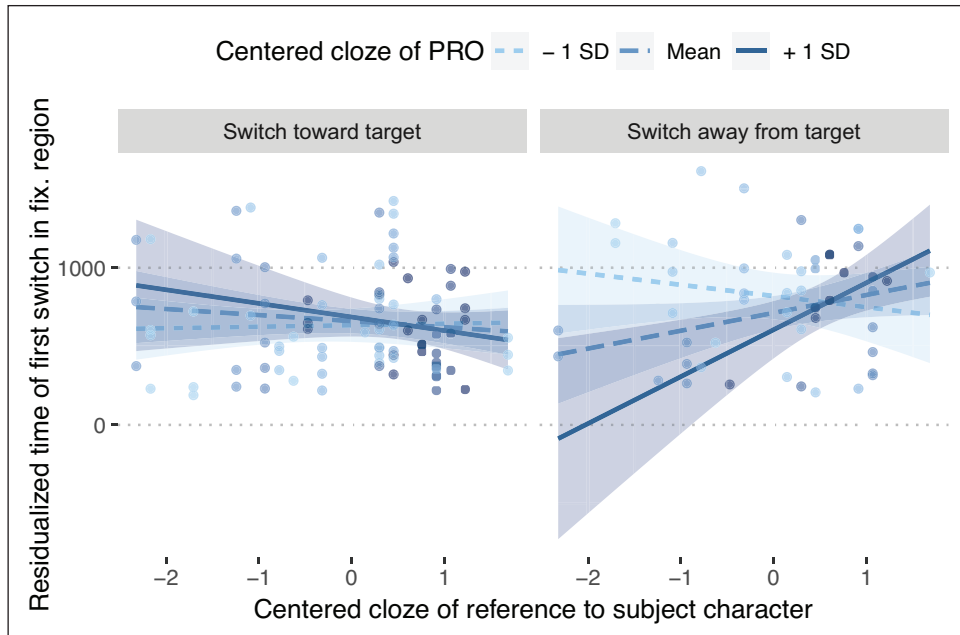
In pron<sub>subj</sub> items, there was a significant three-way interaction between switch type, order, and cloze probability of subject reference. Follow-up analysis revealed that there was a significant effect of order on switches away from the target when cloze of subject reference was high ( $p = 0.03$ ), but not when it was mid to low ( $p > 0.1$ ), as illustrated in Figure 11; participants looked at the target for longer when cloze of subject reference was high as the experiment went on. There was no effect on switches toward the target.

### 3.3 Discussion

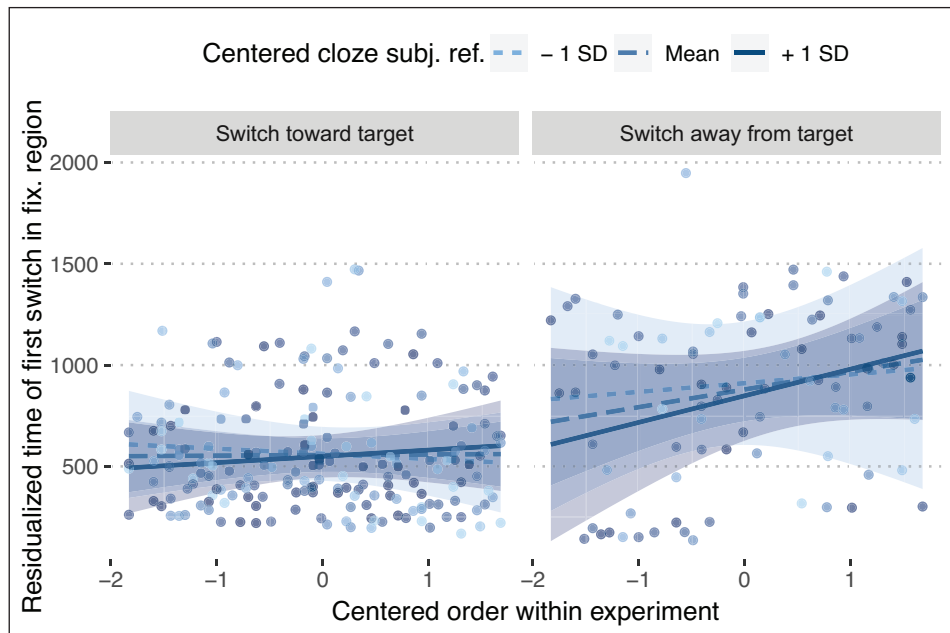
#### 3.3.1 Difficulty in resolving PRO

In Experiment 1, PRO appeared to be interpreted just as fast as overt pronouns. We hypothesized that this was in part due to the high proportion of control structures and/or of items that contain elements in the adjunct clause that make reference to the character named by the main-clause subject; participants may have adjusted quickly to these within-experiment frequencies and begun to predict PRO and resolve its reference earlier than would normally be possible. If that were the case, then we would expect reference resolution to be slower in Experiment 2, in which the within-experiment frequency of both control structures and subject reference was reduced. This prediction was confirmed. In Experiment 2, participants were both slower to look toward the target and faster to look away from it in the PRO condition than in the pron<sub>subj</sub> condition. Switch times in the pron<sub>obj</sub> condition were between the other two, with no significant difference between the pron<sub>obj</sub> condition and either of the others. The fact that reference resolution in the pron<sub>obj</sub> condition may have been somewhat more difficult than in the pron<sub>subj</sub> condition could be due a lingering subject preference for the pronouns. Importantly, despite this possible preference, the PRO condition, which also included reference to the subject character, was still more difficult than the pron<sub>subj</sub> condition.

This result strongly suggests that in Experiment 1, the interpretation of PRO was indeed influenced by the high within-experiment frequency of a control structure. In Experiment 2, participants could no longer depend on the high frequency of control structures to predict the presence of PRO, and instead had to rely more on bottom-up input. The fact that



**Figure 10:** Experiment 2, PRO,  $focus_{obj}$  items: Residualized time of first switch of looking region during critical window, based on cloze probability of reference to the subject character, cloze probability of PRO, and switch type. Each point represents a single observation, with darker points indicating higher cloze of PRO. Regression lines represent the estimated effect of cloze of subject reference across three values of the cloze of PRO: the mean (dashed line), and one standard deviation above (solid line) or below (dotted line) the mean. Shading indicates the standard error of the estimated slope.



**Figure 11:** Experiment 2,  $pron_{subj}$  items: Residualized time of first switch of looking region during critical window based on trial order, cloze probability of reference to the subject character, and switch type. Each point represents a single observation, with darker points indicating higher cloze of subject reference. Regression lines represent the estimated effect of order across three values of the cloze of subject reference: the mean (dashed line), and one standard deviation above (solid line) or below (dotted line) the mean. Shading indicates the standard error of the estimated slope.

reference resolution slowed in comparison to overt pronouns could be due to more difficulty either in using structural information as opposed to morphological/gender information to resolve reference, or simply in recognizing that a referential dependency was necessary, since the bottom-up cue to anaphora was longer in duration for the PRO versus pronoun conditions.

Additionally, although the difference in switch times toward or away from the target in the PRO condition still reached significance, the Bayes factor analysis indicates only weak evidence for that difference, as compared to both pronoun conditions, in which there is extreme evidence for faster looks toward the target than away from it. The exploratory analyses of the PRO items indicate that the small difference that was present was due to both structural and referential predictions. When focus was on the object, participants looked longer at the target only if reference to the subject character was predicted *and* if that reference was predicted to be realized in a control structure. This is different from what was seen in Experiment 1, in which only referential predictions affected PRO's interpretation. This again suggests that in Experiment 1, participants were adjusting to the high frequency of control structures within the experiment; when this frequency was lowered, participants' interpretations were still aided by structural predictions, but only based on individual item contexts.

### 3.3.2 Pron<sub>subj</sub> items

Turning to the exploratory analyses of the pron<sub>subj</sub> items, in Experiment 1, looks away from the target were affected by the cloze probability of reference to the subject character when the focus was on that character. Experiment 2 saw a similar effect, in that participants looked longer at the target the more predictable reference to the subject was, but only later on in the experiment. This may be due to the lower proportion of items in Experiment 2 with reference to the subject character. Participants could not use within-experiment frequency expectations, but did rely more on referential predictions as the experiment went on.

The pron<sub>subj</sub> condition of Experiment 1 also saw an effect of the cloze probability of a control structure in an interaction with other measures. We hypothesized that participants were faster to look toward the target when multiple factors all led to a prediction of subject reference, including the high proportion of items in the experiment with subject reference. When in Experiment 2 this proportion decreased, the effect of item-wise prediction of a control structure disappeared. If it had any effect, it was too small to detect without all the other factors involved.

## 4 General discussion

### 4.1 Summary of findings

The experiments reported here provide evidence for the rapid interpretation of PRO in temporal adjuncts, modulated by the predictability of reference to PRO's antecedent as well as by how likely a structure containing PRO was to occur. Our results show that participants can use the structural information inherent in the control dependency to resolve PRO just as quickly as they can use gender information to resolve an overt pronoun, but only when a structure containing PRO is predicted. The strongest effect of such prediction in these experiments was due to the high within-experiment frequency of such structures seen in Experiment 1. A weaker effect was also seen with a lower within-experiment frequency of PRO in Experiment 2 based on predictions arising from individual item contexts. When neither factor led to prediction of PRO, its interpretation slowed significantly.



#### 4.2 Use of structural information in reference resolution

There are two main tasks a listener faces in interpreting referential expressions: recognition and resolution. Listeners must recognize that a speaker is attempting to refer to someone, and they must also use the information available—from the discourse context, structural and morphological features of the speaker's utterance, etc.—to decide who that someone is, i.e. to resolve reference. The slowdown in the interpretation of PRO seen in Experiment 2, especially for items where PRO or reference to the character corresponding to the main clause subject was not predicted, could in principle be due to difficulties in either of these tasks.

First, the interpretation of PRO may be slowed due to difficulties in using structural information versus morphological/discourse information in reference resolution. However, a large body of literature has argued that structural constraints can apply at the earliest stages of processing for reflexives such as *himself*, which are similar to PRO in the way structure guides resolution (for an overview, see Dillon 2014). Dillon's explanation of this fact is that for reflexives, antecedent retrieval involves serial search of the syntactic structure. Although Dillon argues that such a search is rapid, such a search may be slower due to its serial nature than the cue-based retrieval argued to be involved in pronoun resolution, which searches all potential antecedents simultaneously (Lewis & Vasishth 2005; for discussion, see Kush, Lidz & Phillips 2015). That being said, Parker & Phillips (2017) argue that structural information *can* be encoded as a searchable, direct-access cue, and that it may in fact have a stronger weight than morphological information. If this is true, and listeners are able to use structural information so early in the processing of reflexives, then it would be surprising if they were unable to similarly do so to resolve PRO in the current experiment.

The other possible explanation for the slow interpretation of PRO when it is not predicted is that it simply took listeners longer to recognize that reference resolution was necessary. This seems likely for a number of reasons. First, the bottom-up cue that a referential dependency is needed in the PRO sentences, namely the nonfinite verb, had a longer duration than *he/she*, the cue to anaphora in the pronoun conditions, by an average of 160 ms. Furthermore, it is the final syllable of this cue (the *-ing*) that indicates the presence of a control structure. It is therefore possible that participants simply did not realize that reference resolution was needed as quickly in unpredicted control structures as when there was an overt pronoun or when a control structure was predicted. Second, although the non-finite verb was the cue to anaphora in the PRO condition, this verb does not itself have reference to an individual as its semantic function, unlike a pronoun. Lexically, the verb expresses an event concept, or the concept of a relation to an event. The truth conditions of this concept will entail the role that would be associated with the subject, but the signal of anaphoric *reference* to an individual bearing that relation comes only from the verb's grammatical context. In addition to interpreting the verb, listeners must also find one of its participants in order to establish the referential dependency necessary to interpret PRO. This is one instance where, as Van Berkum (2008: 376) put it, it is not the case that “[f]irst you recognize each of the words, then you look up their meaning in your mental dictionary, and then, using syntax to guide the combination, you simply combine the meanings so that you know what I said.” Instead, recognition of the nonfinite verb in this particular context activates both the event concept lexically expressed by the verb, and the cue to anaphoric resolution of the subject argument. The extra semantic processing required may therefore have delayed the initiation of reference resolution in PRO versus pronoun items.

There are many different kinds of referential expressions and types of anaphora. Each may rely on syntactic, discourse, and conceptual sources of information in different ways.

The interpretation of PRO in adjunct control structures is heavily dependent on structural features of the sentence, while pronouns, although restricted by their “ $\phi$  features” (person, gender, and number), rely more on discourse information. If the slowdown in the resolution of PRO when it was not predicted is due to difficulty in the recognition of the need for reference resolution, and not in the resolution itself, then this would mean that structural information can be used in absence of other features just as efficiently as information guiding the interpretation of overt pronouns, once the cue to anaphora has been identified.

One additional note on the use of structural information as a cue to retrieval is necessary. Although Parker & Phillips (2017) argued that structural information is a searchable cue, they did not elaborate on the nature of that cue. In the ACT-R architecture of Lewis & Vasishth (2005), searchable cues on an NP are said to be encoded in memory chunks representing the head N. For example, the pronoun *he* searches for an NP in memory with the feature [+ masculine]. Whether a potential antecedent is in a proper structural position would be difficult to encode in such a manner.

Kush, Lidz & Phillips (2015) provide a possible solution for bound variable anaphora in sentences such as (9), in which *he* can only be bound by *any janitor* if the pronoun is c-commanded by the quantificational phrase, as in (9a).

- (9) a. Kathi didn't think **any janitor**<sub>i</sub> liked performing his custodial duties when **he**<sub>i</sub> had to clean up messes.  
 b. Kathi didn't think **any janitor**<sub>i</sub> liked performing his custodial duties, but **he**<sub>s<sub>i</sub></sub> had to clean up messes.

Kush, Lidz & Phillips argue that the pronoun triggers retrieval of an antecedent based in part on the cue ACCESSIBLE, which is only present on the memory chunk for *any janitor* in (9a). The memory chunk for *any janitor* in (9b) loses its ACCESSIBLE feature when the sentence reaches the end of *any janitor*'s scope domain, indicated by the conjunction. This does not occur for non-quantificational NPs, however. In (10), *the janitor* does not require c-command in order to corefer with a later pronoun, and so it does not lose its ACCESSIBLE feature. As a result, the pronoun *he* is able to corefer with it.

- (10) Kathi didn't think **the janitor**<sub>i</sub> liked performing his custodial duties, but **he**<sub>i</sub> had to clean up messes.

Although this allows the structural information relevant to bound variable anaphora to be encoded in a content-addressable way, this approach cannot be directly applied to the resolution of PRO. The reason is that the antecedent to PRO need not be a quantificational phrase. Because of this, there is no reason that only c-commanding antecedents would remain accessible.<sup>9</sup> The chunk representing *Minnie* in (11) would still remain active, and there would be no way to distinguish c-commanding from non-c-commanding antecedents.

- (11) Mickey<sub>i</sub> talked to Minnie<sub>j</sub> before PRO<sub>i/\*j</sub> putting on a hat.

If, however, a similar feature to Kush, Lidz & Phillips's ACCESSIBLE were to represent syntactic accessibility for any NP, then the structural information relevant to adjunct control could be encoded. In (11), the chunk for *Minnie* would lose its SYNTACCESS feature

<sup>9</sup> A reviewer notes that the same is true for other instances of bound variable anaphora that do not involve quantificational antecedents, but that still require c-command (e.g. sloppy readings in ellipsis).

as soon as the parse required adjoining a clause higher in the structure—in this case the temporal adjunct. But because the adjunct adjoins lower than *Mickey*, it keeps its SYNTAX feature, and is therefore retrievable once the control structure is recognized if PRO triggers a search only for memory chunks that have that feature.

Regardless of whether the slowdown is due to difficulty in recognizing the referential dependency, as we have argued, or in its resolution, the results of these experiments make clear that when an adjunct control dependency is highly frequent, its resolution is just as easy as the resolution of overt pronouns.

### 4.3 Prediction in the resolution of anaphora

This experiment also has implications for the role of prediction in the processing of anaphora. Effects of prediction in reference resolution have been well documented. For example, Kehler et al. (2008) provide evidence that pronoun interpretation is incrementally influenced by predictions listeners make about what coherence relations are likely to be at play as well as what discourse entities are most likely to be next mentioned (see also Kehler & Rohde 2013). In the examples in (12) from Caramazza et al. (1977), for example, listeners predict that an explanation for the first clause will be given, and their interpretation of the pronoun is biased toward whatever interpretation will satisfy that prediction. In (12a), this explanation is likely to be a description of something Mary did, and in (12b), something Jane did. Listeners may be quick to assign those referents to the pronoun, despite the fact that the rest of the sentence may favor an alternative interpretation, as (13) does.

- (12) a. Jane hit Mary because she had stolen a tennis racket.  
 b. Jane angered Mary because she had stolen a tennis racket.

- (13) Jane hit Mary because she reacts violently to criticism.

The present experiments give evidence that not only conceptual predictions, but also predictions about the upcoming structure of a sentence may affect anaphora resolution. How quickly a listener can resolve the reference of the null subject of an adjunct control structure is impacted by how strongly that control structure was predicted. This is in line with the arguments of Kehler (2008) that reference resolution is not a purely reactive process, with participants always waiting for cues before retrieving potential antecedents. Instead, listeners may actively predict upcoming structure and likely referents before any cue for reference is received.

There were two sources for effects of prediction in the present experiments. In Experiment 1, participants seem to have been affected by the high frequency of control structures within the experiment, which may have led them to predict PRO more often than they otherwise would have. This is in line with a large body of research demonstrating that statistical learning within an experiment can increase reaction times and lead participants to make new predictions (see, e.g. Wells et al. 2009; Misyak, Christiansen & Tomblin 2010; Dale, Duran & Morehead 2012; Karuza et al. 2014). In Experiment 2, when the frequency of control structures was lessened, the prediction of PRO affected resolution times only based on individual item contexts. In either case, the prediction of a control structure led to more rapid resolution of PRO than when such a prediction was likely to be absent.

One remaining question is the extent to which these kinds of structural predictions influence reference resolution in real-world language use. The frequency effect seen in Experiment 1 is unlikely to be seen in every-day situations outside the lab, simply because

adjunct control structures are generally less frequent than what was present in the experiment. In Experiment 2, individual item contexts appeared to favor prediction of control structures. But was that due to properties unique to these items or to the simple image and discourse context? Or are control structures predicted in similar sentences more generally? Although these experiments do not answer this question, it is clear that listeners *can* be influenced by structural predictions in anaphora resolution.

#### 4.4 Incremental sentence processing

These experiments also add to a large body of literature demonstrating incremental parsing and interpretation in sentence processing (e.g. Marslen-Wilson 1975; Altmann & Steedman 1988; Kamide, Altmann & Haywood 2003; Kutas, DeLong & Smith 2011; Poesio & Rieser 2011). Not only do these experiments demonstrate that reference resolution can be completed as soon as the cue to anaphora is heard, but also that comprehenders may initiate reference resolution before the sentence unambiguously indicates that it is needed. Although the non-finite verb in the PRO conditions was the major cue for reference resolution in the PRO items, it itself does not unambiguously indicate that a control relation is necessary. As was seen in (7), repeated in (14), at the non-finite verb, the sentence could still have a continuation that does not require a referential dependency between the null subject of the adjunct and the main clause subject. If (14a) is continued with (14b), the null subject receives an arbitrary interpretation, no anaphora being necessary, but this only becomes clear later in the sentence, long after the non-finite verb.

- (14) a. Mickey talked to Minnie before eating pizza at the park...  
 b. ... was forbidden.

The fact that participants assume a control structure when a continuation such as (14b) is possible is perhaps not surprising, however, since such continuations are likely less frequent. In addition, assuming a control structure rather than a structure where the non-finite verb is part of a gerundival subject may be preferred due to general processing strategies such as Minimal Attachment (Frazier & Rayner 1982), which favors parses with simpler structures.

Additionally, however, a non-finite verb in a temporal adjunct could also be part of a non-obligatory control structure, in which the null subject may refer to an antecedent that is not syntactically represented, as in (15), although this is far less common (Landau 2017; Green 2019a; b).

- (15) The pizza tasted better [after drinking root beer].

Although non-obligatory control in temporal adjuncts is used infrequently compared to obligatory control structures, where PRO is syntactically bound, and although all of the examples in our experiment did require control by the main clause subject, it is still possible that participants in principle would wait to interpret the null subject until it was clearly necessary. But this was not the case; participants quickly looked to the character corresponding to the main clause subject upon hearing the non-finite verb, evidently establishing the anaphoric control dependency at the earliest possible indication that it might be necessary.<sup>10</sup>

<sup>10</sup> Whether obligatory control would be preferred under Minimal Attachment depends on which control theory is adopted. In the Two-tiered Theory of Control (Landau 2015; forthcoming), obligatory control does indeed involve a simpler structure than non-obligatory control. In other theories, such as the Movement Theory of Control (Hornstein 1999; Green 2019a), there are no *structural* differences between obligatory and non-obligatory control, but obligatory control is still preferred.

#### 4.5 Additional implications and questions

The results of these experiments have several other relevant implications. First, they call into question a previous claim that the interpretation of PRO in adjuncts involves a “most-recent filler strategy.” Based on the results of an eyetracking-while-reading study, Betancort, Carreiras & Acuña-Fariña (2006; following Frazier, Clifton & Randall 1983; Nicol & Swinney 1989) suggest that upon encountering PRO, comprehenders first consider the most (linearly) local potential antecedent, even if they must later revise that initial interpretation to establish an anaphoric dependency between PRO and the main clause subject. If such a strategy were active in the current experiment, we would expect initial looks during the critical region in the PRO condition to be to the competitor (the referent of the main clause object), as it was the most recently mentioned potential antecedent. This was not the case in either experiment. At no point were participants more likely to look toward the competitor and away from the target than vice versa. Even when participants did not consistently look more toward the target than away from it, there were at least equal looks to the two characters. When the interpretation of PRO slowed, either both characters were under equal consideration during early stages of anaphora resolution, which on its own would counter the most-recent filler account, or more likely, the interpretation of PRO simply started later, with the subject being quickly recognized as the only potential antecedent thereafter.

Second, previous research on the processing of PRO (e.g. McCourt et al. 2015) has left open the possibility that its reference is not always instantaneously resolved. The current experiments give evidence that at least in some contexts, PRO is resolved quickly during incremental sentence processing. It’s still an open question for future work, however, whether the same retrieval mechanisms hold for control relations involving different cues. It may be the case that the retrieval mechanisms used in the adjunct control structures used here differ from some or all cases of complement control, for example. According to Landau (2015), complement control can either be the result of predication, which is also argued to be involved in obligatory adjunct control, or logophoric variable binding, which is only seen in non-obligatory control adjuncts. Logophoric complement control is especially likely to involve at least somewhat different retrieval mechanisms from what was used in the present experiments, since its resolution requires more than structural information.<sup>11</sup>

Throughout this paper, we have remained neutral with respect to the theoretical representation of what we have been calling “PRO” and the control dependency, and theories differ with respect to how PRO is resolved. In some theories, obligatory control in adjuncts is nothing more than predication (Landau 2017; forthcoming). In others, PRO’s referent is determined through a syntactic dependency akin to binding or movement (e.g. Hornstein 1999). Still others assume that control does not involve any syntactic dependency, but instead is dependent on the semantics or on pragmatic inference (e.g. Jackendoff & Culicover 2003). Although this paper does not directly bear on this debate, future work along this line has the potential to do so. We have demonstrated that PRO has a similar processing profile to overt pronouns, but that it is somewhat less sensitive to referential predictions, consistent with its being more dependent on structural sources of information. To provide evidence on the exact nature of the control dependency, it would be fruitful to directly compare the processing of PRO with that of movement relations such as filler-gap dependencies and with the processing of (secondary) predication relations.<sup>12</sup>

<sup>11</sup> See also Jeffrey, Han & Pappas (2015), which discusses the processing of transitive subject control in complements, which under Landau’s (2015) theory would be logophoric.

<sup>12</sup> A reviewer notes that if the adjunct control dependency is processed similarly to predication or to filler-gap dependencies, then it may be fruitful to examine Gibson’s (2000) Dependency Locality Theory, which considers such dependencies together.



Finally, an additional area for future research would be to compare the processing of obligatory control structures such as those examined in this paper with the processing of non-obligatory control adjuncts. Participants were quick to assume that an anaphoric dependency was needed in these experiments. If participants automatically attempt to retrieve an antecedent to PRO upon encountering the non-finite verb, then in sentences like (15), where the intended referent of PRO is not in the sentence, participants may experience processing difficulty due to retrieval failure. Such a difficulty has been given as one possible explanation for why obligatory control is so much more prevalent than non-obligatory control in adjuncts (Green 2019a; Landau forthcoming), but more empirical work is needed to confirm this.

## 5 Conclusion

This paper used visual-world eyetracking to investigate the processing of the null subject in adjunct control. It has shown that when adjunct control structures are predicted, the reference of the null subject can be resolved just as quickly as that of overt pronouns. Studying different forms of reference can shed light on how different sources of information are implemented during sentence processing, and this study contributes to this agenda by providing evidence that structural information can be immediately utilized in reference resolution, especially when aided by prediction of structures where such information is crucial.

## Ethics and Consent

This research involved human subjects, and was approved by the University of Maryland IRB (approval 00972). All subjects gave written consent prior to participating.

## Acknowledgements

Thank you to the audience at CLS 54 for helpful feedback, and to Jan Edwards and Marissa Barlaz for comments at various stages of this research.

## Funding Information

This material is based upon work supported in part by the National Science Foundation under Grants No.1449815 and No.1749407.

## Competing Interests

The authors have no competing interests to declare.

## Author Contributions

JJG and MM designed the experiments with feedback from AW and EL. JJG and MM collected the data. JJG analyzed the data and wrote the paper. All four authors participated in editing and revising the manuscript.

## References

- Altmann, Gerry & Mark Steedman. 1988. Interaction with context during human sentence processing. *Cognition* 30(3). 191–238. DOI: [https://doi.org/10.1016/0010-0277\(88\)90020-0](https://doi.org/10.1016/0010-0277(88)90020-0)
- Arnold, Jennifer E. 1998. *Reference form and discourse patterns*. Stanford, CA: Stanford University dissertation.
- Arnold, Jennifer E., Janet G. Eisenband, Sarah Brown-Schmidt & John C. Trueswell. 2000. The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition* 76(1). B13–B26. DOI: [https://doi.org/10.1016/S0010-0277\(00\)00073-1](https://doi.org/10.1016/S0010-0277(00)00073-1)



- Arnold, Jennifer E. & Shin-Yi C. Lao. 2015. Effects of psychological attention on pronoun comprehension. *Language, Cognition and Neuroscience* 30(7). 832–852. DOI: <https://doi.org/10.1080/23273798.2015.1017511>
- Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. DOI: <https://doi.org/10.18637/jss.v067.i01>
- Betancort, Moisés, Manuel Carreiras & Carlos Acuña-Fariña. 2006. Processing controlled PROs in Spanish. *Cognition* 100(2). 217–282. DOI: <https://doi.org/10.1016/j.cognition.2005.04.001>
- Blanchard, Harry E. 1987. The effects of pronoun processing on information utilization during fixations in reading. *Bulletin of the Psychonomic Society* 25. 171–174. DOI: <https://doi.org/10.3758/BF03330319>
- Boersma, Paul. 2001. Praat, a system for doing phonetics by computer. *Glott International* 5(9/10). 341–345.
- Boersma, Paul & David Weenink. 2017. *Praat: Doing phonetics by computer*. Version 6.0.29. <http://www.praat.org>.
- Boland, Julie E., Michael K. Tanenhaus & Susan M. Garnsey. 1990. Evidence for the immediate use of verb control information in sentence processing. *Journal of Memory & Language* 29(4). 413–432. DOI: [https://doi.org/10.1016/0749-596X\(90\)90064-7](https://doi.org/10.1016/0749-596X(90)90064-7)
- Caramazza, Alfonso, Ellen Grober, Catherine Garvey & Jack Yates. 1977. Comprehension of anaphoric pronouns. *Journal of Verbal Learning and Verbal Behavior* 16(5). 601–609. DOI: [https://doi.org/10.1016/S0022-5371\(77\)80022-4](https://doi.org/10.1016/S0022-5371(77)80022-4)
- Crawley, Rosalind A., Rosemary J. Stevenson & David Kleinman. 1990. The use of heuristic strategies in the interpretation of pronouns. *Journal of Psycholinguistic Research* 19(4). 245–264. DOI: <https://doi.org/10.1007/BF01077259>
- Dale, Rick, Nicholas D. Duran & J. Ryan Morehead. 2012. Prediction during statistical learning, and implications for the implicit/explicit divide. *Advances in Cognitive Psychology* 8. 196–209. PMID: 22723817. DOI: <https://doi.org/10.5709/acp-0115-z>
- Davies, Mark. 2008–. *The Corpus of Contemporary American English (COCA): 600 million words, 1990–present*. <https://www.english-corpora.org/coca/>.
- Demestre, Josep, Sheila Meltzer, Jose E. Garcia-Albea & Andreu Vigil. 1999. Identifying the null subject: Evidence from event-related brain potentials. *Journal of Psycholinguistic Research* 28(3). 293–312. DOI: <https://doi.org/10.1023/A:1023258215604>
- Dillon, Brian. 2014. Syntactic memory in the comprehension of reflexive dependencies: An overview. *Language and Linguistics Compass* 8(5). 171–187. DOI: <https://doi.org/10.1111/lnc3.12075>
- Dink, J. W. & B. Ferguson. 2015. *eyetrackingR: An R Library for Eye-Tracking Data Analysis*. <http://eyetrackingr.com>.
- Ehrlich, Kate & Keith Rayner. 1983. Pronoun assignment and semantic integration during reading: Eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior* 22(1). 75–87. DOI: [https://doi.org/10.1016/S0022-5371\(83\)80007-3](https://doi.org/10.1016/S0022-5371(83)80007-3)
- Fernald, Anne, Renate Zangl, Ana Luz Portillo & Verginia A. Marchman. 2008. Looking while listening: Using eye movements to monitor spoken language comprehension by infants and young children. In Irina A. Sekerina, Eva M. Fernández & Harald Clahsen (eds.), *Developmental psycholinguistics: On-line methods in children's language processing*, 97–135. Amsterdam: John Benjamins. DOI: <https://doi.org/10.1075/lald.44.06fer>
- Ferreira, Fernanda, Karl G. D. Bailey & Vittoria Ferraro. 2002. Good-enough representations in language comprehension. *Current Directions in Psychological Science* 11(1). 11–15. DOI: <https://doi.org/10.1111/1467-8721.00158>

- Frazier, Lyn, Charles Clifton & Janet Randall. 1983. Filling gaps: Decision principles and structure in sentence comprehension. *Cognition* 13. 187–222. DOI: [https://doi.org/10.1016/0010-0277\(83\)90022-7](https://doi.org/10.1016/0010-0277(83)90022-7)
- Frazier, Lyn & Keith Rayner. 1982. Making and correcting errors during sentence comprehension: Eye movements in the analysis of structurally ambiguous sentences. *Cognitive Psychology* 14. 178–210. DOI: [https://doi.org/10.1016/0010-0285\(82\)90008-1](https://doi.org/10.1016/0010-0285(82)90008-1)
- Garnsey, Susan M., Michael K. Tanenhaus & Robert M. Chapman. 1989. Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18(1). 51–60. DOI: <https://doi.org/10.1007/BF01069046>
- Gerard, Juliana. 2016. *The acquisition of adjunct control: Grammar and processing*. College Park, MD: University of Maryland dissertation. DOI: <https://doi.org/10.13016/M2621K>
- Gibson, Edward. 2000. The Dependency Locality Theory: A distance-based theory of linguistic complexity. In Alec Marantz, Yasushi Miyashita & Wayne O’Neil (eds.), *Image, language, brain: Papers from the first mind articulation project symposium*, chap. 5, 95–134. Cambridge, MA: MIT Press.
- Green, Jeffrey J. 2019a. A movement theory of adjunct control. *Glossa: A Journal of General Linguistics* 4(1). 87. 1–34. DOI: <https://doi.org/10.5334/gjgl.724>
- Green, Jeffrey J. 2019b. Non-subject control of temporal adjuncts. In Eszter Ronai, Laura Stigliano & Yenan Sun (eds.), *Proceedings of the fifty-fourth annual meeting of the Chicago Linguistic Society*, 137–148. Chicago, IL: CLS.
- Grosz, Barbara J., Scott Weinstein & Aravind K. Joshi. 1995. Centering: A framework for modeling the local coherence of discourse. *Computational Linguistics* 21(2). 203–225. DOI: <https://doi.org/10.21236/ADA324949>
- Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1). 67–90. DOI: [https://doi.org/10.1207/s15516709cog0301\\_4](https://doi.org/10.1207/s15516709cog0301_4)
- Hornstein, Norbert. 1999. Movement and control. *Linguistic Inquiry* 30(1). 69–96. DOI: <https://doi.org/10.1162/002438999553968>
- Hornstein, Norbert. 2003. On control. In Randall Hendrick (ed.), *Minimalist syntax*, 6–81. Oxford, UK: Blackwell. DOI: <https://doi.org/10.1002/9780470758342.ch1>
- Jackendoff, Ray & Peter W. Culicover. 2003. The semantic basis of control in English. *Language* 79(3). 517–56. JSTOR: 4489464. DOI: <https://doi.org/10.1353/lan.2003.0166>
- Jeffrey, Meghan I., Chung-hye Han & Panayiotis A. Pappas. 2015. Experimental evidence of variation and gradience in the syntax and semantics of transitive subject control. *Lingua* 164. 100–131. DOI: <https://doi.org/10.1016/j.lingua.2015.06.011>
- Jeffreys, Harold. 1961. *Theory of probability*. 3rd edn. Oxford, UK: Oxford University Press.
- Kamide, Yuki, Gerry T. M. Altmann & Sarah L. Haywood. 2003. The time-course of prediction in incremental sentence processing: Evidence from anticipatory eye movements. *Journal of Memory & Language* 49(1). 133–156. DOI: [https://doi.org/10.1016/S0749-596X\(03\)00023-8](https://doi.org/10.1016/S0749-596X(03)00023-8)
- Karimi, Hossein & Fernanda Ferreira. 2016. Good-enough linguistic representations and online cognitive equilibrium in language processing. *The Quarterly Journal of Experimental Psychology* 69(5). 1013–1040. DOI: <https://doi.org/10.1080/17470218.2015.1053951>
- Karuza, Elisabeth A., Thomas A. Farmer, Alex B. Fine, Francis X. Smith & Florian T. Jaeger. 2014. On-line measures of prediction in a self-paced statistical learning task. In *Proceedings of the annual meeting of the Cognitive Science Society*, vol. 36.
- Kehler, Andrew. 2008. Rethinking the smash approach to pronoun interpretation. In Jeanette Gundel & Nancy Hedberg (eds.), *Reference: Interdisciplinary perspectives*, 95–122. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195331639.003.0005>

- Kehler, Andrew & Hannah Rohde. 2013. A probabilistic reconciliation of coherence-driven and centering-driven theories of pronoun interpretation. *Theoretical Linguistics* 39(1–2). 1–37. DOI: <https://doi.org/10.1515/tl-2013-0001>
- Kehler, Andrew, Laura Kertz, Hannah Rohde & Jeffrey L. Elman. 2008. Coherence and coreference revisited. *Journal of Semantics* 25(1). 1–44. DOI: <https://doi.org/10.1093/jos/ffm018>
- Kush, Dave, Jeffrey Lidz & Colin Phillips. 2015. Relation-sensitive retrieval: Evidence from bound variable pronouns. *Journal of Memory & Language* 82. 18–40. DOI: <https://doi.org/10.1016/j.jml.2015.02.003>
- Kutas, Marta, Katherine A. DeLong & Nathaniel J. Smith. 2011. A look around at what lies ahead: Prediction and predictability in language processing. In Moshe Bar (ed.), *Predictions in the brain: Using our past to generate a future*, 190–207. Oxford: Oxford University Press. DOI: <https://doi.org/10.1093/acprof:oso/9780195395518.003.0065>
- Kuznetsova, Alexandra, Per B. Brockhoff & Rune H. B. Christensen. 2017. lmerTest package: Tests in linear mixed effects models. *Journal of Statistical Software* 82(13). 1–26. DOI: <https://doi.org/10.18637/jss.v082.i13>
- Landau, Idan. 2015. *A two-tiered theory of control* (Linguistic Inquiry Monographs, vol. 71). Cambridge, MA: MIT Press. DOI: <https://doi.org/10.7551/mitpress/9780262028851.001.0001>
- Landau, Idan. 2017. Adjunct control depends on voice. In Claire Halpert, Hadas Kotek & Coppe van Urk (eds.), *A pesky set: Papers for David Pesetsky*. Cambridge, MA: MITWPL.
- Landau, Idan. Forthcoming. *A selectional theory of adjunct control*.
- Lee, Michael D. & Eric-Jan Wagenmakers. 2013. *Bayesian cognitive modeling: A practical course*. Cambridge, MA: Cambridge University Press.
- Lenth, Russell. 2020. *emmeans: Estimated marginal means, aka least-squares means*. R package version 1.4.4. <https://CRAN.R-project.org/package=emmeans>.
- Lewis, Richard L. & Shravan Vasishth. 2005. An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29(3). 375–419. DOI: [https://doi.org/10.1207/s15516709cog0000\\_25](https://doi.org/10.1207/s15516709cog0000_25)
- Long, Jacob A. 2019. *interactions: Comprehensive, user-friendly toolkit for probing interactions*. R package version 1.1.0. <https://cran.r-project.org/package=interactions>.
- Marslen-Wilson, William D. 1975. Sentence perception as an interactive parallel process. *Science* 189(4198). 226–228. DOI: <https://doi.org/10.1126/science.189.4198.226>
- McCourt, Michael, Jeffrey Jack Green, Ellen Lau & Alexander Williams. 2015. Processing implicit control: Evidence from reading times. *Frontiers in Psychology* 6. 1629. DOI: <https://doi.org/10.3389/fpsyg.2015.01629>
- Mirman, Daniel, James A. Dixon & James S. Magnuson. 2008. Statistical and computational models of the visual world paradigm: Growth curves and individual differences. *Journal of Memory & Language* 59(4). 475–494. DOI: <https://doi.org/10.1016/j.jml.2007.11.006>
- Misyak, Jennifer B., Morten H. Christiansen & J. Bruce Tomblin. 2010. Sequential expectations: The role of prediction-based learning in language. *Topics in Cognitive Science* 2(1). 138–153. DOI: <https://doi.org/10.1111/j.1756-8765.2009.01072.x>
- Morey, Richard D. & Jeffrey N. Rouder. 2018. *BayesFactor: Computation of Bayes factors for common designs*. R package version 0.9.12-4.2. <https://CRAN.R-project.org/package=BayesFactor>.
- Nicol, Janet & David Swinney. 1989. The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18(1). 5–19. DOI: <https://doi.org/10.1007/BF01069043>

- Parker, Dan & Colin Phillips. 2017. Reflexive attraction in comprehension is selective. *Journal of Memory & Language* 94. 272–290. DOI: <https://doi.org/10.1016/j.jml.2017.01.002>
- Parker, Dan, Sol Lago & Colin Phillips. 2015. Interference in the processing of adjunct control. *Frontiers in Psychology* 6(1346). DOI: <https://doi.org/10.3389/fpsyg.2015.01346>
- Poesio, Massimo & Hannes Rieser. 2011. An incremental model of anaphora and reference resolution based on resource situations. *Dialogue & Discourse* 2(1). 235–277. DOI: <https://doi.org/10.5087/dad.2011.110>
- R Core Team. 2019. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.Rproject.org/>.
- Stewart, Andrew J., Judith Holler & Evan Kidd. 2007. Shallow processing of ambiguous pronouns: Evidence for delay. *The Quarterly Journal of Experimental Psychology* 60(12). 1680–1696. DOI: <https://doi.org/10.1080/17470210601160807>
- Stewart, Andrew J., Martin J. Pickering & Anthony J. Sanford. 2000. The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory & Language* 42(3). 423–443. DOI: <https://doi.org/10.1006/jmla.1999.2691>
- Tanenhaus, Michael K., Michael J. Spivey-Knowlton, Kathleen M. Eberhard & Julie C. Sedivy. 1995. Integration of visual and linguistic information in spoken language comprehension. *Science* 268(5217). 1632–1634. DOI: <https://doi.org/10.1126/science.7777863>
- Van Berkum, Jos J. A. 2008. Understanding sentences in context: What brain waves can tell us. *Current Directions in Psychological Science* 17(6). 376–380. DOI: <https://doi.org/10.1111/j.1467-8721.2008.00609.x>
- Wells, Justine B., Morten H. Christiansen, David S. Race, Daniel J. Acheson & Maryellen C. MacDonald. 2009. Experience and sentence processing: Statistical learning and relative clause comprehension. *Cognitive Psychology* 58(2). 250–271. DOI: <https://doi.org/10.1016/j.cogpsych.2008.08.002>
- Yurovsky, Daniel & Michael C. Frank. 2015. Beyond naïve cue combination: Salience and social cues in early word learning. *Developmental Science* 20(2). e12349. DOI: <https://doi.org/10.1111/desc.12349>

**How to cite this article:** Green, Jeffrey J., Michael McCourt, Ellen Lau and Alexander Williams. 2020. Processing adjunct control: Evidence on the use of structural information and prediction in reference resolution. *Glossa: a journal of general linguistics* 5(1): 112.1–33. DOI: <https://doi.org/10.5334/gjgl.1133>

**Submitted:** 10 October 2019    **Accepted:** 25 August 2020    **Published:** 17 November 2020

**Copyright:** © 2020 The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CC-BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited. See <http://creativecommons.org/licenses/by/4.0/>.



*Glossa: a journal of general linguistics* is a peer-reviewed open access journal published by Ubiquity Press.

