Processing pronouns in global discourse context

Marisa Nagano & Emily Zane[*]

Abstract. This study examined the interpretation and processing of third-person pronouns when global discourse context supports a less-salient referent as antecedent of a subject pronoun. In particular, we investigated whether such information cancels a default generalized conversational implicature (GCI) biasing a local subject antecedent interpretation for an English overt pronoun. Eye-tracking data was recorded as participants heard four-sentence mini-stories with one of three Contexts: one biasing the subject of the previous clause as antecedent (SB), one biasing another human referent (OB), and one neutral to biasing either referent. Results showed that looking patterns did not diverge in OB and Neutral conditions until after crucial information tying into the larger discourse context was given in the post-pronoun verb. Strong preferences for non-subject referents did not emerge until after the sentence ended, a time-course consistent with participants calculating and then cancelling a default implicature for a subject antecedent. Meanwhile, discourse context reinforcing the default subject implicature in the SB condition facilitated processing, in terms of less time spent looking at either human referent compared the Neutral condition. Overall, results suggest that upon hearing an overt pronoun, English speakers first calculate a GCI that results in a local subject antecedent interpretation, but that, like all implicatures, this GCI can be defeated by contextual factors.

Keywords. pronoun; context; processing; eye-tracking; conversational implicature

1. Introduction. This paper investigates the interpretation and processing of English third-person pronouns when the global discourse context supports a less-salient referent as the pronominal antecedent, In particular, we explore whether such contextual information cancels a default conversational implicature biasing a local subject interpretation for these pronouns. Psycholinguistic studies of pronoun interpretation often refer to the discourse-pragmatic concept of an anaphora hierarchy, in which "lighter" referring expressions (e.g., pronouns) retrieve highly-salient referents, while "heavier" expressions (e.g., full NPs) retrieve less-salient ones, a pattern that has been claimed to be rooted in Gricean conversational implicature (Givon, 1983; Ariel, 1990; Gundel, Hedberg & Zacharski, 1993; Levinson, 2000). Psycholinguistic evidence suggests that for English, antecedent salience depends largely on syntactic position, with a local subject NP preferred as the antecedent of a subject pronoun (Fukumura & van Gompel, 2014). Studies have also shown that manipulating the linguistic properties of non-subject NPs (e.g., increasing their length) can increase their salience in the discourse and therefore their likelihood to be chosen as the pronominal antecedent (Karimi & Ferreira, 2016). However, less research has explored whether and how semantic information in the global discourse, rather than linguistic properties of potential antecedents, could shift pronominal interpretation.

---

* Authors: Marisa Nagano, Long Island University Brooklyn Campus (marisa.nagano@liu.edu); Emily Zane, James Madison University (zaneer@jmu.edu).

2. Background

2.1. ANAPHORA SCALES AND PRONOMINAL INTERPRETATION. English pronouns, as "light" anaphors, are claimed to retrieve referents that are highly salient in the discourse (such as a recent subject) under certain discourse-pragmatic models of anaphora. Such models argue that not just overt pronouns, but all types of referring expressions (e.g., null pronouns, demonstratives, full NPs, etc.) each prefer antecedents of a corresponding level of salience/prominence in the discourse (Givon, 1983; Ariel, 1990; Gundel, Hedberg & Zacharski, 1993). While the specifics vary from model to model, such systems require both (i) a ranking of (or a mechanism for ranking) anaphors themselves, and (ii) a set of constraints on antecedent saliency. Also implicit in such models is (iii) a justification for the coordination between anaphor rank and antecedent saliency.

Rankings of anaphors rely on factors such as phonological, morphosyntactic, and/or semantic weight, resulting in a scale with null pronouns (no phonological or semantic content) on one end and lexical NPs on the other (themselves sometimes ranked by factors such as definiteness), with forms such as overt pronouns (both stressed and unstressed) and demonstratives somewhere in between. Given this ranking, anaphors on the lighter end of the scale (e.g., null and overt pronouns) are claimed to retrieve highly salient antecedents, while those on the heavy end (e.g., full NPs) retrieve less-salient antecedents, including those not yet introduced into the discourse. Exactly which factors determine the saliency of potential antecedents varies from theory to theory, with properties such as topichood, subjecthood, focus, and linear distance often named. For example, Givon (1983)'s Topic-Continuity model focuses on locality as a major constraint, while Ariel (1990)'s Accessibility Model stresses the accessibility of a given referent in memory storage. Gundel, et al.'s (1993) Givenness Hierarchy does not include saliency criteria per se, instead offering discrete "cognitive statuses" ranging from *in-focus* to *type-identifiable*. Recent work has also argued that different types of anaphoric expressions may be sensitive to different constraints on antecedent saliency both within (Kaiser & Trueswell, 2008) and across (Filiaci, Sorace & Carreiras, 2014) languages.

2.2. ROLE OF CONVERSATIONAL IMPLICATURE. Implicit in the anaphora scale models described above is some sort of mechanism that coordinates between anaphors of various weights and their preferred antecedent saliencies. A conversational implicature, such as one rooted in Grice's (1975) Maxim of Quantity,[1] is one possible mechanism. The use of a "heavy" anaphor where a "lighter" one is possible triggers a conversational implicature in which the hearer must figure out why the speaker chose a non-minimal form. Shifting reference to a less-salient antecedent is a plausible reason for doing so. Gundel, et al. (1993) in particular make reference to the Maxim of Quantity as a driving mechanism in their Givenness Hierarchy. Another psycholinguistic theory of anaphora to emphasize the Maxim of Quantity is the Informational Load Hypothesis, which stresses comparative levels of informativeness of the various referring expressions in the discourse as the key factor in the interpretation of anaphora (Almor, 1999).

Levinson (2000) uses a neo-Gricean framework to argue that the connection between a given anaphoric form (e.g., a pronoun) and the type of antecedent that it retrieves is underwritten by a generalized conversational implicature (GCI) that arises by default whenever that form is used. Specifically, the I[nformativeness]-principle (akin to Part 2 of Grice's Maxim of Quantity)

---

[1] From Grice (1975): "(1) Make your contribution [to the conversation] as informative as is required (for the purposes of the exchange). (2) Do not make your contribution more informative than is required" (p. 45).

encourages local co-reference for a light anaphor, while the M[anner]-Principle (akin to Grice's Maxim of Manner) leads to a non-local interpretation when a heavier anaphor is chosen over a lighter one. Crucially, Levinson (2000) argues that the conversational implicatures that result from these principles are *generalized*, i.e., that they arise by default whenever a specific form (e.g., a pronoun) is used. These generalized conversational implicatures contrast with one-off *particularized* conversational implicatures that arise as a result of a specific discourse context:

(1) Distinction between PCIs and GCIs (Levinson, 2000; p. 16, paraphrasing Grice, 1975):

   a. An implicature i from utterance U is ***particularized*** iff U implicates i only by virtue of specific contextual assumptions that would not invariably or even normally obtain.

   b. An implicature i is ***generalized*** iff U implicates i *unless* there are unusual specific contextual assumptions that defeat it.

Although GCIs arise by default, as a kind of conversational implicature, they can be defeated. In the current paper, we investigate whether global discourse context (i) can override the default local subject interpretation for subject pronouns in English, and (ii) whether on-line pronominal processing in such cases reflects the timing of an implicature cancellation.

2.3. PRONOMINAL INTERPRETATION AND PROCESSING IN ENGLISH. Although to our best knowledge, no previous psycholinguistic study has investigated the on-line calculation and cancellation of the GCI proposed to operate on pronominal interpretation, several studies have investigated how antecedent saliency effects affect the interpretation of pronouns in English. Psycholinguistic evidence suggests that syntactic position is a strong determiner of antecedent saliency in English, with a local subject NP preferred as the antecedent of a subject pronoun, regardless of order of mention of potential antecedents (Fukumura & van Gompel, 2015). Furthermore, an eye-tracking study using a visual world paradigm found that manipulating the linguistic properties of non-subject NPs (such as adding a lengthy relative clause) could promote their salience in the discourse, therefore making them more attractive as pronominal antecedents, in line with the predictions of anaphora hierarchy models (Karimi & Ferreira, 2016). Interestingly, Karimi & Ferreira (2016)'s results also suggest that participants may indeed have calculated the default subject antecedent interpretation GCI for the pronoun before *canceling* that implicature in conditions with the heavy non-subject NP, since looks to the non-subject referent did not overtake looks to the subject referent until relatively late in the time course. However, the primary focus of the study was the effects of antecedent saliency, and not the interaction of discourse context and the cancellation of a GCI per se. As such, the relevant manipulation was on the NP that occurred immediately *before* the pronoun was encountered, which could have affected the time course of processing in and of itself (i.e., extra time was necessary to process a heavy NP, which may in turn have slowed down processing of subsequent information, including the pronoun), muddying the implications for a GCI interpretation.

2.4. THIS STUDY. The current study asks how information in the global discourse—rather than the salience of potential antecedents—influences the processing and interpretation of pronouns in English. Specifically, we crafted items such that the critical information needed to cancel the GCI was not available until *after* the pronoun was heard, meaning that participants were predicted to calculate a GCI supporting a local subject interpretation in all conditions before cancelling it in the condition where context supported a non-subject reading. Crucially, the saliency of the non-subject referent was not manipulated to be stronger than that of the local subject referent, so that antecedent saliency would not suffice to explain any such shift in interpretation. If such infor-

mation does affect pronoun interpretation and processing, in line with the time course for calculating and cancelling a GCI, these results would not cast doubt on models that incorporate antecedent saliency, but instead strengthen the idea that such models are rooted in default conversational implicatures that can be canceled by factors in the larger discourse, including but not limited to manipulations of antecedent saliency.

3. Methodology

3.1. PARTICIPANTS. Participants included 20 monolingual English-speaking adults (mean age 22 years) currently enrolled in an undergraduate or graduate program in New York City.

3.2. STIMULI. Participants listened to four-sentence mini-stories, each followed by an interpretation question (see Table 1). Each mini-story included a Test Sentence, composed of a subordinate clause with proper noun subject (either *Julia* or *Cassie*) followed by a matrix clause beginning with the pronoun *she*. Test sentences were preceded by one of three Context Sentences: (1) Other-Bias (OB), which biased someone other than the subordinate clause subject as antecedent of *she*; (2) Subject-Bias (SB), which reinforced the bias for the subordinate clause subject, and (3) Neutral, which did not bias an antecedent. Crucially, in OB, the subject of the subordinate clause remained the most "salient" antecedent (i.e., the most recent subject NP encountered before the pronoun), and therefore the most attractive pronominal antecedent. As such, in order to recognize the other, non-subject human referent as the likely pronominal antecedent, it was necessary for participants to connect the meaning of the Test Sentence verb (e.g., *drinks*) to information in the Context Sentence (e.g., *Cassie is very thirsty*)—a connection that could not be made until after the pronoun was heard.

In addition to the Context Sentence and Test Sentence, each mini-story also contained an Introduction Sentence that gave the setting and presented the two characters in a conjoined NP subject, as well as a Conclusion sentence with no animate referents (i.e., no possible antecedents for *she*). Each mini-story was followed by a question that elicited the participant's interpretation of the pronoun. Participants heard a total of 30 test items (10 in each condition), plus 20 filler items. Half of the filler items (10) were identical to the Neutral condition but asked a question about an aspect of the story unrelated to pronoun interpretation, such as the setting (e.g. *Where were they hiking?*); the other ten included an ambiguous plural pronoun in the test clause.[2]

	Neutral	Other-Bias (OB)	Subject-Bias (SB)
Intro Sentence	Cassie and Julia are hiking in the woods.		
Context Sentence	The trail is very steep.	Cassie is very thirsty.	Julia is very thirsty.
Test Sentence	While Julia watches some birds, she drinks water.		
Closing Sentence	It's a hot day.		
Question	"Who drank water?"		

Table 1. Example test item with four-sentence story and interpretation question

---

[2] Ultimately only 9 out of 10 item types (for a total of 45 items across 3 experimental and 2 filler conditions) were included in the final analysis, due to methodological oversight (disambiguating information occurred at the final noun in the sentence, rather than at the verb itself).

As participants listened to each mini-story, they viewed a computer screen with four images that matched the story (Figure 1). Pictures of Julia and Cassie were always displayed, along with two non-human images. One of the two non-human images was always the setting, mentioned in the Introduction Sentence. The other was always the direct object of the subordinate clause, so that participants' eyes would likely be drawn to the same filler image before going into the pronoun.
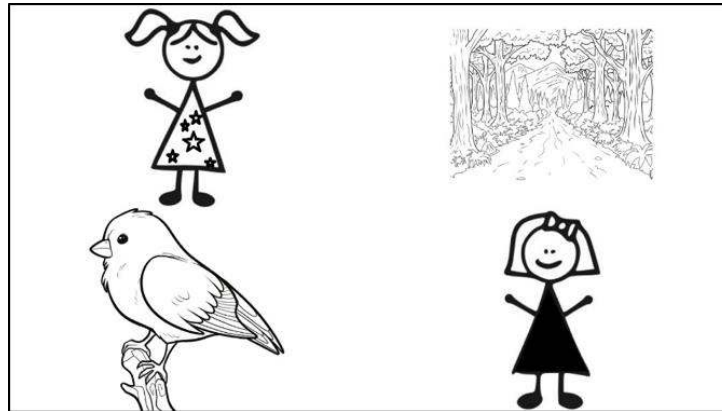

Figure 1. Sample screen with four images from story

3.3. PROCEDURE. After signing a consent form, participants were tested in a soundproof booth with all stimuli presented on a laptop attached to a Gazepoint portable eye-tracker, with audio stimuli played through headphones. Participants were given instructions to minimize movements once eye-tracker calibration was complete. Before the main experiment began, participants were trained on the names of the two characters (*Julia* and *Cassie*), and then completed a practice trial. Each trial consisted of a fixation cross, followed by the visual world presented for 1 second before the onset of the story audio. After participants heard the story in the Neutral, OB, and SB conditions, they indicated their pronoun interpretation by clicking on the answer to a multiple-choice question that appeared on the screen (e.g., *Who drank water?*). Participants were given the choices of both *Cassie* and *Julia* as well as a "Can't tell" option. The trial ended once participants answered the question. Eye-tracking data were recorded at a rate of 60 Hz using a Gazepoint eye-tracker. After the main experiment, participants completed both spatial (using polygons) and verbal (using nonsense words) N-back tasks as metrics of working memory; as we did not discover effects of working memory on either the eye-tracking or interpretation data, we do not present further details about these tasks in this manuscript. Testing took about half an hour, with two opportunities for breaks. Participants were reimbursed $15 for their time.
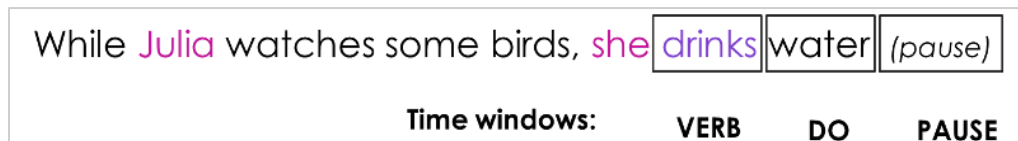

Figure 2. Time windows for eye-tracking data during the Test Sentence: the verb, direct object (DO), and post-sentence pause

3.4. DATA CODING AND ANALYSIS. For each line of eye-tracking data (~1 per 17 ms), a value of 1 was assigned to the image the participant was looking at and 0 to the three other images, creating four possible dependent variables: looks to subject (SUB), looks to the other referent (OTH), looks to the setting (SET), and looks to the last word of the subordinate clause (LW). As an

example, if the participant was looking at the subordinate clause subject, the data for that line would be coded as SUB = 1, OTH = 0, SET = 0, LW = 0. Trials with tracking ratios below 70% were excluded, as well as individual lines of data with track loss. Data was then binned over constituent, with three time windows in the matrix clause of the Test Sentence: the post-pronoun verb (VERB), the direct object (DO), and the pause after the sentence before onset of the final sentence of the story (PAUSE). (See Figure 2).

Data were analyzed using generalized linear mixed effects models (glmer in R) with Context as fixed effect and Participant and Item as random effects (Bates, Mächler, Bolker, and Walker, 2015; R Core Team, 2020). For the interpretation data, the model predicted number of subject interpretations. For the eye-tracking data, two models were fit for each time window, one predicting looks to the subject of the subordinate clause (SUB) and one predicting looks to the other human image (OTH). (Looks to the two non-human images were not analyzed).

3.5. PREDICTIONS. If subject antecedent preference is constrained by a GCI that can be overridden by information in the global discourse, then looking patterns in the OB and Neutral conditions will be identical during the time window from pronoun offset to the end of the verb containing the critical semantic information (VERB), but, once the information in the verb has been encountered, looking patterns will diverge across conditions in the two post-verb time windows (DO, PAUSE). Higher looks to the other referent (i.e., the referent that is *not* the subordinate clause subject) during the post-pronoun window (VERB) in OB vs. Neutral would suggest heightened salience of the other referent *before* participants finish hearing the verb, which would go against prediction for the GCI cancellation account. In other words, elevated looks to the other referent in the OB vs. Neutral immediately after the pronoun would suggest that the default subject implicature is not calculated in the first place in this condition, not that the implicature is cancelled by tying together the information in the verb and the Context Sentence. Looking patterns in SB trials are not predicted to statistically differ from those in the Neutral, since neither condition requires overriding the default interpretation.

4. Results

4.1. INTERPRETATION DATA. Interpretation data in the Neutral condition confirmed a preference for a subject pronoun to retrieve a local subject antecedent in English, with the subordinate clause subject chosen as pronominal antecedent in 94% of the time (see Figure 3). Interpretation data also revealed that the biasing contexts worked as intended: participants chose the subject of the preceding subordinate clause as the pronominal antecedent for only 55% of the Other Bias (OB) trials, statistically less often than in the Neutral condition ($z=-7.58, p<.001$). As expected, Subject Bias (SB) interpretations did not differ statistically from the Neutral (97%, $z=1.04, p=.30$).[3]

4.2. EYE-TRACKING DATA: OTHER BIAS VS. NEUTRAL. For the eye-tracking data, we predicted that if antecedent preference is a kind of default conversational implicature that can be overridden by information in the global discourse, then looking patterns in the OB and Neutral conditions would be identical during the time window from pronoun offset to the end of the verb containing the critical semantic information (VERB), but will begin to diverge during the direct object (DO) and continue to do so in the post-sentence pause (PAUSE). Predictions largely held,

---

[3] Although "Can't tell" was also an option, this accounted for only 15 out of 200 OB trials, 2/200 SB, and 5/200 Neutral trials. That "Can't tell" was higher in OB than the other two conditions is perhaps further testament to the strength of the default subject interpretation preference.

as seen in Figure 4. Participants did not look statistically less often at the subordinate clause in the OB relative to the Neutral until *after* the VERB window, with no statistical differences between conditions in either the VERB ($z=1.08, p=.28$) or the following DO window ($z=1.29, p=.20$). Looks to the subject referent became statistically lower in OB vs. Neutral during the sentence-final PAUSE ($z=-9.65, p<.001$), later than predicted, but still conforming to the timing of cancellation of a default subject antecedent implicature *after* the semantic information in the verb has been processed.
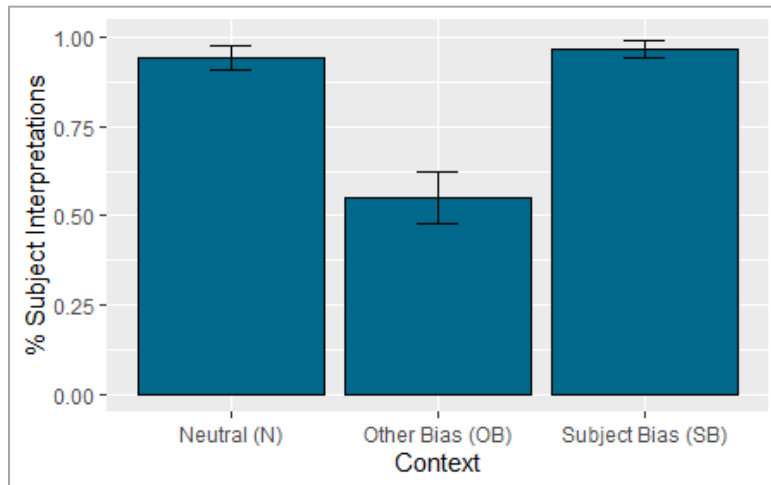

Figure 3. Interpretation Data. Mean percentage of trials in which participants choose the subordinate clause subject as pronominal antecedent
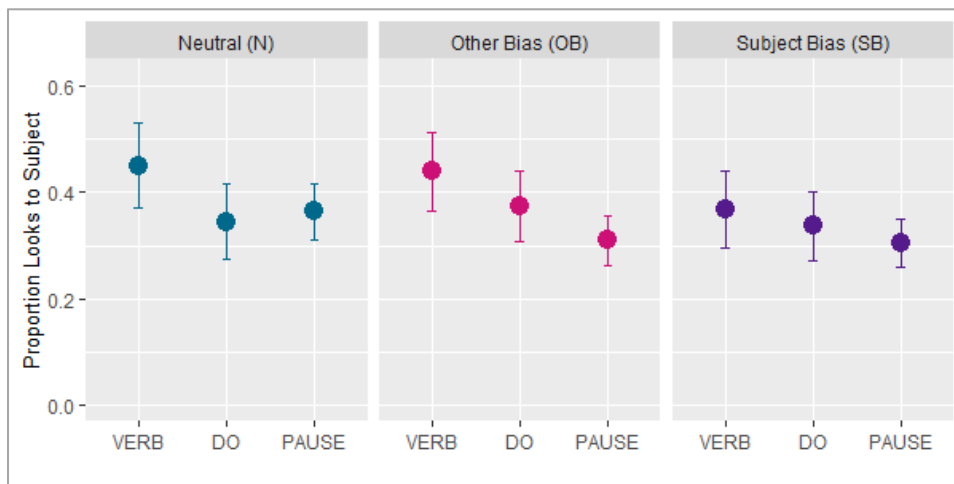

Figure 4. Proportion of looks to the subordinate clause subject across three context conditions

As predicted, participants did not look at the other, non-subject human referent statistically more often in the OB vs. Neutral until the DO window ($z=3.87, p<.001$), continuing into the PAUSE window ($z=20.82; p<.001$), as seen in Figure 5. However, unexpectedly, participants looked at the other referent *less* often in the OB compared to the Neutral during the VERB window ($z=-6.91; p<.001$); we predicted that looks to other would not statistically differ between these two conditions during this time period. While surprising, this result is not inconsistent with the cancellation of a GCI after the verb information is encountered; indeed, it

highlights the effect of the information in the verb, since looks to the other rose so sharply after the verb was heard. Only *greater* looks to the other referent in the window immediately following pronoun offset (VERB) would suggest that participants preferred the other interpretation all along and that the subject implicature was never calculated. This was not found.
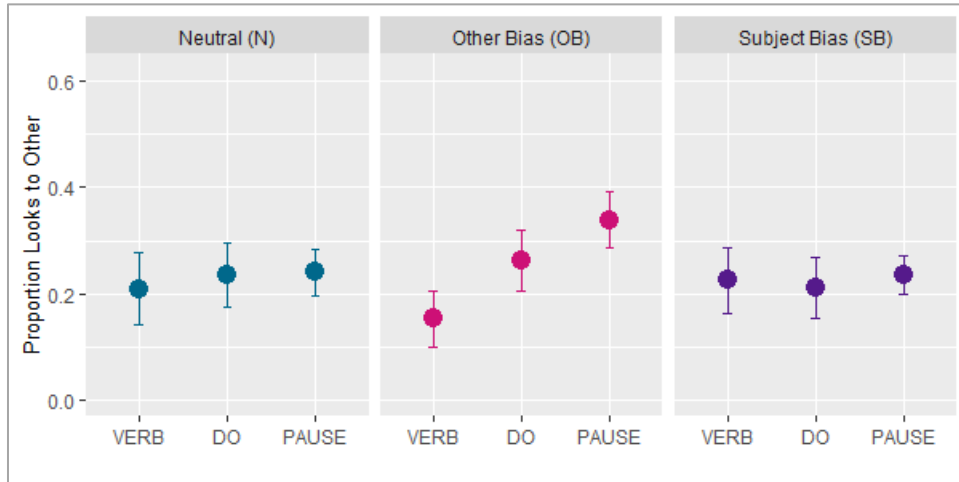

Figure 5. Proportion of looks to the other referent across three context conditions

A follow-up mixed model analysis split the OB eye-tracking trials by ultimate interpretation, comparing the Neutral to the 55% of OB trials in which participants ultimately chose the subordinate clause subject (Other Bias-Chose SUB) and the 45% of OB trials in which they chose the other referent (Other Bias-Chose OTH). Results (Figure 6) showed that the decreased looks to the other, non-subject referent in OB vs. Neutral during VERB reflect those trials in which participants ultimately chose the *subject* as antecedent ($z=-10.18, p<.001$). This suggests that when participants were relatively inattentive to the non-subject referent as they heard the critical verb in OB trials (perhaps due to general low levels of attention during that trial), they ultimately fell back on the default subject interpretation. In OB trials where participants ultimately chose the other referent, they did not differ from Neutral in looks to Other during VERB ($z=0.71, p=.48$), in line with original predictions for the OB condition.
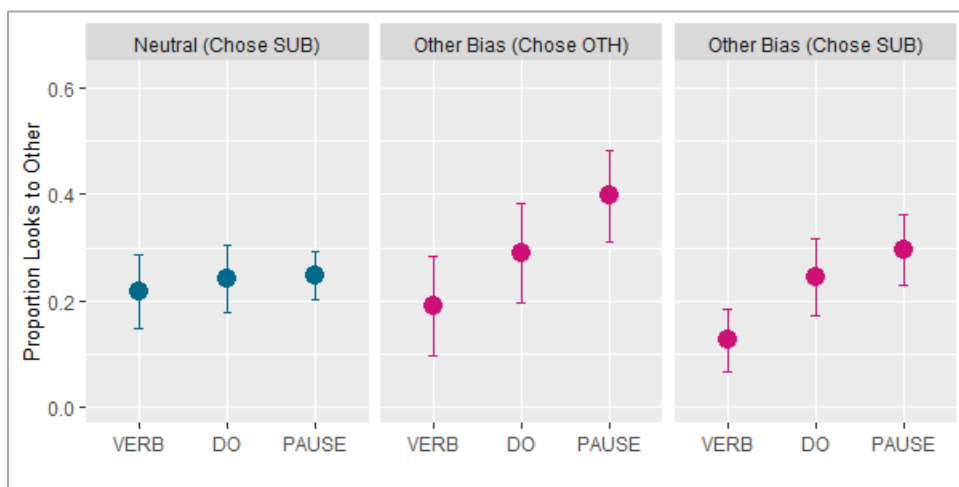

Figure 6. Proportion of looks to the other referent with OB trials split by ultimate interpretation

4.3. EYE-TRACKING DATA: SUBJECT BIAS VS. NEUTRAL. We predicted that Subject Bias (SB) trials would not differ significantly from the Neutral, since both contexts support the default subject antecedent interpretation. Again, predictions largely held, in that participants looked at the two potential referents equally or less often in SB compared to Neutral across VERB (looks to SUB: $z=-5.12, p<.001$; looks to OTH: $z=1.35, p=.18$), DO (looks to SUB: $z=-1.39, p=.17$; looks to OTH $z=-2.67, p<.01$), and PAUSE (looks to SUB $z=-11.01, z<.001$; looks to OTH: $z=1.51, p=.13$). The fact that some windows showed even fewer looks to the human referents in SB vs. Neutral suggests that pronominal processing was facilitated when context reinforced the default interpretation, so that participants were free to look at the other non-human images on the screen.

5. Discussion. Overall, the interpretation and eye-tracking data together suggest that, in English, there is a strong preference for a subject pronoun to retrieve a local subject antecedent, in line with an analysis that such a preference arises by default via a GCI. That the preference for a subject antecedent is strong can be seen in the interpretation data: even in the Other Bias condition, participants only selected the other, non-subject referent as antecedent over the local subject in about half of the trials. In the eye-tracking data, looking patterns did not shift away from the subject and towards the other referent until after the critical verb had been heard; indeed, looks to the subject referent in OB vs. Neutral conditions did not statistically decrease until the pause following the sentence, even as looks to the other referent rose slightly earlier during the direct object.

Results also suggest that although the GCI arises, it *can* be overridden by incorporating information in the global discourse context, even if the critical information is presented *after* the pronoun has been encountered. This is evident in the interpretation data, in which participants chose the other, non-subject referent significantly more often in the condition where context biased such a reading, compared to a condition with no particular contextual bias. Furthermore, eye-tracking results support the predicted timeline for the cancellation of a GCI in the OB condition. After hearing the pronoun—but before processing the verb containing the critical call-back to the earlier semantic information—participants preferred to look at the subordinate clause subject referent an equal amount in the OB and Neutral conditions; indeed, looks to the subject referent did not decrease in OB relative to the Neutral until after the sentence had ended, during the sentence-final pause. Likewise, participants did not begin to look at the contextually-biased non-subject referent at a higher rate until after the critical verb had been heard. In other words, despite the mention of the other referent in the Context Sentence, participants preferred the local subject referent as pronominal antecedent until they heard the critical verb—after the pronoun had been encountered—and, although they began to consider the other referent in the window immediately following the verb, they did not completely disregard an original subject preference until after the sentence had ended.

Interestingly, the choice of a subject interpretation in OB contexts seems associated with relatively low attention to the other, non-subject referent immediately after the pronoun is presented. This could suggest that in trials where participants are not paying close attention to the story (and therefore the discourse context), they fall back on the default subject interpretation for the pronoun, although further investigation is necessary. A related follow-up study might explore how executive function metrics such as working memory and/or attention-switching affect interpretation/processing during OB conditions. While working memory data were recorded in this study, we found no significant relationship between working memory metrics and either behavioral or eye-tracking results. This is somewhat unsurprising, given that the relative simpli-

616

city and brevity of the narratives in this study were unlikely to significantly burden executive functioning among neurotypical adults. However, in a previous study using the same experimental items on children and adolescents with and without ASD, we found that the former group showed a stronger preference for the subject referent than the latter in the eye-tracking but not the interpretation data (Nagano, Grossman & Zane, 2020). We argued that while this result could reflect well-known issues with pragmatics in ASD, the fact that group divergence was found only in processing and not the interpretation data may point to executive functioning differences among the two participant groups. A follow-up to the current study using materials that are more taxing in terms of length or complexity may well elicit executive functioning effects even in neurotypical adults.

6. Conclusion. Overall, our results suggest that hearers are able to override a default preference for a subject pronoun to retrieve a local subject antecedent when another interpretation is supported by global discourse context, even when the competing referent is not more salient than the subject referent (in terms of syntactic position, recency, and/or length). The word *override* is key: the default local subject preference arises regardless of context, and then, in trials where context supports another referent, this preference is set aside later in the time course. Such a process is consistent with the idea that the subject antecedent preference is rooted in a GCI that arises by default but can be cancelled by other factors in the larger discourse.

References

Almor, Amit. 1999. Noun-phrase anaphora and focus: The informational load hypothesis. *Psychological Review* 106(4). 748-765. https://doi.org/10.1037/0033-295X.106.4.748.

Ariel, Mira. 1990. *Accessing noun-phrase antecedents*. New York: Routledge.

Bates, Douglas, Martin Mächler, Ben Bolker, Steve Walker. 2015.. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. https://doi.org/10.18637/jss.v067.i01 .

Filiaci, Francesca, Antonella Sorace & Manuel Carreiras. 2014. Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language, Cognition and Neuroscience* 29(7). 825–843. https://doi.org/10.1080/01690965.2013.801502.

Fukumura, Kumiko & Roger P. G. van Gompel. 2014. Effects of order of mention and grammatical role on anaphor resolution. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 41(2). 501–525. https://doi.org/10.1016/j.jml.2009.09.001.

Givón, Talmy. 1983. Topic continuity in discourse: An introduction. In Talmy Givón (ed.), *Topic continuity in discourse*. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.3.

Grice, H. Paul. (1975). Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Speech acts* (Syntax and semantics 3), 41–58. New York: Academic Press. https://doi.org/10.1163/9789004368811_003.

Gordon, Peter C., Barbara J. Grosz & Laura A. Gilliom. 1993. Pronouns, names, and the centering of attention in discourse. *Cognitive Science* 17(3). 311–347. https://doi.org/10.1207/s15516709cog1703_1.

Gundel, Jeanette, Nancy Hedberg & Ron Zacharski. 1993. Cognitive status and the form of referring expressions in discourse. *Language* 69(2). 274–307. https://doi.org/10.2307/416535.

Kaiser, Elsi & John C. Trueswell. 2008. Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive*

Processes 23(5). 709–748. https://doi.org/10.1080/01690960701771220.

Karimi, Hossein & Fernanda Ferreira. 2016. Informativity renders a referent more accessible: Evidence from eyetracking. *Psychonomic Bulletin & Review* 23(2). 507–525. https://doi.org/10.3758/s13423-015-0917-1.

Levinson, Stephen. 2000. *Presumptive meanings: The theory of generalized conversational implicature.* Cambridge, MA: MIT Press.

Nagano, Marisa, Emily Zane & Ruth B. Grossman. 2020. Structural and contextual cues in third-person pronoun interpretation by children with autism spectrum disorder and their neurotypical peers. *Journal of Autism and Developmental Disorders* 1–22. https://doi.org/10.1007/s10803-020-04645-7.

R Core Team (2020). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/.