UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Social Cues Modulate Cognitive Status of Discourse Referents

Permalink
https://escholarship.org/uc/item/9fd5398n

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 38(0)

Authors
Hawthorne, Kara
Arnhold, Anja
Sullivan, Emily
et al.

Publication Date
2016

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Social Cues Modulate Cognitive Status of Discourse Referents

Kara Hawthorne (khawthor@olemiss.edu)
Department of Communication Sciences and Disorders, 304 George Hall
University of Mississippi, University, MS 38677 USA

Anja Arnhold (anja.arnhold@uni-konstanz.de)
Department of Linguistics, Post Box D 186
University of Konstanz, 78457 Konstanz GERMANY

Emily Sullivan (esulliva@ualberta.ca)
Department of Communication Sciences and Disorders, Corbett Hall
University of Alberta, Edmonton, AB T6G 2G4 CANADA

Juhani Järvikivi (jarvikiv@ualberta.ca)
Department of Linguistics, 4-55 Assiniboia Hall
University of Alberta, Edmonton, AB T6G 2E7 CANADA

Abstract

We use visual world eye-tracking to test if a speaker's eye gaze to a potential antecedent modulates the listener's interpretation of an ambiguous pronoun. Participants listened to stories that included an ambiguous pronoun, such as "The dolphin kisses the goldfish… He…." During the pre-pronominal context, an onscreen narrator gazed at one of the two characters. As expected, participants looked more at the subject character overall. However, this was modulated by the narrator's eye gaze and the amount of time the participant spent looking at the gaze cue. For trials in which participants attended to the narrator's eye gaze for > 500ms, participants were significantly more likely to interpret the pronoun as referring to the object if the narrator had previously looked at the object. Results suggest that eye gaze – a social cue – can temper even strong linguistic/cognitive biases in pronoun resolution, such as the subject/first-mention bias.

Keywords: Ambiguous pronoun resolution, visual world paradigm, eye-tracking, reference, social cues, eye gaze.

Introduction

In this paper, we test if a social cue – the speaker's eye gaze to potential referents – impacts the listener's interpretation of ambiguous pronouns in a discourse context where there are two characters who could serve as the pronoun's antecedent. For example, in "The dolphin kisses the goldfish behind the lake. He…" *he* could refer to either the dolphin or the goldfish. Unlike previous work on the effect of a narrator's attention on offline pronoun interpretation (Nappa & Arnold, 2014), we manipulate the narrator's eye gaze during the pre-pronominal context, which is the time period in which speakers naturally look at their intended referent (Griffin & Bock, 2000). We test if the narrator's eye gaze modulates the listener's assumptions about the narrator's focus of attention by using the visual world eye-tracking paradigm to monitor online processing of ambiguous pronouns.

Linguistic and Cognitive Biases

Pronoun resolution is constrained by the relative salience or prominence of potential antecedents in the discourse representation (e.g., Gordon, Grosz, & Gilliom, 1993; Gundel, Hedberg, & Zacharski, 1993). Salience modulates the listener's attention to potential referents and the degree to which s/he expects each referent to be talked about in the upcoming discourse (Ariel, 1990; Arnold, 1998).

There are several linguistic and cognitive biases that rapidly affect referent salience during online pronoun processing. Linguistic factors include pronoun gender (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000), verb type (Koornneef & Van Berkum, 2006; Pyykkönen & Järvikivi, 2010), and parallel syntactic structure and syntactic function. Subjecthood, in particular, strongly increases the salience of a discourse entity (Järvikivi, Hyönä, Bertram, & Van Gompel, 2005; Kaiser & Trueswell, 2008).

Perhaps the most widely known cognitive bias concerns the order in which the referents are introduced in the pre-pronominal context: first-mentioned entities are preferred over later-mentioned entities. This finding has been replicated often since Gernsbacher & Hargreaves (1988; e.g., Carreiras, Gernsbacher, & Villa, 1995), including in visual world studies of pronoun resolution (Arnold, et al., 2000; Järvikivi et al., 2005). Unlike in languages with freer word order (Järvikivi et al., 2005; 2014), subjecthood and first-mention are difficult to tease apart in English; we will primarily use the term *subject bias* in this paper, acknowledging that order-of-mention and syntactic function both contribute to the observed effects.

Social Biases

In addition to linguistic and cognitive effects, social cues impact discourse processing (e.g., Van den Brink et al., 2012; Jiang & Zhou, 2015) and conversational success. For

example, two people engaged in a cooperative task are much slower when they are not able to use social cues such as pointing, eye gaze, and head nodding (Clark & Krych, 2004). Social cues also impact pronoun resolution: a coreferential gesture at pronoun onset tempers the subject bias when each of the potential antecedents has previously been associated with a particular location in the speaker's gestural space (Goodrich & Hudson Kam, 2012).

Eye gaze is a salient way for a speaker to signal their attention to the listener (Langton, Watt, & Bruce, 2000). Interlocutors attend closely to each other's faces (Argyle & Cook, 1976), and a speaker will even restart an utterance if the listener is not visually attending (Goodwin, 1981). In production, speakers fixate characters before they name them when describing an image (Griffin & Bock, 2000), so attending to the speaker's eye gaze could have processing payoffs. For example, Hanna and Brennan (2007) report that listeners can use a speaker's gaze to figure out which object in a hidden array the speaker is referring to even before the speaker has reached the point of linguistic disambiguation. Moreover, the attentional effects of eye gaze are reflexive and occur even when eye gaze is manipulated to not be an informative cue (e.g., Firesen & Kingstone, 1998).

Only recently have researchers begun to explore the role of eye gaze in pronoun resolution. Nappa and Arnold (2014) tested the influence of several social cues in the moment when the pronoun is heard. They found that when a narrator turns her head and looks at a character while producing an ambiguous pronoun (with or without a pointing gesture), listeners were more likely to interpret the pronoun as the character being deictically cued.

When a speaker turns to look at a character right at the moment of pronoun production, it is perhaps not surprising that it influences pronoun interpretation, since the speaker is directly highlighting a potential referent. However, gaze cues are typically less overt, and they usually occur during the preceding discourse context rather than during the pronoun itself. Griffin and Bock (2000) report that speakers look at a character close to a full second before referring to it with a full noun phrase. Speakers look at the referent before producing a pronoun as well, though at a somewhat reduced rate (van der Meulen, Meyer, & Levelt, 2001).

While eye gaze serves as a visual cue that increases the salience of one potential referent, recent evidence suggests that not all visual cues have such an effect on ambiguous pronoun interpretation. Arnold and Lao (2015, Exp. 2) had participants listen to stories of two characters, e.g., "Birdy picked apples with Doggy near the farmhouse. He…." while they briefly (200ms) flashed a halo around one of the characters. This cue did not affect participants' gaze behaviour or their antecedent selection preferences. A further study (Järvikivi & Pyykkönen-Klauck, submitted), shows that absence of one of the referents at the pronoun onset (i.e., if a potential referent had walked out of the visual scene) did not affect adult listeners' pronoun resolution preferences. This suggests that visual cues that are coincidental with linguistic information but that are not social – in other words, cues that are not indicative of the speaker's intentions or attention to discourse referents – do not automatically impact language comprehension.

In this paper, we ask if a speaker's eye gaze during the pre-pronominal context influences how listeners interpret an ambiguous pronoun.

Methods

Participants

Participants were 86 native English speakers. Data from additional participants were excluded because of poor calibration (n = 27), corrupted results file (n = 5), or experimenter/equipment/participant error (n = 1/8/2).

Materials

Experimental stimuli consisted of 20 mini-stories involving one animal character performing an action on the other at a particular location (Table 1), as well as a visual display (Figure 1). The visual display contained both characters, the location, and a hedgehog narrator, who introduced herself as such at the beginning of the experiment (cf. Staudte & Crocker, 2011 for evidence that listeners attend to an artificial speaker's gaze similarly to a human's gaze).

Table 1: Example mini-story.
The letters (a)-(d) represent the four experimental conditions.

Sentence	Audio	Narrator's Gaze
Intro (a,b)	There are the dolphin and the goldfish.	forward
Intro (c,d)	There are the goldfish and the dolphin.	forward
Action (a)	The dolphin kisses the goldfish	**dolphin (*subject*)**
Action (b)	The dolphin kisses the goldfish	**goldfish (*object*)**
Action (c)	The goldfish kisses the dolphin	**dolphin (*object*)**
Action (d)	The goldfish kisses the dolphin	**goldfish (*subject*)**
Location	behind the lake.	forward
Pronoun	He wants to play on the playground.	forward
Probe	Who wants to play on the playground?	n/a

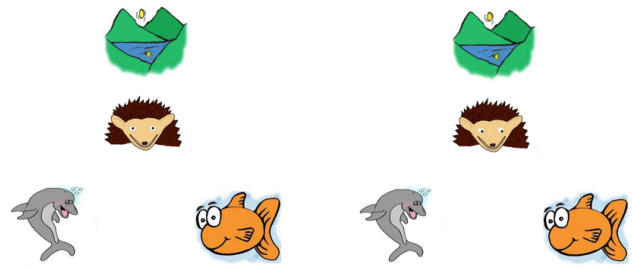

Figure 1. Example visual display. The animal in the center of each image is the narrator throughout the experiment. The image on the left appeared during the intro, location, and pronoun sentences (see Table 1). The image on the right, in which the narrator is gazing at the goldfish, appeared only during the action sentence. During the probe question, only the two animal characters (e.g., the dolphin and the goldfish) appeared on the screen.

Each story began with an introduction to the two characters. The subject of the action sentence was always named first to control for effects of the first-mention and subject biases. Next was the action sentence, in which one animal performed an action on the other. This sentence was

manipulated in two ways – which animal was the subject of the sentence (e.g., dolphin or goldfish) and which animal was gazed at by the narrator (subject or object) – to create four versions of the story (a-d, Table 1). After the action, the narrator's gaze returned to the front for the location sentence; the location was mentioned to draw participants' eyes away from the animals before the ambiguous pronoun in the pronoun sentence. Finally, the participants heard a new voice asking for an overt judgment on the referent of the pronoun (the probe sentence).

An additional ten mini-stories were recorded as filler items. The fillers were structurally the same as the experimental stories, except they did not have an ambiguous pronoun. Instead, one of the characters was referred to by name. The named animal was the subject of the action sentence half of the time and the object of the action sentence half of the time.

The stories were recorded by a 21-year-old female, and the probe questions were recorded by a 19-year-old male. Both were native English speakers. The speakers were asked to read the stories in a happy, animated voice, and care was taken that neither animal character was prosodically more prominent. Recordings were done in a sound-attenuated booth using a head-mounted CountryMan microphone and Korg MR-2000S Studio Recorder. Each story was recorded individually, and a 1-second pause was inserted at the sentence boundaries.

The visual displays were created using Adobe Photoshop CS5.1 software. Most of the images were previously used in Pyykkönen et al. (2010) and Järvikivi et al. (2014) and others were drawn by hand to match the style of the existing images. Each image fit into an area of 426 by 341 pixels. The images were counterbalanced, such that the subject appeared on the left side of the screen half of the time.

Procedure

The session began with a brief familiarization to the animals using Microsoft Office PowerPoint. Each animal was displayed one at a time, and the participant was asked to label it. If the participant provided a label different from the one used in the mini stories, s/he was told the label that would be used in the experiment.

The experiment used eye-tracking and the visual world paradigm (Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995; Arnold et al., 2000). The experimenter first calibrated the eye-tracking equipment. Next, the narrator hedgehog appeared at the center of the screen and introduced herself by saying "Hi, my name is Hailee! I'm going to tell you some stories about the animals you just saw. Are you ready?" This was done so that it was clear that the hedgehog was the one telling the stories.

A drift correction was performed before each trial to ensure that the equipment remained properly calibrated. Then the participant heard one of the stories while his or her eye gaze was tracked. After the probe question, the participant was asked to indicate which animal was referred to in the pronoun sentence by pressing a key on the keyboard.

Equipment

Participants were tested on either an EyeLink 1000 or 1000+ eye-tracker. The experiment was run using Experiment Builder (SR-Research Ltd) and a 500 Hz sampling rate. Stimuli were played through Bose SoundLink Mini speakers, and testing was done in a quiet room.

Design

Each participant was assigned to one of four experimental lists, with twenty experimental and ten filler trials per list. Each list included each story in one of the four conditions (e.g., 2a-d, Table 1). The lists were counterbalanced, so that the narrator's gaze was on the subject and object of the action sentence for an equal number of trials. In order to maximize the number of trials in which the participant noticed the narrator's change of eye gaze during the action sentence, we included all eligible participants, even though the number of participants on each list varied from 20-25.

Results

Offline responses and eye-tracking data from all 1718 trials were analyzed with linear mixed effects modelling, using *R* (R Core Team, 2013) and the *glmer* and *lmer* functions from the package *lme4* (Bates, Maechler, & Bolker, 2012). For all analyses, the best fit model was determined using backwards stepwise model comparisons. Models were compared with likelihood ratio tests; only factors that significantly improved the model fit at a $p < .05$ level were retained for the fixed and random intercept effects (see Baayen, 2008, Bates et al. 2015). Random slopes were checked but omitted from our final analyses due to convergence errors.

Gazed at role (whether the narrator looked to the subject or object character) and earlier attention to the narrator's gaze (how long the participant looked at the narrator during the action sentence) were the factors of interest. Trials were considered to have no attention to the narrator if the participant looked at the narrator < 200ms during the action sentence (57% of trials). If the participant looked at the narrator between 200 and 500ms during the action sentence, this was considered short attention (25% of trials), while looks longer than 500ms were considered long attention (18%). Results from offline responses and eye-tracking analyses are summarized in Tables 2 and 3. For the intercept, "no," "short," and "long" indicate duration of fixation to the narrator and "S" and "O" indicate whether the narrator was gazing at the subject or object animal. The intercept values were releveled and models were re-run to examine all possible comparisons. Redundant comparisons have been omitted.

Offline Responses

The dependent variable for the offline analysis was whether the participant selected the subject or object character in response to the probe question. Responses are presented in Figure 2.

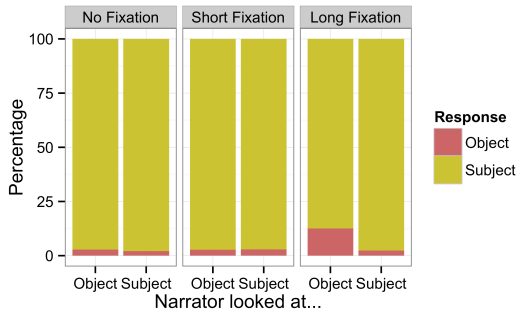

Figure 2. Percentage of object and subject responses, split by participants' earlier attention to the narrator.

Data were evaluated with binomial generalized linear mixed effects models. A model containing an interaction between gazed at role (subject versus object) and earlier attention to narrator's gaze (long, short, no) significantly outperformed a model containing only main effects ($\chi 2(1) = 7.43$, $p = .024$), so the full model was retained as the final model. When comparing the short versus no attention to narrator trials, there were no interaction or simple effects (p-values > .1). Results from the long versus short attention to narrator trials are presented in Table 2. If the participant had previously attended for a long time to the narrator while the narrator was looking to the object, s/he was significantly more likely to make an "object" response than if the narrator was looking at the subject or if the participant had only paid short attention to the narrator.

Table 2. Response results for the model: response ~ GazedAtRole* EarlierAttentnToNarr + (1|participant) + (1| item) + (1|trial).

Offline Responses: Short vs. Long Attention to Narrator				
Fixed Effect	**Intercept**	**Estimate (*SE*)**	***z*- value**	***p*-value**
2-way interaction	short, S	-2.05 (.87)	-2.34	.019
GazedAtRole	short, S	.060 (.62)	-.10	n.s
GazedAtRole	long, S	-1.99 (.62)	-3.22	.0013
EarlierAttentnToNarr	short, S	.34 (.69)	.49	n.s
EarlierAttentnToNarr	short, O	-1.71 (.55)	-3.14	.0017

Eye-Tracking Analysis

Eye-tracking data for each trial were aggregated into 300ms windows, starting at the onset of the ambiguous pronoun through 1500ms post-onset. Aggregation mitigates the auto-correlation that is inherent to time-course data. Time window was included as a fixed factor in the models, along with gazed at role (subject vs. object) and previous attention to the narrator's gaze cue (long, short, no). The dependent variable was the logit transformed proportion of looks to the subject divided by the proportion of looks to the subject plus object. This allows us to look at the strength of the subject bias while controlling for different amounts of time spent looking at the narrator and location. Data points in which the participant did not look at either the subject or object during the time window were dropped.

There were no three way interactions of time by gazed at role by earlier attention to narrator's gaze (p-values > .1). We next tested a model with all 2-way interactions. There were no gazed at role by attention to narrator's gaze or time by gazed at role interactions, so these were checked and removed one at a time (p-values > .1).

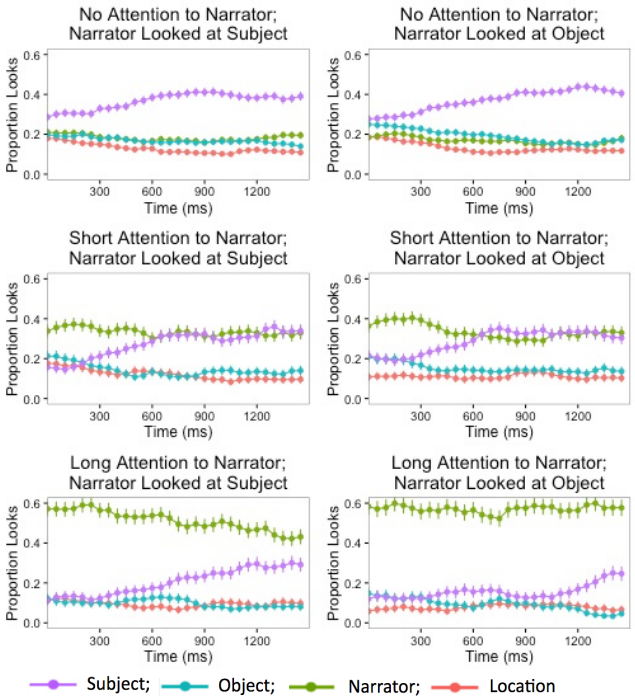

Figure 3. Proportion of participant looks to the interest areas from pronoun onset.

The final model contains a main effect of gazed at role, with participants spending less time looking at the subject relative to the subject plus object for trials in which the narrator had previously gazed at the object (Figure 3). This suggests that the narrator's earlier eye gaze cue influenced participants' eye movements after the onset of the ambiguous pronoun. The model also contains an interaction of time by earlier attention to the narrator's gaze, suggesting that the participants' previous attention or lack of attention to the narrator influenced the timing with which they settled on a referent for the ambiguous pronoun.

The effects of time for no, short, and long attention to narrator's gaze are presented in Table 3. Simple effects of attention to narrator are not included, since they were not significant at any time window. However, looking behavior across time was different across these conditions. For trials with no earlier attention to the narrator, participants looked increasingly more at the subject through the 300-599ms window, then leveled off. For trials with short attention to the narrator, looks to the subject leveled off even earlier, with no significant differences in proportion looks to the subject for 300-599ms with any subsequent time windows. The proportion of looks to the subject increased most slowly

for trials with long earlier attention to the narrator, with significant differences between 600-899ms and the subsequent time windows.

Table 3. Response results for the model: Looks to S/(S+O) ~ GazedAtRole + EarlierAttentnToNarr *Time + (1|participant) + (1| item) + (1|trial).

Eye-Tracking Results: Effects of GazedAtRole for adjacent time windows, split by degree of attention to the narrator's earlier gaze cue			
Fixed Effect	**Intercept**	**Est. (*SE*)**	***t*-value**
GazedAtRole	long	-.18 (.09)	-2.03
NO ATTENTION TO NARRATOR'S GAZE CUE			
Time (vs 300-599ms)	0-299ms	.50 (.17)	2.85*
Time (vs 600-899ms)	0-299ms	.78 (.17)	4.51*
Time (vs 900-1199ms)	0-299ms	.94 (.17)	5.46*
Time (vs 1200-1500ms)	0-299ms	.99 (.17)	5.71*
Time (vs 600-899ms)	300-599ms	.39 (.17)	1.68
Time (vs 900-1199ms)	300-599ms	.45 (.17)	2.65*
Time (vs 1200-1500ms)	300-599ms	.50 (.17)	2.92*
Time (vs all later windows)	600-899ms	--	*all < 2*
Time (vs 1200-1500ms)	900-1199ms	.05 (.17)	0.30
SHORT ATTENTION TO NARRATOR'S GAZE CUE			
Time (vs 300-599ms)	0-299ms	.88 (.30)	2.96*
Time (vs 600-899ms)	0-299ms	1.25 (.29)	4.35*
Time (vs 900-1199ms)	0-299ms	1.16 (.29)	3.99*
Time (vs 1200-1500ms)	0-299ms	1.21 (.28)	4.24*
Time (vs all later windows)	300-599ms	--	*all < 2*
Time (vs all later windows)	600-899ms	--	*all < 2*
Time (vs 1200-1500ms)	900-1199ms	.05 (.28)	0.18
LONG ATTENTION TO NARRATOR'S GAZE CUE			
Time (vs 300-599ms)	0-299ms	.56 (.43)	1.32
Time (vs 600-899ms)	0-299ms	.85 (.42)	2.02*
Time (vs 900-1199ms)	0-299ms	1.66 (.41)	4.02*
Time (vs 1200-1500ms)	0-299ms	1.82 (.41)	4.48*
Time (vs 600-899ms)	300-599ms	.28 (.41)	.69
Time (vs 900-1199ms)	300-599ms	1.10 (.40)	2.71*
Time (vs 1200-1500ms)	300-599ms	1.25 (.40)	3.15*
Time (vs 900-1199ms)	600-899ms	.81 (.39)	2.06*
Time (vs 1200-1500ms)	600-899ms	.97 (.39)	2.49*
Time (vs 1200-1500ms)	900-1199ms	.16 (.38)	0.41

Visual inspection of Figure 3 reveals a potential effect of earlier attention to the narrator on the proportion looks to the narrator after the pronoun onset, despite the fact that the location was mentioned last before the pronoun. To examine this effect, we conducted a secondary analysis in which the dependent variable was logit transformed proportion looks to the narrator, aggregated into 300ms windows.

There were several three way interactions of time by gazed at role by earlier attention to the narrator (*t*-values > 2), so the 2-way interactions of gazed at role by earlier to narrator were tested separately for each time window. There were no 2-way interactions or effects of gazed at role for the 0-1199ms time windows (all *t*-values < 2). The proportion of looks to the narrator during each of those time windows was greater the longer the participant had attended to the narrator during the action sentence (long > short > no). For the 1200-1500ms window, there was a 2-way interaction of gazed at role by earlier attention to the narrator. Participants who had previously paid long attention to the narrator while the narrator gazed at the object looked more to the narrator during the 1200-1500ms window than if the narrator had looked at the subject.

Discussion

Results suggest that a narrator's eye gaze to a character impacts the listener's resolution of a subsequent ambiguous pronoun, but that the effect is modulated by the listener's attention to the narrator's gaze. As predicted, when the narrator looked at the character that served as the grammatical object, the listener was more likely to interpret the pronoun as referring to the object, but only if the participant had attended closely to the narrator's gaze cue. In contrast to previous work, in which the narrator both looked at and turned her head toward one of the referents during production of the pronoun itself (Nappa & Arnold, 2014), we found (1) that eye gaze alone and (2) that an eye gaze cue that occurs during the pre-pronominal discourse context are sufficient to temper the subjecthood and first-mention biases. The latter is an important finding, since the timing with which the narrator gazed at one of the potential referents in the present study closely approximates the time period in which speakers are likely to provide such cues in a real discourse context (Griffin & Bock, 2000). Thus, a social cue – the speaker's gaze to a potential referent *before the listener even hears the pronoun* – can modulate other linguistic and cognitive biases at the point of pronoun disambiguation.

Another notable finding relates to attention to the narrator herself after the pronoun was produced. The longer the participant attended to the narrator during the time when her eyes were cueing one of the potential referents, the more the participant continued to look at the narrator after the onset of the pronoun, perhaps anticipating further informative eye movements that could help disambiguate it. Indeed, for trials in which the participant paid long attention to the narrator's gaze cue, the participant looked even longer at the narrator after pronoun onset if the narrator had previously looked at the object (vs. the subject), even though the location was the last item mentioned before the pronoun. Thus, participants appear to particularly look to the narrator for 'help' when she had gazed at the object and the participant was therefore more likely to be entertaining the object as a potential referent for the pronoun.

The present results, together with other work on social and other visual cues to pronoun resolution (Arnold & Lao, 2015; Järvikivi & Pyykkönen-Klauck, submitted; Nappa & Arnold, 2014) suggest that visual context impacts discourse processing, but only when that visual information is relevant to language comprehension. In that vein, many studies have demonstrated that the visual environment can help to rapidly resolve temporary referential ambiguities in sentences, even overriding linguistically-based parsing preferences (e.g., Tanenhaus et al., 1995; Chambers, Tanenhaus, & Magnuson, 2004; Knoeferle & Crocker, 2006). The present study shows that listeners are sensitive to visual cues in reference resolution as well, at least when the cue is social (i.e., eye gaze) and is therefore potentially informative about the speaker's intentions. Social visual cues can increase the salience of a potential referent even when the cue occurs in the pre-pronominal discourse context.

Acknowledgments

Thank you to University of Alberta Support for the Advancement of Scholarship (SAS/EFF) and Undergraduate Research Initiative (URI) for funding that supported this project. Thank you also to Lauren Rudat and the other members of the Centre for Comparative Psycholinguistics.

References

Argyle, M. & Cook, M. (1976). *Gaze and mutual gaze.* Cambridge, MA: Cambridge University Press.

Arnold, J. E. (1998). *Reference form and discourse patterns* (Doctoral dissertation). Stanford University, California.

Arnold, J.E., Eisenband, J.G., Brown-Schmidt, S., & Trueswell, J.C. (2000). The rapid use of gender information: Evidence of the time course of pronoun resolution from eyetracking. *Cognition, 76*, B13–B26.

Arnold, J. E. & Lao, S. Y. C. (2015). Effects of psychological attention on pronoun comprehension. *Language, Cognition and Neuroscience, 30*(7), 832-852.

Bates, D., Maechler, M., & Bolker, B. (2012). lme4: Linear mixed-effects models using S4 classes. R package version 0.999999-0. http://CRAN.R-project.org/package=lme4.

Bates, D., Kliegl, R., Vasishth, S., & Baayen, R. H. (2015). Parsimonious Mixed Models, Retrieved from http://arxiv.org/abs/1506.04967

Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mentioned in Spanish. *Psychonomic Bulletin Review, 2,* 124-129.

Crawley, R., Stevenson, R., & Kleinman, D. (1990). The use of heuristic strategies in the interpretation of pronouns. *J. of Psycholinguistic Research, 19,* 245-264.

Friesen, C. K. & Kingstone, A. (1998). The eyes have it! Reflexive orienting is triggered by nonpredictive gaze. *Psychonomic Bulletin and Review*, *5*(3), 490-495.

Gernsbacher, M. A., & Hargreaves, D. J. (1988). Accessing sentence participants: The advantage of first mention. *J. of Memory and Language, 27*, 699–717.

Goodrich Smith, W. & Hudson Kam, C. L. (2012). Knowing 'who she is' based on 'where she is': The effect of co-speech gesture on pronoun comprehension. *Language and Cognition, 4*(2), 75-98.

Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.

Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention in discourse. *Cognitive Science, 17,* 311-347.

Griffin, Z. M., & Bock, K. (2000). What the eyes say about speaking. *Psych. Science*, *11*(4), 274-279.

Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language, 69,* 274-307.

Hanna, J. E. & Brennan, S. E. (2007). Speakers' eye gaze disambiguates referring expressions early during face-to-face conversation. *J. of Memory and Language*, *57*(4), 596-615.

Järvikivi, J., Van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. *Psych. Science, 16,* 260-264.

Järvikivi, J., & Pyykkönen- Klauck (submitted). Visual context influences the representation of discourse referents in 4-year-olds.

Järvikivi, J., Pyykkönen-Klauck, P., Schimke, S., Colonna, S., & Hemforth, B. (2014). Information structure cues for 4-year olds and adults: Tracking eye movements to visually presented anaphoric referents. *Language, Cognition and Neuroscience*, 29, 877- 892.

Jiang, X. & Zhou, X. (2015). Who is respectful? Effects of social context and individual empathic ability on ambiguity resolution during utterance comprehension. *Frontiers in Psychology*, *6*, 1588.

Kaiser, E. & Trueswell, J. C. (2008). Interpreting pronouns and demonstratives in Finnish: Evidence for a form-specific approach to reference resolution. *Language and Cognitive Processes, 23,* 709-748.

Konopka, A. E. (2012). Planning ahead: How recent experience with structures and words changes the scope of linguistic planning. *J. of Memory and Language*, *66*(1), 143-162.

Koornneef, A. W. & Van Berkum, J. J. A. (2006). On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. *J. of Memory and Language, 54*, 445-465.

Kuznetsova, A., Brockhoff, P.B., & Christensen, R.H.B. (2015). lmerTest: Tests in linear mixed effects models. R package version 2.0-29.

Langton, S. R., Watt, R. J., & Bruce, V. (2000). Do the eyes have it? Cues to the direction of social attention. *Trends in Cognitive Sciences*, *4*(2), 50-59.

Nappa, R. & Arnold, J. E. (2014). The road to understanding is paved with the speaker's intentions: Cues to the speaker's attention and intentions affect pronoun comprehension. *Cognitive Psych.*, *70*, 58-81.

Pyykkönen, P. & Järvikivi, J. (2010). Activation and persistence of implicit causality information in spoken language comprehension. *Experimental Psych., 57, 5-16.*

Snodgrass, J. G. & Yuditsky, T. (1996). Naming times for the Snodgrass and Vanderwart pictures. *Behavior Research Methods, Instruments, & Computers*, *28*(4), 516-536.

Staudte, M. & Crocker, M. W. (2011). Investigating joint attention mechanisms through spoken human–robot interaction. *Cognition*, *120*(2), 268-291.

Van den Brink, D., Van Berkum, J. J., Bastiaansen, M. C., Tesink, C. M., Kos, M., Buitelaar, J. K., & Hagoort, P. (2012). Empathy matters: ERP evidence for inter-individual differences in social language processing. *Social Cognitive and Affective Neuroscience*, *7*(2), 173-183.

Van der Meulen, F. F., Meyer, A. S., & Levelt, W. J. (2001). Eye movements during the production of nouns and pronouns. *Memory & Cognition*, *29*(3), 512-521.