Published in final edited form as:

Lang Cogn Neurosci. 2015 August ; 30(7): 832–852. doi:10.1080/23273798.2015.1017511.

Effects of Psychological Attention on Pronoun Comprehension

Jennifer E. Arnold and **Shin-Yi C. Lao**

Department of Psychology, University of North Carolina at Chapel Hill

Abstract

Pronoun comprehension is facilitated for referents that are focused in the discourse context. Discourse focus has been described as a function of attention, especially shared attention, but few studies have explicitly tested this idea. Two experiments used an exogenous capture cue paradigm to demonstrate that listeners' visual attention at the onset of a story influences their preferences during pronoun resolution later in the story. In both experiments trial-initial attention modulated listeners' transitory biases while considering referents for the pronoun, whether it was in response to the capture cue or not. These biases even had a small influence on listeners' final interpretation of the pronoun. These results provide independently-motivated evidence that the listener's attention influences the on-line processes of pronoun comprehension. Trial-initial attentional shifts were made on the basis of non-shared, private information, demonstrating that attentional effects on pronoun comprehension are not restricted to shared attention among interlocutors.

Keywords

pronoun resolution; attention; discourse processing; accessibility

Understanding language frequently requires listeners to find referents for pronouns, for example in sentences like *Athena went skiing with Sylvia, and she fell down.* Listeners typically do so quite rapidly, despite the frequent ambiguity of pronouns, in part because they can draw on extra-linguistic information from the previous discourse and nonlinguistic context. Some entities (e.g., recently or prominently mentioned ones) are perceived as better referents. These are called **focused** or **salient** in the discourse, and are typically more accessible during the interpretation of subsequent referring expressions (Ariel, 1990; Bock & Irwin, 1980; Chafe, 1994; Grosz, Joshi, & Weinstein, 1995; Gundel, Hedberg, & Zacharaski, 1993; see Arnold, 1998, 2008, 2010 for reviews). The link between cognitive status and pronoun comprehension is widely accepted, but there are many open question about the mechanisms by which information becomes mentally privileged, and how this status affects pronoun comprehension.

One view suggests that some information is represented in memory in such a way that it is easier to access, and that this facilitates the comprehension of reference, in particular underspecified forms like pronouns. This mental status has been described as "salience", "accessibility", "activation", "givenness", "topicality" or "prominence" (for example, Ariel,

Address all correspondence to: Jennifer Arnold, UNC Chapel Hill, Dept. of Psychology, Davie Hall #337B, CB #3270, Chapel Hill, NC 27599-3270, jarnold@email.unc.edu; 919-843-5737.

1990, 2001; Arnold, 1998, 2008, 2010; Bower & Morrow, 1990; Brennan, 1995; Chafe, 1994; Foraker & McElree, 2007; Garnham, Traxler, Oakhill, & Gernsbacher, 1996; Gundel, Hedberg, & Zacharski, 1993; Givón, 1983; Grosz, Joshi, & Weinstein, 1995; Kaiser & Trueswell, 2004). Although existing proposals differ in important details, they share the use of nonlinguistic mental representations as explanations for both speakers' choices in production, and listeners' preferences in comprehension. This mental status is often assumed to be a gradient representation (e.g., Ariel, 1990; Arnold, 1998; Chafe, 1994), although many theories suggest that there is also a single most highly focused referent (Gundel, Hedberg, & Zacharski, 1993; Givón, 1983; Grosz, Joshi, & Weinstein, 1995; Stevenson et al., 2000). For example, following the sentence *Sofia jumped high on the trampoline*, the character *Sofia* is highly salient, and the most likely referent of a matching pronoun, e.g. … *and she laughed.* By contrast, entities that are unrelated to the current situation are extremely low in salience, and are unlikely to be considered as potential referents for pronouns.

The question we ask here is whether pronoun comprehension is influenced by the listener's attention. This is a question worth asking, because there is uncertainty in the literature about whether the mechanism behind pronoun comprehension relies on actual, psychological attention, or a language-specific category that is called "in focus". On one hand, some scholars discuss referential salience in terms of how it relates to the speaker's assumptions about the listener's attention (Gundel et al., 1993; Chafe, 1994) or memory (Ariel, 1990, 2001). On the other hand, much of the literature uses a more discourse-specific notion, termed simply "the focus" (e.g., Dahan et al., 2002; McKoon, Greene, & Ratcliff, 1993; Marslen-Wilson, Tyler, & Koster, 1993; Stephenson, Crawley, & Kleinman, 1994). The implication is that "in focus" information is also that which is attended by the discourse participants, but in some cases the link between attention and discourse focus is not explicit. Moreover, there have been few explicit tests of whether attention is actually involved in the representation of discourse entities, or how. Is it that pronoun comprehension is driven by a language-specific category that we call "focus", or by actual fluctuations in attention?

The answer to this question is not straightforward. The linguistic context has been proposed to define what is in focus, and what is not. We do know that the linguistic context affects reference comprehension generally, and pronoun comprehension specifically. Listeners tend to perceive as accessible those things that were recently mentioned, especially those mentioned in prominent linguistic positions like subject or first-mentioned position (Ariel, 1990; Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Chafe, 1976; Clark & Sengul, 1979; Givón, 1983, Brennan, 1995; Gordon, Grosz, & Gilliom, 1993), or the focus of a cleft (Almor, 1999; Arnold, 1998; Cowles, Walenski, & Kleunder, 2007; Foraker & McElree, 2007). Yet the causal relationship between discourse focus and attention is unclear. It may be that discourse context affects both pronoun comprehension and attention separately, or it may be that the discourse context affects the listener's attention, and it is these attentional fluctuations that drive pronoun comprehension. Which, then, guides pronoun comprehension – the discourse context or the attended state of some referents? Despite the frequent use of attention to explain discourse status effects during processing, there has been little direct assessment of its psychological role.

The one study that directly tested this question was Foraker & McElree (2007). They tested the hypothesis that linguistic focus leads to the active maintenance of a single entity in focal attention, using clefts to manipulate linguistic focus (e.g., *It was the new foreman who unrolled the latest blueprint. He…*). Following McElree's work on memory in the n-back task (2001), they argued that focal memory should lead to faster retrieval of the antecedent. However, they found that clefting did not affect speed, and instead only affected the likelihood of successfully retrieving the antecedent for the pronoun. They concluded that clefting does not result in actively maintaining the referent representation in focal attention, but instead affected the strength with which the representation of the character was encoded. Nevertheless, their findings are consistent with the broader proposal that discourse representations vary in their cognitive status. Moreover, they do not answer the larger question of whether this cognitive status is purely a linguistic constraint, or whether it is modulated by attentional processes that are not comparable to the processing needed for n-back tasks.

In this paper we tackle this question with a new approach, focusing on nonlinguistic, nonpublic modulations of the listener's attention. We hypothesize that the listener's discourse representation is modulated by their private attentional fluctuations, in addition to known effects of the linguistic context. In particular, when two entities are mentioned in an utterance, we hypothesize that attention to one of them can increase its mental accessibility, and thus its availability during later pronoun comprehension. Here we use the term "referential accessibility" to represent the privileged cognitive status (aka salience, prominence, etc.) that is predicted to facilitate pronoun comprehension. We assume that accessibility varies along a continuum (following Ariel, 1990; 2001; Arnold, 1998; Arnold & Griffin, 2001), and thus attention to an entity might result in a more activated or elaborated representation (Almor, 1999; Gernsbacher, 1990; McDonald & MacWhinney, 1990), for example making it easier to retrieve.

In order to dissociate the effects of attention from the linguistic context itself, we examine participants' allocation of visual attention to discourse entities at the onset of the trial, and examine the relationship between this initial attention and later pronoun processing. Our critical question is thus the following: in a story that mentions two characters at the beginning (e.g., "Doggy picked apples with Birdy… "), does increased attention on one character affect later pronoun processing?

We examine this question by measuring listeners' visual attention to each character as the story begins. In our task, visual attention at the start of the story is driven by a variety of item- and participant-specific constraints, as well as our visual capture manipulation. We focus on listeners' immediate interpretations of the pronoun, as it is encountered and immediately afterward – that is, online processing. We expect that in this domain we are most likely to observe any effects of attention over and above discourse biases. These on-line biases are then compared to measures of final, off-line interpretations of the pronoun.

Does attention underlie accessibility? Three hypotheses

The literature on reference processing and discourse representation suggests that there are three ways in which attention may be involved in the accessibility of discourse information. These hypotheses differ in their dependence on linguistic vs. nonlinguistic information, and shared vs. unshared information.

Accessibility as shared attention

One possibility is that pronoun comprehension is sensitive to the salience of only shared information. Here we use the term **shared attention** to denote information that one person considers likely to be attended by their discourse partner – that is, the attended portion of common ground. This is a somewhat weaker definition than **joint attention**, which may require both participants to be aware that the other is attending (Tomasello, 1995).

The importance of shared attention emerges in several accounts of referent accessibility. For example, Gundel et al.'s (1993) Givenness Hierarchy explains the cognitive status of discourse entities in terms of "assumptions that a cooperative speaker can reasonably make regarding the addressee's knowledge and attention state" (p. 275; see also Bard, Anderson & Sotillo, 2000; Brennan, 1995; Chafe, 1994; Levelt, 1989). This approach suggests that reference processing is only affected by evidence about the attention of one's interlocutor (see also Clark & Marshall, 1981).

The characterization of accessibility in terms of shared attention fits well with known discourse effects. The linguistic context is a public record of the current task. Interlocutors have good reason to assume that linguistic co-presence (Clark & Marshall, 1981) is a strong indicator of shared knowledge. Thus, the discourse record is a good estimate of what is mutually attended by discourse participants. In keeping with this, the effects of linguistic context on pronoun comprehension are often explained in terms of attention.

However, information that is in shared attention doesn't have to stem from linguistic sources. You could imagine watching a juggler perform on the street and saying "Wow, he's amazing" to the stranger next to you. The pronoun would be fully interpretable, based on evidence of physical co-presence, and physical cues like posture and eyegaze that indicate shared attention on the juggler -- even if the juggler had not been previously mentioned.

Nevertheless, there is little work that directly examines the effects of nonlinguistic shared attention on the accessibility of discourse entities. Instead, what work exists in this domain looks at nonlinguistic cues that co-occur with the referring expression itself. For example, eyegaze and pointing gestures facilitate the comprehension of both definite NPs (Hanna & Brennan, 2008), and pronouns (Goodrich & Hudson Kam, 2008, 2009, Nappa & Arnold, 2014). These findings leave open the question of whether shared attention affects the discourse representation itself. Thus, claims about the importance of shared attention are best supported by evidence about the importance of the linguistic context, making it difficult to distinguish accounts of accessibility as shared attention vs. accounts of accessibility as a linguistic construct.

Accessibility as egocentric attention

On the other hand, some researchers emphasize the attention of the comprehender. Bower and Morrow (1990; see also Morrow, Bower, & Greenspan, 1987, 1989) suggest that readers use mental models to represent the objects and characters in a discourse, and as they move through the discourse their attention shifts to the physical location in their mental representation. This raises a second possibility: it is individual, non-joint attention that guides accessibility. Likewise, discourse cues like first-mentioned status or recency could directly modulate the listeners' focus of attention without requiring representations of joint attention (Arnold, 2008). Because of this possibility, it is difficult to tease out the role of the listener's own attention from evidence of social cues to attention, like eyegaze.

Note that our question here is about whether egocentric attention affects the accessibility of discourse representations, and not about the listener's attention at the moment of encountering the pronoun, which does not appear to be influenced by noncommunicative visual cues (Nappa & Arnold, 2014)

Accessibility as a language-specific construct

A third possibility is that discourse information affects reference processing because there are discourse-specific rules that make some entities preferred as referents (e.g., first-mentioned ones). In fact, there must be some contribution of language-specific rules, since languages like Finnish have different pronouns that respond to different aspects of the linguistic context (Kaiser & Trueswell, 2008). More generally, the textual information in the discourse goes a long way toward predicting the preferred referent for a pronoun (Arnold, 1998). Centering Theory is a computational model that aims to use the text to make such predictions (Grosz, et al., 1995; also see Walker, Prince, & Joshi 1998 and papers within). In Centering, the highest-ranked entity referred to in the next utterance is the "center of attention." But it is unclear whether "center" is just a discourse category, or a description of a psychological state of the discourse participants. In fact, Grosz and Sidner (1986) state explicitly that in their model (a precursor to Centering theory), "The attentional state is a property of the discourse itself, not of the discourse participants," (p. 179). While this model implies that there is a relationship between discourse focus and psychological attention, the proposed mechanism for representing discourse status is an abstraction. In sum, the importance of the linguistic context is well-established, but these models are consistent with both accessibility as a language-specific construct, and accessibility as shared attention.

In this paper we examine whether pronoun comprehension is sensitive to fluctuations in accessibility that go beyond both shared information and linguistic rules about accessibility. To do so, we cannot simply examine the effects of the linguistic context, because as we have argued, these could guide reference comprehension either directly, or by affecting attention to discourse elements. We therefore turn to a nonlinguistic domain: visual attention.

Visual attention as a window onto referential accessibility

Visual attention offers a promising window onto questions about where referential accessibility comes from, because it is provides a measure of the listener's attention that can be both nonlinguistic and stem from sources that are not a part of the public discourse. For

example, imagine that a worker is giving a tour to a new secretary, and says *Here's your office. It's nice and sunny*. If the secretary happens to be gazing at the desk during this conversation, it would not change the fact that the desk is an inappropriate interpretation of the pronoun *it*. However, the linguistic context can sometimes include more than one referent, and the relative accessibility of each is somewhat ambiguous. For example, the worker may say *The phone is kept under the desk. It's pretty small, and we need to keep it clear for paperwork*. In this case, the pronoun *it* refers to the desk, even though it is moderately less accessible than the phone, which was mentioned first in the preceding sentence. The listener's own attention to the desk (for whatever reason) might modulate its relative accessibility, and facilitate processing of the pronoun *it*.

We therefore use visual attention as a way to examine whether attention is involved in discourse accessibility. The strongest support for an attentional role in accessibility would be evidence that accessibility effects are modulated by attentional variation that comes from nonlinguistic, nonshared sources. For example, if the sun happens to light up the desk and attract the listener's attention, it shouldn't overturn the effects of the public, shared discourse. But can it push the listener's attention to one of the task-relevant objects? Similarly, if the listener's attention is attracted to the desk for idiosyncratic reasons, does it change pronoun processing?

Measures of visual attention provide a good testing ground for the nature of linguistic focus of attention, because visual attention does not clearly map onto the linguistic category of focus. The "discourse focus" (or "center of attention") is typically used to describe backgrounded, given, topical referents -- i.e., those that are highly expected to be mentioned again. By contrast, visual attention is often attracted to objects that are new and need further processing, (see Bock, Irwin, & Davidson, 2004, for discussion) – e.g., those that appear suddenly (e.g. Jonides & Yantis, 1988; Remington, Johnston & Yantis, 1992), are unfamiliar (e.g. Brockmole & Boot, 2009; Theeuwes, 1994), or do not fit thematically with the scene (De Graef, Christiaens, & d'Ydewalle, 1990; Henderson, Weeks, & Hollingworth., 1999; Loftus & Mackworth, 1978, van der Muelen, Meyer, & Levelt, 2001).

There is reason to believe that visual attention might impact speakers' choices during reference production, given research on the speaker's choices about how to begin an utterance. For example, speakers tend to choose syntactic structures that allow them to begin their utterances with information that is more accessible conceptually or lexically (Bock, 1982; 1986; Bock et al., 2004; Ferreira & Dell, 2000; Ferreira, 2003; Gleitman et al., 2007; Tomlin, 1997). This accessibility is influenced by the same kinds of things as referential accessibility, e.g. the preference to order given before new (Arnold, Wasow, Losongco, & Ginstrom, 2000; Hawkins, 1994; Bock & Irwin, 1980). In addition, speakers tend to choose structures that allow them to begin their utterance with the referent that they initially fixated in a scene (Gleitman et al., 2007). Even more striking, Gleitman et al. were able to modulate the speaker's choice of starting point by manipulating their attention with an exogenous attentional capture cue. Speakers described scenes with two characters. Immediately before the picture appeared, a black square appeared for 80 ms, which increased the likelihood of starting the sentence with the cued character. These results suggest that visual and linguistic attention overlap at some level, despite their differences (but for a different interpretation

see Bock et al. 2004). The current study examines a parallel question: does attention (as measured by visual attention) contribute to accessibility effects in comprehension?

Does visual attention modulate the accessibility of discourse entities?

Two experiments tested the role of nonshared attention on pronoun comprehension. In both experiments, we monitored listeners' eye movements as they listened to a short story about two same-gendered characters, which included a pronoun that referred to one of them. We were particularly interested in the role that attention plays at the start of the discourse, as participants set up a representation of the story. In a story like *Doggy picked apples with Birdy…*, the story itself only provides partial information about who the discourse focus is. The first-mentioned/subject referent is generally considered more accessible (Arnold, 2010; Arnold, 2001; Brennan, 1995; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989; Järvikivi et al., 2005; Kaiser & Trueswell, 2004), but this is only a partial constraint. That is, the linguistic cues to accessibility are ambiguous. In this situation, the listener's attention at story onset may modulate the strength of the first-mention bias. This study is the first (to our knowledge) that examines the relationship between attention at the onset of a story and later pronoun comprehension.

As a measure of listeners' attention at story onset, we measured their fixations to discourse characters in the first second of the story. Although visual fixation is not isomorphic with visual attention (Mack, 2003), we did not restrict participants from making eye movements, which means that their looks were good indicators of their visual attention (see discussion in Chelazzi et al., 1995). In addition, our task encouraged participants to attend to the visual display by asking them to indicate whether the story matched the text (Arnold et al., 2000). This meant that visual fixations were a reasonable approximation of their attention to discourse characters.

We specifically wanted to examine how discourse processing was influenced by attentional modulations that were not limited to the linguistic input. We expected the listener's attention to fluctuate as a result of the visual stimulus, for example the tendency to view a picture from left to right (Dahan, Tanenhaus, & Salverda, 2007), as well as other idiosyncratic item properties. In addition, we introduced a modulation of the listener's attention via an attentional capture cue. Several studies have shown that visual attention is automatically captured by sudden onset cues, for example a black square on a red background (Folk, Remington, & Johnson, 1992; Folk et al., 1992; Jonides, 1981; Hopfinger & Mangun, 1998; Müller & Rabbit, 1989). This kind of capture cue is clearly not a part of the discourse itself, and should not affect pronoun comprehension – especially when the listener was not aware of it.

On the other hand, people are not always adept at ignoring information that is egocentrically available when participating in a joint activity like conversation. For example, expressions like "the blue triangle" are interpreted by considering both what information is available to the speaker (i.e., what is in common ground), and what objects in one's own ground are good matches for the input (Hanna, Tanenhaus, & Trueswell, 2003; Hanna & Tanenhaus, 2004; Keysar, Barr, Balin & Paek, 1998; Keysar, Barr, Balin, & Brauner 2000).

Furthermore, any representations of joint attention must be processed by the listener's internal cognitive system. Egocentrically attended information may therefore leak into calculations of shared attention. Such "leakage" effects may be especially apparent when they are partially consistent with shared information (as in the desk example). Nevertheless, effects of nonshared attention would provide the strongest evidence that pronoun comprehension is driven by attention per se, and not just discourse context.

Our use of a story context provided a fairly natural use of pronouns. It also meant that the effects of the listener's attention were tested in the context of discourse cues, specifically the known preference to interpret a pronoun as co-referential with a first-mentioned/ subject referent, rather than a second-mentioned/ nonsubject referent (Arnold et al., 2000; Gordon et al., 1993). We predicted strong effects of the discourse context. The critical question was whether the participant's attention at trial onset would also predict pronoun processing measures.

Experiment 1

Method

Participants—A total of 86 psychology students at UNC Chapel Hill participated in exchange for course credit. Data from 13 participants were not analyzed at all, because the participants reported on a voluntary questionnaire that they had noticed the flashes (n = 11) or that they had attention deficit disorder (n=2). An additional 21 participants were excluded because of problems with their fixations. The fixations for each participant were visualized using McMurray's EyeAnal software (2002), and those with excessive drifting were excluded. Participants were also excluded if more than 30% of their trials were unusable due to track loss, or if the participant had not fixated either the target or competitor during the entire trial (i.e. they used peripheral vision). This left 52 participants in the analysis.

Stimuli—The basic task paradigm was adapted from that used by Arnold et al. (2000). Participants viewed a picture on a PC computer (see Figure 1) while listening to a spoken story. All experimental verbal stimuli contained two sentences, e.g. *Birdy picked apples with Doggy near the farmhouse. He was wearing a hat to protect himself from the sun.* The first sentence introduced two same-gender characters, either Doggy and Birdy (male), or Kitty and Bunny (female). The introductory sentence mentioned one character in subject, first-mentioned position (N1), and the other in nonsubject, second-mentioned position (N2). Each character (Birdy, Doggy, Kitty, Bunny) was N1 five times. The second sentence began with a pronoun that referred to one of the characters; this was the target word. The referent of the pronoun was either a) the first-mentioned (N1) or b) the second-mentioned (N2) character from sentence 1. The pronoun referent was manipulated by changing the picture so that the disambiguating characteristic (e.g., wearing a hat) coincided with the intended referent. The disambiguating word (e.g., *hat*) always occurred after the verb; for example, in this picture both characters are wearing something, but only one is wearing a hat. This created an ambiguous period during which the scene was consistent with both pronoun interpretations (*He was wearing a)*. The critical question was where participants would look in during the period of time immediately following the pronoun. During the ambiguous period, the

proportion of looks to a character are assumed to reflect a consideration of that character as a referent, all other things being equal.

The experimental verbal stimuli used two verb types in the first sentence: 1) Joint-action, e.g. *Birdy picked apples with Doggy*, or *Kitty went sightseeing with Bunny*; and 2) Source-Goal, e.g. *Bunny showed a sculpture to Kitty*, or *Kitty floated over to Bunny*. Eight items were source-goal, and 12 were joint action. Verb type was controlled because it is known to influence the relative accessibility of its arguments; for example goal arguments are slightly more accessible than source arguments (Arnold, 2001; Rohde, Kehler, & Elman, 2007; Stevenson, Crawley, & Kleinman, 1994). However (see below), in this experiment it had little effect on our outcome measures.

The experimental visual stimuli contained two characters, one on each side of the screen. These were always in the same ports, equidistant from the center of the screen. A third object always appeared in the top center (e.g., the barn). These three objects were equidistant from the center and from each other. In some pictures there was also a fourth object in the center (e.g., the tree). The displayed picture took up the entire 1280 × 1024 resolution screen. The characters filled left and right ports that were 352 × 402 pixels in size.

Immediately preceding the story picture, a brief visual capture screen appeared for 200 msec. This screen pictured a 38 × 38-pixel black square in the location of one of the characters, as a manipulation of visual attention (cf. Gleitman et al., 2007). On filler trials the capture cue appeared in other locations.

Figure 2 illustrates the time course of stimuli for each trial. Participants clicked a GO button to initiate the trial, which was followed by a white screen with a black square in the location of one of the two characters. After 200 msec the square disappeared and the picture appeared, and 200 msec later the story began to play. The goal of this experiment was to test whether attention would modulate pronoun processing, in a situation where it was clearly not a part of the discourse record. The strongest way to achieve this was to ensure that participants would not be aware of the presence of the capture cue. Pilot testing revealed that a 200-msec sudden onset cue was not noticed by most participants. Post-experimental questionnaires confirmed this, and anyone who reported any suggestion of awareness of the cue was excluded from analysis (N=11).

Each trial was followed by a screen asking participants to press a button to indicate if the story and picture matched. A following screen prompted participants to provide an explanation for No-Match responses. This allowed us to identify rejections of the story that revealed a strong preference to identify the pronoun with the other character. For example, in the second-mentioned condition of the example, the story said *He's wearing a hat*, the picture showed Doggy in a hat. If the participant said *Birdy isn't wearing a hat*, it was taken as evidence that they couldn't accept Doggy as the pronoun referent.

Design—The experimental design was 2 (pronoun target N1 vs. N2) × 2 (target cued vs. target not cued). Several properties of the stimuli were balanced across the stimuli set, including verb type (joint action vs. source-goal), target location (left vs. right side of the

screen), and N1 character (Doggy, Birdy, Kitty, Bunny). 20 experimental items were pseudo-randomized with 8 filler items in four lists, each of which had both a forwards and a backwards version. Of the 8 fillers, 6 had stories that were not intended to match the picture for reasons that had nothing to do with pronominal reference (e.g., *Doggy went boating one summer*… when the picture showed Doggy snorkeling). There were two practice items, one that matched and one that did not match. The practice and filler items had one, two, or three characters, but used different structures and had no ambiguous pronouns. There was an attentional capture cue on all fillers, but in different locations.

Apparatus—We monitored participants' eye movements with a head-mounted Eyelink II eyetracker as they viewed the display and listened to the story. Eye position was sampled at a rate of one datapoint every 4 msec (250 Hz), but the data were converted to a granularity of one data point for every 16 msec (62.5 Hz) prior to analysis to speed processing (McMurray, 2002). Corneal reflection monitoring was used when possible (n=37); the pupil-only monitoring mode was used if we could not achieve adequate calibration with the corneal reflection mode (n=27). We analyzed only one eye, using the Eyelink automatic procedure for choosing the eye with better calibration. Calibration always received the highest rating of "good" by the Eyelink program (max deviation: x = 26.1 pixels, y = 118.4 pixels), otherwise the participant was re-calibrated. At analysis, fixations were allowed to deviate by 100 pixels to account for possible calibration error or calibration drift over the course of the experiment. This did not result in any overlap in the critical character ports.

The visual and auditory stimuli were presented on a PC computer running the ExBuilder software (an in-house software created at the University of Rochester; Longhurst, 2006), running on a PC computer with a 19″ monitor (resolution: 1280 × 1024 pixels, refresh rate: 75 Hz).

Procedure—We gave participants a general description of the task. Then participants were calibrated using Eyelink's 9-point calibration and validation procedure. Participants sat at the display computer so they could comfortably reach the mouse; viewing distance was approximately 22–34 inches. After calibration, participants were introduced to the characters in the stories and asked to name sample pictures of each character. Participants then performed two practice trials and were given a chance to ask questions. Then the experimental session began. Participants clicked on a crosshair at the center of the screen to begin each trial. During each trial, participants viewed a picture and listened to the story. When each story finished playing, another screen appeared which asked participants to indicate whether the story matched the picture by clicking a YES or NO button. A following screen asked participants to explain any NO responses out loud to the experimenter. A post-experiment questionnaire was used to assess participants' knowledge of the purpose behind the experiment, and whether they were aware of the capture cue.

Logic of analyses and statistical methods—Of primary interest in this study are the mechanisms that occur during language comprehension, as listeners attempt to identify a referent for ambiguous pronouns. Our interest in listener's **on-line** (i.e., momentary) biases stems from the assumption that they are related to the mechanism of arriving at a final interpretation. To this end, we first focus on evidence of participants' gaze immediately after

the pronoun, and consider how this relates to their final interpretation. As a secondary analysis, we also present the results of the off-line picture/story match task. In both cases, the purpose of this study is to examine the relationship between visual attention at the start of the story and later pronoun interpretation. We therefore also present analyses of eye gaze at the onset of the trial.

Eye gaze analysis procedure

Dependent variable: Eye movement data throughout the paper are presented in terms of **looks**, where a look is defined as a fixation grouped together with the prior saccade. This categorization is often used in language studies, which aim to measure the period of time during which listeners were directing their attention toward one referent and not another (e.g., Arnold, Hudson-Kam & Tanenhaus, 1997; McMurray, Tanenhaus, & Aslin, 2009; Sedivy, Tanenhaus, Chambers, & Carlson, 1999 and J. Sedivy, personal communication, March 2009; Trueswell, Sekerina, Hill, & Logrip, 1999 and J. Trueswell, personal communication; March 2009). The onset of the saccade represents the decision to begin looking at that object, and this look continues until the launch of a new saccade. Saccades were identified using Eyelink's on-line parser, which uses a velocity and acceleration-based detection algorithm. Using McMurray's (2002) EyeLinkAnal program in Microsoft Access, fixations were grouped by area of interest (target, competitor, object 1, object 2, other). In all eye movement analyses, trials were excluded if there was more than 33% track loss during the ambiguous period following the pronoun or if the participant failed to fixate either target or competitor for the entire trial (2.3% of the data)

We are particularly interested in the proportion of time spent looking at the target and competitor characters immediately following the pronoun (300 msec following pronoun onset until 300 msec following the disambiguation point). We therefore used a variant of Barr (2009)'s empirical logit, calculating the empirical logit of the ratio of the duration of target looks to the duration of competitor looks during this region: log ((# samples target looks + .5)/# samples competitor looks + .5)). This region was 690 msec long on average.

Analysis procedure and control variables: The effects of experimental conditions were evaluated in a mixed-effects linear regression model, using SAS proc mixed, including random effects for both participants and items, as well as random slopes for manipulated variables with respect to both participants and items (Barr, Levy, Scheepers, & Tily, 2013). As noted for each analysis, in some cases the random effects were estimated to be zero and were thus removed from the model. For each analysis reported, we first built a control model to assess the effects of potential control variables (verbtype; List (1,2,3,4); forward vs. backward order; first vs. second half; target location (left/right); participant gender (female/male); whether the participant was fixating the target at the onset of the pronoun (yes/no); whether the participant was fixating the competitor at the onset of the pronoun (yes/no). Any contributions that were significant at a level of t=1.5 or greater were retained for the final analysis.

The final analysis included the predictors of interest; referent (first vs. second mention), our metric of trial-initial attention, and visual capture cue. Each of these was also centered.

Those control predictors that reached criterion were also included. Control variables were deemed important, even though they were also controlled by experiment design, because our final analyses used an outcome measure as a predictor (see below). The inclusion of fixations at pronoun onset was a conservative choice, with the purpose of ensuring that the post-pronominal looks were not simply a function of gazes prior to pronoun processing (i.e., baseline effects).

Response analysis procedure: The task that participants performed was to decide if the story matched the picture. The primary motivation for this task was to encourage time-locking between eye movements and the linguistic input (Arnold et al., 2000). An additional advantage is that the responses provided some information about listeners' final interpretation of the story. On those trials where the participant said that the story and picture did not match, they also provided a verbal explanation of their response. This allowed us to identify items where the participant provided a reason that had something to do with the pronoun referent (e.g., *Birdy isn't wearing a hat*). These were categorized as "Pronoun Mismatch" responses. All match responses and mismatches with non-pronoun related reasons (e.g., *It looks more like a barn, not a farmhouse*) were categorized as "Pronoun Match" responses. A potential concern with this coding method is that difficulty with processing the pronoun in some conditions (e.g., for N2 referents) might lead to more Mismatch responses, but if participants cannot articulate the problem with the pronoun, our measure would underestimate differences between conditions in responses. However, there was no evidence that this was a problem: participants were unlikely to provide non-pronoun-related explanations for both N1 (4%) and N2 (2%) conditions. Note that this response measure is not the same as asking them what they thought the pronoun referred to. Rather, it is a measure of whether they were willing to accept the picture as a valid representation of the story.

Because match responses were binary, we analyzed them in a multilevel logistic regression, using SAS proc glimmix (version 9.1) with a Penalized Quasi-Likelihood (PQL) estimator, with a binary distribution and a logit link. The same procedure for including control variables was used as for the eye movement analyses.

Results

The goal of this study is to test the hypothesis that listeners' attention as a discourse develops can influence their biases as they seek out potential referents for a pronoun, where we manipulated attention at the onset of the story with a brief visual capture. We do this by examining the relationship between a) fixations at the start of the trial, and b) fixations when the ambiguous pronoun is encountered. We begin by examining fixations at the start of each trial.

Trial-initial attention: eye gaze at the start of the trial—We first examined looks to the animate characters as the trial began, as a way to assess how trial-initial attention was influenced by item features, both manipulated and control characteristics. We expected listeners to visually attend to the character who appeared at the location of the visual capture cue, which appeared for 200 msec immediately prior to the onset of the stimulus picture. We

also expected attention at trial onset to be constrained by a bias toward the left-hand character (Dahan, Tanenhaus, & Salverda, 2007).

Figure 3 illustrates the average allocation of visual attention during the first sentence of the story. Looks are plotted as the proportion of looks to N1 out of both N1 and N2, and as a function of both screen location (left vs. right) and capture cue (N1 cued vs. not cued). As the figure shows, the flash (i.e., capture cue) attracted attention before picture onset. When N1 was cued, proportion N1 looks were very high. When N2 was cued, looks to N2 were high, which shows up on this graph as a negative value for proportion of N1 looks. In addition, participants spent more time fixating the left-side character than the right-side character. After the first second of the trial, the effects of the physical properties of the stimulus diminished, and instead looks were influenced by the verbal input. The fact that all lines are above the 50% line at this point reveals a general bias toward the first-mentioned character.

We therefore chose the first second as the region for measuring the participant's attention at the start of the trial. This "Trial-initial attention" metric was calculated as the proportion of looks to the target out of all looks to both characters[1]. While this measure is rough, it provides an approximate handle on the participant's attention during the first second of the trial. To assess its relationship to predictors of interest, we built a model of trial-initial target looks, as predicted by condition (N1 vs. N2), flash (N1 vs. N2), and target location (left vs. right), plus control variables (List, forward, and participant gender)[2]. Trial-initial target looks were greater when the target flashed (β=.24; t=9.67, p<.001), when it was on the left (β =.34, t=8.0, p<.001). When two-way interactions between the condition, flash, and target location predictors were added to the model, they did not significantly predict trial-initial looks, and did not change the pattern of results.

A related question is whether looks at the onset of the trial reflect general viewing preferences for that picture, as opposed to processes related to the establishment of the discourse context. If so, we might expect Trial-initial attention to predict gazing throughout the trial. To examine this question, we analyzed participants' fixations between the sentences, i.e., from the offset of sentence 1 up to the onset of sentence 2 (avg. 430 msec). At this point, the "target" was not distinctive, because the pronoun had not yet been mentioned. As expected, a model of the target empirical logit revealed that it was not predicted by any of the critical predictors (condition: β = .099, p = .68; target flash: β = −0.13, p = .54; trial-initial attention: β = .13, p = .62) or control variables (List)[3]. This demonstrated that trial-initial attention did not have widespread effects on gaze preferences.

The effect of the trial-initial attention on pronoun resolution—The critical question was whether trial-initial attention affected gaze preferences in the ambiguous region following the pronoun. The effect of trial-initial attention is shown in Table 1 and Figure 4. For presentational purposes, we binned trials into either "High attention" (greater

[1]For the 25 trials on which there were no looks to either target or competitor in the first second, trial-initial attention to target was scored as 0.
[2]This model had a random slope for condition by items only, and a random slope by capture cue by both subjects and items.
[3]This model had a random slope for condition by both subjects and items, and a random slope by capture cue by both subjects only.

than 50% looks in the first second) or “Low attention” (fewer than 50% looks in the first second), although Trial-initial attention was entered into the statistical model as a gradient predictor. High attention trials were often those where the target was cued, but not always.

As Figure 4 shows, there were more looks to N1 than N2, for both N1 and N2 target conditions. This reflects the well-known N1 advantage. In addition, there was a boost in post-pronominal fixations for characters that had been attended at trial onset. For the N1-target trials (top panels), trial-initial attention to the target supported the N1 bias, resulting in little competition from the N2 character. For the N2-target trials (bottom panels), trial-initial attention to the target served to eliminate the on-line bias to N1.

A multilevel mixed effects model was used to assess the effects of pronoun target (N1 and N2) and trial-initial attention on character looks, where the dependent variable was the empirical logit of target looks during the post-pronominal ambiguous region. For the model, trial-initial attention was included as a continuous predictor, using the proportion looks to the target out of all character looks. The attentional cueing manipulation was also included as a critical predictor, since it was manipulated, and provided a critical comparison for the trial-initial attention predictor. As shown in Table 2, there was a significant effect of pronoun target, such that there were greater looks to N1 than N2 characters in the ambiguous region. There was also a significant effect of trial-initial attention, such that there were more post-pronominal target looks when the target had also been attended at trial onset. The attentional cueing manipulation itself did not contribute to the model. Note that the same critical pattern obtains even if all the control predictors are excluded from the model.

Responses: We also asked whether trial-initial attention affected responses. As shown in Table 2, there was a significant effect of target referent (N1 vs. N2), reflecting the finding that Match responses were more common for N1 than N2 referents. In addition, there was a small effect of trial-initial attention for responses in the N2 condition, where Match responses were higher when N2 had been attended. This emerged as a marginally significant effect of trial-initial attention in the model. Even though trial-initial attention had a greater effect on N2 targets, the interaction term (target x trial-initial attention) was not significant when added to the model. Nonetheless, the marginal effect was clearly driven by the N2 target items. A separate analysis of just N2 target trials revealed a significant effect of trial-initial attention ($t = 2.25$; $p = .025$).

Discussion

Experiment 1 demonstrated that participants’ attention at the onset of the trial was systematically related to two measures of pronoun resolution: 1) looks during the ambiguous region following the pronoun, and 2) responses to the offline picture verification task. These effects occurred despite the fact that participants were not aware of the visual capture cue that temporarily attracted their attention to one side of the screen.

At the same time, the effects of visual attention were secondary to the strong effects of order of mention. As expected, first-mentioned (N1) characters were preferred as the referents of the pronoun, both online and offline. The strength of order-of-mention as a discourse cue

was especially apparent in the offline findings. This suggests that even though initial attention modulated on-line consideration of referents, public discourse cues had the strongest effects on final interpretation.

Nevertheless, it is notable that we found even a small effect of egocentric attention. The participant's focus of attention at the onset of the trial was driven primarily by two sources of information, both of which were external to the discourse context: the capture cue, and being on the left side of the picture. Since none of the participants were aware of the capture cue, its effect is tantamount to any other pressure on the listener's private attention. That is, there is no public information that the cued character should be attended – for all the listeners know, they just happened to attend to the cued character. This provides clear evidence that the effects of attention on discourse accessibility go beyond the linguistic discourse context, and also go beyond shared attention more generally.

Experiment 2 examines the extent of this effect, by asking whether the effects of trial-initial attention persist in the presence of a public, shared attentional capture cue. If the cue is more visually salient, will listeners view it as a part of the discourse record and ignore their own attentional biases?

Experiment 2

Experiment 2 pits the listeners' own attentional biases at the onset of the story against a visually salient exogenous capture cue that occurred as the second character was mentioned. This experiment permits us to dissociate the effects of attention at trial onset from the effects of shared attentional cues. The task was identical to that of Exp. 1, except the capture cue was a yellow halo around one character, which appeared in the middle of the first sentence. This manipulation had three key properties. First, it was not concurrent with trial onset, which meant that trial-initial attention could be assessed separately from the capture cue. Second, it occurred immediately preceding the mention of the second character in the first sentence of the story. This tests an alternate hypothesis that referential accessibility is affected by attention at the moment that the characters are mentioned. The second-mentioned character has the most ambiguous accessibility (i.e., it is somewhat accessible by virtue of having been mentioned, but not highly accessible like N1 is), and thus has the most potential to be affected by an attentional manipulation. Third, the capture cue was more salient than the cue in Experiment 1, and all participants were aware of it. This awareness was necessary, since any visual capture cue that disrupts the static scene would be noticed. This created a further contrast between the participant's own attention at trial onset, and the salient, public visual capture cue.

Method

Participants—A total of 51 psychology students at UNC Chapel Hill participated for course credit. Data from three were not analyzed at all (1 reported having attention deficit disorder; and data for 2 participants data were lost due to technical problems. An additional 13 were excluded because of problems with track calibration, using the same criteria as for Exp. 1.

Design, Materials, and Procedure—The task, apparatus and procedure were identical to Experiment 1 except for the nature of the attention capture cue. A yellow halo around either N1 or N2 appeared for 200 msec, beginning 200 msec prior to the mention of the second character (see Figure 5). Since it takes about 200 msec to program and launch a saccade (Matin, Shao, & Boff, 1993), this should result in saccades to N2 around the time its name is mentioned. All other aspects of the design and procedure were the same, except that participants viewed the picture for 400 msec before the story began, not 200 msec.

Results and Discussion

As before, trials were excluded if there was more than 33% track loss during the post-pronominal ambiguous period, or if the participant failed to fixate either target or competitor for the entire trial (3.7% of all trials). We again categorized items by trial-initial attention during the first second of the trial, using the same procedure as in Exp. 1 (proportion target looks out of all target and competitor looks). Since the capture cue did not occur until between 1253 and 2244 msec into the trial, these early fixations were unaffected by the capture cue, and instead were driven by the participants' own decisions about where to fixate. In this experiment participants spent more time fixating the left character in the first second than the right. An analysis of looks in the first second revealed that it was influenced by target location (left/right; $\beta = 0.44$, $p < .001$); first mention ($\beta = .09$, $p = .007$); but not by the capture cue, which had not occurred at that point ($\beta = -0.02$, $p = .4$)[4].

All participants reported after the experiment that they were aware of the visual capture. This meant that the results provided a good contrast between the effects of unconscious, internally-driven attention at the onset of the story, and a visual capture cue that was obvious yet not really a part of the story task itself.

Eye movement analyses—Again we analyzed the dependent variable of empirical logit of target looks during the ambiguous region after the pronoun, i.e. from 300 after pronoun onset until 300 msec after the disambiguating word. Attention at trial onset was calculated in the same way as for Experiment 1. As shown in Table 3 and Figure 6, target looks after the pronoun were more likely if the target had been attended at the onset of the trial. In addition, as expected, target looks after the pronoun were greater for N1 than N2 targets.

To assess the statistical significance of these patterns, these data were again submitted to linear mixed effects models, following the same procedure as for experiment 1. The control predictors were the same, except for the fact that target location (right/left) was excluded as a predictor. The reason for this is that trial-initial looks were primarily driven by screen location, where left-side targets were fixated more often. As shown in Table 4, looks following the pronoun were influenced by both target referent (N1 vs. N2), and trial-initial attention. The capture cue manipulation had no effect on post-pronominal looks.

As in Experiment 1, we examined the possibility that trial-initial attention revealed general gazing preferences for the picture, and not pronoun comprehension specifically. If so, trial-

[4]This analysis included random intercepts for both subjects and items, and a random slope for first mention by items only and for capture cue by subject only.

initial attention should predict target looks throughout the story. We again examined gazes during the period between sentence 1 and sentence 2, and found no effect of trial-initial attention (β = .01, p = .98), capture cue (β =−.36, p = .11), or condition (β = .22, p = .31)[5]. If target location is added to this model, it has no effect and does not change the pattern of results.

Response Data Analysis—As shown in Table 3, participants were more likely to provide Match responses for N1 than N2. They were also numerically more likely to respond Match when N2 had been attended at trial onset (31%) than when it had not (23%). However, the effect of trial-initial attention was not significant in the model (See table 4). Instead, there was a marginal effect of the capture cue on responses. Match responses were higher for N2 targets that were cued (31%) than those that were not cued (24%), while Match responses were high for N1 targets overall (cued: 95%; uncued: 93%).

Discussion

Experiment 2 provided further evidence that on-line pronoun resolution is influenced by the participants' attention at trial onset. This effect emerged in a task in which visual attention was cued publically, with an obvious yellow flash cue in the middle of the story. Despite the salience of the visual cue, trial-initial attention had a greater effect than the capture cue on where participants looked following the pronoun. As in Experiment 1, trial-initial attention was an egocentric experience: participants were aware of the later flash, but initial attention was guided by stimuli characteristics that are not usually considered a part of the discourse record, such as being on the left side of the screen. This early attention did not have effects on the entire trial, but did emerge later when the ambiguous pronoun was encountered. Its effects on postpronominal looks is consistent with the hypothesized link between pronoun processing and the role of attention in establishing discourse representations of the characters.

A potential concern is that in this experiment, trial-initial attention was heavily influenced by the bias toward left-side characters. This means that we must consider two possible interpretations of the effect of trial-initial attention. One possibility is that post-pronominal looks were biased toward the attended character precisely because it had been attended at story onset, which resulted in a more developed discourse representation. A second possibility is that the bias toward the attended character in this experiment is really a bias toward the left-side character. We know that the left-side bias cannot be a general viewing preference, in that left-sidedness did not affect looks prior to the pronoun, in the inter-sentential region. Yet it is still possible that the effect of initial attention may be merely a preference to look at the left character at the moment of pronoun resolution. While the data for experiment 2 do not distinguish these possibilities, the effect of the left-side bias is inconsistent across experiments. This favors an interpretation in which gazing preferences at story onset affect discourse accessibility, which then affects online biases during pronoun processing. We will return to this in the general discussion.

[5]This model included random intercepts for both subjects and items, and a random slope for pronoun target but subjects only, and for capture cue by items only.

In contrast with Experiment 1, however, in this experiment the participants' offline responses were somewhat guided by the capture cue, and not by trial-initial attention. This is likely to result from the salient nature of the capture cue in this experiment. Participants were all aware of the yellow halo, and they may have taken it as an intended signal about the importance of the N2 character to the story. Other studies have shown that listeners can use social cues, like pointing, to help identify referents for ambiguous pronouns (Nappa & Arnold, 2014; Goodrich & Hudson Kam, 2009; Goodrich Smith & Hudson Kam, 2012).

Nevertheless, the attentional effects in this experiment were again secondary to the effects of the discourse. The predominant finding was that listeners preferred to link pronouns with N1 targets. This replicates the well-established finding that entities in subject or first-mentioned position tend to be salient, and preferred as pronoun targets (Arnold et al., 2001; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989; Jarvikivi et al., 2005; Kaiser & Trueswell, 2004).

General Discussion

Two experiments demonstrated that when listeners consider potential referents for a pronoun, they are initially biased toward the character that they attended at the beginning of the story. In Experiment 1, the listener's attention to one character at the start of the story increased the likelihood of considering that character as a potential referent for a pronoun that occurred later in the story, as indexed by fixations to that character following the pronoun. Attention at story onset was driven by the listener's idiosyncratic biases, as well as character location and our visual capture cue. Neither of these determinants of attention were significant predictors of post-pronominal looks, supporting the conclusion that it is the actual attention to the character that matters, and not the reason for that attention (see Gleitman et al., 2007, for a similar effect on production). In experiment 2, we again observed the effect of trial-initial attention, and the later capture cue did not modulate post-pronominal fixations – despite the fact that participants were more likely to notice it. The trial-initial biases were not indicators of general gaze preferences, in that they did not correlate with looks in the inter-sentential region. Rather, they specifically predicted looks during the post-pronominal period.

These findings suggest that attention at the beginning of a story can modulate the effects of the linguistic context on referent accessibility. If so, this would predict that similar effects of attention should occur in an experiment where there were no attentional capture cues whatsoever. This is precisely what we found in another study (Arnold, 2015). This experiment used the same paradigm as the experiments reported here, without capture cues of any sort, for the purposes of investigating variation in the strength of the N1 bias between men and women. The exact same measure of trial-initial attention was included as a control measure, and it yielded the same effect as reported here. In that study, target location again guided trial-initial attention, but not as strongly as in Experiment 2. Importantly, while trial-initial attention significantly predicted post-pronominal gazes, target location did not.

Thus, across three experiments we have consistent evidence that attention at the onset of a story increases the likelihood of considering the attended character when an ambiguous

referential expression is encountered. The sources of these attentional modulations include attentional capture, and properties of the stimuli, like left-side location – neither of which are properties that are typically considered a part of the discourse representation. At the broadest level, these findings support claims from the literature that attention is related to referential accessibility. Although this hypothesis is well accepted in the literature, it has received little explicit examination. Thus, this is the first solid evidence that referential accessibility is related to the comprehender's attention, and not limited to either "focus" as a linguistic construct, or cues to shared attention.

In contrast with the clear effects of egocentric attention on on-line processing, we found less consistent evidence that it affected participants' final Match responses. These findings lead to two conclusions. On one hand, the minor effects of attention on responses suggest that attentional modulations of accessibility can have real effects for discourse understanding. On the other hand, the relative strength of the discourse cue over the listener's attention highlights the strong contrast between public and private sources of information about referent interpretation. The strongest constraints on discourse comprehension stem from public, task-related sources of information, such as the linguistic discourse context.

The timecourse of attentional effects during reference processing

The results reported here also highlight the importance of attention at the story onset. The target's accessibility was boosted by comprehender's attention to the target in the first second of the story – but not at other points of the story. This boost occurred regardless of whether the comprehender's trial-initial attention was influenced by a capture cue (exp. 1), or not (exp. 2, and Arnold, 2015).

We propose that our findings are consistent with the idea that attention is important in the establishment of a discourse representation, in which some characters are represented with greater accessibility. We know that discourse processing involves the construction of non-linguistic representations of discourse entities and events (e.g., van Dijk & Kintsch, 1983; Kintsch, 1988; Johnson-Laird, 1983; Bransford, Barclay & Franks, 1972, Bower & Morrow, 1990; Sanford & Garrod 1981; Zwaan & Radvansky 1998). These representations may be more or less activated, where activation is gradient and influenced by multiple sources of information (Ariel, 1990, 2001). We hypothesize that trial-initial attention increases the activation of the representations of discourse entities, making them more available when the pronoun is encountered. It is at story onset that the listener is instantiating a discourse representation, and the relatively more attended entities are represented with more activation. Under this view, discourse cues should also influence the activation of each representation. While we do not have specific evidence that attentional mechanisms also mediate the influence of discourse information, it is possible that they do.

We conjecture that trial-initial attention was relevant in these experiments precisely because it preceded the linguistic information itself. The linguistic first-mentioned bias is inherently probabilistic (e.g., Arnold et al., 2000): while pronouns do tend to refer to first-mentioned characters, they can also refer to second-mentioned characters. At the onset of the story, listeners need to establish their perspective on the story. Their private attentional shifts may influence the way they interpret subsequent, probabilistic linguistic input, either

strengthening or weakening the first-mentioned bias. Thus, trial-initial attention was important because it was directed towards characters that were relevant to the later discourse. Once participants attended to a character, it may have led to further processing of the character and its role in the story. To the extent that the capture cue modulated trial-initial attention, our findings are consistent with evidence that endogenous (task-related) goals mitigate the effect of exogenous cues (Folk et al., 1992).

Thus, the current evidence is consistent with models in which the search for the referent of a pronoun is guided by the accessibility of the discourse representation. Pronouns are pragmatically specialized, in that they favor accessible referents. Thus, once a pronoun is encountered, it triggers the use of accessibility representations to guide the consideration of potential referents. Thus, character accessibility does not influence discourse comprehension in general, but specifically when relevant, such as when a referring expression is encountered.

Under this view, the listener's attention is more important in the construction of a discourse representation than at the moment of encountering a pronoun. This explains the constrast between the current data and those of Nappa and Arnold (2014, Exp. 2), who report that egocentric attention at the moment of hearing a pronoun (as opposed to at story onset) does **not** affect pronoun comprehension. Nappa and Arnold had participants watch videos of a speaker telling a story about two co-present puppets. At pronoun onset, a capture cue (a black square) appeared at the location of one puppet. Even though the capture cue increased judgments about character location, it did not influence either judgments about the pronoun referent or speed of response. By contrast, the only relevant attentional cues at pronoun onset were social ones: pointing, gazing, or a novel condition in which the speaker claimed that she controlled the black squares on screen. These findings are consistent with the idea that the role of egocentric attention only has its effect as character representations are being formed, early in the story. Once these representations are established, later modulations of egocentric attention are irrelevant.

The task-relevance of the cued characters in our experiment may also explain another apparent paradox: visual attention is often attracted to new information, whereas linguistic discourse cues typically attract attention to old or topical information – yet both facilitated on-line pronoun resolution. In our study, this may have occurred because the critical moment of attention occurred immediately preceding the onset of the story. All task information at this point was new. But this was also the point at which listeners began to assess exactly how central each mentioned entity was to the story. Thus, both visual and linguistic cues may have influenced the expectancy of hearing a later reference to that character (Arnold, 1998; 2008).

Our interpretation of the results is consistent with a characterization of accessibility as a gradient property of entity representations. This is consistent with Foraker & McElree's (2007) claim that discourse accessibility does not involve putting a single entity into a privileged, focal category. In our experiments, attention only subtly modulated referential accessibility, showing no evidence that attention led to the privileged focusing of the attended character. It also seems unlikely that participants were aware of their visual

attention at trial onset, suggesting their attention was not being actively directed or maintained.

Shared vs. egocentric attention

Our study used measures of the listener's private attentional biases, at the onset of the story. These provided the clearest evidence that the kind of attention that affects pronoun processing is more than just a linguistic construct. There was no sense in which the capture cue or the stimulus-specific properties of a character could be construed as a part of the discourse record.

This finding speaks against a strong interpretation of models in which pronoun comprehension is **only** guided by information in common ground. Proponents of such models might argue that our data are consistent with a vaguer sense of shared attention, in which the listener assumes that *anything* that they experience is also available to their interloctutor. However, this extreme "you know what I know" assumption is much weaker than the common view is that referential accessibility specifically emerges from assumptions about the attention or knowledge of one's discourse partners (e.g., Brennan, 1995; Chafe, 1994; Gundel et al., 1993). We have demonstrated that on-line biases during pronoun comprehension are influenced by information that is not public, and certainly not established as shared information. We observed a more general effect of participants' own decisions about where to attend in the first second of the trial, which can hardly be taken as evidence of shared attention. The capture cue in Experiment 1 contributed to the initial-attention effect, but participants were not aware of it. This provides striking evidence that the effects of attention on referential accessibility are not limited to situations where attention is clearly shared.

Nevertheless, we do not take these results as evidence that pronoun comprehension is primarily the result of egocentric representations of accessibility. While we found significant effects of trial-initial attention, they emerged most strongly in our measures of on-line consideration of potential referents, i.e. in the eye movement analyses. Responses were overwhelmingly driven by linguistic status: pronouns referring to first-mentioned characters were accepted far more often than pronouns referring to second-mentioned characters. This is consistent with claims that reference should be interpreted with respect to shared information (e.g., Brennan, 1995; Gundel et al., 1993; Clark, 1996; Chafe, 1994).

However, listeners may not always keep track of the source of their attentional modulations. Whether they attend to a character because it is important to the story, because it was cued, or because it occurred on the left side of the screen, this attention still contributes to later pronoun processing. This is consistent with mounting evidence the mechanisms of language processing do not categorically ignore entities that are unknown to their speaker when identifying referents for nominal expressions (Barr & Keysar, 2006; Hanna, Trueswell, & Tanenhaus, 2003; Hanna & Tanenhaus, 2004; Keysar, Barr, Balin & Brauner, 2000).

Conclusions

In sum, we have provided evidence that private shifts of attention influence both listener's transitory biases during pronoun resolution, and to a lesser extent, listener's final interpretation of the pronoun. This study provides the first evidence about the relationship between attention at trial onset, and later reference resolution, in a visual-world eyetracking study. It also supports widespread claims that pronoun comprehension is driven by modulations in the attention of discourse participants, which make some referents more accessible than others.

Acknowledgments

We are grateful to Florian Jaeger, Daniel Serrano, Daniel Bauer, and Chris Weisen for discussion of statistical methods. Many thanks also to Cathleen Sparks, Glenn Kern, and Jordan Todd for their help preparing stimuli and collecting data. This research was partially supported by NIH grant R01 HD-41522 and partially supported by NSF grant BCS-0745627.

References

Almor A. Noun-phrase anaphora and focus: The informational load hypothesis. Psychological Review. 1999; 106(4):748–765.10.1037/0033-295X.106.4.748 [PubMed: 10560327]

Ariel, M. Accessing noun-phrase antecedents. London: Routledge; 1990.

Ariel, M. Accessibility theory: An overview. In: Sanders, T.; Schilperoord, J.; Spooren, W., editors. Text representation: Linguistic and psycholinguistic aspects. Amsterdam: Benjamins; 2001. p. 29-87.

Arnold JE. Reference form and discourse patterns (Doctoral dissertation, Stanford University, 1998). Dissertation Abstracts International. 1998; 59:2950.

Arnold JE. The effect of thematic roles on pronoun use and frequency of reference continuation. Discourse Processes. 2001; 31(2):137–162.10.1207/S15326950DP3102_02

Arnold JE. Reference production: Production-internal and addressee-oriented processes. Language And Cognitive Processes. 2008; 23(4):495–527.10.1080/01690960801920099

Arnold JE. How speakers refer: The role of accessibility. Language And Linguistics Compass. 2010; 4(4):187–203.10.1111/j.1749-818X.2010.00193.x

Arnold JE. Women and men have different discourse biases for pronoun interpretation. Discourse Processes. 2015; 52:77–110.10.1080/0163853X.2014.946847

Arnold JE, Brown-Schmidt S, Trueswell J. Children's use of gender and order-of-mention during pronoun comprehension. Language And Cognitive Processes. 2007; 22(4):527–565.10.1080/01690960600845950

Arnold JE, Eisenband JG, Brown-Schmidt S, Trueswell JC. The rapid use of gender information: Evidence of the time course for pronoun resolution from eyetracking. Cognition. 2000; 76(1):B13–B26.10.1016/S0010-0277(00)00073-1 [PubMed: 10822045]

Arnold JE, Wasow T, Losongco T, Ginstrom R. Heaviness vs. Newness: the effects of structural complexity and discourse status on constituent ordering. Language. 2000; 76:28–55.

Baldwin DA. Infants' contribution to the achievement of joint reference. Child Development. 1991; 62(5):875–890.10.2307/1131140 [PubMed: 1756664]

Baldwin DA. Infants' ability to consult the speaker for clues to word reference. Journal Of Child Language. 1993; 20(2):395–418.10.1017/S0305000900008345 [PubMed: 8376476]

Bard EG, Anderson AH, Sotillo C, Aylett M, Doherty-Sneddon G, Newlands A. Controlling the intelligibility of referring expressions in dialogue. Journal Of Memory And Language. 2000; 42(1):1–22.10.1006/jmla.1999.2667

Barr DJ. Analyzing 'visual world' eyetracking data using multilevel logistic regression. Journal Of Memory And Language. 2008; 59(4):457–474.10.1016/j.jml.2007.09.002

Barr, DJ.; Keysar, B. Perspective-taking and the coordination of meaning in language use. In: Traxler, MJ.; Gernsbacher, MA., editors. Handbook of psycholinguistics. 2. Amsterdam: Elsevier; 2006. p. 901-938.

Barr DJ, Levy R, Scheepers C, Tily HJ. Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal Of Memory And Language. 2013; 68(3):255–278.10.1016/j.jml.2012.11.001

Bock JK. Toward a cognitive psychology of syntax: Information processing contributions to sentence formulation. Psychological Review. 1982; 89(1):1–47.10.1037/0033-295X.89.1.1

Bock JK. Meaning, sound, and syntax: Lexical priming in sentence production. Journal Of Experimental Psychology: Learning, Memory, And Cognition. 1986; 12(4):575–586.10.1037/0278-7393.12.4.575

Bock JK, Irwin DE. Syntactic effects of information availability in sentence production. Journal of Verbal Learning & Verbal Behavior. 1980; 19:467–484.

Bock, JK.; Irwin, DE.; Davidson, DJ. Putting first things first. In: Henderson, JM.; Ferreira, F., editors. The integration of language, vision, and action: Eye movements and the visual world. New York: Psychology Press; 2004.

Bower GH, Morrow DG. Mental models in narrative comprehension. Science. 1990; 247(4938):44–48.10.1126/science.2403694 [PubMed: 2403694]

Bransford JD, Barclay JR, Franks JJ. Sentence memory: A constructive versus interpretive approach. Cognitive Psychology. 1972; 3(2):193–209.

Brennan SE. Centering attention in discourse. Language And Cognitive Processes. 1995; 10(2):137–167.10.1080/01690969508407091

Brockmole JR, Boot WR. Should I stay or should I go? Attentional disengatement from visually unique and unexpected items at fixation. Journal of Experimental Psychology: Human Perception and Performance. 2009; 35:808–815. [PubMed: 19485693]

Chafe, WL. Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In: Li, CN., editor. Subject and topic. New York: Academic Press; 1976. p. 25-56.

Chafe, WL. Discourse, consciousness, and time. Chicago: Chicago University Press; 1994.

Chelazzi L, Biscaldi M, Corbetta M, Peru A, Tassinari G, Berlucchi G. Oculomotor activity and spatial attention. Behavioural Brain Research. 1995; 71:81–88. [PubMed: 8747176]

Clark, HH. Using language. Cambridge: Cambridge University Press; 1996.

Clark HH, Sengul CJ. In search of referents for nouns and pronouns. Memory & Cognition. 1979; 7:35–41.

Clark, HH.; Marshall, CR. Definite reference and mutual knowledge. In: Joshi, AK.; Webber, BL.; Sag, IA., editors. Elements of discourse understanding. Cambridge: Cambridge University Press; 1981. p. 10-63.

Cowles W, Walenski M, Kleunder R. Linguistic and cognitive prominence in anaphor resolution: topic, contrastive focus and pronouns. Topoi. 2007; 26:3–18.10.1007/s11245-006

Dahan D, Tanenhaus MK, Chambers CG. Accent and reference resolution in spoken language comprehension. Journal of Memory and Language. 2002; 47:292–314.

Dahan, D.; Tanenhaus, MK.; Salverda, AP. How visual information influences phonetically-driven saccades to pictures: effects of preview and position in display. In: van Gompel, R.; Fischer, M.; Murray, W.; Hill, R., editors. Eye-movements: a window on mind and brain. Oxford, England: Elsevier; 2007.

De Graef P, Christiaens D, d'Ydewalle G. Perceptual effects of scene context on object identification. Psychological Research. 1990; 52:317–329. [PubMed: 2287695]

Ferreira VS. The persistence of optional complementizer production: Why saying 'that' is not saying 'that' at all. Journal of Memory and Language. 2003; 48:379–398.

Ferreira VS, Dell GS. Effect of ambiguity and lexical availability on syntactic and lexical production. Cognitive Psychology. 2000; 40:296–340. [PubMed: 10888342]

Folk CL, Remington RW, Johnston JC. Involuntary covert orienting is contingent on attentional control settings. Journal of Experimental Psychology: Human Perception and Performance. 1992; 18:1030–1044. [PubMed: 1431742]

Foraker S, McElree B. The role of prominence in pronoun resolution: Active vs. passive representations. Journal of Memory and Language. 2007; 56:357–383.

Garnham A, Traxler M, Oakhill J, Gernsbacher MA. The locus of implicit causality effects in comprehension. Journal of Memory and Language. 1996; 35:517–543.

Gernsbacher, MA. Language comprehension as structure building. Hillsdale, NJ: Erlbaum; 1990.

Gernsbacher MA, Hargreaves DJ. Accessing sentence participants: the advantage of first mention. Journal of Memory and Language. 1988; 27:699–717. [PubMed: 25520541]

Gernsbacher MA, Hargreaves DJ, Beeman M. Building and accessing clausal representations: the advantage of first mention versus the advantage of clause recency. Journal of Memory and Language. 1989; 28:735–755. [PubMed: 25505819]

Givón, T. Topic continuity in discourse: An introduction. In: Givón, T., editor. Topic continuity in discourse: A quantitative cross-language study. Amsterdam: John Benjamins Publishing; 1983. p. 1-42.

Gleitman LR, January D, Nappa R, Trueswell JC. On the *give* and *take* between event apprehension and utterance formulation. Journal of Memory and Language. 2007; 57:544–569. [PubMed: 18978929]

Goodrich Smith W, Hudson Kam CL. Pointing to 'her': The effect of co-speech gesture on pronoun resolution. Language and Cognition. 2012; 4:75–98.

Goodrich W, Hudson Kam CL. Co-speech gesture as input in verb learning. Developmental Science. 2009; 12:81–87. [PubMed: 19120415]

Gordon PC, Grosz BJ, Gilliom LA. Pronouns, names, and the centering of attention in discourse. Cognitive Science. 1993; 17:311–347.

Grosz BJ, Joshi AK, Weinstein S. Centering: A framework for modeling the local discourse. Computational Linguistics. 1995; 21:203–225.

Grosz B, Sidner C. Attention, intentions, and the structure of discourse. Computational Linguistics. 1986; 12:175–204.

Gundel JK, Hedberg N, Zacharski R. Cognitive status and the form of referring expressions in discourse. Language. 1993; 69:274–307.

Hanna JE, Brennan SE. Speakers' eyegaze disambiguates referring expressions early during face-to-face conversation. Journal of Memory and Language. 2008; 57:596–619.

Hanna JE, Tanenhaus MK. Pragmatic effects on reference resolution in a collaborative task: Evidence from eye movements. Cognitive Science: A Multidisciplinary Journal. 2004; 28(1):105–115.

Hanna JE, Tanenhaus MK, Trueswell JC. The effects of common ground and perspective on domains of referential interpretation. Journal of Memory and Language. 2003; 49(1):43–61.

Hawkins, JA. A performance theory of order and constituency. Cambridge: Cambridge University Press; 1994.

Henderson JM, Weeks PA, Hollingworth A. The effects of semantic consistency on eye movements during complex scene viewing. Journal of Experimental Psychology: Human Perception and Performance. 1999; 25:210–228.

Hopfinger JB, Mangun GR. Reflexive attention modulates processing of visual stimuli in human extrastriate. Psychological Science. 1998; 9:441–447.

Järvikivi J, van Gompel RPG, Hyönä J. Ambiguous pronoun resolution: Contrasting the first-mention and subject-preference accounts. Psychological Science. 2005; 16(4):260–264. [PubMed: 15828971]

Johnson-Laird, PN. Mental models: Toward a cognitive science of language, inference and consciousness. Cambridge, MA: Harvard University Press; 1983.

Jonides, J. Voluntary vs. automatic control over the mind's eye's movement. In: Long, JB.; Baddeley, AD., editors. Attention and performance IX. Hillsdale, NJ: Erlbaum; 1981. p. 187-203.

Jonides J, Yantis S. Uniqueness of abrupt stimulus onset in capturing attention. Perception & Psychophysics. 1988; 43:346–354. [PubMed: 3362663]

Kaiser E, Trueswell JC. Interpreting pronouns and demonstratives in Finnish: evidence for a form-specific approach to reference resolution. Language and Cognitive Processes. 2008; 23:709–748.

Kaiser E, Trueswell JC. The role of discourse context in the processing of a flexible word-order language. Cognition. 2004; 94:113–147. [PubMed: 15582623]

Keysar B, Barr DJ, Balin JA, Brauner JS. Taking perspective in conversation: The role of mutual knowledge in comprehension. Psychological Science. 2000; 11:32–38. [PubMed: 11228840]

Keysar B, Barr DJ, Balin JA, Paek TS. Definite reference and mutual knowledge: Process models of common ground in comprehension. Journal of Memory and Language. 1998; 39:1–20.

Kintsch W. The role of knowledge in discourse comprehension: A constructive-integration model. Psychological Review. 1988; 95:163–182. [PubMed: 3375398]

Klein RM. Inhibition of return. Trends in Cognitive Sciences. 2000; 4:138–147. [PubMed: 10740278]

Koorneef AW, van Berkum JJA. On the use of verb-based implicit causality in sentence comprehension: Evidence from self-paced reading and eye tracking. Journal of Memory and Language. 2006; 54:445–465.

Levelt, WJM. Speaking. Cambridge: MIT Press; 1989.

Loftus GR, Mackworth NH. Cognitive determinants of fixation location during picture viewing. Journal of Experimental Psychology: Human Perception and Performance. 1978; 4:565–572. [PubMed: 722248]

Longhurst, E. ExBuilder. 2006. Computer program

Marslen-Wilson WD, Tyler LK, Koster J. Integrative processes in utterance resolution. Journal of Memory and Language. 1993; 32:647–666.

Matin E, Shao K, Boff K. Saccadic overhead: Information-processing time with and without saccades. Perception and Psychophysics. 1993; 53:372–380. [PubMed: 8483701]

McElree B. Working memory and focal attention. Journal of Experimental Psychology: Learning, Memory, and Cognition. 2001; 27:817–835.

Mack A. Inattentional blindness: Looking without seeing. Current Directions in Psychological Science. 2003; 12 (5):179–184.10.1111/1467-8721.01256

McDonald JL, MacWhinney B. The time course of anaphor resolution: Effects of implicit verb causality and gender. Journal Of Memory And Language. 1995; 34(4):543–566.10.1006/jmla.1995.1025

McKoon G, Greene SB, Ratcliff R. Discourse models, pronoun resolution, and the implicit causality of verbs. Journal of Experimental Psychology: Learning, Memory and Cognition. 1993; 18:1040–1052.

McMurray, B. Analysis scripts written in Microsoft Access. 2002. EyelinkAnal, version 1.5.

McMurray B, Tanenhaus MK, Aslin RN. Within-category VOT affects recovery from “lexical” garden paths: Evidence against phoneme-level inhibition. Journal of Memory and Language. 2009; 60:65–91. [PubMed: 20046217]

Morrow DG, Bower GH, Greenspan SL. Accessibility and situation models in narrative comprehension. Journal of Memory and Language. 1987; 26:165–187.10.1016/0749-596X(87)90122-7

Morrow DG, Bower GH, Greenspan SL. Updating situation models during narrative comprehension. Journal of Memory and Language. 1989; 28:292–312.

Müller HJ, Rabbit PMA. Reflexive and voluntary orienting of visual attention: Time course of activation and resistance to interruption. Journal of Experimental Psychology: Human Perception and Performance. 1989; 15:215–330.

Nappa R, Arnold JE. The road to understanding is paved with the speaker’s intentions: Cues to the speaker’s attention and intentions affect pronoun comprehension. Cognitive Psychology. 2014; 70:58–81.10.1016/j.cogpsych.2013.12.003 [PubMed: 24534295]

Posner, MI.; Cohen, Y. Components of visual orienting. In: Bouma, H.; Bouwhuis, D., editors. Attention and Performance. Vol. X. London, England: Erlbaum; 1984. p. 531-556.

Remington RW, Johnston JC, Yantis S. Involuntary attentional capture by abrupt onsets. Perception and Psychophysics. 1992; 51:279–290. [PubMed: 1561053]

Rohde, H.; Kehler, A.; Elman, J. Pronoun interpretation as a side effect of discourse coherence. Proceedings of the 29th Annual Conference of the Cognitive Science Society; 1997. p. 617-622.

Samuelson LK, Smith LB. Memory and attention make smart word learning: An alternative account of Akhtar, Carpenter, and Tomasello. Child Development. 1998; 69:94–104. [PubMed: 9499560]

Sanford, AJ.; Garrod, SC. Understanding written language. Chichester: Wiley; 1981.

Sedivy JC, Tanenhaus MK, Chambers CG, Carlson GN. Achieving incremental semantic interpretation through contextual representation. Cognition. 1999; 71:109–147. [PubMed: 10444906]

Stevenson RJ, Crawley RA, Kleinman D. Thematic roles, focus and the representation of events. Language and Cognitive Processes. 1994; 9:473–592.

Stevenson R, Knott A, Oberlander J, McDonald S. Interpreting pronouns and connectives: Interactions among focusing, thematic roles, and coherence relations. Language and Cognitive Processes. 2000; 15:225–262.

Theeuwes J. Stimulus-driven capture and attentional set: Selective search for color and visual abrupt onsets. Journal of Experimental Psychology: Human Perception and Performance. 1994; 20:799–806. [PubMed: 8083635]

Tily, H.; Piantadosi, S. Refer efficiently: use less informative expressions for more predictable meanings. Proceedings of the Workshop on the Production of Referring Expressions: Bridging the Gap between Computational and Empirical Approaches to Reference; 2009.

Tomasello, M. Joint attention as social cognition. In: Moore, C.; Dunham, PJ., editors. Joint attention, its origins and role in development. Hillsdale, NJ: Lawrence Erlbaum; 1995. p. 103-130.

Tomlin, RS. Mapping conceptual representations into linguistic representations: the role of attention in grammar. In: Nuyts, J.; Pederson, E., editors. Language and Conceptualization. Cambridge: Cambridge University Press; 1997. p. 162-189.

Trueswell JC, Sekerina I, Hill NM, Logrip ML. The kindergarten path effect: studying on-line sentence processing in young children. Cognition. 1999; 73:89–134. [PubMed: 10580160]

van der Meulen FF, Meyer AS, Levelt WM. Eye movements during the production of nouns and pronouns. Memory & Cognition. 2001; 29(3):512–521.10.3758/BF03196402 [PubMed: 11407428]

Van Dijk, TA.; Kintsch, W. Strategies of discourse comprehension. New York: Academic Press; 1983.

Walker, MA.; Joshi, AK.; Prince, EF. Centering theory in discourse. Oxford: Oxford University Press; 1988.

Zwaan RA, Radvansky GA. Situation models in language comprehension and memory. Psychological Bulletin. 1998; 123:162–185. [PubMed: 9522683]

Appendix

Verbal stimuli for Experiments 1 and 2

Source-Goal items

Birdy brought some mail to Doggy during a big rainstorm. He was carrying an umbrella, which was a lucky thing.

Bunny showed a sculpture to Kitty outside the art museum. She had brought a camera to take pictures of the artwork.

Doggy gave some batting tips to Birdy on a sunny day at the park. He held the bat carefully to ensure a good grip.

Doggy read a story to Birdy near a tree outside. He sat on a log because the ground was wet.

Kitty floated over to Bunny at the pool party. She had a bag of chips for everyone to eat.

Kitty sang a song to Bunny at the school car wash. She scrubbed the window carefully so it would sparkle in the sun.

Kitty taught a song to Bunny at the neighborhood cookout. She sang really badly and ruined the song.

Kitty taught the hula to Bunny near some palm trees. She took a sip of her soda because it was a hot day.

Joint-Action items

Birdy picked apples with Doggy near the farmhouse. He was wearing a hat to protect himself from the sun.

Birdy played board games with Doggy during the snowstorm. He picked up the dice because it was his turn to roll.

Birdy walked around the amusment park with Doggy in the rain. He got an ice cream cone even though it was cold and rainy.

Birdy went rollerskating with Doggy at the playground. He brought a frisbee so they could play with it later.

Bunny baked some cookies with Kitty with a brand new recipe. She took out a carton of milk to go with the cookies.

Bunny drew pictures with Kitty beneath the rainbow. She used a paintbrush to capture the vibrant colors.

Bunny threw a party with Kitty with a big piñata. She took a piece of cake and accidentally knocked the rest of it off the table.

Bunny went shopping with Kitty at the new mall. She wanted to buy a new watch because her old one was broken.

Doggy ate a snack with Birdy at the fruit stand He chose an apple because it looked really fresh.

Doggy made soup with Birdy in a big pot on the stove. He picked up the spoon to taste the soup.

Doggy went kite flying with Birdy near a bench in the park. He stood on a rock so that his kite would fly highest.

Kitty went sightseeing with Bunny at the old Spanish castle. She had brought a guide book for more information.

Figure 1.

Sample visual stimuli for Experiment 1, in a) first-mentioned pronoun condition, and b) second-mentioned pronoun condition. A sample first-mentioned capture cue is given in c).

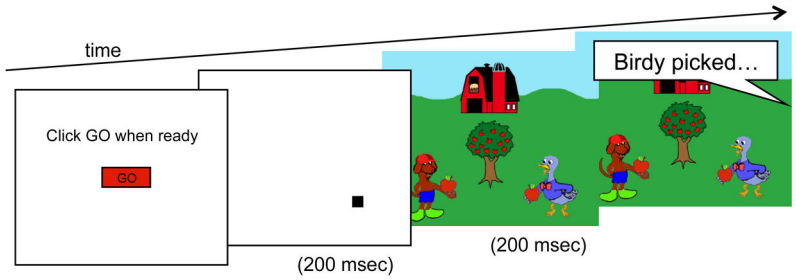

Figure 2.

Example timing of visual stimuli for Exp. 1 (N2 target, N1 cued condition). After participants initiated the trial, a capture cue appeared at the location of one character for 200 msec. The picture appeared for 200 msec before the story begins. (The story was heard only and not shown on screen).

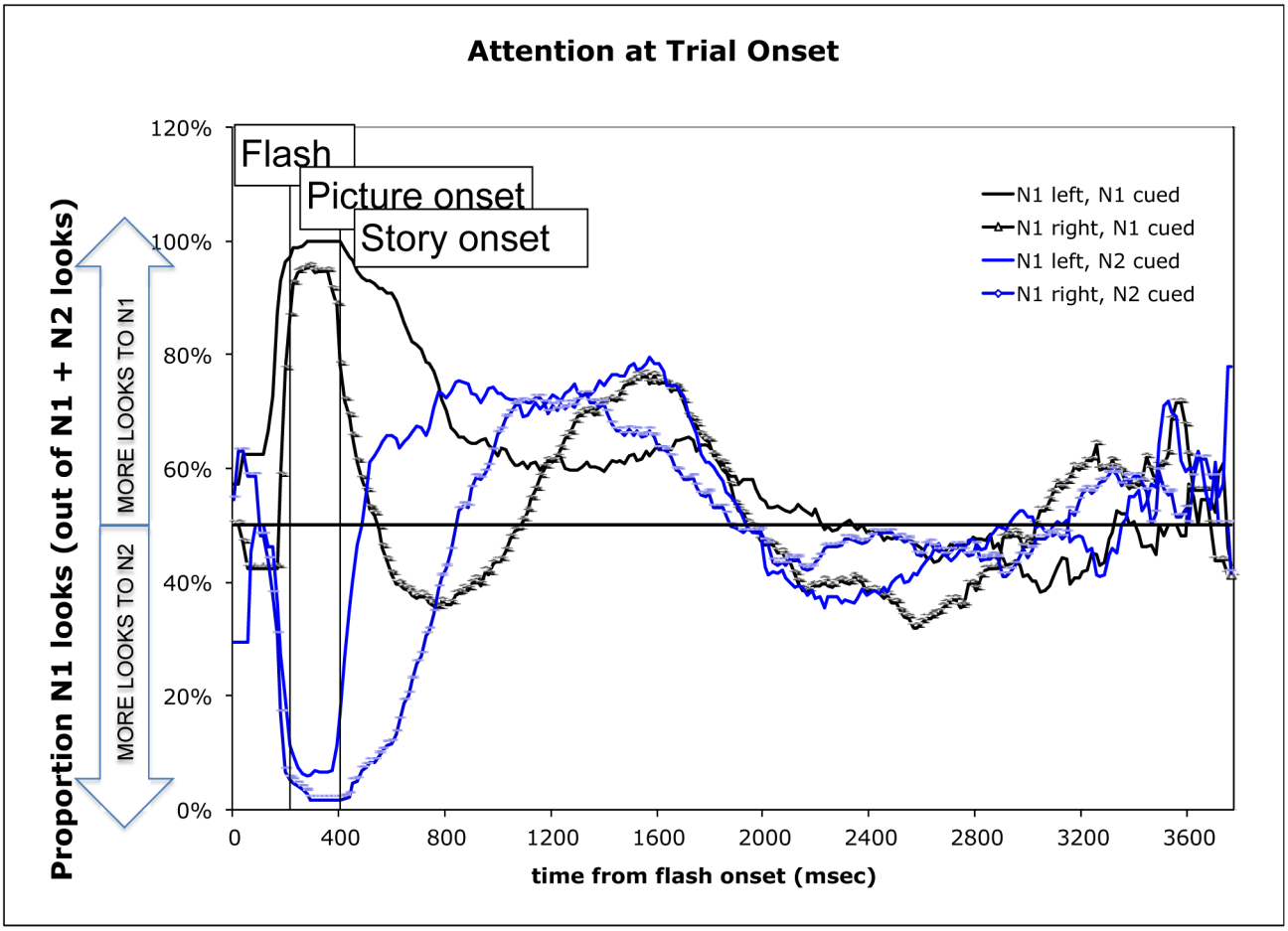

Figure 3.

Proportion of looks to N1 (the first-mentioned character) out of all looks to the two characters (i.e., N1/(N1+N2) at each time point during the first sentence of the story, as a function of the cueing and left/right position of N1. Note that when N1 was not cued, N2 was cued. The first character's name began at story onset; the second character's name began on average 1298 ms later. The pronoun began on average 430 msec after the end of sentence 1. Lines toward the top of the graph indicate greater looks to N1 than N2; lines toward the bottom of the graph indicate greater looks to N2 than N1.

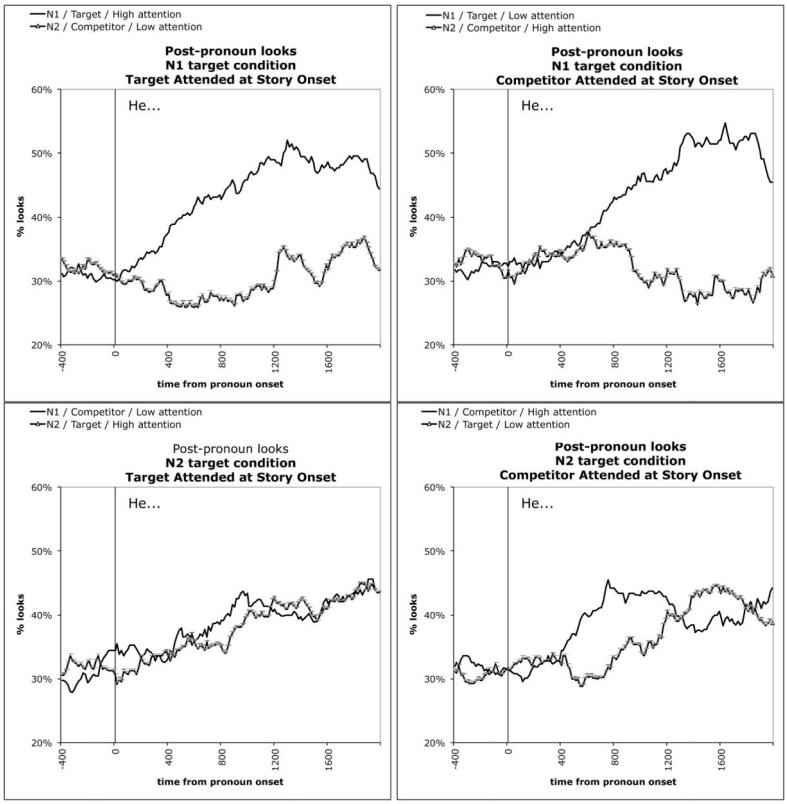

Figure 4.

Experiment 1 Results. Percentage looks (average participant means) to the N1 and N2 characters over time, where 0 marks the onset of the pronoun. The top panels depict N1-target conditions; the bottom panels depict N2-target conditions. The left panels represent those trials on which the target had been attended at trial onset; the right panels represent those trials on which the competitor (rather than the target) had been attended at trial onset.

a.  b.  c. 

Figure 5.

Sample visual stimulus for Experiment 2. The main picture (a) is followed by a 200 ms presentation of either (b) or (c), followed again by (a).

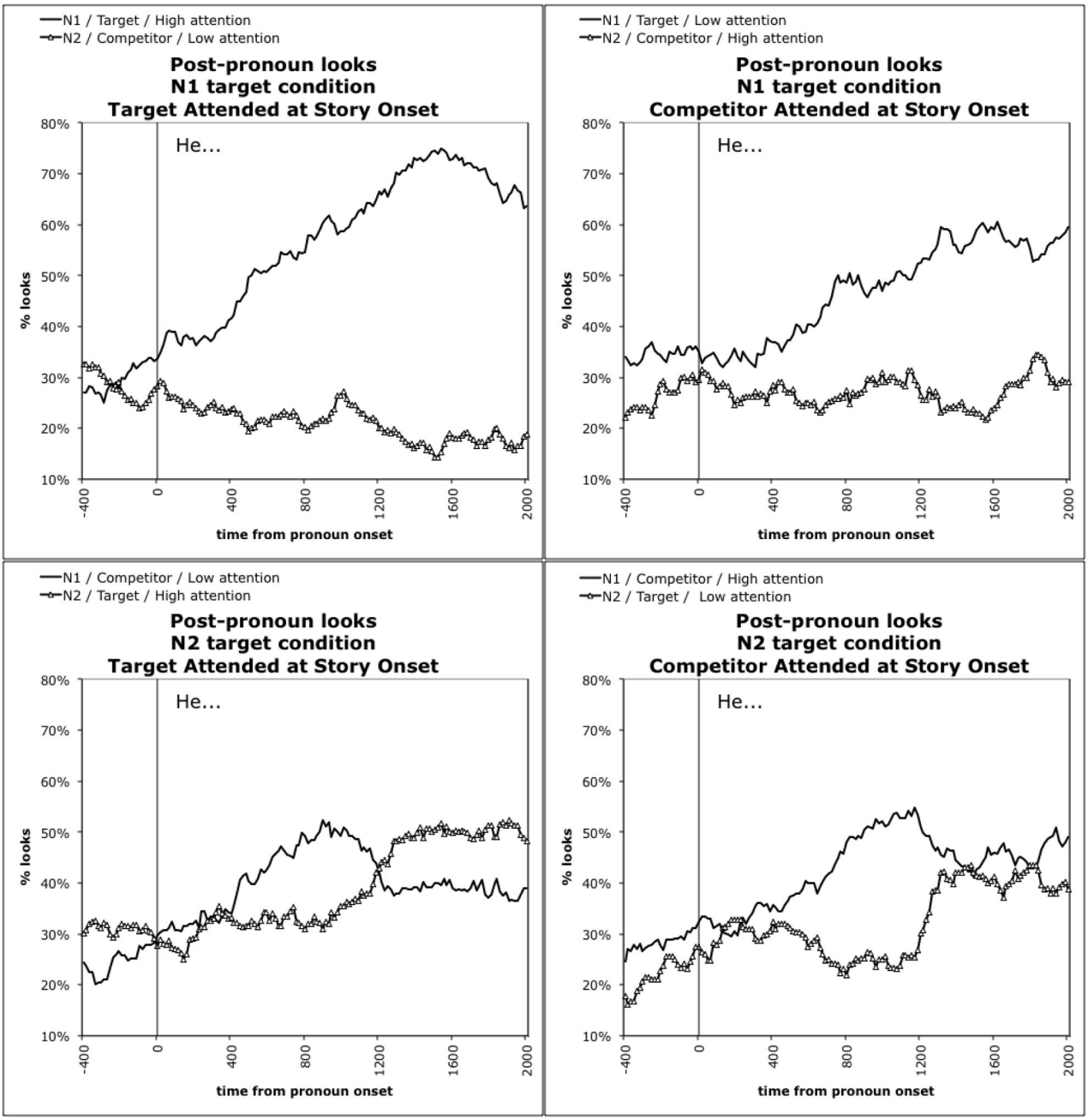

Figure 6.

Experiment 2 Results. Percentage looks (average participant means) to the N1 and N2 characters over time, where 0 marks the onset of the pronoun. The top panels depict N1-target conditions; the bottom panels depict N2-target conditions. The left panels represent those trials on which the target had been attended at trial onset; the right panels represent those trials on which the competitor (rather than the target) had been attended at trial onset.

Table 1

Experiment 1 results, Analysis 1: A) Average looks over the ambiguous region following the pronoun and B) Percentage Match responses. Trials are categorized by both condition and trial-initial attention (i.e., visual fixations during the first second of the trial).

Condition	Trial-initial attention	% N1 looks	% N2 looks	Target Empirical logit	%Match response	N
N1 target	High attention to target	0.56	0.18	0.79	97%	273
	Low attention to target	0.47	0.31	0.33	98%	231
N2 target	High attention to target	0.42	0.36	−0.11	21%	248
	Low attention to target	0.52	0.27	−0.49	11%	264

Lang Cogn Neurosci. Author manuscript; available in PMC 2015 August 01.

Table 2

Parameter estimates, t-statistics and p-values for critical and control predictors in the analysis of post-pronominal looks. 1016 observations included; data excluded for track loss. Models include random intercepts for participants and items, as well as slopes for pronoun referent by both participants and items. In the eye movements model, there was a random slope for capture cue by items only. The other random slopes for capture cue were estimated to be zero and therefore removed from the models. Dashes indicate that the variable did not reach criterion in the control model and was not included in the final model.

	EYE MOVEMENTS				**RESPONSES**			
Variable	**Estimate (error)**	**DF**	**t**	**p**	**Estimate (error)**	**DF**	**t**	**p**
PREDICTORS OF INTEREST								
N1 vs. N2 pronoun referent	0.85(0.19)	1004	4.54	<.0001	6.17(0.53)	29.54	11.58	<.0001
Trial initial target attention	0.31(0.13)	1004	2.31	0.021	0.73(0.43)	932.8	1.69	0.09
Target vs. competitor cued	0.04(0.09)	1004	0.44	0.663	0.16(0.27)	1009	0.61	0.54
CONTROL PREDICTORS								
List 1	0.21(0.14)	1004	1.5	0.134				
List 2	0.17(0.08)	1004	2.1	0.036	-	-	-	-
List 3	0.08(0.05)	1004	1.66	0.097	-	-	-	-
Forward List direction	−0.27(0.09)	1004	−3.02	0.003	-	-	-	-
Target on left vs. right side	0.27(0.18)	1004	1.52	0.128	-	-	-	-
Target fixated at pronoun onset	0.29(0.09)	1004	3.16	0.002	0.62(0.31)	1009	1.99	0.047
Competitor fixated at pronoun onset	0.13(0.09)	1004	1.49	0.136	0.39(0.30)	1009	1.31	0.189
First vs. second half	-	-	-	-	0.91(0.26)	1009	3.46	0.0006
Female vs. male participant	0.27(0.12)	1004	2.14	0.033	-	-	-	-
Verb bias	-	-	-	-	-	-	-	-

Table 3

Experiment 2 results: A) Average looks over the ambiguous region following the pronoun and B) Percentage Match responses. Trials are categorized by both condition and trial-initial attention (i.e., visual fixations during the first second of the trial).

Condition	Trial-initial attention	%N1 looks	%N2 looks	Target Empirical logit	%Match	N
N1 target	target attended at trial onset	0.52	0.22	0.61	0.93	180
	competitor attended at trial onset	0.43	0.25	0.35	0.96	159
N2 target	target attended at trial onset	0.43	0.33	−0.20	0.31	162
	competitor attended at trial onset	0.43	0.27	−0.33	0.23	173

Table 4

Parameter estimates, t-statistics and p-values for critical and control predictors in the analysis for Experiment 2. 674 observations included. Models include random intercepts for participants and items. The eye movements analysis included a random slope for pronoun referent by subjects, and a random slope for capture cue by items. The responses analysis included a random slope for pronoun referent by both subjects and items. Other random slopes were estimated to be zero. Dashes indicate that the variable did not reach criterion in the control model and was not included in the final model.

	EYE MOVEMENTS				**RESPONSES**			
Variable	**Estimate (error)**	**DF**	**t**	**p**	**Estimate (error)**	**DF**	**t**	**p**
PREDICTORS OF INTEREST								
N1 vs. N2 pronoun referent	0.71(0.12)	664	5.69	<.0001	4.38(0.54)	38.79	8.05	<.0001
Trial initial attention on target	0.4(0.13)	664	3.14	0.002	0.47(0.38)	329.5	1.23	0.221
Target vs. competitor cued	−0.04(0.13)	664	−0.29	0.77	0.46(0.24)	668	1.89	0.059
CONTROL PREDICTORS								
List 1	0.24(0.14)	664	1.71	0.088	-	-	-	-
List 2	−0.13(0.16)	664	−0.85	0.398	-	-	-	-
List 3	0.09(0.14)	664	0.61	0.54	-	-	-	-
Forward List direction	0.2(0.1)	664	2.01	0.044	-	-	-	-
Target fixated at pronoun onset	0.2(0.11)	664	1.76	0.078	0.13(0.27)	668	0.46	0.644
Competitor fixated at pronoun onset	0.22(0.11)	664	1.97	0.05	-	-	-	-
First vs. second half	-	-	-	-	0.39(0.24)	668	1.58	0.114
Female vs. male participant	-	-	-	-	-	-	-	-
Verb bias	-	-	-	-	-	-	-	-