

The role of prominence in pronoun resolution: Active versus passive representations [☆]

Stephani Foraker ^{*}, Brian McElree

Department of Psychology, New York University, 6 Washington Place, New York, NY, 10003, USA

Received 24 January 2006; revision received 25 July 2006

Available online 12 September 2006

Abstract

A prominent antecedent facilitates anaphor resolution. Speed-accuracy tradeoff modeling in Experiments 1 and 3 indicated that clefting did not affect the speed of accessing an antecedent representation, which is inconsistent with claims that discourse-focused information is actively maintained in focal attention [e.g., Gundel, J. K. (1999). On different kinds of focus. In P. Bosch & R. van der Sandt, (Eds.), *Focus: Linguistic, cognitive, and computational perspectives*. Cambridge: Cambridge University Press]. Rather, clefting simply increased the likelihood of retrieving the antecedent representation, suggesting that clefting only increases the strength of a representation in memory. Eye fixation measures in Experiment 2 showed that clefting did not affect early bonding of the pronoun and antecedent, but did ease later integration. Collectively, the results indicate that clefting made antecedent representations more distinctive in working memory, hence more available for subsequent discourse operations. Pronoun type also affected resolution processes. Gendered pronouns (*he* or *she*) were interpreted more accurately than an ungendered pronoun (*it*), and in one case, earlier in time-course. We argue that both effects are due to the greater ambiguity of *it*, as a cue to retrieve the correct antecedent representation.

© 2006 Elsevier Inc. All rights reserved.

Keywords: Pronoun resolution; Syntactic clefting; Working memory; Speed-accuracy tradeoff

[☆] Portions of this research appeared as part of SF's dissertation, and were presented at the CUNY Sentence Processing conference, 2005 and 2006. This research was supported by an NSF grant (BCS-0236732) awarded to B.M., and by a Summer Predoctoral Fellowship from NYU's Graduate School of Arts and Science to S.F. We thank Mike Allerhand, Chuck Clifton, Julie van Dyke, and Steven Frisson for programming advice and support, and the McElree lab for help with data collection. We also thank three anonymous reviewers for valuable comments.

^{*} Corresponding author. Present address: Department of Psychology, 5848 S. University Ave., Green 201, Chicago, IL 60637, USA. Fax: +1 773 702 0886.

E-mail address: sforaker@uchicago.edu (S. Foraker).

Introduction

Pronoun resolution requires access to previously processed representations. The antecedent representation that a pronoun refers back to has to be accessed in the comprehender's discourse model, and then aligned and integrated with the pronoun so that a coherent representation of the text or conversation can be formed. For example, when someone asks "Where is my toothbrush? Have you seen it?" a representation of the antecedent *toothbrush* must be made available in the comprehender's mind and integrated with the pronoun *it*. The reported research examines two accounts of how

antecedent representations for pronouns may be mentally represented in a comprehender's discourse model.

The cognitive state of a prominent representation

In general, pronouns are used to refer to psychologically more prominent entities, while forms with more semantic information are preferred for less prominent ones. Pronouns, such as *it*, tend to refer back to concepts that are highly predictable or salient in a discourse situation, such as *toothbrush* in the example above. For less prominent concepts (e.g., *toothbrush* in “Where is my black toiletries kit with the toothbrush and toothpaste in it?”), we typically use more lexically specific forms, such as a simple noun phrase (“Have you seen *my toothbrush?*”) or a more descriptive noun phrase (“Have you seen *my Panasonic power one?*”). Presumably, this occurs because these forms are better for selecting a particular entity than a pronoun would be (e.g., *toothbrush* for “Have you seen *it?*”). Based on such observations, Ariel (1990) suggests that there is a direct connection between the lexical form used for a referent in the discourse and the cognitive state of that concept in the comprehender's discourse model (see also Chafe, 1994; Givón, 1983).

Discourse prominence is controlled by several factors other than referential form, such as syntactic-semantic role and repetition of coreference (see Garnham, 2001 for a review). Hence, even for a particular type of coreferencing form—we examined pronouns in this study—researchers find that coreference with a more prominent antecedent appears to be easier than with a less prominent one. In particular, processing load is reduced when an antecedent is more prominent (see Garnham, 2001; Garrod & Sanford, 1994; Nicol & Swinney, 2003 for reviews).

We tested two accounts of how the cognitive state of a prominent antecedent can be characterized. Several researchers have suggested that prominent antecedent representations are more active in a comprehender's discourse model than less prominent ones. For example, approaches such as the Focus Memory Framework (Garrod, Freudenthal, & Boyle, 1994; Stewart, Pickering, & Sanford, 2000) and the Structure Building Framework (Gernsbacher, 1990; see also Garnham, Traxler, Oakhill, & Gernsbacher, 1996) draw on this metaphor in asserting that representations vary in strength along a continuum of energy, or activation strength. A continuum of strength is often assumed even when these approaches draw categorical distinctions between antecedent types (e.g., explicit versus implicit focus, Garrod et al., 1994).

Alternative accounts propose that the most prominent antecedent is maintained in a special cognitive state. Gundel (1999; Gundel, Hedberg, and Zacharski, 1993) argues that discourse factors that increase prominence place the most salient item in the psychological

focus of attention. Centering Theory, while not making as explicit a psychological claim, is in spirit also designed to capture this type of cognitive distinction (Grosz & Sidner, 1986). In a coherent discourse, the highest-ranked entity is essentially a forward-looking force, predicting the topic (Grosz, Joshi, & Weinstein, 1995). The key underlying assumption of these approaches is that if a prominent antecedent representation is in psychological focus, then it will be immediately on hand for operations such as bonding and resolution upon encountering the pronoun (Garrod & Terras, 2000; see also Garnham et al., 1996; Stewart et al., 2000; Greene, McKoon, & Ratcliff, 1992 for focusing accounts similar in spirit).

Based upon findings from studies of memory retrieval, we argue that accounts such as Gundel (1999) predict that prominent antecedents will differ from less prominent ones in the *speed* with which an antecedent representation can be accessed upon encountering a pronoun. In contrast, approaches which assume that prominent antecedents simply have a stronger representation in the discourse model predict that comprehenders will be more *likely* to access a prominent antecedent, resulting in an interpretation of higher quality. In the reported experiments, we measured the speed and accuracy of pronoun resolution as a function of prominence to test these two accounts of how prominent antecedents might be cognitively represented in the comprehender's discourse model.

Evidence for focal attention

Studies of memory retrieval indicate that information that is actively maintained in focal attention can be discriminated from information stored in a more passive memory state by the respective speed at which the two types of information can be accessed (McElree, 2001, 2006). For example, when no activity intervenes between the study and test of an item in a standard probe recognition task (e.g., Sternberg, 1975), the item is accessed at an exceptionally fast rate, from 30 to 50% faster than other items on the list (McElree, 1996, 1998; McElree & Doshier, 1989, 1993; Wickelgren, Corbett, & Doshier, 1980). Several lines of evidence indicate that such an advantage is due to this item having been actively maintained in focal attention at test time (for a review, see McElree, 2006). Studies using procedures where participants were challenged to actively maintain several items in focal attention (McElree, 1998, 2006) or procedures that encouraged participants to actively maintain items from different list positions (McElree, 2001, 2006) have likewise found a comparable processing rate advantage.

The same type of processing advantage appears to be evident in sentence comprehension, as well. McElree, Foraker, and Dyer (2003; see also McElree, 2000) investigated the memory operations that underlie resolving

syntactic dependencies within a sentence. Interpretation of a subject-verb dependency occurred at a measurably faster rate when the dependent elements were adjacent to one another in the input (e.g., *The editor laughed.*), compared to a slower rate that was common to all cases where other material intervened between the dependent elements (e.g., *The editor that the book amused laughed*, *The editor that the book that won the award amused laughed*, and *The editor that the book that the journalist wrote amused laughed*). This pattern directly mirrors the processing advantage seen in standard memory tasks for the recognition of an item in which no other chunk of information has intervened between study and test. Hence, findings in both domains suggest that processing is fast when information is active in focal attention, and, by hypothesis, when retrieval is not necessary.

Perhaps most relevant to coreference resolution are studies measuring the speed of access for an item that participants attempt to maintain in focal attention, while concurrently processing other information. McElree (2001) investigated this process with an *n*-back working memory task, in which participants were required to determine whether an item at a random position on a list matched the *n*th-item back in the list, either 1-, 2-, or 3-back. Analysis of the time-course data indicated that processing speed was determined by a mixture of two states: When the *n*th-back target item was successfully maintain in focal attention, processing speed was exceptionally fast, consistent with the idea that participants were able to match the test item directly to the contents of focal attention; when the *n*th-back target was displaced from focal attention, participants used a slower retrieval process to recover the *n*th-back target. This task likely involves many of the basic cognitive operations that are required in online discourse processing. That participants can successfully maintain a target item in focal attention across the presentation of intervening information demonstrates that approaches such as Gundel (1999), which assume active maintenance of a prominent antecedent in focal attention, are at least within the bounds of known cognitive abilities.

Passive versus active representations

To test the two accounts of how prominent antecedents are cognitively represented, we exploited the documented speed differences between accessing representations that are actively maintained in focal attention and retrieving representations from a more passive memory state. A focal attention account proposes that the most prominent representation is in a *qualitatively* different representational state than other discourse entities. Such representations are *active*, in the sense that they are actively maintained in current awareness, immediately on hand for further processing. For example, a clefted structure such as *It was his toothbrush that the traveler forgot to pack in the suitcase* would place a represen-

tation of the clefted noun phrase, *his toothbrush*, in focal attention. Hence, a later occurring pronoun (e.g. *it*) can be directly matched to “toothbrush.” The crucial prediction of this account is that the speed of accessing this representation should be markedly faster than the time required to access a representation outside of focal attention (e.g., *it* accessing a representation of *the suitcase*), because the pronoun can be directly matched to the contents of focal attention. In short, having the representation actively maintained in focal attention circumvents the need for retrieval processes to restore a passive representation to active processing. We refer to the speed of identifying and resolving a pronoun’s referent as the *accessibility* of the pronoun’s antecedent.

No such advantage is predicted for accounts which propose that the most prominent antecedent representation differs from other discourse entities in only a *quantitative* way, by having greater strength in memory. In this type of approach, all representations are *passively* stored in memory, in the sense that a retrieval process must be used to restore the item to current awareness for ongoing processing.

The likelihood that a representation can be retrieved from passive memory is a function of both encoding context and retrieval context (cues). A representation may be encoded into memory in a relatively distinctive or salient manner if the antecedent is linguistically or structurally marked (Almor, 1999; Garrod et al., 1994; Gernsbacher, 1989; Gernsbacher, Hargreaves, & Beeman, 1989; Gordon, Grosz, & Gilliom, 1993; MacDonald & MacWhinney, 1990), or if it is semantically richer and more connected to other discourse information, compared to other elements in the discourse (e.g., Garrod & Terras, 2000; McKoon & Ratcliff, 1980; McKoon, Gerrig, & Greene, 1996). Elaborative encoding processes increase an item’s distinctiveness in memory, and in doing so increase the probability that the representation can be retrieved (Nairne, 1996). Although all passive memory representations are subject to interference (e.g., Nairne, 1996), which decreases the likelihood they can be successfully retrieved from memory as new material is processed, a more richly encoded representation may be less susceptible to decay and interference over time, and there may be less interference from other potential competitors in memory at retrieval. A representation may also be retrieved more efficiently if the cues present at retrieval resonate more strongly with the antecedent (Greene et al., 1992; Ratcliff, 1978). For example, a pronoun such as *it* could refer to either *toothbrush* or *suitcase*, while one such as *he* would match with the antecedent *traveler* less ambiguously.

Crucially, accounts that do not assume that prominence places the antecedent in a special cognitive state predict that prominent antecedents will differ from less prominent ones only in terms of the likelihood that the antecedent can be retrieved and successfully matched

to the pronoun. We refer to this property as a difference in the *availability* of the antecedent for coreferential processing. Because this account assumes that a retrieval process must be used to restore *any* antecedent to active processing, we suggest that it does not predict an effect on the speed of resolving a pronoun. Our reasoning is based on studies of memory retrieval, outlined below, which have found that distinctiveness (strength or analogous properties) does not affect retrieval speed.

Discriminating between active and passive representations

Unfortunately, simple timing measures, including response time and reading time, provide measures of the relative difficulty of processing but do not clearly identify the source of the difficulty. Shorter processing times in these tasks could result from a prominent antecedent being *either* more accessible or more available than a less prominent one. Shorter processing times could indicate greater accessibility if that antecedent had been in an active state. But, they could also indicate greater availability. If a prominent antecedent had a stronger representation in memory, reading times could be shorter for several reasons. On some proportion of times, a less prominent antecedent representation may not be successfully retrieved when the pronoun is read. In the limit, this could cause interpretation to fail, or it could require the comprehender to initiate reanalysis to recover the antecedent. Even if the differences in the probability of successful retrieval are not substantial, the recovered representation for a poorly encoded antecedent may not support interpretive operations as well as a richly encoded antecedent. Again, this could cause interpretation to fail or could result in a less meaningful interpretation. For these reasons, any difference in reading time between a prominent and nonprominent antecedent could reflect differences in only the strength of two passive representations.

We used speed-accuracy tradeoff (SAT) modeling to discriminate between these two accounts of prominence. The primary benefit of this procedure is that the speed and the accuracy of processing can be measured jointly within a single task (e.g., Doshier, 1979; Reed, 1973, 1976; Wickelgren, 1977). In our application of the procedure, participants read sentences presented phrase by phrase and, at some point, were asked to decide (yes/no) whether the passage was sensible. Participants were trained to respond to an auditory response signal presented at 14 times after the onset of a crucial expression: a pronoun coreferent with either a prominent or nonprominent antecedent. Hence, the sampled times (0–5250 ms) enabled us to fully measure how the interpretation of the coreferring pronoun unfolded over time. For each sampled point, we constructed a d' measure of accuracy by scaling correct responses to sensible pronoun expressions (hits) against incorrect responses to control expressions with nonsensical pronoun interpre-

tations (false alarms). This scaling provided a measure of the ability of participants to discriminate acceptable from unacceptable pronoun interpretations.

To quantify how the interpretation of a pronoun unfolded over time, the full time-course data (observed d' points across t response lags) for each participant were fit with an exponential approach to a limit:

$$d' = \lambda(1 - e^{-\beta(t-\delta)}) \text{ for } t > \delta, \quad \text{otherwise } t = 0.$$

Fitting this equation to the time-course data allowed us to estimate how conditions varied in different phases of processing.

The asymptotes of the functions, reflected in the parameter λ , provided a measure of the highest level of discrimination reached with maximal processing time. If conditions differ in asymptote, it indicates that they differ in the likelihood that a meaningful interpretation was assigned to each type of expression or that the interpretation of the expressions differed in their degree of acceptability. The asymptote, therefore, indexes how successful comprehenders are at retrieving an antecedent for the pronoun, which is a precursor to constructing a meaningful interpretation of the expression. The asymptote will be lower if comprehenders fail to retrieve the correct antecedent more often for one condition than another, either because sometimes the antecedent was lost from memory or because a more salient antecedent interfered with the retrieval process. Whether asymptotic differences reflect failures to recover the antecedent or the inherent quality of the retrieved information (or a mixture of both), we interpret lower asymptotic performance as reflecting differences in the *availability* of information in memory essential to forming coherent interpretations of the anaphoric expression.

A primary advantage of the speed-accuracy tradeoff procedure is that it enables one to measure the speed with which an interpretation is computed. Hence, we can measure and compare the speed of interpretation for conditions that may also differ in overall accuracy (availability). The intercept (δ) and rate (β) of the function provide joint measures of the speed of processing, indexing how quickly accuracy accrues to its asymptotic level. The parameter δ estimates the intercept of the function, the point at which participants are first sensitive to information necessary to discriminate acceptable from unacceptable pronoun coreference (i.e. d' departs from 0, chance performance). The parameter β estimates the rate at which accuracy grows from chance to asymptote. If one condition can be interpreted more quickly than another, the SAT functions will differ in rate, intercept, or some combination of the two parameters (e.g., Bornkessel, McElree, Schlesewsky, & Friederici, 2004; McElree, 1993; McElree & Nordlie, 1999; McElree, Pylkkänen, Pickering, & Traxler, 2006). Whether speed differences are expressed in rate or intercept can be important in some theoretical contexts (e.g., McElree

& Doshier, 1993). However, the predictions we tested are based on general differences in speed of processing, so we treated δ and β as a single construct, and will refer to a difference in either parameter as a difference in the dynamics of processing. The observed dynamics of the function provide a measure of the *accessibility* of the antecedent information.

Accounts which assume that prominent antecedents have a stronger, more distinctive representation in the comprehender's discourse model predict higher asymptotic accuracy, but not differences in the dynamics of the time-course functions. This prediction follows directly from time-course studies examining the retrieval of sequentially presented items from working memory. For example, recency can be assumed to affect the availability of items in memory, as the quality of a memory trace will decrease as a function of the time since study (viz., decay) and the amount of information interpolated between study and test (viz., interference). Recency affects asymptotic accuracy, with lower performance for less recent items, but it does not affect the dynamics of the time-course functions (McElree, 1996, 1998, 2001, 2006; McElree & Doshier, 1989, 1993; Wickelgren et al., 1980). Other variables assumed to affect the quality of the memory representation, such as study time and the number of study trials (Doshier, 1984), also impact asymptotic performance only. Collectively, these studies are consistent with memory retrieval being mediated by a content-addressable mechanism, where cues in the retrieval context enable direct access to relevant memory representations, without the need to search through extraneous ones. The quality of both the memory representation and cues at retrieval affect the probability of retrieving a representation, but not the speed at which the representation is retrieved.

Crucially, if a prominent antecedent is actively maintained in focal attention (Gundel, 1999), then, extrapolating from the various memory studies mentioned above, it should be more accessible than other antecedents. Consequently, conditions that involve coreference to a prominent antecedent should engender faster dynamics (earlier intercept δ , or faster rate β) Than conditions that involve coreference to a less prominent antecedent, whether or not those conditions also differ in asymptotic accuracy.

Experimental manipulations

Prominence is controlled by several factors that guide coreference and contribute to matching a coreferring expression with an antecedent (see Garnham, 2001 for a review). We manipulated two factors in our experiments: whether the pronoun's antecedent was syntactically clefted or not, and the type of coreferring pronoun. Our main focus in these experiments was

how an antecedent made prominent by syntactic clefting was represented. A secondary factor, which arose from the clefting manipulations, was whether the pronoun was gendered, such as *he* or *she* used to refer to an animate antecedent, or not, such as *it* used to refer to an inanimate antecedent.

Syntactic clefting

Clefting is argued to increase prominence by placing focus, or contrast, on the fronted item (Carpenter & Just, 1977; Rochemont & Culicover, 1990). Clefting can preferentially direct a reader's or listener's attention to the clefted entity (Birch & Rayner, 1997; Carpenter & Just, 1977; Engelkamp & Zimmer, 1982; Klin, Weingartner, Guzmán, & Levine, 2004; Zimmer & Engelkamp, 1981), produce a stronger memory representation than other, nonclefted constituents (Birch, Albrecht, & Myers, 2000; Birch & Garnsey, 1995), and enable the clefted concept to be easily integrated with subsequent material (Morris & Folk, 1998).

Almor (1999) found that a noun phrase that coreferred with a clefted antecedent was easier to resolve than one that coreferred with a nonclefted antecedent. Using the same type of clefting structures, shown below in (1) and (2), Foraker (2004) found a similar clefting advantage during pronoun coreference. The pronoun occurred in a second sentence, either (a) *he* coreferring with "foreman," or (b) *it* coreferring with "blueprint." Italics denote the noun phrase made more prominent by clefting.

1. It was *the new foreman* who unrolled the latest blueprint.
2. What the new foreman unrolled was *the latest blueprint*.
 - (a) He squinted at the lines of the paper.
 - (b) It curled at the edges of the paper.

Self-paced reading times on the verb following the pronoun were shorter when the antecedent had been syntactically clefted, as in (1) (a) and (2) (b). These findings show that, for both types of cleft constructions, the antecedent of a pronoun has a more prominent representation when it is syntactically clefted. We used these same *It-* and *What-* clefting structures in Experiments 1 and 2, and used *It-* and *Whom-* clefts in Experiment 3.

Pronoun type

By using *It-* and *What-*clefts, a second issue emerged in the design: "what" cannot index an animate entity, which ruled out the possibility of using gender (*he* versus *she*) to uniquely identify the pronoun's antecedent. Therefore, we used the pronouns *he/she* to corefer with an animate noun that was stressed with an *It-*cleft, and *it* to corefer with an inanimate noun that was stressed with a *What-*cleft.

Gendered pronouns may be easier to resolve than the *it*-pronoun, for two reasons. First, *it* is potentially more ambiguous than *he* or *she*. *It* has several potential antecedent classes and linguistic functions, beyond referring to the explicitly mentioned antecedent, such as *blueprint*: *It* can also index the whole event in the first sentence; *it* can refer to a property of the situation or event, (cf. *It was amusing to watch*); and *it* can be used nonreferentially, as in a cleft construction (e.g., *It was the greedy salesman who eyed him*). Second, the *he/she*-pronoun conditions included other factors that increase the prominence of an antecedent: The animate noun *foreman* is the subject of the sentence (Arnold, Eisenband, Brown-Schmidt, & Trueswell, 2000; Gordon et al., 1993; Grosz et al., 1995; Hudson, Tanenhaus, & Dell, 1986), *it* is the first-mentioned item in the sentence (Carreiras, Gernsbacher, & Villa, 1995; Gernsbacher, 1989; Gernsbacher & Hargreaves, 1988; Gernsbacher et al., 1989; Järvikivi, van Gompel, Hyönä, & Bertram, 2005; Matthews & Chodorow, 1988; Speelman & Kirsner, 1990), and its coreferring pronoun is in a parallel syntactic and semantic role (Chambers & Smyth, 1998; Grober, Beardsley, & Caramazza, 1978; Stevenson, Nelson, & Stenning, 1995; Streb, Roesler, & Hennighausen, 1999; Vonk, 1984). In Experiments 1 and 2, these facilitating factors co-occurred with *he*-pronouns. In Experiment 3, we crossed these factors with the pronoun type by having them co-occur with *it*-pronouns.

Experiment 1: Speed-accuracy tradeoff modeling

Pronominal coreference with a clefted antecedent appears to be easier than coreference with a nonclefted antecedent. At issue in this experiment was whether the advantage was due to the clefted antecedent being actively maintained in focal attention or just being retrieved more often from a passive state. Both accounts assume that clefting should affect the availability of the antecedent for on-going operations, which would be reflected in higher asymptotic levels. If clefting increases

the distinctiveness of a representation in memory, then a clefted representation should be *more available* than a nonclefted one. Likewise, if a clefted representation is maintained in focal attention, then *it*, too, should be more available than other nonfocused, potential antecedent representations in memory. The accounts differ in that the focal attention account also predicts that a clefted antecedent should be *more accessible* than other representations by virtue of being maintained in focal attention. Differences in accessibility would be reflected in the dynamics of the fitted exponential function, in either rate or intercept parameters.

We also reasoned that *he*-pronoun conditions might show greater availability than *it*-pronoun conditions. First, the gendered pronouns provide a less ambiguous cue to recover the antecedent. Second, properties of the *he*-pronoun's antecedent (NP1: subject, first-mention, parallel roles) should increase prominence, while those of the *it*-pronoun's antecedent (NP2: object, second-mention, nonparallel roles) should decrease prominence. If *he*-pronouns also show greater accessibility, this would be consistent with two possibilities. One is that the more ambiguous *it*-pronoun conditions actually have a slower speed of resolution, perhaps because of reanalysis on occasions when the antecedent was not retrieved the first time, or not correctly interpreted. The other is that the antecedent in the *he*-pronoun conditions is actively maintained in focal attention.

Method

Participants

Participants in all three experiments were monolingual native American-English speakers with no history of reading difficulties. They were students at New York University, receiving partial course credit or pay for participating. Twelve participated in Experiment 1.

Materials

Table 1 shows the 2 × 2 design of clefting (Clefted versus Nonclefted antecedent) and pronoun type (He

Table 1
Example materials from Experiment 1

Condition	Antecedent and pronoun sentences
Clefted-He (NP1)	It was/the ardent boyfriend/who contemplated/the engagement ring. He stared. # He sparkled.
Clefted-It (NP2)	What/the ardent boyfriend/contemplated was/the engagement ring. It sparkled. # It stared.
Nonclefted-He (NP1)	What/the ardent boyfriend/contemplated was/the engagement ring. He stared. # He sparkled.
Nonclefted-It (NP2)	It was/the ardent boyfriend/who contemplated/the engagement ring. It sparkled. # It stared.

Note. Slashes indicate presentation regions for the first sentence. # Indicates the unacceptable continuation paired with the first sentence.

versus It). The clefting label refers to the pronoun's correct antecedent (i.e., clefted or nonclefted antecedent), as there was always a cleft structure present in the first sentence. The unacceptable versions of each condition (indicated with #) were constructed by switching the verbs in the pronoun sentence, producing a selectional restriction violation at the verb, based on animacy.

We did not include an explicit contrast set for the antecedents in a discourse context before the clefting sentence (following Almor, 1999; Birch & Garnsey, 1995; Birch & Rayner, 1997; Engelkamp & Zimmer, 1982). But since clefting plays a highlighting or contrastive function, we did include a descriptive adjective as part of the antecedents to encourage formation of an implicit contrast set.

A pronoun does not itself uniquely identify a referent; rather, it relies on other constraints to do so. We attempted to minimize ambiguity of reference by introducing two plausible referents in the first sentence, and having participants judge the acceptability of the two sentences together as one scenario (see Procedure). We attempted to minimize the ambiguity of the pronoun-antecedent combinations, with one referent being a person (a stereotypically male or female occupation or kinship term), and the other being an inanimate thing or object. The pragmatic information at the verb following the pronoun made the intended referent unambiguous.

To produce stable fits to the SAT function for each participant, it is necessary to collect more data per participant than in other tasks such as self-paced reading or eye-tracking. In memory and attention experiments, the design strategy is often to expose a participant to each item in each condition, distributed across different sessions. However, the standard in language processing experiments is to have no or minimal repetition of items, as the gist of a sentence can be recognized, even after days. Therefore, a total of 156 items were developed for the four conditions and two continuations, and the resulting 1248 stimuli were distributed across eight counterbalanced lists. Each list was divided into three sessions, comprising 104 stimuli per session. The critical trials constituted 26% of the trials in each session. The remaining 74% were fillers of various one and two clause constructions, with an equal number of acceptable and unacceptable sentences (unacceptable regions are underlined e.g., *The carpenter built the table/monsoon*; *The climber imagined the fall/acuity survivable*; *The hair stylist blow-dried the straight hair/vision*; *Before jumping into a taxi, the reporter and the candidate had a quick interview/greenhouse*). Across the three sessions, participants read each item only twice, in counterbalanced combinations of the four conditions and two continuations. A representative subset of the materials appears in Appendix A.

Procedure

The materials were randomized within a session and presented on a personal computer running E-Prime (Schneider, Eschman, & Zuccolotto, 2002), which recorded button press responses and latencies. We used the multiple-response SAT procedure (Bornkessel et al., 2004; McElree, 1993; McElree et al., 2006; Wickelgren et al., 1980) to conjointly measure processing speed and accuracy.

Participants were first presented with a screen reminding them of the keys indicating acceptable and unacceptable judgments. For each trial, one of these keys was randomly designated to start the trial. The antecedent sentence was presented phrase-by-phrase across one line in the center of the screen, remaining for a duration of 335 ms/word. Three hundred milliseconds before onset of the pronoun sentence, a series of 14 auditory response cues (100 ms duration) occurred every 350 ms, with a total time span of 5250 ms. The first tone served as a warning, alerting participants to begin pressing the randomly designated key (that they started the trial with) in sync with each subsequent auditory response cue. The complete second sentence (*He stared.*) appeared on the screen until the tones concluded. Participants were instructed to read the sentences as normally as possible and indicate whether they currently made sense together (whether the second sentence was an acceptable continuation of the first) by pressing the appropriate response key, modulating their responses if their assessment changed. We recorded the accuracy of a response for each sampled time, which enabled us to fully measure how the interpretation of the pronoun sentence unfolded over time.

Participants first completed a 1 h practice session to familiarize them with the concurrent tasks. They were trained on pressing and switching responses rhythmically across the sampling period to ensure that they were practiced at modulating their responses, and became comfortable with the continuation judgments. The experimental sessions consisted of 3 1 h–10 min sessions on subsequent days. Between-trial intervals were participant controlled, and there were two mandatory breaks each session.

Data analysis

Comprehension accuracy was calculated using a standard d' measure ($d' = z(\text{hits}) - z(\text{false alarms})$), where a "hit" was an "acceptable" response to an acceptable pair of sentences and a "false alarm" was an "acceptable" response to an unacceptable pair. A common false alarm rate from all four unacceptable continuations was used in the d' scaling. This made d' for each condition sensitive to differences in the hit rate, or acceptability of the continuation, rather than reflecting sources of unacceptability (such as the local violation between the pronoun and verb in the second sentence, or temporary differences in the unacceptable continuations such as *She crinkled* being at least temporarily compatible with a verb particle

continuation such as *She crinkled up her nose*, while *It scribbled* is not). Note that if acceptability judgments were made based only on the local mismatch between pronoun and verb in the second sentence, no differences between conditions are predicted.

A hierarchical model-testing scheme was used to determine whether the four conditions differed in asymptote (λ), rate (β), or intercept (δ). For each participant and the averaged data, separate parameters were allotted to the different conditions and tested for whether they systematically improved the fit of the SAT function to the observed d' data. The exponential function (see Discriminating between active and passive representations, above) was fit to the data with an iterative hill-climbing algorithm (Reed, 1976), similar to STEPIT (Chandler, 1969), which minimized the squared deviations of predicted values from observed data. Fit quality was assessed by an adjusted- R^2 statistic—the proportion of variance accounted for by the fit, adjusted by the number of free parameters (Judd & McClelland, 1989)—and by an evaluation of the consistency of the parameter patterns across the individual participant fits.

Additionally, we performed inferential tests of significance computed over individual participants' d' data, and the fitted parameter estimates for each of the candidate models detailed in the Results section (repeated-measures 2 [clefting] \times 2 [pronoun] ANOVAs, or paired t -tests). We provide a 95% confidence interval (CI) around the mean difference for a priori comparisons.

Averaged data and fits to that data are used to illustrate consistent patterns across participants. This experiment was designed to compute stable SAT model fits for individual participants, using many items to limit repetition (see Materials). Unfortunately, this did not allow meaningful item analyses, as there was too little data per item. However, the important findings were replicated with SAT modeling in Experiment 3 and the eye-tracking measures in Experiment 2, both of which report item analyses. The items used in those experiments are a subset of this one, providing little reason to believe that the effects reported here do not generalize across items.

Results

The best-fitting $4\lambda-1\beta-2\delta$ model is shown as the smooth function in Fig. 1, which summarizes the findings of this experiment. We found that clefted antecedent conditions produced higher asymptotic accuracy than nonclefted conditions, as predicted. As well, we found that *he*-pronouns produced higher asymptotic accuracy than *it*-pronouns. There was no interaction between the factors. Crucially, clefting had no effect on the speed of processing. We did find, however, that *it*-pronoun conditions were resolved at a slower rate than *he*-pronoun conditions. Below, we document the findings in detail.

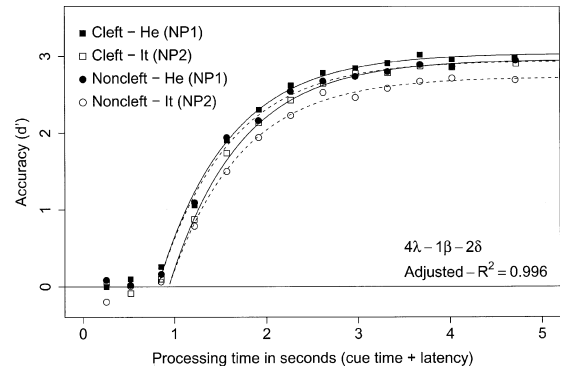


Fig. 1. The accuracy of discriminating acceptable from unacceptable continuations (in d' units) is plotted as a function of processing time (seconds). This model fit uses the average parameters listed in Table 2, from Experiment 1. Clefted antecedent conditions are shown as square symbols, and Nonclefted as circles. Gendered pronoun conditions (i.e., *he* or *she* coreferring with NP1) are shown as filled symbols, and ungendered pronoun conditions (i.e., *it* coreferring with NP2) are shown as open symbols. The smoothed lines show the best-fitting model (see text), with solid lines for clefted conditions and dashed lines for nonclefted conditions.

Differences in availability

The symbols in Fig. 1 show the judgment data averaged over participants' d' values. As an initial means of investigating asymptotic differences, we used the mean of the last 5 d' values for each condition as an empirical estimate of asymptotic accuracy. For the averaged data (Fig. 1), the Clefted-He condition yielded the highest asymptote (2.94 d' units), followed by Clefted-It (2.84) and Nonclefted-He (2.85) conditions, and finally the Nonclefted-It condition (2.62). Across individual participants, clefted conditions had higher values than nonclefted ones, $F(1,11) = 5.83$, $p = .03$, but there was no main effect of pronoun type, $p = .20$, nor an interaction, $F < 1$.

Competitive fits of the SAT equation began with a null $1\lambda-1\beta-1\delta$ model (adjusted- R^2 for the averaged data = .986), assigning one common asymptote, one common rate, and one common intercept to all the conditions. Next, based on differences in the d' estimates of terminal accuracy, we tested two-asymptote models, one assigning the two asymptotes to clefted versus nonclefted asymptotes ($2\lambda-1\beta-1\delta$.989) and one assigning 2 asymptotes to the *he*- versus *it*-pronoun conditions ($2\lambda-1\beta-1\delta$.991), both of which improved the adjusted- R^2 over the null model. A four-asymptote model ($4\lambda-1\beta-1\delta$), assigning one asymptote to each condition, further increased the adjusted- R^2 (.995), and was superior to one or both two-asymptote models for 75% of participants. For the four-asymptote model, asymptotic estimates were higher for clefted conditions than nonclefted ones, $F(1,11) = 7.07$, $p = .02$, and they were higher for *he*-pronoun than *it*-pronoun conditions, $F(1,11) = 5.11$, $p = .04$. The interaction was not significant, $F < 1$.

Differences in time-course

Next, we tested whether systematic differences in the dynamics parameters (rate or intercept) would produce a better fit. Clefting did not affect either the rate or intercept of the SAT functions. Models that allocated separate rate ($4\lambda-2\beta-1\delta$) or intercept ($4\lambda-1\beta-2\delta$) parameters to the clefted and nonclefted conditions (a) did not improve the adjusted- R^2 values for the averaged data, and improved for only 42% of participants in either model, (b) did not yield a consistent ordering of rates or intercepts across participant fits, and (c) statistical tests on the dynamics parameters were not significant (intercepts: $p = .79$, mean difference = 5 ms, CI = $-35-45$; rates: $p = .42$, mean difference = .10, CI = $-.16-.37$).

However, models that allocated separate rate or intercept parameters to *he*- and *it*-pronoun conditions did improve the adjusted- R^2 values in the averaged data ($4\lambda-2\beta-1\delta$.995, $4\lambda-1\beta-2\delta$.997), and for 83% of participants in at least one model. They produced consistently ordered parameters for 75% of participant fits, with dynamics being faster for *he*-pronouns than *it*-pronouns, intercepts: $t(11) = 2.81$, $p = .02$, mean difference = 74 ms, CI = 16–132; rates: $t(11) = 2.49$, $p = .03$, mean difference = .27, CI = .03–.50. Equal numbers of participants preferred to have this difference expressed in rate or intercept, and those adjusted- R^2 values were very similar. In the averaged data the difference was better fit when cast in terms of intercept. Further tests of models with four rates or four intercepts allotted to each condition produced lower adjusted- R^2 values, and the only significant dynamics difference was for *he*-pronoun conditions being faster than *it*-pronoun conditions ($4\lambda-1\beta-4\delta$ Pro-

noun type intercepts: $F(1,11) = 6.75$, $p = .02$; $4\lambda-4\beta-1\delta$ Pronoun type rates: $F(1,11) = 4.75$, $p = .05$).

Best-fit model

The best-fitting $4\lambda-1\beta-2\delta$ model, with one asymptote for each condition and two intercepts based on pronoun type, is shown as the smooth function in Fig. 1. Table 2 presents the parameter values from the fitted function for the averaged data, illustrated in Fig. 1, and for the individual participant fits, which were entered into the analyses below. The asymptotic clefting advantage remained significant in this model, $F(1,11) = 7.90$, $p = .02$, although asymptotes for *he*-pronouns were only marginally higher than for *it*-pronouns, $F(1,11) = 3.46$, $p = .09$. There was no interaction between factors, $F < 1$. The *he*-pronoun advantage in intercept was significant across individual participant fits, $t(11) = 2.81$, $p = .02$, mean difference = 74 ms, CI = 16–132.

Discussion

Syntactic clefting increased the likelihood that an antecedent representation was successfully retrieved and the dependency with its pronoun resolved, as reflected by higher asymptotes for clefted conditions. As well, conditions with a gendered *he* or *she* pronoun were more likely to be successfully resolved than those with an *it*-pronoun. Hence, both clefting and pronoun type increase the availability, or distinctiveness, of an antecedent representation in working memory.

A difference in time-course dynamics was also found, but it was due to pronoun type, not clefting. Resolution occurred earlier for *he/she*-pronouns than for the *it*-pronoun. We found no evidence that syntactic clefting

Table 2

Experiment 1: parameter values from the best fitting model, $4\lambda-1\beta-2\delta$, for the averaged data (over participants) and individual participants

	Adjusted- R^2	Asymptotes (d' units)				Rate	Intercepts (ms)		
		Cl-He	Cl-It	Non-He	Non-It		He	It	He-advantage
Average	0.996	3.04	2.96	2.94	2.73	1.29	814	936	122
P1	0.955	1.87	1.59	1.91	1.27	1.82	781	814	33
P2	0.988	2.79	3.16	2.58	3.15	2.08	933	868	-65
P3	0.974	3.43	2.65	2.97	2.91	0.91	1067	1086	19
P4	0.975	2.51	3.22	2.56	2.21	0.77	799	961	162
P5	0.979	4.00	3.55	4.13	3.47	2.28	776	900	124
P6	0.978	2.72	3.11	3.29	2.74	0.93	1067	1242	175
P7	0.931	2.35	2.15	2.37	2.24	1.97	1045	1180	135
P8	0.973	3.85	3.77	3.87	3.01	1.41	823	1034	211
P9	0.977	3.80	2.54	3.71	3.47	2.50	773	760	-13
P10	0.979	3.35	3.09	2.95	3.39	1.76	1069	1075	6
P11	0.954	3.14	2.43	2.31	2.42	1.06	847	822	-25
P12	0.984	3.00	3.45	3.03	2.54	1.12	761	890	129

Note. Cl, clefted; non, nonclefted.

generated an actively maintained antecedent representation, in focal attention.

Clefting increases distinctiveness

The higher observed *d'* values and fitted asymptotes for pronouns referring to a clefted antecedent suggest that clefting made an antecedent representation more available in memory. Because clefting does not alter the retrieval cues provided by the pronoun, it likely affects how the antecedent was first encoded in memory. One of the linguistic functions of clefting is contrastive, either between explicit entities, or in the present experiments, between an explicitly introduced entity and an implicit contrast set. Our asymptotic findings are consistent with syntactic clefting making the memory representation more distinctive at encoding, or even with clefting tagging a mental entity as more important or central to the discourse.

Gendered pronouns increase availability

We found that conditions with the pronouns *he* or *she* yielded higher asymptotes than conditions with the pronoun *it*. Based on a match of the grammatical and semantic features of an anaphor to each entity in the current discourse model, Greene et al. (1992) proposed a pronoun-as-cue framework, where the pronoun acts as a cue to the most likely antecedent in the discourse model. This approach provides a principled explanation of the asymptotic differences between these conditions. Both *s/he* and *it* provided the minimally necessary information, such as gender, number, and animacy, for a correct bond with the antecedent presented in the first sentence. However, our data suggest that the gendered pronouns were better able to pick out or resonate with their antecedent compared to *it*. There are two explanations for this finding. First, *it* is more ambiguous, having more potential antecedent classes and linguistic functions than *he* or *she*. Although the verb following *it* forced the correct antecedent to be the inanimate noun, on the pronoun itself, the referential and functional ambiguity of *it* would have made recovering the correct antecedent less likely compared to *he*. Second, higher accuracy in the *he*-pronoun conditions is probably also due to several correlated properties of its antecedent that increase prominence. The antecedent of *s/he* was also the subject, the first-mentioned noun, and occurred in a parallel role with the pronoun. In contrast, the antecedent of *it* was the object, second-mentioned, and in a nonparallel position.

Why might gendered pronouns increase accessibility?

Resolving *he*-pronouns was not only more likely, but also showed a faster time-course, than resolving *it*-pronouns. The intercept difference we found might indicate an architectural distinction between active and passive antecedent representations. That is, the NP1 subject

representation (to which *he* referred) might have been actively maintained in focal attention, while the NP2 object representation (to which *it* referred) required time-consuming reactivation from memory to current awareness. However, such an interpretation is problematic because the later intercept for *it*-pronoun conditions is fully consistent with ambiguity or reanalysis. The set of possible referents and interpretations for *it* is greater than for *he*, so greater competition between referents could have caused a slower resolution time-course. The later intercept for *it*-pronoun conditions could also reflect reanalysis of an incorrect bond between pronoun and antecedent on a proportion of trials. We suggest that gendered pronouns led to a faster time-course of resolution because they were less ambiguous and did not lead to as much competition between antecedents or require reanalysis as often as *it*-pronouns did. The properties of NP1 may have contributed to greater accessibility of the antecedent representation, but we do not think the faster time-course was solely because of them. Experiments 2 and 3 examine the extent to which the time-course difference we found is due to ambiguity and reanalysis in the *it*-pronoun conditions.

Experiment 2: Eye-tracking

In Experiment 2, we examined eye-tracking measures as a means of exploring how these speed-accuracy trade-off differences are expressed in more natural reading situations. First, we attempted to confirm with eye-tracking methodology the clefting advantage found in Experiment 1. Second, we examined how the time-course findings align with more conventional eye-tracking markers of difficulty due to ambiguity and reanalysis.

Different properties of the eye movements that occur during reading have been argued to reflect when information or particular constraints are operative during comprehension (see Liversedge, Paterson, & Pickering, 1998; Pickering, Frisson, McElree, & Traxler, 2004; Rayner, 1998 for reviews). Particularly, researchers have suggested that effects in early measures, such as how long a region is first fixated, might reflect different aspects of processing than later measures, such as how many times or for how long a region is refixated. In pronoun and other types of anaphoric resolution, Garrod and colleagues have used such a difference as a basis for proposing a two-stage model (Garrod & Sanford, 1990; Garrod & Terras, 2000; Sanford, Garrod, Lucas, & Henderson, 1983). *Bonding* is argued to be an initial, automatic and low-level process, driven by lexical information, which establishes a superficial, tentative link between a pronoun and its referent. Difficulty in bonding should be reflected in early eye-tracking measures. *Resolution*, which we refer to as *integration*, is a more

strategic operation during which the bond is evaluated in light of the discourse context and overall discourse model. If necessary, information is reanalyzed to form a new bond and the discourse context is updated, or may be reconstrued. Difficulty in integration is argued to appear in later eye movement measures.

To our knowledge, no one has investigated the effect of clefting on pronoun resolution using eye-tracking. However, Birch and Rayner (1997) found longer reprocessing times for information that was clefted compared to not clefted, and Morris and Folk (1998) found that clefting led to easier integration of related information encountered later on. Based on these differences in later measures, and our finding in Experiment 1 that clefting did not increase the speed of resolution, we expected that the advantage should be apparent in later measures such as second-pass (re-reading time for a region, after it has been fixated once prior) and total time (the sum of all fixations within a region), rather than early measures, such as first-pass (the first fixations in a region before exiting).

Pursuing an explanation for the time-course difference between *it* and *he* pronouns in Experiment 1, we will concentrate on whether a similar effect emerges in early eye movement measures. The time-course difference we found could be due to the ambiguity of the *it* pronouns or to *he*-antecedents being actively maintained. If it was due to the ambiguity of an *it* pronoun, we expected that the differences between pronoun conditions should emerge in later measures. If comprehenders experience difficulty in interpreting *it* pronouns, there should be increased immediate regressions from the coreference region, longer second-pass times and regression-path durations (all of the time from when a reader first fixates within a target region until the reader fixates anything to the right of the region; Brysbaert & Mitchell, 1996; Traxler, Bybee, & Pickering, 1997). These measures are often taken to signal reanalysis. However, if the time-course difference was due to *he*-antecedents being actively maintained, we might expect to see an advantage in early measures in the coreference region for *he*-pronoun conditions, such as shorter first-pass times.

Method

Participants

Forty people from the same population as Experiment 1 participated in Experiment 2. All had normal or corrected vision (contacts).

Materials

We used the same 2 (clefting) × 2 (pronoun type) design as Experiment 1. We included a spillover region following the verb in the pronoun sentence, as shown in Table 3. To increase the likelihood of the pronoun being fixated and to be sure that effects observed on or directly after the pronoun were not contaminated by the beginning of a new line, an adverb such as *reassuringly* was inserted at the beginning of the second sentence before the pronoun.

A subset of 48 items from Experiment 1 were selected and revised as necessary, so that the verbs following the pronoun were equated for frequency and length. Following *he/she*, verbs had a mean root frequency of 26.9 observations per million, a mean past tense form frequency of 11.4 per million, and mean length of 7.3 letters. Those following *it* had a mean root frequency of 27.2, a mean past tense form frequency of 8.0, and mean length of 7.1. There were no statistical differences between the verbs for these three measures, $ps > .26$. The full set of materials appears in Appendix A.

The 48 scenarios × 4 conditions were distributed among 8 counterbalanced lists (there were four other conditions which are not presented here). Participants were assigned randomly to a list. Each one saw 24 trials, with 6 trials per condition, constituting 22% of the total trials. These appeared pseudo-randomly among two-sentence filler trials. The fillers included an equal number of cleft structures with a pronoun in a different position (e.g., *It was the tired mother that cooked/What the tired mother cooked was the daily oatmeal. The preschooler clutched her/it tightly in the morning before kindergarten.*), and the remaining 66% of discourses did not manipulate pronoun use (e.g., *My nephew's bride was cold at the wedding reception. I heard that she snubbed all the guests and upset her family/had to put on a sweater*

Table 3
Example materials from Experiment 2

Condition	Antecedent and pronoun sentences
Clefted-He	It was the cheerful waitress who made the decaffeinated coffee. Reassuringly, she gossiped behind the counter of the diner.
Clefted-It	What the cheerful waitress made was the decaffeinated coffee. Reassuringly, it brewed behind the counter of the diner.
Nonclefted-He	What the cheerful waitress made was the decaffeinated coffee. Reassuringly, she gossiped behind the counter of the diner.
Nonclefted-It	It was the cheerful waitress who made the decaffeinated coffee. Reassuringly, it brewed behind the counter of the diner.

and her lace gloves.; Until the bell rang for a break, the man stamped/recited the material in his office. He was always in a hurry to finish the job.) A yes–no comprehension question followed half of the trials, which enquired about all parts of the two sentences equally often, with one third of the questions for the experimental items (4/12) querying the coreference dependency in some manner. Half the answers were “yes” and half were “no,” balanced across conditions, items, and lists.

Procedure

Participants were run individually using a Sensori-Motor Instruments EyeLink II head-mounted eye-tracker apparatus and presentation software. The eye cameras recorded eye movements and fixations binocularly every 4 ms. Participants had their head stabilized on a chin rest, and 1° of visual angle subtended 2.7 characters, approximately the size of a pronoun. Calibration and validation of the recording apparatus was performed before beginning the experiment.

At the beginning of a trial, a fixation dot appeared on the left of the screen at the position where the first sentence would begin. Once fixation was stable and any correction made, the two sentences were presented all at once on two double-spaced lines. They remained on the screen until the participant was finished reading and had pressed a button to proceed to the next trial. Participants were instructed to read for comprehension in a natural way. When a yes–no comprehension question appeared, they answered by pressing an appropriate button. No feedback was given. Participants began with five practice trials to ensure familiarity with the procedure. The reading portion of the experiment took approximately 30 min. Participants could ask for a break at any time; when necessary, a full calibration was performed before continuing the experiment.

Data analysis

For each participant, data from the eye with the least error at calibration was analyzed (27 right eye and 13 left eye). An automatic routine combined fixations that were less than 80 ms with a previous or subsequent fixation that was within one character. Following that, fixations less than 80 ms or greater than 1300 ms (as the regions reported here are one longer word or two word regions) were excluded.

Two regions in the pronoun sentence were chosen for analysis, denoted by square bracketing: *Reassuringly*, [*she gossiped*] [*behind*] *the counter of the diner*. Due to the very low probability of a first pass fixation on the pronoun alone (21–26%) or on the pronoun plus four characters to the left (42–49%, Ehrlich & Rayner, 1983; Garrod et al., 1994), the pronoun and following verb were defined as the critical region and the following preposition or preposition plus determiner was the spillover region (e.g., *during*, or *in the*). Across conditions,

the probability of a first-pass fixation in the pronoun + verb region was .96–.98, and in the spillover region .74–.79. Within each region, these probabilities did not differ between conditions (all $F_s < 1$).

Six eye-tracking measures are reported. *First-pass time* is the sum of all fixations inside a region beginning with the first fixation inside until the gaze travels outside the region (either to the left or right), given that the reader has not previously fixated subsequent text. This is the primary early measure. *First-pass regressions out* is the probability of leaving a region on the saccade following a first-pass fixation to regress to earlier parts of the sentence(s). This measure is sometimes referred to as an early measure, but it is also an indication of difficulty early on in processing, perhaps signaling reanalysis. The next four measures we refer to as later measures. *Regression path duration* is all of the time from when a reader first fixates within a target region until the reader fixates anything to the right of the region, which includes first-pass time and reinspection of prior regions (Brysbaert & Mitchell, 1996; Traxler et al., 1997). *Second-pass time* is the sum of all re-reading in a region after having previously exited the region to the right. *Total time* is the sum of all fixations in a region. *Regressions back* into a region is the probability of coming back to a region from later areas after having passed through it.

Participants' overall accuracy on the comprehension questions was 91%, which was very similar for the four conditions (89–93%). No trials were subsequently excluded based on accuracy. Trials with major tracker loss or excessive blinking were excluded (1.3% of trials). On the basis of the first-pass measure, trials with skips in (a) consecutive regions, (b) the beginning of the pronoun sentence, (c) the critical region, or (d) either one of the two antecedents were excluded for all measures (3.3% of the data), as we assumed that the information necessary for pronoun resolution would not be available. In total, 4.6% of the data were excluded. Any remaining skips of a region in a measure were treated as missing data points, except for regression path times and second-pass, which included zero times (Birch & Rayner, 1997; Pickering et al., 2004).

Repeated-measures 2 (clefting) \times 2 (pronoun type) ANOVAs were calculated with participants as a random factor (F_1) and with items as a random factor (F_2) for each of the measures. *MinF'* values are reported for analyses that are significant or marginal by participants or items (Clark, 1973). Participant ANOVAs include the between-participants term for the counterbalancing list.

Results

Table 4 shows the six eye-tracking measures (participant means) for the critical region and spillover region.

Table 4
Eye-tracking results for Experiment 2

	Pronoun + verb <i>she/it + gossiped/ brewed</i>	Spillover <i>behind</i>
<i>First-pass (ms)</i>		
Clefted-He	327 (14)	292 (13)
Clefted-It	329 (14)	278 (16)
Nonclefted-He	326 (16)	285 (14)
Nonclefted-It	340 (15)	285 (15)
<i>First-pass regressions out (%)</i>		
Clefted-He	16 (2)	7 (2)
Clefted-It	20 (3)	8 (2)
Nonclefted-He	15 (2)	5 (1)
Nonclefted-It	17 (3)	12 (2)
<i>Regression path duration (ms)</i>		
Clefted-He	458 (25)	338 (20)
Clefted-It	491 (32)	329 (22)
Nonclefted-He	443 (27)	339 (33)
Nonclefted-It	445 (26)	355 (24)
<i>Second-pass (ms)</i>		
Clefted-He	162 (31)	144 (24)
Clefted-It	164 (27)	154 (20)
Nonclefted-He	203 (24)	155 (18)
Nonclefted-It	199 (32)	159 (22)
<i>Total time (ms)</i>		
Clefted-He	538 (36)	421 (27)
Clefted-It	555 (33)	418 (24)
Nonclefted-He	583 (32)	415 (23)
Nonclefted-It	589 (38)	444 (25)
<i>Regressions back in (%)</i>		
Clefted-He	29 (4)	33 (4)
Clefted-It	23 (3)	29 (4)
Nonclefted-He	31 (4)	33 (3)
Nonclefted-It	34 (3)	30 (4)

Note. Standard error of the mean appears in parentheses. See text for an explanation of the measures.

Table 5 summarizes the main inferential tests, for main effect analyses over participant means, item means, and for *MinF'*. No interactions between clefting and pronoun type were found for any region or measure.

Early measures

First-pass reading times showed no differences due to clefting or pronoun type, for the critical pronoun + verb or spillover regions. The only evidence for any early effect was that *it*-pronoun conditions produced a marginally greater proportion of first-pass regressions out to earlier regions of the sentences. As shown in Table 5 (Pronoun Type), we found that the pronoun + verb region showed a weak trend for *it*-pronouns to produce more regressions out than *he*-pronouns, and in the following spillover region, this trend became significant. First-pass regressions were not affected by clefting.

Table 5
Inferential statistics for Experiment 2

Source of effect	F_1	F_2	$MinF'$	df for $MinF'$
Clefting				
<i>First-pass regressions out (%)</i>				
Pronoun + verb	<1	<1		
Spillover	<1	<1		
<i>Regression path duration</i>				
Pronoun + verb	2.78 [†]	<1		
Spillover	<1	<1		
<i>Second-pass</i>				
Pronoun + verb	5.01*	3.22 [†]	1.96	1, 79
Spillover	<1	<1		
<i>Total time</i>				
Pronoun + verb	4.28*	2.64	1.63	1, 79
Spillover	<1	<1		
<i>Regressions back in (%)</i>				
Pronoun + verb	4.52*	5.71*	2.52	1, 72
Spillover	<1	<1		
Pronoun type				
<i>First-pass regressions out (%)</i>				
Pronoun + verb	1.41	2.80 [†]	<1	1, 62
Spillover	3.18+	5.01*	1.94	1, 67
<i>Regression path duration</i>				
Pronoun + verb	<1	1.09		
Spillover	<1	<1		
<i>Second-pass</i>				
Pronoun + verb	<1	<1		
Spillover	<1	<1		
<i>Total time</i>				
Pronoun + verb	<1	<1		
Spillover	<1	<1		
<i>Regressions back in (%)</i>				
Pronoun + verb	<1	<1		
Spillover	3.65 [†]	1.20	<1	1, 72

[†] $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

For participant analyses, $df = 1, 32$, for Item analyses, $df = 1, 47$. No interactions between Clefting and Pronoun type were significant.

Later measures

In the spillover region, there were no reliable effects for any of the four later measures. In the pronoun + verb region, nonclefted conditions showed more re-reading than clefted ones. Regression path times, second-pass times, and total reading times were all longer for nonclefted conditions (significant by participants). As well, the probability of regressing back into the critical region was significantly greater for nonclefted conditions, by both analyses.

There was no main effect of pronoun type nor an interaction with clefting for regression path time, second-pass, total time, or regressions back in.

Discussion

Clefting did not produce any effects in early eye-tracking measures. Rather, the advantage for clefted antecedents occurred in several later measures. These findings suggest that clefting aided integration into the discourse model, a late phase in coreference resolution (e.g., Garrod & Sanford, 1990; Garrod & Terras, 2000; Sanford et al., 1983). That clefting might aid integration is generally consistent with previous eye-tracking studies of cleft constructions (Birch & Rayner, 1997; Morris & Folk, 1998). Overall, these findings accord well with the speed-accuracy tradeoff results of Experiment 1. There, we found that clefting produced no differences in accessibility (intercept or rate of the time-course function), but rather an advantage in availability, reflected in the asymptotes of the function. This pattern suggests that clefting an antecedent produced a stronger or more distinctive memory representation. Together, the experiments suggest that when an antecedent representation is less distinctive, it is more difficult to integrate.

In Experiment 1, we found that *it*-pronouns were associated with a slower resolution time-course than *he*-pronoun conditions, which we argued might reflect a higher proportion of incorrect bonding and reanalysis operations. In Experiment 2 we found that conditions with the pronoun *it* caused more first-pass regressions from the spillover region than those with *he/she*. However, this effect was marginal and short-lived, as there were no differences for any of the later measures, including regression-path duration. We interpret this as evidence that the *it*-pronoun conditions were in fact functionally ambiguous, although the ambiguity was quite easily overcome. When an incorrect bond was formed, the verb and prepositional phrase following the pronoun provided diagnostic information immediately that helped to identify and repair the coreferent bond (e.g., Fodor & Inoue, 2000). This interpretation is in the spirit of bonding being driven largely by lexical properties of the anaphor and antecedent (Garrod & Terras, 2000). For example, following the ambiguous pronoun *it*, at the verb *brewed*, nonreferential interpretations of *it* would

be ruled out, and reference is narrowed to inanimate things that brew (intransitive) or can brew something else (transitive). At the spillover region, the preposition, *behind*, further constrains interpretation to an intransitive structure, aiding convergence on *coffee* as the antecedent of *it*.

An alternative interpretation of the time-course difference in Experiment 1 was that an antecedent representation in the *he*-pronoun conditions was actively maintained, causing a faster speed of resolution. However, we did not find any advantage in first-pass times (or other early measures not reported here), which would constitute the most convincing evidence, to our minds, of this account. Therefore, given the ambiguous nature of *it*, Experiments 1 and 2 provide converging evidence that *it* was the locus of the time-course difference, implicating slower resolution.

The correspondence between eye-tracking and speed-accuracy tradeoff measures is not direct, and we do not suppose there is a one-to-one mapping between these measures. For example, we do not think that later eye-tracking measures uniquely provide an index of the referent's availability in memory, or that an advantage in first-pass reading time would necessarily correspond to a faster time-course function. However, we do think that the observed effects in eye-tracking provide an additional means of validating our interpretation of the parameters in the time-course functions, provided one accepts the standard interpretation of differences in early and late measures in eye-tracking. Minimally, the strong correspondence between measures demonstrates that observed speed-accuracy tradeoff differences have consequences in natural reading situations.

Experiment 3: Antecedent position switched

In Experiment 3, we again used speed-accuracy tradeoff modeling to further investigate the clefting advantage. We examined whether the effects observed in Experiment 1 generalized to other cleft structures. As shown in Table 6, we investigated a different pseudo-cleft in Experiment 3:

Table 6
Example materials from Experiment 3

Condition	Antecedent and pronoun sentences
Clefted-He (NP2)	The one whom/the lead paint/annoyed was/the safety inspector. He grimaced. # He flaked.
Clefted-It (NP1)	It was/the lead paint/that annoyed/the safety inspector. It flaked. # It grimaced.
Nonclefted-He (NP2)	It was/the lead paint/that annoyed/the safety inspector. He grimaced. # He flaked.
Nonclefted-It (NP1)	The one whom/the lead paint/annoyed was/the safety inspector. It flaked. # It grimaced.

Note. Slashes indicate presentation regions for the first sentence. # Indicates the unacceptable continuation paired with the first sentence.

The one whom the lead paint annoyed was the safety inspector. A *Whom*-cleft highlights a direct object referent and requires an animate antecedent for which we used the gendered coreferring pronouns *he* or *she*. To retain a design similar to Experiments 1 and 2, we used an *It*-cleft construction to highlight the inanimate antecedent, which appeared as the subject and was referred to with the ungendered pronoun *it*. Based on Experiment 1, we expected that coreference with clefted antecedents would increase the likelihood of antecedent retrieval and subsequent resolution, due to the greater distinctiveness that clefting affords a representation.

We also investigated whether increasing the prominence of an *it*-pronoun's antecedent representation would alter its availability or accessibility, eliminating the disadvantages for this type of pronoun observed in Experiment 1. The eye-tracking results of Experiment 2 suggested that the ambiguity of *it* was short-lived and easily overcome. Therefore, we reasoned that reference may not be as ambiguous for the reader if discourse constraints made the antecedent more prominent. In Experiment 3, the position of the antecedents was switched, such that structural and discourse factors increased the prominence of the inanimate antecedent. For example, *paint* was now the subject and the first-mentioned noun, and the pronoun *it* was now in parallel (subject) position with that antecedent. In Experiment 1, we attributed the lower availability of *it*-pronoun antecedents to both a less distinctive antecedent at encoding (object, second-mentioned noun, nonparallel) and ambiguity over the identity of the antecedent at retrieval. Similarly, we speculated that the slower speed of resolution for the ambiguous *it*-pronoun conditions in Experiment 1 was due to some combination of competition among candidate antecedents and reanalysis of the coreferent bond. Increasing the prominence of an inanimate antecedent might counteract the ambiguity of an *it*-pronoun by increasing the prominence of the correct representation (and/or decreasing the prominence of competitors) to a point that is effectively less ambiguous.

Lastly, we pursued a more stringent test of what caused the difference in the speed of pronoun resolution in Experiment 1. The alternative interpretation of the time-course difference was that the *he*-antecedent was actively maintained in focal attention, increasing the speed of resolution, rather than *it*-pronouns being ambiguous and decreasing the speed of resolution. We reasoned that if the time-course advantage was due to properties of the NP1 antecedent (subject, first-mentioned noun, parallel with pronoun), then we would expect to see such an advantage in this experiment for the *it*-antecedent as NP1, despite the greater ambiguity of *it*-pronouns. If an NP1 antecedent representation is actively maintained, then the *it*-pronoun should bond immediately with that foregrounded, actively maintained representation with little chance of error.

Method

Participants

Twenty-two people from the same population as Experiments 1 and 2 participated in the experiment.

Materials

We used the same 2 (clefting) \times 2 (pronoun) design, with acceptable and unacceptable pronoun continuations. As shown in Table 6, the position of the antecedents was switched compared to Experiments 1 and 2. The inanimate noun appeared as NP1 (subject, first-mentioned position, parallel with pronoun) and the animate noun appeared as NP2 (direct object, second-mentioned position, not parallel with pronoun).

In the interest of performing analyses with items as a random factor, we limited the number of items to 32, which necessitated repeated exposure. In each of 4 sessions, a participant read 64 experimental scenarios, two conditions per item, counterbalanced within and across sessions. The critical trials constituted 18% of the experiment, and were presented randomly among the remaining 82% of trials, none of which contained cleft structures. The fillers consisted of multi-clause sentences, with equal numbers of acceptable and unacceptable (underlined) versions: *The architect admired the blueprint, but the funding did not materialize/the entrepreneur did not melt.*; *The janitor believed in good manners, but the principal at the school was upset to discover that the librarian/the meeting did not.*; *The reporter who had hated the commercial knew that the candidate on the radio lied/fused.*; *The scientist that the climate fascinated responded/derailed.* The full set of experimental materials appears in Appendix A.

Procedure

The same procedure as Experiment 1 was employed with one minor change. Instead of starting a trial with an experimenter-generated random key and the participant changing keys if necessary, participants were instructed to begin a trial by registering an "undecided" response by pressing both keys simultaneously. Once they had enough information to make an acceptable-unacceptable decision, they pressed only the appropriate key. They were again trained to modulate their responses if their judgment changed during the trial, as in Experiment 1.

Data analysis

We used the same method as Experiment 1 to compute d' and test for the best-fitting model of the speed-accuracy tradeoff equation. We report the findings with participants (F_1, t_1) and items (F_2, t_2) as a random factor, as well as $MinF'$ when F_1 or F_2 is significant. ANOVAs over participant variance include the between-list counterbalancing term. For a priori paired comparisons,

we report the 95% confidence interval (CI) around the mean difference between participant means. When the t_1 is a post hoc, follow-up comparison, we report the 95% confidence interval surrounding the mean difference using the mean square error (MSE) term from the interaction in the overall ANOVA (i.e., based on all four cells of the design), and the associated critical t -value at $\alpha = .05$ (two-tailed) for the denominator degrees of freedom. This provides a common source for evaluating follow-up differences (Masson & Loftus, 2003).

Results

The best-fitting $4\lambda-1\beta-1\delta$ model is shown in the smooth function in Fig. 2. First, we found that clefted antecedent conditions produced higher asymptotic accuracy than nonclefted antecedents, as predicted. This clefting advantage was qualified by an interaction, such that *it*-pronouns (NP1) showed a clefting advantage while *he*-pronouns (NP2) did not. Second, we found that *he*-pronouns produced higher asymptotic accuracy than *it*-pronouns, despite the fact that *he*-antecedents were in a less prominent discourse position in this experiment. Third, we did not find any reliable difference in time-course, with only a weak indication that *it*-pronoun (NP1) conditions were processed more slowly than *he*-pronoun (NP2) conditions. Overall, clefting had no effect on the speed of processing.

Table 7 presents the parameter estimates from the best-fitting $4\lambda-1\beta-1\delta$ model for the averaged (over participants) data, also illustrated in Fig. 2, and for the

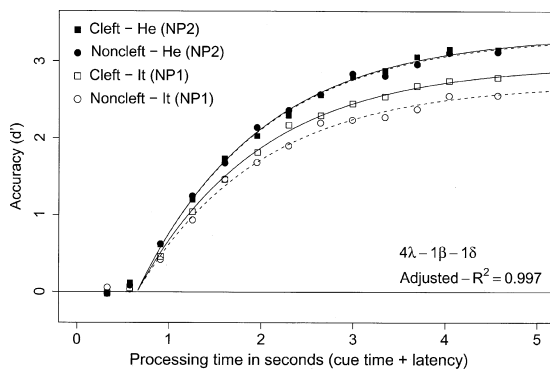


Fig. 2. The accuracy of discriminating acceptable from unacceptable continuations (in d' units) is plotted as a function of processing time (seconds). This model fit uses the average parameters listed in Table 7, from Experiment 3. Clefted antecedent conditions are shown as square symbols, and Nonclefted as circles. Gendered pronoun conditions (i.e., *he* or *she* coreferring with NP2) are shown as filled symbols, and ungendered pronoun conditions (i.e., *it* coreferring with NP1) are shown as open symbols. The smoothed lines show the best-fitting model (see text), with solid lines for clefted conditions and dashed lines for nonclefted conditions.

individual participant fits. Inferential statistics, computed over individual participant fits and item fits for each of the candidate models detailed below, appear in Table 8.

Differences in Availability

The symbols in Fig. 2 show the observed d' data averaged over participants. We computed the mean of the last 4 d' values as an empirical measure of asymptotic accuracy. For the averaged data, the Clefted-He condition was highest (by participants 3.06; by items 2.93 d' units), followed by Nonclefted-He (3.00; 2.82), Clefted-It (2.69; 2.57), and finally, Nonclefted-It was lowest (2.44; 2.24). This pattern prevailed across participants and items, albeit more consistently for pronoun type than for clefting. As indicated in Table 8 (Terminal d' Accuracy), clefted conditions had significantly higher d' accuracy levels than nonclefted ones, and *he*-pronoun conditions were higher than *it*-pronouns. There was a marginal interaction. Follow-up comparisons showed that clefting did not affect the accuracy of *he*-pronoun conditions, mean difference by participants = .06 d' units, CI = $-.08-.21$, while clefting increased the accuracy of retrieval and interpretation for *it*-pronouns, mean difference = .25, CI = $.10-.39$.

Consistent with differences between the d' values, a fitted model with four asymptotes ($4\lambda-1\beta-1\delta$), assigning one asymptote to each condition, increased the adjusted- R^2 (by participants .997; by items .994) over null and two-asymptote models in fits of the averaged data. As well, the adjusted- R^2 values for 73% of participant and 87% of item fits also increased for the four-asymptote model. In this model, clefted conditions had significantly higher asymptotes than nonclefted ones, and *he*-pronouns had higher estimates than *it*-pronouns (Table 8). There was a significant interaction, which, like for the empirical d' values, was driven by *he*-pronoun conditions showing no effect of clefting, mean difference = .03, CI = $-.15-.20$, but for the *it*-pronoun pair, clefting produced higher asymptotes, mean difference = .31, CI = $.13-.48$.

Differences in time-course

We report the fits separately for participants and items analyses, as the item fits were less stable, requiring a different set of fits. The bottom half of Table 8 provides the inferential statistics for each model discussed below.

Fits by participants. Clefting did not affect time-course. When separate rate or intercept parameters were assigned to the clefted and nonclefted conditions, the adjusted- R^2 values for the averaged data did not improve over the $4\lambda-1\beta-1\delta$ model, and increased for only 36% of participant fits, in either the 2-rate or 2-intercept model. Furthermore, the two rates or intercepts were not ordered consistently across fits, and, thus, did not differ statistically.

Table 7

Experiment 3: parameter values from the best fitting model, $4\lambda-1\beta-1\delta$, for the averaged data (over participants) and individual participants

	Adjusted- R^2	Asymptotes (d' units)				Rate	Intercept (ms)
		Cl-He	Non-He	Cl-It	Non-It		
Average	.997	3.35	3.34	2.96	2.72	0.75	655
P1	.944	2.91	2.92	2.42	2.31	0.93	819
P2	.953	2.03	2.41	2.06	2.01	1.23	735
P3	.961	3.41	4.01	2.80	3.18	2.09	780
P4	.924	2.38	2.43	3.00	2.58	1.07	1094
P5	.979	4.98	4.70	4.00	3.41	0.40	793
P6	.904	4.70	4.35	4.15	2.86	0.33	498
P7	.971	3.90	4.28	3.90	3.69	1.05	576
P8	.966	3.59	3.93	3.73	3.88	1.22	646
P9	.963	3.76	4.03	4.16	3.74	1.11	984
P10	.957	3.87	4.42	3.59	3.51	0.44	1011
P11	.970	3.66	3.96	3.05	2.76	0.38	1570
P12	.961	4.58	4.42	2.95	3.10	0.64	848
P13	.945	4.74	4.29	3.78	3.57	0.38	731
P14	.970	3.88	3.50	4.05	3.72	1.61	423
P15	.908	1.47	1.10	0.68	0.50	2.75	577
P16	.844	4.70	4.52	3.40	3.27	0.34	620
P17	.942	2.36	2.80	1.79	0.50	0.53	1134
P18	.981	3.77	3.51	3.22	3.20	2.04	779
P19	.957	4.16	3.88	3.24	3.47	3.04	487
P20	.971	3.07	2.66	3.00	1.84	0.71	1348
P21	.979	3.48	3.24	2.30	2.82	0.87	873
P22	.966	3.22	3.06	3.33	2.09	1.04	1130

Note. Cl, clefted; non, nonclefted.

It-pronoun conditions showed marginally slower time-course dynamics than *he*-pronoun conditions. The adjusted- R^2 value in the averaged data improved slightly when two intercepts were allotted ($4\lambda-1\beta-2\delta = .998$), but did not improve for 2 rates ($4\lambda-2\beta-1\delta = .997$). *It*-pronoun conditions had a slower rate for 73% of fits, a marginally significant difference, mean difference = .21, $CI = -.04-.47$. *It*-pronouns similarly had a later intercept for 68% of participant fits, but the difference was not significant, mean difference = 48 ms, $CI = -14-110$.

Tests of models with four rates or four intercepts (one per condition) produced similar findings. In the 4-rate fit, adjusted- R^2 increased for the averaged data (compared to a 411 model), and clefted rates did not differ from nonclefted ones. *It*-pronoun conditions did have significantly slower rates than *he*-pronouns, but only 14% of participants showed an increase in adjusted- R^2 and such an ordering of rates, and the rate difference occurred at the expense of a speed-accuracy trade off with asymptote parameters, making the model not viable. The 4-intercept fits produced no significant differences.

Fits by items. For the item fits, we tested models with fixed asymptotic values. When asymptote parameters were allowed to vary freely, they were not appropriately constrained by terminal accuracy of the d' data for some items. Hence, we fixed the asymptotic parameter values

(λ) to an empirical estimate of observed d' performance in each condition (see above, Differences in availability, terminal d' accuracy). Rates or intercepts were allowed to vary freely, and are notated with those parameters only, $x\beta - x\delta$. Fits with freely varying asymptotes showed a similar pattern of results concerning dynamics as the fixed ones reported here.

Clefting did not affect time-course. For the averaged data, fits that allocated separate rate or intercept parameters based on clefting did not show an improvement in adjusted- R^2 ($2\beta - 1\delta = .987$, $1\beta - 2\delta = .987$) over a similarly fixed $1\beta - 1\delta$ model (.987). Less than half of the individual item fits showed improvement in adjusted- R^2 values for their 2-rate or 2-intercept model. For the averaged data, clefted conditions showed numerically slower dynamics, but differences across individual item fits were nonsignificant: rates mean difference = .06, $CI = -.03-.14$, intercepts mean difference = 30 ms, $CI = -20-80$.

Pronoun type did not affect time-course, either. The adjusted- R^2 value in the averaged data did not improve when 2 intercepts ($1\beta - 2\delta = .987$) or 2 rates ($2\beta - 1\delta = .987$) were allotted, and just over half of the items led to an increase in adjusted- R^2 . For the averaged data, the *it*-pronoun conditions had a numerically faster rate, but across item fits the difference was not significant, mean difference = .04, $CI = -.03-.11$. There

Table 8

Experiment 3: inferential statistics assessing differences in availability (asymptotes) and time-course dynamics (rate or intercept)

Source and direction of effect	F_1 or t_1	F_2 or t_2	$MinF'$	df for $MinF'$
Differences in availability				
<i>Terminal d' accuracy</i>				
Clefted > Nonclefted	5.69*	17.21***	4.28*	1, 34
<i>He</i> -pronouns > <i>It</i> -pronouns	30.43***	28.51***	14.72*	1, 51
Interaction	2.76	4.48†	1.71	1, 43
<i>He</i> -pronouns: Clefted = Nonclefted	.97	1.46		
<i>It</i> -pronouns: Clefted > Nonclefted	2.39*	4.47***		
<i>4λ-1β-1δ model: Asymptotes</i>				
Clefted > Nonclefted	4.66*	9.71**	3.15	1, 40
<i>He</i> -pronouns > <i>It</i> -pronouns	30.55***	24.00***	13.44*	1, 52
Interaction	5.25*	4.23*	2.34	1, 52
<i>He</i> -pronouns: Clefted = Nonclefted	.33	1.03		
<i>It</i> -pronouns: Clefted > Nonclefted	2.65**	3.70***		
Differences in time-course				
<i>Clefted faster (slower) than Nonclefted</i>				
<i>4λ-2β-1δ model: rates</i>	1.19	1.29		
<i>4λ-1β-2δ model: intercepts</i>	0.98	1.21		
<i>4λ-4β-1δ model: rates</i>	<1	(2.34)		
<i>4λ-1β-4δ model: intercepts</i>	<1	(1.41)		
<i>It-pronouns slower (faster) than He-pronouns</i>				
<i>4λ-2β-1δ model: rates</i>	1.74†	(1.24)		
<i>4λ-1β-2δ model: intercepts</i>	1.62	(0.47)		
<i>4λ-4β-1δ model: rates</i>	4.75*	(1.73)	1.27	1, 48
<i>4λ-1β-4δ model: intercepts</i>	1.44	<1		

Degrees of freedom (df) for F_1 are 1, 21, for t_1 are 21, for F_2 are 1, 31, and for t_2 are 31.† $p \leq .10$, * $p \leq .05$, ** $p \leq .01$, *** $p \leq .001$.

was no difference between intercepts, mean difference = 12 ms, CI = -42–67.

Tests of models with 4 rates or 4 intercepts did not show significant effects of clefting or pronoun type, nor an interaction.

Discussion

Experiment 3 generalized our previous findings to another pseudo-cleft structure (*The one whom...*). Again, we found that clefted antecedent representations were successfully resolved more often than nonclefted ones, supporting the conclusion that clefted representations are more available in memory. As before, clefting did not increase the speed of resolution, providing no evidence that prominent antecedents are actively maintained in focal attention. Clefting only increased asymptotic accuracy, suggesting that it affects the strength or distinctiveness of an antecedent representation in memory.

Once again, we found that the ambiguous pronoun *it* was less likely to be correctly resolved than the gendered pronouns *he* and *she*, despite the more favorable grammatical and structural properties of *its* antecedent (subject, first-mention, pronoun in parallel role). This finding accords well with our stance that greater availability

reflects the greater likelihood of a *he*-pronoun “cueing” recovery of an antecedent representation compared to the more ambiguous *it*-pronoun. In the General Discussion, we consider the role that animacy of the antecedent might play, in light of the finding that *he*-pronoun conditions showed greater availability in both speed-accuracy tradeoff experiments.

In this experiment, we also found a significant interaction in asymptotic performance between clefting and pronoun type, whereas in the previous two experiments there was no interaction of the factors. When the pronoun was more ambiguous, clefting increased the distinctiveness of the antecedent, significantly increasing the likelihood of retrieving the antecedent. But clefting had no facilitating effect when there was a more specific, gendered pronoun. We suggest that, in this experiment, the *he*-pronoun conditions showed no enhancement from the *Whom*-cleft construction because of a strong recency effect: the NP2 antecedent occurred at the end of the first sentence, just-recent to the coreferencing *he*-pronoun (Duffy & Rayner, 1990; O'Brien, 1987).

Unlike Experiment 1, we did not observe any reliable difference in time-course. There were some indications that *he*-pronoun conditions might still be resolved faster than *it*-pronouns. Hence, although the structural and grammatical properties of the NP1 antecedent appeared

to help counteract the effect of *it*-pronoun ambiguity observed in Experiment 1, they may not have been sufficient to completely eliminate the inherent ambiguity of this type of pronoun. However, it may also be the case that the time-course advantage for *he*-pronouns, if reliable, is partly due to another factor. McElree, Foraker, and Dyer (2003) found that subject-verb dependencies were processed with a faster time-course when arguments were adjacent to one another, rather than separated by intervening information. The faster rate might be explained by the fact that the antecedent for the *he*-pronoun occurred in the NP2 position, which was adjacent to the *he*-pronoun. Hence, it is possible that the NP2 antecedent might have remained active in focal attention when the *he*-pronoun was encountered, enabling it to be directly matched to the pronoun. We did not observe a comparable advantage in Experiment 1, but in that case the NP2 antecedents were coreferent with *it*-pronouns, and the ambiguity of these pronouns likely masked any advantage based on adjacency. Whether adjacency confers any time-course advantage in processing remains to be determined, and is best examined by comparing conditions that do not differ in ambiguity.

General discussion

Clefting

We investigated the memory states that underlie prominent antecedent representations. To do so, we manipulated prominence by syntactic clefting to test whether syntactic focusing causes an antecedent representation to be actively maintained in focal attention (Gundel, 1999; Gundel et al., 1993). No evidence was found to support this view. Studies of memory retrieval have demonstrated that representations maintained in focal attention are more accessible than representations in a more passive memory state, because they can be matched to a recognition probe without requiring retrieval (McElree, 1996, 1998, 2006; McElree & Doshier, 1989, 1993; Wickelgren et al., 1980). However, in two separate speed-accuracy tradeoff studies, there was no indication that pronouns corefering to clefted antecedents were resolved faster than those referring to nonclefted antecedents. The speed-accuracy time-course profiles for the clefted and nonclefted antecedents did not differ in the two parameters, rate and intercept, that would reflect differences in the accessibility of antecedent information.

Importantly, the absence of an effect of clefting on processing speed cannot be easily attributed to the potential insensitivity of the speed-accuracy tradeoff procedure. In Experiment 1, we found a slower speed of resolution for pronouns that were ambiguous (*it*) and coreferred with a backgrounded antecedent (NP2). That this effect was due to ambiguity was supported

by eye-tracking measures in Experiment 2, which indicated that these ambiguous pronoun conditions produced more regressions during a first-pass reading. Collectively, these findings indicate that the speed-accuracy tradeoff procedure is sensitive to and capable of capturing differences associated with the time-course of coreference resolution. The absence of an accessibility difference, therefore, provides evidence against the idea that comprehenders actively maintain a clefted antecedent in focal attention, challenging claims such as Gundel (1999; Gundel et al., 1993). Minimally, these results indicate that claims concerning the psychological basis of prominence must be revisited.

Rather than affecting processing speed, clefting appears to increase only the likelihood of successful pronoun resolution. Rather than increasing the accessibility of an antecedent representation through active maintenance, our data indicate that clefting simply served to make the antecedent prominent by amplifying the strength of the representation in the comprehender's discourse model, increasing its availability for on-going operations such as coreference resolution and integration. Several theorists have referred to such a change in availability as increasing the activation of the antecedent in memory (e.g., Almor, 1999; Garrod et al., 1994; Garrod & Terras, 2000; Gernsbacher, 1989; Gernsbacher et al., 1989; MacDonald & MacWhinney, 1990). Our findings are generally consistent with an activation metaphor, although we note that there are many ways, other than clefting, by which distinctiveness or salience may be encoded into a memory representation.

The eye-tracking results of Experiment 2 appear to be fully consistent with this interpretation. Clefting did not engender any "early" effects in eye-tracking, as might be expected if a pronoun could be resolved by matching it directly to the contents of focal attention. Rather, we observed an advantage for clefted antecedent conditions in several "late" eye-tracking measures. To the degree that effects on these measures can be uniquely attributed to late stages in anaphor resolution, such as integrating the pronoun-antecedent bond into the discourse (e.g., Garrod & Terras, 2000; Sanford et al., 1983), the speed-accuracy tradeoff and eye-tracking measures converge in suggesting that clefting does not affect the speed of accessing a memory representation so much as the quality of a representation. This is consistent with findings from other experiments indicating that clefted constituents produce stronger memory representations than nonclefted controls (e.g., Birch et al., 2000; Birch & Garnsey, 1995).

There are several reasons for why differences in memory strength could result in the observed differences in both "late" eye-tracking measures and asymptotic levels of performance in the speed-accuracy tradeoff procedure. Asymptotic differences could indicate that less prominent antecedent representations may not be retrieved as successfully as prominent antecedents. As noted, any failure

to retrieve the correct antecedent, either because the antecedent was lost from memory or a more salient antecedent interfered with the retrieval process, will lower the asymptote. Presumably, a failure to recover the correct antecedent would also engender late effects in eye-tracking measures, such as the observed increases in second-pass time, total time, and the proportion of regressions, particularly if readers attempted reanalysis to recover the antecedent. The differences in both measures could also indicate that the recovered representation for a poorly encoded antecedent did not support interpretive operations as well as a richly encoded antecedent. As we have noted, this could cause interpretation to fail or could result in a less meaningful interpretation.

We suspect that the effects observed in both asymptotic accuracy and late eye-tracking measures reflect a mixture of retrieval failure and less successful interpretation. Unfortunately, it is difficult to cleanly isolate unique sources of the differences. One possibility might be to collect plausibility ratings on the discourses, and then use these ratings to determine to what extent differences in asymptote are predicted, or accounted for, by plausibility of the two sentences together. Such an approach assumes, however, that plausibility ratings reflect differences in plausibility solely. This assumption is probably unwarranted. For example, if readers fail to retrieve the correct antecedent, then they will obviously rate the expression as less plausible than one for which the antecedent was successfully retrieved. In short, various offline measures are no more informative than the online asymptotic differences in resolving this matter. Nonetheless, our interpretation of these differences is generally consistent with approaches which argue that less prominent antecedents are less active in memory and therefore less likely to be retrieved (Birch et al., 2000; Birch & Garnsey, 1995).

Pronoun type

In addition to effects of clefting, we found clear differences between gendered (*he* and *she*) and ungendered pronouns (*it*). In both speed-accuracy tradeoff experiments, conditions with *it*-pronouns were less likely to be correctly resolved, producing lower asymptotes when the antecedent was either foregrounded (NP1 in Experiment 3) or backgrounded (NP2 in Experiment 1). When the *it*-pronouns coreferred with NP2 (nonsubject, second-mention, nonparallel structural role), analysis of the speed-accuracy tradeoff functions also showed that interpreting the *it*-pronoun had a slower time-course than *he*-pronouns coreferring with NP1 (subject, first-mention, parallel structural role).

Pronoun ambiguity

A likely interpretation of both the accuracy and time-course effects is that the *it*-pronouns in our constructions

did not provide cues strong enough to uniquely recover the intended antecedent (Greene et al., 1992). In our materials, coreference was not ambiguous in terms of the choice between explicitly mentioned entities (e.g., Järvikivi et al., 2005; MacDonald & MacWhinney, 1990; McDonald & MacWhinney, 1995), as, in all cases, the pronoun in the second sentence had only one possible explicit antecedent, and the verb following the pronoun forced just that interpretation. Rather there was ambiguity about other possible interpretations of *it*: *It* has a nonreferential use (e.g., pleonastic), as well as a referential use in which *it* could refer to either an inanimate noun, as intended, or to an event expressed in the immediate discourse.

For example, consider the following discourse from Experiment 1 (Table 1), in (3).

(3) *What the ardent boyfriend contemplated was the engagement ring. It sparkled.*

On some proportion of trials, comprehenders can initially interpret *it* as nonreferential or as referring to the event of contemplating the ring. In either case, the interpretation will clash with the verb *sparkle*, and comprehenders will need additional time to recover the appropriate coreference relation. Lower asymptotes than those for gendered pronouns will arise if the recovery is not fully successful. A slower time-course will arise because of the additional time required for reanalysis.

In Experiment 3, discourse properties of the antecedent that signal prominence, like subject, first-mention, and parallel relation with the pronoun, appear to have moderated the slower time-course for *it*-pronouns found in Experiment 1. In Experiment 3, time-course functions for the *it*-pronouns had marginally slower rate parameters than the functions for *he*-pronouns, but these differences were not reliable. Whether these types of discourse constraints can completely eliminate the ambiguity associated with *it*-pronouns is unclear, but our data suggest that they may have narrowed the interpretation of *it*-pronouns sufficiently so that inappropriate interpretations did not compete and delay the correct interpretation of the pronoun as much. We did not test directly whether these constraints enhanced the correct antecedent representation or, conversely, suppressed competing interpretations of the *it*-pronoun (e.g., Gernsbacher, 1990), but the net effect may be the same in both cases.

Animacy of the antecedent

We have attributed the processing differences between gendered and ungendered pronouns to the fact that the latter are inherently more ambiguous than the former. However, they also differ in animacy: Gendered pronouns coreferred with animate antecedents, typically people, whereas ungendered pronouns in our experiments coreferred with inanimate antecedents.

Conceptually, animate entities appear to be encoded more richly and strongly than inanimate ones. Keil's

(1979) investigation of category structures has shown that we are more likely to have highly elaborated knowledge about animates, suggesting that the associated knowledge structures are more complex than those for inanimate concepts.

Increasingly, researchers are finding that animacy has measurable effects on language processing. In language production, Bock and colleagues have argued that animacy is a conceptual feature which is reliably associated with the subject argument, because even with less than optimal conditions, animates are produced in subject slot (Bock & Loebell, 1990; Bock, Loebell, & Morey, 1992, see also Bates & MacWhinney, 1982). Recent comprehension research has also shown that animacy modulates the difficulty of understanding relative clause structures (Mak, Vonk, & Schriefers, 2002, 2006; Traxler, Morris, & Seely, 2002; Traxler, Williams, Blozis, & Morris, 2005). Object-relative clauses (e.g., *The musician whom the accident frightened phoned the police*) are more difficult to process than subject-relative clauses (e.g., *The musician who witnessed the accident phoned the police*), but only when the subject is animate and the object is inanimate. If the typical roles are reversed, so that the subject is then inanimate and the object animate (e.g., subject-relative: *The accident that the musician witnessed caused a lot of injuries* versus object-relative: *The accident that frightened the musician caused a lot of injuries*), the difficulty with the object-relative structure is reduced or eliminated.

All the differences we found between gendered and ungendered pronouns probably cannot be explained by animacy without appealing to inherent differences in pronoun ambiguity. However, the higher asymptotic levels that we found for animate antecedents in both of the speed-accuracy tradeoff experiments might reflect their greater availability in memory, if animate concepts, or especially humans, are by and large more strongly encoded in memory than inanimate ones.

It is less clear whether the faster speed of resolution in Experiment 1 should be attributed to the *increased* speed of accessing animate antecedents as opposed to the *delayed* processing of inherently ambiguous *it*-pronouns, as we have suggested. Mak, Vonk, and Schriefers (2006) proposed that when animacy coincides with topicality (givenness), it is used at the earliest stage of analysis. Their proposal is at least partially consistent with the pattern of our time-course results. Consider that in Experiment 1, the animate antecedent occurred in a foregrounded, topic position (NP1), and exhibited a faster time-course than the inanimate antecedent (NP2). In Experiment 2, when the animate antecedent occurred in a backgrounded, nontopic position (NP2), there was no reliable time course advantage over the inanimate antecedent (NP1). However, we maintain that the ambiguity of the *it*-pronouns, and the need for reanalysis (Traxler et al., 2002), provides a natural explanation for why the speed of processing was delayed

for inanimate antecedent cases in Experiment 1. As well, we note that the more prominent position of the inanimate antecedent in Experiment 3 attenuated this time-course delay, but perhaps not entirely.

In sum, animacy might be responsible for some of the observed differences between pronoun types in our experiments, but we did not factorially manipulate pronoun animacy and ambiguity, so it is not possible for us to cleanly isolate what effect animacy might have had. Indeed, it may not be possible to investigate these issues in pronoun resolution because, as we have noted, pronouns such as *it* that refer to inanimate antecedents are inherently more ambiguous. Consequently, researchers may need to investigate the role of animacy in coreference resolution with other anaphor devices, such as coreferencing noun phrases.

Generality of the effects

We found no evidence that prominent antecedents were maintained within the focus of attention, and hence were more accessible than other discourse entities, even though our studies examined short segments (Grosz et al., 1995), which should have maximized the comprehender's potential to maintain the prominent antecedent in focal attention. Nonetheless, it is possible that differences in accessibility could arise under other circumstances. Repeated coreference may help keep an antecedent representation active. Garrod et al. (1994) found an early eye-tracking advantage for coreference with a discourse-focused entity, which was established by a proper name at the beginning of the discourse and then maintained with repeated coreference. Cataphors, which introduce the anaphor before the referent, may establish active maintenance of the anaphoric phrase. van Gompel and Liversedge (2003) found early eye-tracking effects of gender agreement for cataphoric pronouns and their referents. Another factor that may lead an antecedent representation to be actively maintained is whether it is emphasized prosodically. Both production and comprehension studies of speech suggest that processing is facilitated by the accenting of new or focused information and the deaccenting of old, or given information (Fowler & Housum, 1987; see Cutler, Dahan, & van Donselaar, 1997 for a review). For example, Gernsbacher and Jescheniak (1995) showed that prosodic stress on a nonanaphoric entity preserved its ease of recognition (probe verification) across a second, potentially interfering noun.

Although such studies are suggestive, findings from several basic memory tasks indicate that focal attention is extremely limited (McElree, 2006). Comprehenders may not be able to actively maintain the relevant representation in focal attention with a high degree of success while concurrently processing new information that is introduced into the discourse. Hence, it remains to be

determined whether circumstances such as these, where there is repeated coreference across a discourse, cataphoric coreference, or prosodic accent on an antecedent, might yield evidence for the active maintenance of a prominent antecedent representation.

Finally, we note that we did not explicitly manipulate the number or kind of possible antecedents for a pronoun in our experiments. Preceding the pronoun, there were always two nouns, which were the same in all conditions. Several studies have found that nouns which should not be considered as antecedents due to their structural or syntactic position in a sentence do, nonetheless, slow processing time (Badecker & Straub, 2002; Kennison, 2003; Sturt, 2003; van Gompel & Liv-ersedge, 2003). One explanation for such findings is that competition still occurs between the correct antecedent and structurally disallowed candidates. In our present results, we do consider competition between antecedents or interpretations to be greater for the *it*-pronouns than the *he*-pronouns. But we attribute this to the pronoun type, and not the nature of candidate referents, because they were always the same across conditions. It may be that as the number of competing referents explicitly mentioned increases, the slower the speed of processing, resulting in systematic slowing of the rate or intercept of the speed-accuracy tradeoff functions. Contrasting the success or ease of coreference with possible referents in discourse available positions varying in discourse focus, with possible referents in structurally disallowed positions, may help to further clarify the nature of competition among entities during pronoun resolution.

Appendix A

Throughout the appendix, the antecedent sentence is shown without a clefting structure. Tables 1, 3, and 6 show the sentence structures used in each of the three experiments.

Materials from Experiment 1

The unacceptable versions of each item are formed by switching the verbs following the pronouns. See Table 1 for an example. A subset of materials is shown below. As well, Experiment 1 included the items for Experiments 2 and 3, with minimal changes.

1. The incompetent accountant figured out the tax payment. He scribbled./It tripled.
2. The bustling midwife treated the recurrent pain. She hurried./It subsided.
3. The diligent secretary copied the worn letter. She typed./It crinkled.
4. The self-confident suitor presented the scrawny bouquet. He winked./It withered.
5. The rookie policewoman discovered the terrorist bomb. She squealed./It ticked.

6. The clumsy stewardess spilled the red wine. She tripped./It stained.
7. The bratty nephew knocked over the block tower. He yelled./It toppled.
8. The pretty witness listened to the cassette tape. She testified./It rewound.
9. The red-faced plumber detested the grimy drain. He wheezed./It clogged.
10. The graying manager demonstrated the zip drive. He mumbled./It ejected.
11. The successful actor lit the Cuban cigar. He puffed./It smoldered.
12. The chamber maid delivered the breakfast tray. She knocked./It tipped.
13. The skillful carpenter repaired the antique dresser. He hammered./It creaked.
14. The jolly nanny warmed the chocolate milk. She chuckled./It curdled.
15. The conscientious doctor was concerned about the massive gash. He muttered./It oozed.
16. The temperamental sister kicked the bedroom door. She whined./It slammed.
17. The doting aunt created the woolen sweater. She knitted./It shrank.
18. The city lady tried to light the camp fire. She yelped./It crackled.
19. The devout nun prepared the baptismal water. It trickled./She prayed.
20. The balding gambler banged the slot machine. He cursed./It malfunctioned.

Materials from Experiment 2

1. The new foreman examined the latest blueprint. Annoyingly, he squinted at the lines/it curled at the edges of the paper.
2. The eldest princess wore the prettiest necklace. Enchantingly, she curtsied in the ballroom/it glittered in the candlelight at the palace.
3. The teenage boy built the pipe bomb. Offensively, he snickered/it fizzled by the river in the woods.
4. The head priest delivered the weekly sermon. Relentlessly, he prayed/it extended into the afternoon at the church.
5. The novice seamstress cut the delicate fabric. Irritatingly, she fussed/it snagged at the wrinkles near the middle.
6. The meticulous surgeon studied the cesarean incision. Expectedly, he grumbled/it oozed way too much about/from the infection.
7. The rookie fireman fought the hazardous inferno. Continuously, he coughed/it blazed through the night without any relief.
8. The cheerful waitress made the decaffeinated coffee. Reassuringly, she gossiped/it brewed behind the counter of the diner.
9. The impertinent pilot presented the overstuffed suitcase. Regrettably, he nagged at the co-pilot/it ruptured at the zipper during the flight.

10. The sleepy daughter ignored the alarm clock. Regularly, she snored/it rang for five minutes under the covers.
11. The exhausted girl fled the violent tornado. Unfortunately, she fainted/it gusted along the road leading into town.
12. The overworked nurse unclamped the wrong artery. Undesirably, she panicked/it spurting in the operating room of the hospital.
13. The ailing butler sounded the dinner bell. Insensitively, he eavesdropped/it clanged in the hallway outside the door.
14. The voluptuous actress eyed the movie camera. Quietly, she whispered/it whirred during the scene at the studio.
15. The bald salesman sold the baby carriage. Afterward, he celebrated/it rattled on the sidewalk near a restaurant.
16. The star quarterback threw the tattered football. Aggressively, he jumped/it spiraled into the air with great force.
17. The brave marine carried the national flag. Auspiciously, he marched in the parade/it rippled in the wind down/along the street.
18. The responsible girl scout tended the evening campfire. Persistently, she sneezed/it smoldered in the twilight of the day.
19. The lion tamer raised the wooden chair. Suddenly, he shouted/it splintered inside the cage of the cat.
20. The capable grandma entered the recipe contest. Finally, she experimented/it opened in the fall before the holidays.
21. The tired mother cooked the cinnamon oatmeal. Gradually, she yawned in the morning near the stove/it thickened in the pot on the stove.
22. The nervous debutante opened the expensive champagne. Subsequently, she giggled from the thrill/it bubbled from the bottle at the party.
23. The talkative salesgirl sprayed the spicy perfume. Eventually, she smiled/it leaked behind the counter of the boutique.
24. The devastated widow received the sympathy arrangement. Consequently, she grieved for many months/it bloomed for many weeks after the funeral.
25. The dishonest husband skimmed the retirement money. Ultimately, he lied/it diminished throughout the year very subtly.
26. The motocross stuntman admired the flashy bike. Surprisingly, he died/it shattered in an accident two weeks later.
27. The picky model rejected the cheap dress. Predictably, she sneered at the design/it ripped at the seams in the showroom.
28. The jaded prostitute applied the gaudy lipstick. Frequently, she shivered/it smudged in the rain outside the nightclub.
29. The proud fisherman used the trawling net. Effortlessly, he boasted on the docks/it swirled in the water about/around the catch.
30. The friendly barmaid served the strong cocktail. Reliably, she chattered to the guys at the bar/it fizzed to the rim of the glass.
31. The boring uncle flipped the greasy hamburger. Unpleasantly, he complained on the patio/it burned on the grill in the backyard.
32. The handsome cowboy twirled the frayed rope. Distractingly, he muttered/it unraveled during the event at the rodeo.
33. The finicky tailor completed the pinstriped suit. Accordingly, he sighed/it creased after much work throughout the day.
34. The reclusive woman donated the tropical island. Annually, she disappeared/it flooded during the spring without any warning.
35. The stylish gentleman identified the priceless vase. Inopportunistically, he gasped/it cracked during the bidding at the auction.
36. The belly dancer sported the coin belt. Suggestively, she wiggled/it jingled at the restaurant by the sea.
37. The rambunctious son smashed the lego tower. Promptly, he laughed/it broke in the playroom during the afternoon.
38. The popular newswoman revealed the ransom note. Prophetically, she spoke/it tore before the release of the victim.
39. The prison warden discovered the escape tunnel. Inevitably, he swore/it lengthened into the darkness of the hole.
40. The prima ballerina tried the lightweight shoe. Gracefully, she danced/it supported at the theater during a performance.
41. The devoted grandson inherited the ramshackle house. Progressively, he paid/it improved over the summer due to repairs.
42. The beauty contestant championed the charity fund. Fortunately, she won/it grew after some setbacks the previous year.
43. The bashful groom envisioned the wedding dance. Tenderly, he practiced/it finished on the day of the wedding.
44. The seasoned congresswoman carried the digital organizer. Punctually, she arrived/it beeped on the hour for a meeting.
45. The stock broker transported the metal briefcase. Awkwardly, he tripped/it clunked over the chair near the door.
46. The fairy godmother bestowed the magic wand. Mysteriously, she levitated/it sparkled in the moonlight outside the window.
47. The peevish headmaster abhorred the paper airplane. Invariably, he yelled/it landed at the front of the classroom.
48. The demanding queen planned the momentous banquet. Festively, she triumphed/it succeeded with much fanfare from the court.

Materials from Experiment 3

The unacceptable versions of each item are formed by switching the verbs following the pronouns. See [Table 6](#) for an example.

1. The digital organizer interested the seasoned congresswoman. She nodded./It beeped.
2. The sympathy arrangement comforted the devastated widow. She grieved./It bloomed.

3. The hearty soup revitalized the tired mother. She hummed./It boiled.
4. The center spotlight left the prima ballerina. She breathed./It dimmed.
5. The light perfume invigorated the talkative salesgirl. She chatted./It dissipated.
6. The momentous banquet appeased the stately queen. She feasted./It ended.
7. The diamond tiara heartened the beauty contestant. She smiled./It fit.
8. The expensive champagne enlivened the nervous debutante. She giggled./It overflowed.
9. The carved plaque honored the committee chairwoman. She retired./It warped.
10. The jasmine tea pleased the geisha girl. She shuffled./It steeped.
11. The cheap dress exasperated the picky model. She sneered./It ripped.
12. The unsafe food distressed the investigative newswoman. She spoke./It spoiled.
13. The alarm clock awakened the sleepy daughter. She stretched./It rang.
14. The colorful flag excited the enthusiastic cheerleader. She strutted./It fluttered.
15. The tropical island amazed the naïve heiress. She sunbathed./It flooded.
16. The rickety vault unsettled the graceful gymnast. She tripped./It clattered.
17. The hazardous inferno overwhelmed the senior fireman. He coughed./It blazed.
18. The scary nightmare spooked the little boy. He cried./It recurred.
19. The heavy bookcase jarred the furniture mover. He cursed./It splintered.
20. The fire alarm alerted the security guard. He flinched./It clanged.
21. The corruption charge tormented the powerful executive. He fretted./It stuck.
22. The rusty boiler perturbed the school janitor. He frowned./It leaked.
23. The lead paint annoyed the safety inspector. He grimaced./It flaked.
24. The gold treasure astounded the balding archeologist. He gulped./It glimmered.
25. The well-marked trail reassured the bearded mountaineer. He hiked./It narrowed.
26. The greasy hamburger cheered the hungry uncle. He joked./It sizzled.
27. The endless tide mesmerized the retired grandfather. He mused./It ebbed.
28. The plaintive bugle saddened the mourning general. He saluted./It resounded.
29. The savory stew impressed the uncompromising chef. He sniffed./It thickened.
30. The engagement ring captivated the ardent boyfriend. He stared./It sparkled.
31. The old flashlight guided the careful prison warden. He tiptoed./It flickered.
32. The cold wind chilled the grizzled pioneer. He trudged./It gusted.

References

- Almor, A. (1999). Noun–phrase anaphora and focus: the informational load hypothesis. *Psychological Review*, *106*, 748–765.
- Ariel, M. (1990). *Accessing noun–phrase antecedents*. London: Routledge.
- Arnold, J. E., Eisenband, J. G., Brown-Schmidt, S., & Trueswell, J. C. (2000). The rapid use of gender information: evidence of the time course for pronoun resolution from eyetracking. *Cognition*, *76*, B13–B26.
- Badecker, W., & Straub, K. (2002). The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 748–769.
- Bates, E., & MacWhinney, B. (1982). Functionalist approaches to grammar. In E. Wanner & L. R. Gleitman (Eds.), *Language acquisition: The state of the art* (pp. 173–218). Cambridge, England: Cambridge University Press.
- Birch, S. L., Albrecht, J. E., & Myers, J. L. (2000). Syntactic focusing structures influence discourse processing. *Discourse Processes*, *30*, 285–304.
- Birch, S. L., & Garnsey, S. (1995). The effect of focus on memory for words in sentences. *Journal of Memory and Language*, *34*, 232–267.
- Birch, S. L., & Rayner, K. (1997). Linguistic focus affects eye movements during reading. *Memory & Cognition*, *25*, 653–660.
- Bock, J. K., & Loebell, H. (1990). Framing sentences. *Cognition*, *35*, 1–39.
- Bock, J. K., Loebell, H., & Morey, R. (1992). From conceptual roles to structural relations: bridging the syntactic cleft. *Psychological Review*, *99*, 150–171.
- Bornkessel, I., McElree, B., Schlesewsky, M., & Friederici, A. D. (2004). Multi-dimensional contributions to garden path strength: dissociating phrase structure from relational structure. *Journal of Memory and Language*, *51*, 495–522.
- Brysbaert, M., & Mitchell, D. C. (1996). Modifier attachment in sentence parsing: evidence from Dutch. *Quarterly Journal of Experimental Psychology*, *49A*, 664–695.
- Carpenter, P. A., & Just, M. A. (1977). Reading comprehension as eyes see it. In M. A. Just & P. A. Carpenter (Eds.), *Cognitive processes in comprehension*. Hillsdale, NJ: Erlbaum.
- Carreiras, M., Gernsbacher, M. A., & Villa, V. (1995). The advantage of first mention in Spanish. *Psychonomic Bulletin & Review*, *2*, 124–129.
- Chandler, J. P. (1969). Subroutine STEPIT—finds local minimum of a smooth function of several parameters. *Behavioral Science*, *14*, 81–82.
- Chafe, W. L. (1994). *Discourse, consciousness, and time. The flow and displacement of conscious experience in speaking and writing*. Chicago: Chicago University Press.
- Chambers, C. G., & Smyth, R. (1998). Structural parallelism and discourse coherence: a test of centering theory. *Journal of Memory and Language*, *39*, 593–608.
- Clark, H. H. (1973). The language-as-fixed-effect fallacy: a critique of language statistics in psychological research. *Journal of Verbal Learning and Verbal Behavior*, *12*, 355–359.

- Cutler, A., Dahan, D., & van Donselaar, W. (1997). Prosody in the comprehension of spoken language: a literature review. *Language and Speech*, 40, 141–201.
- Dosher, B. A. (1979). Empirical approaches to information processing: speed-accuracy tradeoff or reaction time. *Acta Psychologica*, 43, 347–359.
- Dosher, B. A. (1984). Degree of learning and retrieval speed: Study time and multiple exposures. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 10, 541–574.
- Duffy, S. A., & Rayner, K. (1990). Eye-movements and anaphor resolution: effects of antecedent typicality and distance. *Language and Speech*, 33, 103–119.
- Ehrlich, K., & Rayner, K. (1983). Pronoun assignment and semantic integration during reading: eye movements and immediacy of processing. *Journal of Verbal Learning and Verbal Behavior*, 22, 75–87.
- Engelkamp, J., & Zimmer, H. D. (1982). The interaction of subjectivization and concept placement in the processing of cleft sentences. *Quarterly Journal of Experimental Psychology*, 34A, 463–478.
- Fodor, J. D., & Inoue, A. (2000). Syntactic features in reanalysis: positive and negative symptoms. *Journal of Psycholinguistic Research*, 29, 25–36.
- Foraker, S. (2004). The mechanisms involved in the prominence of referent representations during pronoun coreference (Doctoral dissertation, New York University, 2004). *UMI ProQuest Digital Dissertations*.
- Fowler, C. A., & Housum, J. (1987). Talkers' signaling of "new" and "old" words in speech and listeners' perception and use of the distinction. *Journal of Memory and Language*, 26, 489–504.
- Garnham, A. (2001). *Mental models and the interpretation of anaphora*. East Sussex: Psychology Press.
- Garnham, A., Traxler, M., Oakhill, J. V., & Gernsbacher, M. A. (1996). The locus of implicit causality effects in comprehension. *Journal of Memory and Language*, 35, 517–543.
- Garrod, S., Freudenthal, D., & Boyle, E. (1994). The role of different types of anaphor in the on-line resolution of sentences in a discourse. *Journal of Memory and Language*, 33, 39–68.
- Garrod, S., & Sanford, A. J. (1990). Referential processing in reading: Focusing on roles and individuals. In D. A. Balota, G. B. Flores d'Arcais, & K. Rayner (Eds.), *Comprehension processes in reading* (pp. 465–486). Hillsdale, NJ: Erlbaum.
- Garrod, S., & Sanford, A. J. (1994). Resolving sentences in a discourse context: How discourse representation affects language understanding. In M. A. Gernsbacher (Ed.), *Handbook of psycholinguistics* (pp. 675–698). San Diego, CA: Academic Press.
- Garrod, S., & Terras, M. (2000). The contribution of lexical and situational knowledge to resolving discourse roles: Bonding and resolution. *Journal of Memory and Language*, 42, 526–544.
- Gernsbacher, M. A. (1989). Mechanisms that improve referential access. *Cognition*, 32, 99–156.
- Gernsbacher, M. A. (1990). *Language comprehension as structure building*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Gernsbacher, M. A., & Hargreaves, D. (1988). Accessing sentence participants: the advantage of first mention. *Journal of Memory and Language*, 27, 699–717.
- Gernsbacher, M. A., Hargreaves, D., & Beeman, M. (1989). Building and accessing clausal representations: the advantage of first mention versus the advantage of clause recency. *Journal of Memory and Language*, 28, 735–755.
- Gernsbacher, M. A., & Jescheniak, J. D. (1995). Cataphoric devices in spoken discourse. *Cognitive Psychology*, 29, 24–58.
- Givón, T. (Ed.). (1983). *Topic continuity in discourse: A quantitative crosslanguage study*. Amsterdam: Benjamins.
- Gordon, P. C., Grosz, B. J., & Gilliom, L. A. (1993). Pronouns, names, and the centering of attention. *Cognitive Science*, 17, 311–347.
- Greene, S. B., McKoon, G., & Ratcliff, R. (1992). Pronoun resolution and discourse models. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 18, 266–283.
- Grober, E. H., Beardsley, W., & Caramazza, A. (1978). Parallel function strategy in pronoun assignment. *Cognition*, 6, 117–133.
- Grosz, B. J., Joshi, A. K., & Weinstein, S. (1995). Centering: a framework for modelling the local coherence of discourse. *Computational Linguistics*, 21, 203–226.
- Grosz, B. J., & Sidner, C. L. (1986). Attention, intentions, and the structure of discourse. *Computational Linguistics*, 12, 175–204.
- Gundel, J. K. (1999). On different kinds of focus. In P. Bosch & R. van der Sandt (Eds.), *Focus: Linguistic, cognitive, and computational perspectives*. Cambridge: Cambridge University Press.
- Gundel, J. K., Hedberg, N., & Zacharski, R. (1993). Cognitive status and the form of referring expressions in discourse. *Language*, 69, 274–307.
- Hudson, S. B., Tanenhaus, M. K., & Dell, G. S. (1986). The effect of the discourse center on the local coherence of a discourse. In *Proceedings of the eighth annual conference of the Cognitive Science Society* (pp. 96–101). Hillsdale, NJ: Lawrence Erlbaum.
- Järvikivi, J., van Gompel, R. P. G., Hyönä, J., & Bertram, R. (2005). Ambiguous pronoun resolution: contrasting the first-mention and subject-preference accounts. *Psychological Science*, 16, 260–264.
- Judd, C. M., & McClelland, G. H. (1989). *Data analysis: A model-comparison approach*. San Diego: Harcourt Brace Jovanovich.
- Keil, F. C. (1979). *Semantic and conceptual development: an ontological perspective*. Cambridge, MA: Harvard University Press.
- Kennison, S. M. (2003). Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language*, 49, 335–352.
- Klin, C. M., Weingartner, K. M., Guzmán, A. E., & Levine, W. H. (2004). Readers' sensitivity to linguistic cues in narratives: How salience influences anaphor resolution? *Memory & Cognition*, 32, 511–522.
- Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Eye movements and measures of reading time. In G. Underwood (Ed.), *Eye guidance in reading and scene perception* (pp. 55–75). Oxford: Elsevier.
- Mak, W. M., Vonk, W., & Schriefers, H. (2002). The influence of animacy on relative clause processing. *Journal of Memory and Language*, 47, 50–68.

- Mak, W. M., Vonk, W., & Schriefers, H. (2006). Animacy in processing relative clauses: The hikers that rocks crush. *Journal of Memory and Language*, 54, 466–490.
- Matthews, A., & Chodorow, M. S. (1988). Pronoun resolution in two-clause sentences: effects of ambiguity, antecedent location, and depth of embedding. *Journal of Memory and Language*, 27, 245–260.
- MacDonald, M. C., & MacWhinney, B. (1990). Measuring inhibition and facilitation from pronouns. *Journal of Memory and Language*, 29, 469–492.
- Masson, M. E. J., & Loftus, G. R. (2003). Using confidence intervals for graphically based data interpretation. *Canadian Journal of Experimental Psychology*, 57, 203–220.
- McDonald, J. L., & MacWhinney, B. (1995). The time course of anaphor resolution: effects of implicit verb causality and gender. *Journal of Memory and Language*, 34, 543–566.
- McElree, B. (1993). The locus of lexical preference effects in sentence comprehension: a time-course analysis. *Journal of Memory and Language*, 32, 536–571.
- McElree, B. (1996). Accessing short-term memory with semantic and phonological information: A time-course analysis. *Memory & Cognition*, 24, 173–187.
- McElree, B. (1998). Attended and non-attended states in working memory: accessing categorized structures. *Journal of Memory and Language*, 38, 225–252.
- McElree, B. (2000). Sentence comprehension is mediated by content-addressable memory structures. *Journal of Psycholinguistic Research*, 29, 111–123.
- McElree, B. (2001). Working memory and focal attention. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 27, 817–835.
- McElree, B. (2006). Accessing recent events. In B. H. Ross (Ed.), *The psychology of learning and motivation* (Vol. 46). San Diego: Academic Press.
- McElree, B., & Doshier, B. A. (1989). Serial position and set size in short term memory: the time course of recognition. *Journal of Experimental Psychology: General*, 118, 346–373.
- McElree, B., & Doshier, B. A. (1993). Serial retrieval processes in the recovery of order information. *Journal of Experimental Psychology: General*, 122, 291–315.
- McElree, B., Foraker, S., & Dyer, L. (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language*, 48, 67–91.
- McElree, B., & Nordlie, J. (1999). Literal and figurative interpretations are computed in equal time. *Psychonomic Bulletin & Review*, 6, 486–494.
- McElree, B., Pyllkkänen, L., Pickering, M. J., & Traxler, M. (2006). The time course of enriched composition. *Psychonomic Bulletin & Review*, 13, 53–59.
- McKoon, G., Gerrig, R., & Greene, S. B. (1996). Pronoun resolution without pronouns: some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 22, 919–932.
- McKoon, G., & Ratcliff, R. (1980). The comprehension processes and memory structures involved in anaphoric reference. *Journal of Verbal Learning and Verbal Behavior*, 19, 668–682.
- Morris, R. K., & Folk, J. R. (1998). Focus as a contextual priming mechanism in reading. *Memory & Cognition*, 26, 1313–1322.
- Nairne, J. S. (1996). Short-term/working memory. In E. L. Bjork & R. A. Bjork (Eds.), *Memory*. San Diego: Academic Press.
- Nicol, J. L., & Swinney, D. A. (2003). The psycholinguistics of anaphora. In A. Barss (Ed.), *Anaphora: A reference guide* (pp. 72–104). Oxford, UK: Blackwell.
- O'Brien, E. J. (1987). Antecedent search processes and the structure of text. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13, 278–290.
- Pickering, M. J., Frisson, S., McElree, B., & Traxler, M. J. (2004). Eye movements and semantic composition. In M. Carreiras & C. Clifton (Eds.), *The on-line study of sentence comprehension: Eyetracking, ERP and beyond*. New York: Psychology Press.
- Reed, A. V. (1973). Speed-accuracy tradeoff in recognition memory. *Science*, 181, 574–576.
- Reed, A. V. (1976). The time course of recognition in human memory. *Memory & Cognition*, 4, 16–30.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85, 59–108.
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124, 372–422.
- Rochemont, M. S., & Culicover, P. W. (1990). *English focus constructions and the theory of grammar*. New York: Cambridge University Press.
- Sanford, A. J., Garrod, S., Lucas, A., & Henderson, R. (1983). Pronouns without explicit antecedents. *Journal of Semantics*, 303–318.
- Schneider, W., Eschman, A., & Zuccolotto, A. (2002). E-Prime (Version 1.1). Pittsburgh, PA: Psychology Software Tools.
- Spelman, C. P., & Kirsner, K. (1990). The representation of text-based and situation-based information in discourse comprehension. *Journal of Memory and Language*, 29, 119–132.
- Sternberg, S. (1975). Memory scanning: new findings and current controversies. *Quarterly Journal of Experimental Psychology*, 27, 1–32.
- Stevenson, R., Nelson, A. W. R., & Stenning, K. (1995). The role of parallelism in strategies of pronoun comprehension. *Language and Speech*, 38, 393–418.
- Stewart, A. J., Pickering, M. J., & Sanford, A. J. (2000). The time course of the influence of implicit causality information: Focusing versus integration accounts. *Journal of Memory and Language*, 42, 423–443.
- Streb, J., Roesler, F., & Hennighausen, E. (1999). Event-related responses to pronoun and proper name anaphors in parallel and nonparallel discourse structures. *Brain and Language*, 70, 273–286.
- Sturt, P. (2003). The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language*, 48, 542–562.
- Traxler, M. J., Bybee, M. D., & Pickering, M. J. (1997). Influence of connectives on language comprehension: eye-tracking evidence for incremental interpretation. *Quarterly Journal of Experimental Psychology*, 50A, 481–497.
- Traxler, M. J., Morris, R. K., & Seely, R. E. (2002). Processing subject and object-relative clauses: evidence from eye-movements. *Journal of Memory and Language*, 47, 69–90.
- Traxler, M. J., Williams, R. S., Blozis, S. A., & Morris, R. K. (2005). Working memory, animacy, and verb class in the

- processing of relative clauses. *Journal of Memory and Language*, 53, 204–224.
- van Gompel, R. P. G., & Livsersedge, S. P. (2003). The influence of morphological information on cataphoric pronoun assignment. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29, 128–139.
- Vonk, W. (1984). Eye movements during comprehension of pronouns. In A. G. Gale & F. Johnson (Eds.), *Theoretical and Applied Aspects of Eye Movement Research*. North-Holland: Elsevier Science.
- Wickelgren, W. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41, 67–85.
- Wickelgren, W. A., Corbett, A. T., & Doshier, B. A. (1980). Priming and retrieval from short-term memory: a speed-accuracy tradeoff analysis. *Journal of Verbal Learning and Verbal Behavior*, 19, 387–404.
- Zimmer, H. D., & Engelkamp, J. (1981). The given-new structure of cleft sentences and their influence on picture viewing. *Psychological Research*, 43, 375–389.