Journal of Experimental Psychology:
Learning, Memory, and Cognition
2007, Vol. 33, No. 5, 914–930

0278-7393/07/$12.00 DOI: 10.1037/0278-7393.33.5.914

If You Say *Thee uh* You Are Describing Something Hard: The On-Line Attribution of Disfluency During Reference Comprehension

Jennifer E. Arnold
University of North Carolina at Chapel Hill

Carla L. Hudson Kam
University of California at Berkeley

Michael K. Tanenhaus
University of Rochester

Eye-tracking and gating experiments examined reference comprehension with fluent (*Click on the red . . .*) and disfluent (*Click on* [*pause*] *thee uh red . . .*) instructions while listeners viewed displays with 2 familiar (e.g., ice cream cones) and 2 unfamiliar objects (e.g., squiggly shapes). Disfluent instructions made unfamiliar objects more expected, which influenced listeners' on-line hypotheses from the onset of the color word. The unfamiliarity bias was sharply reduced by instructions that the speaker had object agnosia, and thus difficulty naming familiar objects (Experiment 2), but was not affected by intermittent sources of speaker distraction (beeps and construction noises; Experiments 3). The authors conclude that listeners can make situation-specific inferences about likely sources of disfluency, but there are some limitations to these attributions.

Keywords: reference comprehension, attribution inferences, disfluency, eye-tracking

Language comprehension is efficient (Marslen-Wilson, 1973, 1975). An illustration of this is reference comprehension, in which listeners typically begin to evaluate the referent of expressions like *the camel* within 200 ms of the onset of the noun (e.g., Allopenna, Magnuson, & Tanenhaus, 1998). Reference resolution begins long before the end of the word, when the input (e.g., *the ca—*) is still ambiguous. People achieve this efficiency by rapidly coordinating multiple sources of information. They recognize phonetic patterns as they unfold over time and match them with consistent lexical information (Marslen-Wilson, 1987). This is not a blind process, however; listeners pay particular attention to hypotheses that are more likely given the phonetic input, for example, preferring more frequent words (Dahan, Magnuson, & Tanenhaus, 2001), preferring objects that match the type of accenting on the expression (Dahan, Tanenhaus, & Chambers, 2002), rapidly identifying referents that are constrained by the linguistic and discourse context (Chambers, Tanenhaus, Eberhard, Carlson, & Filip, 2002), and generating expectations about which entities are likely to be mentioned (Altmann & Kamide, 1999). Reference resolution is even impacted by the presence of disfluency (Arnold, Tanenhaus, Altmann, & Fagnano, 2004; Bailey & Ferreira, 2003; Barr, 2001), such as words like *um* or *uh*, or pronouncing "the" as *thee* (/thiy/) instead of *thuh* (Fox Tree & Clark, 1997). For example, Arnold et al. (2004) found that when listeners heard a disfluent utterance (*Click on* [*pause*] *thee uh . . .*), they were more likely to initially consider an object that had not been previously mentioned than with fluent utterances.

Although the comprehension system can clearly make use of an impressive set of information, it is important to understand the nature of the underlying mechanisms. Here we focus on the role of inferential processing. Do listeners use variations in the input to draw rapid and situation-specific inferences about the speaker's meaning, intentions, or thought processes? For example, disfluency could affect reference comprehension by triggering an inference that the speaker is having difficulty as a result of referring to something that has not previously been mentioned. On the other hand, the appearance of such an inference might occur through the relatively automatic use of distributional information about disfluency, such as the fact that it occurs more often in references to discourse–new objects. Such statistical information could be used relatively automatically, allowing the listener to bypass speaker-specific or utterance-specific inferences.

In this article, we examine the on-line effects of disfluency on reference comprehension, with the goal of understanding the contribution of situation-specific inferences. In particular, we seek evidence of flexibility in the disfluency effect. We begin by extending the established disfluency bias toward unmentioned objects by showing that listeners also expect a referential expres-

Jennifer E. Arnold, Department of Psychology, University of North Carolina at Chapel Hill; Carla L. Hudson Kam, Psychology Department, University of California at Berkeley; Michael K. Tanenhaus, Department of Brain and Cognitive Sciences, University of Rochester.

This research was partially supported by National Institutes of Health Grants HD-41522 to Jennifer E. Arnold and HD-27206 to Michael K. Tanenhaus. We thank the following people for their help creating the stimuli, collecting, and coding the data: Shin-Yi Lao, Alex Christodoulou, Tatiana Meteleva, Glenn Kern, and Selena Beckman-Harned (at the University of North Carolina at Chapel Hill); and Dana Subik, Jill Thorson, and Becky Altmann (at the University of Rochester). We also thank Bob McMurray for help writing the presentation scripts for Experiment 1 and for the use of his analysis software for all the experiments.

Correspondence concerning this article should be addressed to Jennifer E. Arnold, Department of Psychology, University of North Carolina at Chapel Hill, CB #3270 Davie Hall, Chapel Hill, NC 27599. E-mail: jarnold@email.unc.edu

sion that contains a disfluency (*Click on thee uh* . . .) to refer to an unfamiliar rather than a familiar object. We then ask whether this unfamiliarity effect is dependent on the inference that unfamiliar objects are harder to refer to than familiar objects, and whether it persists in the face of alternate explanations for the disfluency.

The ability to draw inferences is not mutually exclusive from the automatic use of distributional information. The question is whether speakers bypass the need for high-level inferences through the use of lower level mechanisms. If speaker-specific and situation-specific information modulates the way listeners respond to disfluency, it would provide evidence for a more sophisticated integration of information than can be achieved by low-level automatic processes alone.

Background

Despite the desire to deliver well-formed and fluent utterances, speakers are often disfluent, saying *um* or *uh*, repeating or repairing words, hesitating, lengthening words, or pronouncing *the* as *thee* (/thiy/) instead of *thuh*. Although fluency has often been considered to be outside the linguistic system, it has been shown to impact listeners' on-line hypotheses about what the speaker is referring to. In a visual world experiment (Cooper, 1974; Tanenhaus, Spivey-Knowlton, Eberhard, & Sedivy, 1995), Arnold et al. (2004) tracked participants' eye movements as they viewed scenes of objects and followed instructions like *Put the camel below the grapes. Now put the* . . . . Each visual scene contained two objects that were consistent with the first part of the target word (e.g., camel and candle), creating a temporary ambiguity. Only one of these objects was "given," meaning that it was mentioned on the first instruction. With a fluent instruction, listeners were more likely to look at an object that was given in the discourse than an object that was new to the discourse, consistent with the general assumption that given information is more accessible than new information (e.g., Chafe, 1976; Gundel, 1988; Prince, 1992). By contrast, with a disfluent instruction (*Now put thee uh* -), listeners were more likely to look at an object that had not been previously mentioned. The observed biases to new and given objects were integrated with the unfolding lexical information very rapidly; listeners were more likely to look at the preferred object beginning at 200 ms after the onset of the target word (e.g., *camel*).

One plausible explanation for Arnold et al.'s (2004) result is that disfluency is the trigger for an inference that the speaker is having difficulty. Speakers tend to hesitate, use fillers like *um* or *uh*, or be otherwise disfluent when they are experiencing some kind of cognitive load (e.g., Bortfield, Leon, Bloom, Schober, & Brennan, 2001; Goldman-Eisler, 1968; Siegman, 1979). Thus, disfluency often occurs when speakers are planning utterances during language production (e.g., Beattie, 1979; Clark & Fox Tree, 2002; Clark & Wasow, 1998). Similarly, disfluency tends to occur when speakers are unsure of the answer to a question (Brennan & Williams, 1995; Smith & Clark, 1993). Disfluency also tends to occur at major breaks in discourse structure (Swerts, 1998; Swerts & Geluykens, 1994), where there is likely to be an increase in the speaker's planning load. Critically for our study, disfluency also occurs when speakers are likely to be having trouble with lexical retrieval, for example, when the word is low in frequency or contextual probability (Beattie & Butterworth, 1979; Goldman-Eisler, 1968), or when the speaker must choose between a number of plausible alternatives (Schachter, Christenfeld, Ravina, & Bilous, 1991; Schachter, Rauscher, Christenfeld, & Tyson Crone, 1994).

Even though cognitive load is not the only source of disfluency, it is plausible for listeners to take disfluency as a sign of likely speaker difficulty. Indeed, listeners cite production difficulty as a potential cause of disfluency when they are explicitly asked (Fox Tree, 2002). When disfluency occurs in the context of answering a question, listeners infer that the speaker is having difficulty coming up with the answer (Brennan & Williams, 1995; Krahmer & Swerts, 2005; Swerts & Krahmer, 2005). It is also plausible that listeners might infer the cause of the difficulty, insofar as difficulty is more likely to occur in some situations than others. For example, it is harder to retrieve the names of new objects, compared with objects that have been previously mentioned. Therefore, listeners might plausibly infer that the disfluency occurred because of difficulty referring to an object, which is more likely to be one that was not just mentioned.

Figure 1 presents a hypothetical string of inferences that could underlie the disfluency bias toward new objects that was found by Arnold et al. (2004). The presence of disfluency could lead to an inference about the cause of the disfluency (gray arrows), for example, that the speaker is describing something difficult. This in turn leads to an inference about what the speaker is likely to be referring to (black arrows), for example, something previously unmentioned.

Although such inferences may be plausible, they also might be slow and computationally expensive. This raises the question of whether the same effect might take place through a simpler mechanism. For example, the same disfluency effects may occur as the result of a learned association between disfluency and particular kinds of stimuli. People have likely encountered disfluent references to discourse–new or unfamiliar objects more often than disfluent references to given or to familiar objects. Indeed, an analysis of naturalistic production data revealed a higher rate of disfluency for reference to objects that had not been mentioned in the preceding utterance (Arnold & Tanenhaus, 2007). If listeners are sensitive to these distributional patterns, then, like other probabilistic distributional cues, they could be automatically exploited in real-time language processing. MacDonald and colleagues (e.g., Gennari & MacDonald, 2006; MacDonald, 1999; Race & MacDonald, 2003) have emphasized that production demands can conspire to create many of the distributional patterns that underlie listeners' expectations in real-time comprehension. This raises the

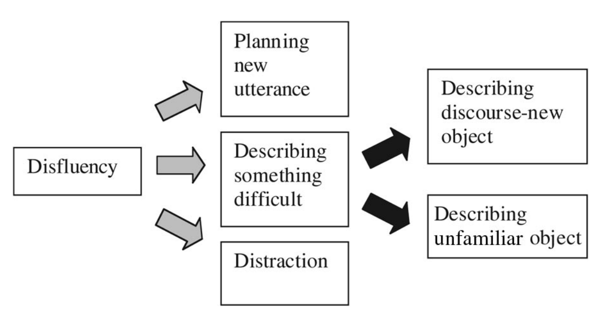

Figure 1. A hypothetical chain of inferences or associations between disfluency and the mental activity of the speaker.

question of whether some or all of the inferences in Figure 1 might be the result of prelearned, probabilistic associations (e.g., the fact that disfluency frequently co-occurs with reference to certain types of objects). People could have learned either a direct association between disfluency and the type of object being described, or a chain of associations like the inferences presented in Figure 1. The inference and association accounts can be viewed as two ends of a continuum, where disfluency effects may involve some degree of both.

In this way, disfluency effects on comprehension provide a natural testing ground for questions about the degree to which language comprehension involves on-line inferences about why a speaker's utterance takes the form it does. We know that listeners can use disfluency to make judgments about speaker knowledge (Brennan & Williams, 1995; Krahmer & Swerts, 2005; Swerts & Krahmer, 2005). The experiments presented here address two further questions: (a) Does disfluency information guide listeners' probabilistic expectations about what the speaker is trying to say, and (b) to what extent do disfluency effects rely on learned associations between disfluency and a type of production event, or by inferences about why the speaker is being disfluent?

The Current Studies

If on-line comprehension is driven by an inference about the cause of disfluency, we would expect disfluency effects to extend to other difficult referring conditions. We begin by examining the comprehension of references to another kind of difficult-to-name object: unfamiliar, complex shapes that do not have a conventional name. Such objects should be difficult to describe and therefore ought to be a reasonable cause of disfluency. If listeners are sensitive to this, we should observe a bias toward unfamiliar objects in a disfluent instruction. We then introduce alternate potential sources of the disfluency to see whether they modulate or eliminate the unfamiliarity preference.

In all experiments, participants viewed visual displays like that in Figure 2, with two familiar and two unfamiliar objects. One of the familiar and one of the unfamiliar objects appeared in one color (e.g., red), and the others appeared in a different color (e.g., black). Participants heard instructions that were either fluent (*Click on the* . . .), or disfluent (*Click on thee uh* . . .), and referred to either the familiar or unfamiliar object. As in Arnold et al.'s (2004) study, the instructions contained a temporary ambiguity: With two objects of each color, the instruction *Click on* {*the/thee uh*} *red* . . . is temporarily ambiguous. Previous work has shown that listeners incrementally identify referents using information from the unfolding input (Eberhard, Spivey-Knowlton, Sedivy, & Tanenhaus, 1995; Tanenhaus, Spivey- Knowlton, Eberhard, & Sedivy, 1996). Thus, the onset of the word *red* should direct listeners' attention to the two red objects. We hypothesized that disfluency should bias the listener to expect reference to the unfamiliar item at this time. That is, given a disfluent instruction, at the onset of the color word, the listener is expected to look preferentially at the unfamiliar object of the named color, but given a fluent instruction, we expect no bias toward either object.

We used two methods to assess listeners' biases as they heard fluent and disfluent instructions. With a gating task, we asked whether disfluent instructions make unfamiliar objects more expected to listeners, when they are explicitly asked what they think the speaker is about to refer to. Listeners heard fragments of instructions (e.g., *Click on thee uh red* . . .) and chose the most likely referent from a display like Figure 2. We then asked whether the expectancy of an unfamiliar referent influenced on-line hypotheses about the unfolding expression, by monitoring listeners' eye movements as they followed intact instructions.

Experiment 1 establishes that, indeed, disfluent fragments make unfamiliar objects more expected, and this influences listeners' on-line hypotheses by biasing attention to unfamiliar objects from the beginning of the color word. Experiments 2 and 3 investigate the flexibility of the unfamiliarity bias, to see whether it disappears when the disfluency could be attributed to some other source of difficulty. In Experiment 2, we presented participants with an alternative explanation for why a particular speaker might often be disfluent when describing familiar objects. In Experiment 3, we introduced intermittent background noise that was timed so that it might plausibly interrupt the speaker at the point where she was preparing her referring expression. Before reporting these experiments, we present two small norming studies: (a) a rating study to establish that the unfamiliar objects are perceived as harder to name than the familiar objects, and (b) a production study to confirm that naming the unfamiliar objects in natural production would result in a higher rate of disfluency than naming the familiar objects.

Norming Studies

Ten native English speakers from the University of Rochester rated 96 familiar and 59 unfamiliar pictures on a scale of 1–7 to indicate how *difficult* (1) or *easy* (7) it would be to name the picture. The study also included an additional 146 unfamiliar pictures that were considered easier to describe, but these were not included in the on-line studies and so are not discussed further.[1] The objects were divided into two lists to avoid fatiguing the participants; 5 participants completed each list. One participant was replaced because he said that some of the unfamiliar objects were engineering symbols and Hebrew letters that were familiar to him. As predicted, the unfamiliar objects were rated as harder to describe ($M = 2.7$) than the familiar objects ($M = 6.8$), $F_1(1, 9) = 584$, $p < .001$; $F_2(1, 153) = 854$, $p < .001$.

A subset of the rated items (24 familiar, 24 unfamiliar) were presented to 3 participants, who named each one in the phrase *Click on the* . . .. They were told that their instructions would be used with later participants who would have to choose the object out of a group. The mean ratings for the unfamiliar items used ranged from 1.2 to 2.8, and all the familiar items had a rating of 7. The naming study also included 24 of the easier unfamiliar items (which were not analyzed), 16 fillers, and 4 practice items, which were pseudorandomized with the experimental stimuli and presented in a single list.

On each trial, the picture to be described was presented with Psyscope (Cohen, MacWhinney, Flatt, & Provost, 1993) and accompanied by a beep. Participants were told to begin speaking at the beep. Responses were transcribed and coded for the number of

[1] The reason for not including these objects was to use the greatest contrast in naming difficulty between categories. However, all unfamiliar objects were rated as more difficult to name than the familiar objects.

words in the target phrase, and the presence of disfluent elements before the onset of the target noun phrase (uh, um, hmm, elongation, the pronunciation of *the* as *thee* [/thiy/], or a laugh). The sound files were analyzed with the phonetics software Praat (Boersma & Weenink, 2003, 2006) for the duration and pitch of the carrier phrase (*Click on the* . . .).

Results revealed a higher rate of disfluent elements for unfamiliar objects (24%) than familiar objects (3%). We also observed longer latencies to begin speaking for unfamiliar objects (5,322 ms) compared with familiar objects (1,255 ms), as well as longer durations for each portion of the carrier phrase for unfamiliar objects (*Click on*: 1,457 vs. 704; *the*: 384 vs. 116). The differences in latency to begin speaking could not be attributed to the length of the target noun phrase: Even descriptions of familiar objects that used five or more words[2] (1,749 ms) were prepared more quickly than descriptions of unfamiliar objects that used only two words (2,201 ms). There was also a relatively high rate of disfluency for both short unfamiliar descriptions (3/8) and long ones (14/63).

Descriptions of unfamiliar objects also had a higher pitch during the initial *Click on the* segment of the instruction (see Table 1). The contrast between unfamiliar and familiar objects emerges in an analysis of variance over items as a main effect of unfamiliarity for maximum pitch,[3] $F_2(1, 46) = 5.11$, $p < .05$, and a near significant effect of mean pitch, $F_2(1, 46) = 3.29$, $p = .076$. The analysis also included word segment (*Click on* vs. *the*), revealing that *Click on* was consistently higher than *the* (Fs > 60 for both pitch average and pitch maximum); there was no interaction between familiarity and word segment (Fs < 1). Participant analyses were not performed because of the small number of speakers.

The results of the naming study show a cluster of characteristics for references to the unfamiliar, more difficult-to-describe objects: more disfluent elements, longer durations for *Click on* and *the*, and higher pitch for both *Click on* and *the*. This pattern reveals the diverse nature of disfluency (for related findings, see Krahmer & Swerts, 2005). Because planning can at least sometimes occur well before a particular word is uttered (e.g., Beattie, 1979; Clark & Wasow, 1998; Ferreira & Swets, 2002), difficulty in production is likely to have multiple effects on a particular region of speech as speakers try to resolve a momentary difficulty. Disfluent items tend to cluster (Clark & Fox Tree, 2002) and correlate with lengthening of nearby words (Bell et al., 2003; Clark & Fox Tree, 2002). Furthermore, as shown by the naming study, more difficult utterances may also involve higher pitch on initial words.

Table 1
Pitches for "Click on" and "the" During Instructions to Click on Familiar and Unfamiliar Objects —Norming Study

Variable	Unfamiliar object instructions	Familiar object instructions
Click on segment	Max. pitch = 288 Hz	Max. pitch = 272 Hz
	Avg. pitch = 230 Hz	Avg. pitch = 223 Hz
the segment	Max. pitch = 237 Hz	Max. pitch = 225 Hz
	Avg. pitch = 207 Hz	Avg. pitch = 204 Hz

Experiment 1

Method

Participants (Gating Task)

A total of 24 students from the University of North Carolina at Chapel Hill participated in exchange for course credit. In this and all subsequent experiments, all participants reported themselves as native speakers of English with normal or corrected-to-normal vision.

Participants (Eye-Tracking Task)

A total of 18 people from the community at the University of Rochester participated in Experiment 1b (eye-tracking task) in exchange for pay; 2 were excluded from the analysis because of technical problems ($n = 1$) and for speaking a non-American dialect of English ($n = 1$).

Design and Materials (Eye-Tracking Task)

On each trial, participants viewed a set of four pictures on a computer screen (see Figure 2). Each experimental item contained two familiar objects (e.g., ice cream cones) and two unfamiliar objects (e.g., the squiggly shape in Figure 2), where each object type was pictured in two colors. The familiar pictures were drawn from Snodgrass and Vanderwart's (1980) data set, and the unfamiliar pictures were taken from clip-art data sets with black and white drawings. All pictures were modified in Adobe Photoshop to replace the black lines and shading with colored lines and shading. Color pairs for a particular item were chosen so as to be maximally distinguishable, both visually and linguistically (e.g., black with red, or yellow with blue, but not black with brown).

Participants heard an instruction to click on one of the objects that was either fluent (*Click on the* . . .) or disfluent (*Click on thee uh* . . .). Instructions referred to either the familiar or unfamiliar object. Thus, the design was 2 (fluent vs. disfluent) × 2 (unfamiliar vs. familiar). See Table 2 for an example of instructions in the four conditions. The 24 experimental stimuli were randomly presented with 24 filler stimuli. The filler items contained either all familiar or all unfamiliar pictures, to de-emphasize the familiar–unfamiliar contrast. Half of each type of filler was fluent; half contained some sort of disfluency.

The auditory stimuli were recorded by the first author. Each condition was recorded in its entirety. Then the *Click on the/thee uh* [*color*] segment was cross-spliced from one of the fluent conditions into the other fluent condition, and from one disfluent condition into the other disfluent condition. Similarly, the unfamiliar and familiar target phrases were cross-spliced so they were identical across the fluent and disfluent conditions. The cross-spliced segment was taken from each condition equally.

The contrast between disfluent and fluent stimuli was achieved by manipulating several characteristics of the stimuli. The article was either fluent (*thuh*) or disfluent ([*pause*] *thee uh*). The instruc-

[2] For example, *the black arrow pointing up*; *the clock that shows 3:00*.

[3] Estimates of maximum pitch over 400 Hz were probably errors of the estimation algorithm and were therefore capped at 400 Hz.

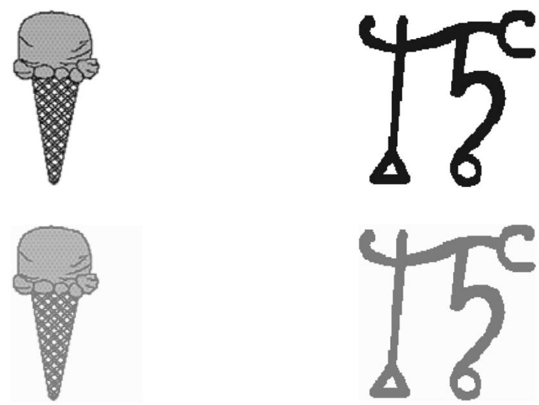

Figure 2. Example unfamiliar and familiar visual stimuli used in all experiments. The top two were shown in black, the bottom two in red.

tions were recorded to sound natural, thus following the tendency for disfluent elements like *thee* and *uh* to co-occur with prosodic characteristics like those observed in the naming study: lengthening of surrounding words and higher pitch. Thus, the disfluent *Click on thee uh* [*color*] in our stimuli was longer and had a higher average pitch than the fluent *Click on the* [*color*]. Table 2 shows the prosodic properties of the auditory stimuli.

Design and Materials (Gating Task)

The materials were identical to those used in the eye-tracking task, with the following exceptions. Only half the items were included to reduce fatigue. They only heard fragments of the instructions, which were truncated to form either a short fragment (*Click on*), a medium fragment (*Click on* {*the*/*thee uh*}), or a long fragment (*Click on* {*the*/*thee uh*} *red*). In none of the experimental stimuli did the sound file contain any information about the target phrase; where there was coarticulation at the end of the color word, we truncated early enough to avoid it. Thus, the stimuli for the gating task followed a 3 (short vs. medium vs. long fragment length) × 2 (disfluent vs. fluent) design. The filler stimuli were truncated later in the instruction so that partial information about the target word was available. This ensured that participants had to continue paying attention to the auditory stimuli throughout the experiment.

Apparatus and Procedure

Gating task. Participants viewed displays, such as Figure 2, and heard fragments of the instructions as described above. Participants were asked to make their best guess as to the object the instruction was about to refer to. They indicated their answer by circling one of the four objects on a printed sheet, of which two were familiar and two unfamiliar.

Eye-tracking task. We monitored participants' eye movements with a head-mounted Eyelink I eye-tracker (SensoMotoric Instruments, Boston, MA) as they viewed the display and followed the auditory instructions. After calibration, participants performed three practice trials and were given a chance to ask questions. The visual and auditory stimuli were presented with PsyScope (Cohen et al., 1993) on a Macintosh computer. On each trial, the four visual stimuli appeared with a blue circle in the center. As in Figure 2, the two unfamiliar objects were always on one side, and the two familiar objects on the other side; the two color-matched objects were always horizontal from each other. The blue circle turned red after 500 ms, at which point the participant clicked on the dot to hear the instruction. This encouraged participants to start each trial fixating in the center of the screen. A drift correction event occurred after every 12 trials.

In both the gating and eye-tracking tasks, as well as all other experiments reported here, we followed the procedure of Arnold et al. (2004), telling participants that the instructions had been recorded by another participant in an earlier phrase of the study. This was done to justify the presence of disfluency, which would not be expected in pre-prepared speech, and to encourage participants to react to the disfluency as they would in a natural setting. We explained that the speaker had seen the same display as the participant would see, except that one object had been highlighted to indicate the object that the speaker should refer to. This meant that listeners could plausibly attribute the disfluency to difficulty describing an object but not to problems deciding what to refer to.

In this and all experiments, the task was followed by an oral questionnaire that assessed participants' awareness of the experimental manipulations. The questions started out very general (e.g.,

Table 2
Example Auditory Stimuli and Mean Duration and Pitch for Each Segment—Experiment 1

	Disfluent		Fluent	
	Familiar: *Click on thiy uh red ice cream cone*		Familiar: *Click on the red ice cream cone*	
	Unfamiliar: *Click on [pause] thiy uh red funny squiggly shape . . .*		Unfamiliar: *Click on the red funny squiggly shape . . .*	
Segment	*M* duration[a] (ms)	*M* max pitch (Hz)	*M* duration (ms)	*M* max pitch (Hz)
Click on	886	373	463	328
the/thee uh	1,278	255	186	240
color	599	282	484	284

[a] Each segment is measured from its onset to the onset of the following segment.

what did you think about the experiment?) and got increasingly more specific (e.g., did you notice the speaker was sometimes disfluent, saying um or uh?). We excluded from analysis participants who guessed that the speech was pre-prepared or that the speaker had been told when to sound disfluent, as well any participants who did not believe the cover story. No participants in Experiment 1 fell into this category, but a few were excluded for these reasons in other experiments.

Results

Gating Task

The responses (see Figure 3) were analyzed in terms of the percentage of items on which the participant chose one of the unfamiliar objects, out of all responses for that condition. The results revealed a tendency to choose an unfamiliar object for disfluent stimuli and a familiar object for fluent stimuli, particularly for the two conditions that contained a longer fragment. This pattern emerged as a main effect of disfluency, $F_1(1, 23) = 58$, $p < .001$; $F_2(1, 11) = 68$, $p < .001$, and an interaction between disfluency and fragment length, $F_1(1, 23) = 5.61$, $p < .01$; $F_2(1, 11) = 3.73$, $p < .05$.

Recall that the disfluency manipulation began at the words *Click on*, which were longer and had a higher pitch in the disfluent condition. If these prosodic differences were sufficient to cause an unfamiliarity bias, we would expect to see a difference between conditions emerging with the short gate. Consistent with this, we found numerically more unfamiliar responses in the disfluent condition, but in a planned comparison, the difference failed to reach significance, $F_1(1, 23) = 2.88$, $p = .10$; $F_2(1, 11) = 0.27$, $p = .60$. This suggests that although prosodic correlates of disfluency may contribute to the impression that the speaker is having difficulty, the bias emerges most strongly after hearing further evidence of difficulty, such as the *thee uh* instead of *the*.

Eye-Tracking

In this and all experiments in this article, the eye movement data were analyzed from the onset of the color word. Recall that the color word was the earliest point of disambiguation and, thus, the earliest point at which we would expect to see looks toward any of the objects. Saccades were grouped with the following fixation to form a *look* at an object. Saccades are ballistic and therefore the onset of the saccade represents the decision to begin looking at that object, and the decision to continue looking at it lasts until the onset of the following saccade. We analyzed looks to the unfamiliar and familiar objects that matched the color word in the input. If disfluency creates a bias toward unfamiliar objects, we should see more looks to the unfamiliar color-matched object than the familiar color-matched object in the disfluent condition but not the fluent condition.

The results are presented in terms of the proportion of looks to the unfamiliar color-matched object, out of both color-matched objects (see Figure 4). That is, the lines represent the relative preference of looking at the unfamiliar color-matched object, starting at the onset of the color word. These results show a clear difference between fluent and disfluent conditions. When the instruction was disfluent (thick line), listeners looked at the unfamiliar object more than twice as often as the familiar object. This preference emerged shortly after the onset of the color word and extended well after the onset of the object description. By contrast, in the fluent condition (thin line), looks to the unfamiliar object hover around the 50% line, revealing an equal tendency to look at the unfamiliar and familiar color-matched objects.

We tested the robustness of this pattern by analyzing the average absolute looks (not proportion looks) to each object from 200 to 1,000 ms after the onset of the color word. The first 200 ms are not included in the analysis because it typically takes around 200 ms to program and launch an eye movement. There were significant interactions between disfluency (disfluent vs. fluent) and the target

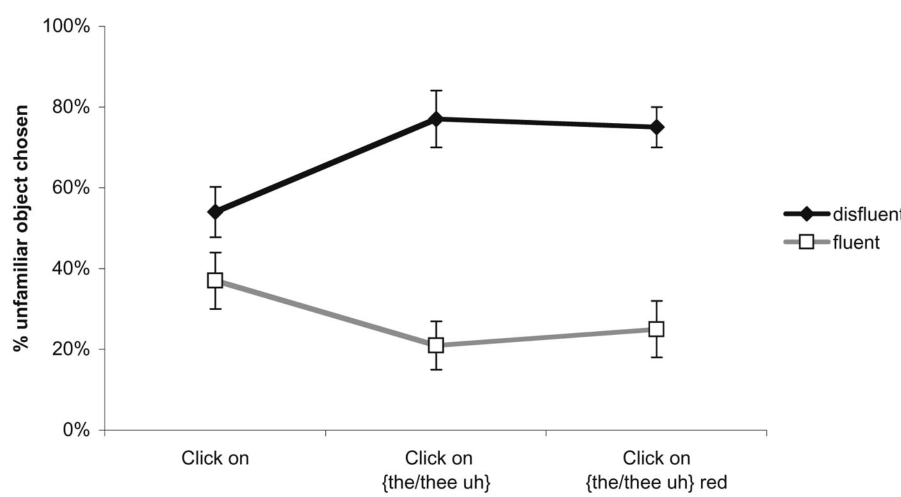

Figure 3. Results for Experiment 1 (gating task): mean percentage of unfamiliar objects chosen in a gating task following short (*Click on . . .*), medium (*Click on the/thee uh . . .*), or long (*Click on the/thee uh red . . .*) fragments. Each item had four objects, of which two were unfamiliar and two familiar. Error bars represent the standard error of the mean.

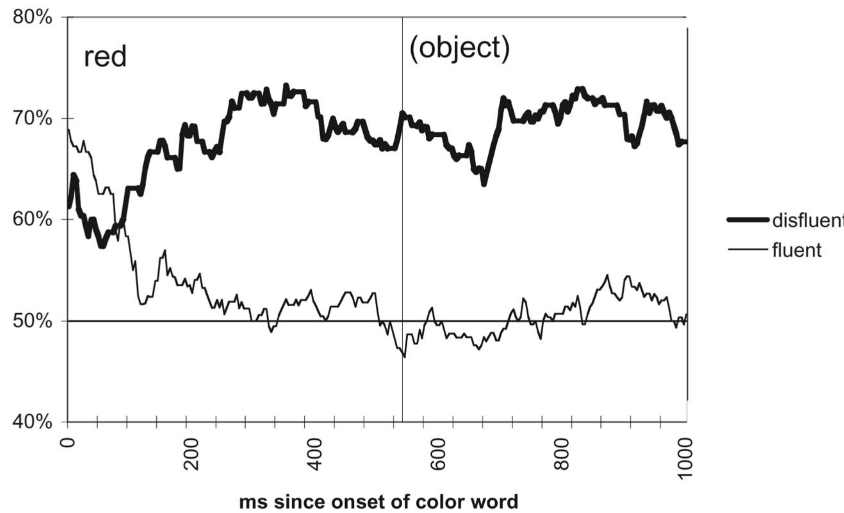

Figure 4. Results for Experiment 1 (eye-tracking task): Proportion looks to the unfamiliar color-matched object, from 0 to 1,000 ms after the onset of the color word. Proportion looks are calculated out of all looks to both familiar and unfamiliar color-matched objects. A horizontal line at 50% represents equal looks to familiar and unfamiliar color-matched objects. The vertical line represents the average onset of the target object description.

object (unfamiliar vs. familiar), reflecting the bias toward the target object for the disfluent/unfamiliar condition, and the bias toward the competitor object for the disfluent/familiar condition. Table 3 presents the condition means for all eye-tracking experiments, and Table 4 presents the statistics for the critical interactions for all eye-tracking experiments. Note that here and throughout the article, we only report the statistical effects of interest for the sake of readability. Some experiments also show main effects of disfluency and unfamiliarity or both, but these are considered secondary to the interaction. For example, this experiment had a main effect of target word (unfamiliar vs. familiar), but this was driven entirely by the increased looks to the unfamiliar object in the disfluent condition.

Discussion

The results from Experiment 1 establish that disfluency creates a bias toward interpreting an initially ambiguous referential expression as referring to an unfamiliar object. Given that unfamiliar objects are harder to describe than familiar objects, and that disfluency typically occurs when the speaker has some kind of production difficulty, it is likely that the perception of disfluency leads listeners to expect reference to difficult-to-name objects, such as unfamiliar ones. Fluent instructions, by contrast, do not lead to any particular bias. This is not surprising because speakers do not always experience difficulty referring to unfamiliar objects, so a fluent instruction does not provide good information about the type of object being referred to. An increase in looks to the unfamiliar object emerged 200 ms after the onset of the color word, which is the same time that listeners begin to exploit information about the unfolding lexical input (e.g., Allopenna et al., 1998). This shows that the expectation of reference to unfamiliar objects affects listeners' on-line hypotheses about what the speaker is referring to, just as was the case with the expectation of reference to discourse–new objects (Arnold et al., 2004).

The unfamiliar bias with disfluent stimuli complements the previously established bias toward discourse–new objects compared with given ones (Arnold et al., 2004). It seems likely that these effects are related because both unmentioned and unfamiliar objects are harder to describe than given and familiar objects. Thus, we interpret the current results as probable evidence that disfluency information guides listeners because of its link with difficulty in naming objects.

However, we must consider an alternative interpretation of these results. Perhaps disfluency instead creates an expectation of reference to something with a long name, which results in a bias toward objects that are likely to have lengthy descriptions. This would be consistent with the fact that disfluency tends to occur more at the onset of complex than simple constituents (Clark & Wasow, 1998; see also Oviatt, 1995). Although we cannot rule out this account on the basis of the current data, there are several reasons to think that the unfamiliarity bias stems from the fact that the object is difficult to name rather than its tendency to require long descriptions per se. First, the length explanation would not allow for a parsimonious account of the unfamiliarity bias and the newness bias because both given and new stimuli in that experiment were felicitously described with simple noun phrases (e.g., *the candle*; *the camel*). Second, our naming study revealed that references to familiar objects tended to be more fluent than both short and long references to unfamiliar objects (see Arnold & Tanenhaus, 2007, for a similar finding for given and new references). Thus, production difficulty seems to result from the properties of the stimulus rather than length per se—despite the fact that givenness and length are often correlated (Arnold, Wasow, Losongco, & Ginstrom, 2000). Third, it would be computationally difficult for listeners to compute the unfamiliarity bias on the basis

Table 3
Results for all Eye-Tracking Experiments

Variable	Disfluent/familiar target	Disfluent/unfamiliar target	Fluent/familiar target	Fluent/unfamiliar target
		Experiment 1		
Eye-tracking				
% target looks[a]	16 (2.7)	35 (4.7)	32 (3.2)	31 (3.3)
% competitor looks	33 (3.9)	14 (2.9)	32 (3.5)	29 (3.5)
% unfamiliar looks	69		51	
% familiar looks	31		49	
		Experiment 2		
Typical				
% target looks	30 (2.2)	35 (2.1)	35 (2.6)	34 (1.7)
% competitor looks	37 (2.4)	26 (1.8)	34 (2.7)	38 (1.8)
% unfamiliar looks	56		48	
% familiar looks	44		52	
Agnosia				
% target looks	29 (3.1)	32 (2.4)	35 (2.4)	35 (2.4)
% competitor looks	30 (2.0)	32 (3.1)	36 (2.0)	28 (2.0)
% unfamiliar looks	50		53	
% familiar looks	50		47	
		Experiment 3		
No distraction				
% target looks	29 (2.0)	41 (2.1)	34 (1.8)	37 (2.1)
% competitor looks	38 (2.2)	26 (1.7)	32 (1.9)	34 (1.7)
% unfamiliar looks	59		50	
% familiar looks	41		50	
Distraction				
% target looks	29 (1.8)	38 (2.1)	32 (1.9)	34 (1.8)
% competitor looks	37 (1.8)	24 (1.6)	32 (2.0)	33 (1.6)
% unfamiliar looks	59		50	
% familiar looks	41		50	

Note. Standard errors are in parentheses.
[a] Percentage target and competitor looks report the average looks to the color-matched objects from 200 to 1,000 ms after the onset of the color word. Percentage unfamiliar and familiar looks report the percentage looks to either unfamiliar or familiar objects out of all target and competitor looks.

of the likely length of the upcoming description. This would require listeners to generate possible descriptions of the objects in the scene; yet listeners do not tend to do this (Dahan & Tanenhaus, 2005; Dahan, Tanenhaus, & Salverda, 2007).

Even if the unfamiliarity bias does stem from a link between disfluency and long constituents, it would be consistent with our more general point that disfluency results in an on-line bias toward certain types of referents. For rhetorical purposes, the rest of this article uses the naming difficulty explanation because we see this as the more plausible one. However, the length explanation could conceivably be substituted (e.g., in Figure 1, the link to "Describing something difficult" could be replaced with a link to "Describing something long"). Future studies are needed to distinguish these two possibilities.

Table 4
Inferential Statistics for Critical Factors for all Experiments

Dependent measure	Factor	F_1 *df*'s	F_1	F_1 *p*	F_2 *df*s	F_2	F_2 *p*
		Experiment 1					
Target looks	Disfluency × Unfamiliarity	1, 15	5.3	<.05	1, 23	9.20	<.01
Competitor looks	Disfluency × Unfamiliarity	1, 15	13.5	<.005	1, 23	13.3	<.005
		Experiment 2					
Target looks	Disfluency × Unfamiliarity × Speaker	1, 46[a]	0.317[a]	*ns*[a]	1, 23[a]	0.481[a]	*ns*[a]
Competitor looks	Disfluency × Unfamiliarity × Speaker	1, 46	16.81	<.001	1, 23	16.31	<.005
		Experiment 3					
Target looks	Disfluency × Unfamiliarity	1, 70	9.19	<.005	1, 23	10.28	<.005
Competitor looks	Disfluency × Unfamiliarity	1, 70	30.50	<.001	1, 23	28.7	<.001

[a] Indicates that effects failed to reach significance.

Association or Inference?

Under either the naming difficulty or length explanations, we still are faced with the question of the extent to which the unfamiliarity bias with disfluent stimuli involves a high-level explicit inference, or whether it can be fully explained by an automatic prelearned association between disfluency and certain referent types. Just as for the discourse–new bias, the unfamiliarity bias could be the result of an inference about the cause of the speaker's production difficulty. That is, listeners may hear disfluency and make a rapid, on-line inference that (a) the speaker is having trouble, (b) the difficulty is with naming, and (c) the most probable source of difficulty is either newness or unfamiliarity (see Figure 1). On the other hand, production data show that there are regularities in the occurrence of disfluent speech, such that it tends to co-occur with descriptions of both unfamiliar and discourse–new objects more than with descriptions of familiar and given objects. These associations could be utilized in a relatively automatic way, yielding the same result in Experiment 1, without requiring a potentially more computationally complex utterance-specific inference. One way to assess the degree of automaticity is to ask whether the unfamiliarity effect can be modulated by situational variables, especially where the relevance of those situational variables is available only through inferences about the speaker's cognitive state.

Experiments 2 and 3 investigate the flexibility of the disfluency effect. If listeners have learned rigid associations between disfluency and both discourse–new and unfamiliar objects, we would expect to see these biases in all situations, for all speakers. By contrast, if listeners make inferences about the cause of the disfluency dynamically, then they may demonstrate sensitivity to the difficulty of naming an object for particular situations or particular speakers. For example, when a linguist is speaking, a spectrogram may be considered easy to name but not an arthroscope, whereas the opposite may be assumed for a doctor. In this case, one would expect that the unfamiliarity bias would only occur if the listener thinks the speaker should have more difficulty naming unfamiliar objects than familiar objects. Experiment 2 investigates whether disfluency causes biases toward different kinds of objects for different speakers by creating a situation in which the listener is given a plausible reason to expect the speaker will be disfluent even when she describes a familiar object. Experiment 3 introduces intermittent noise that could have distracted the speaker and thus caused her to become disfluent when generating a specific referring expression.

Experiment 2

For most people, unfamiliar objects with no conventional name are considered more difficult to describe than familiar objects. Instead of just recalling a frequent and familiar label, the speaker must plan and produce a description of the object that would be understandable to the listener. However, familiar objects may not be as easy to label for individuals who suffer from object agnosia. People with this condition cannot recognize simple objects and often must resort to describing them (e.g., Farah, 1990). Thus, it is plausible that such a speaker might find familiar objects just as hard to describe as unfamiliar objects and, therefore, be equally disfluent for unfamiliar and familiar objects.

If the listener believes that the speaker has object agnosia, how will this change the interpretation of disfluency? Unfamiliar and familiar objects may be perceived as equally difficult for an object agnosic speaker to name, in which case the bias toward unfamiliar objects might be reduced or eliminated. Note that this experiment does not actually measure the characteristic speech of individuals with object agnosia, which is likely to vary considerably. Instead, we are interested in what the *listener* believes about the speaker.

Method

Participants: Gating Task

A total of 37 students from the University of North Carolina at Chapel Hill participated in the gating task as partial fulfillment of a requirement for an introductory psychology class. Data from 5 students were excluded—2 because they were nonnative speakers of English and 3 because of technical problems. This left 32 participants in the analysis.

Participants: Eye-Tracking Task

A total of 58 students participated in the eye-tracking task, also for course credit. Data from 10 students were excluded—5 because they said something in their answers to the postexperiment questionnaire to suggest that they did not believe the cover story, 4 because of track loss, and 1 because he fell asleep during the experiment. This left 48 participants in the analysis.

Design and Materials

The same basic design was used as in Experiment 1, except half the participants were told that the speaker had object agnosia.

The auditory stimuli were very similar to those of Experiment 1, in that the disfluency manipulation included both prosodic differences and the contrast between *the* and *thee uh*. In addition, the following changes were made, and all stimuli were re-recorded. Some of the unfamiliar descriptions in Experiment 1 made reference to familiar objects, for example, *bubbles shaped like a crab*. Such descriptions were less plausible for a speaker with object agnosia, so they were changed to be more physically descriptive (for both typical and object agnosic speaker conditions), for example, *bunch of little circles clumped together*. In addition, the pairings between unfamiliar and familiar objects were shuffled, so that names/descriptions within a given item began with the same or similar phonemes. This facilitated the cross-splicing of the segment *Click on the* [*color*] onto the unfamiliar and familiar descriptions. Two familiar objects were replaced to make a better match with the unfamiliar object initial phoneme. Apart from these two objects, the visual stimuli were the same as in Experiment 1. All instructions were recorded in their entirety and then cross-spliced so that for each fluent or disfluent condition of each item, the same *Click on* {*the/thee uh*} was used.

To support the story that the speaker had object agnosia, the fillers that referred to familiar objects used descriptions rather than conventional labels. For example, the description of a ball was *black, I don't know, um, a circle I guess* (see Grodner & Sedivy, 2007, for a similar design).

Procedure and Apparatus

Gating task. We conducted a gating task using the same basic procedure as the gating task in Experiment 1, except that all items were included. Also, because Experiment 1 found no difference between the middle and long fragments, we only included short (*Click on* . . .) and long fragments (*Click on the/thee uh* [*color*]).

Eye-tracking. Eye movements were monitored with the Eyelink II eye-tracker (SR Research, Osgoode, Ontario, Canada). The visual and auditory stimuli were presented on a computer running the ExBuilder software (Longhurst, 2006). The procedure was similar to that used in Experiment 1, except that every trial was followed by a screen with a single dot in the middle. Participants had to fixate and click on the dot to move on to the next item. This fixation was used by the eye-tracker to perform a drift correction, to maintain good calibration throughout the experiment. As soon as the participants clicked on the dot, the visual stimuli were presented, and the sound file began to play simultaneously. There were about 200 ms of silence at the beginning of each sound file before the spoken instruction began.

The only substantive difference in procedure was the cover story about how the stimuli had been recorded. All participants were told that we were studying communication with both typical and disordered populations. Half the participants (typical speaker condition) were told they would be hearing instructions by a speaker with no disability. The other half were told that the speaker had object agnosia. They read a sheet with detailed information about object agnosia, including a passage from Sacks (1985) in which a man with object agnosia described a glove as "A continuous surface . . . infolded in itself. It appears to have . . . five outpouchings, if that is the word" (p. 14). We stressed that speakers with object agnosia sometimes cannot recognize objects but can be normal in other respects, such as intelligence and language abilities.

Results

Both the gating and on-line experiments revealed a difference between the agnosia and typical speaker conditions. With a typical speaker, the unfamiliarity bias was again observed. With an object agnosic speaker, the bias was significantly reduced (gating task) or absent (on-line task).

Gating Task

The gating task (see Figure 5) revealed a higher rate of choosing the unfamiliar object following a longer disfluent fragment than a longer fluent fragment. However, this pattern was attenuated in the object agnosic speaker condition (dotted lines), resulting in a three-way interaction between disfluency, fragment length, and speaker condition, $F_1(1, 30) = 5.06$, $p < .05$; $F_2(1, 23) = 18.25$, $p < .001$.

Eye-Tracking

The on-line experiment also revealed a contrast between typical and agnosic speakers (see Table 3 and Figure 6). In the typical-speaker condition (black lines), the results replicated the pattern found in Experiment 1. For disfluent instructions (thick line), there was a bias toward the unfamiliar object starting at the onset of the color word. By contrast, the fluent instructions (thin line) displayed no such bias. However, in the agnosic speaker condition (gray lines), the unfamiliarity bias with disfluent stimuli is completely absent—looks to the unfamiliar color-matched object are around the 50% line. There is even a small trend toward the opposite pattern, where fluent stimuli (thin line) have numerically more looks to the unfamiliar color-matched object than the familiar color-matched object. It may be that agnosic speakers are expected to have an easier time with unfamiliar stimuli; perhaps unfamiliar objects are perceived as easier to describe physically than familiar objects.

These effects were evaluated in an analysis of variance with looks to competitor objects as the dependent measure, which revealed a significant three-way interaction between disfluency, referent (unfamiliar vs. familiar), and agnosia (see Table 4). This interaction did not reach significance with target looks as the dependent measure. However, analyses that collapsed over referent type revealed robust interactions between disfluency and agnosia for both unfamiliar looks and familiar looks (see Table 5).

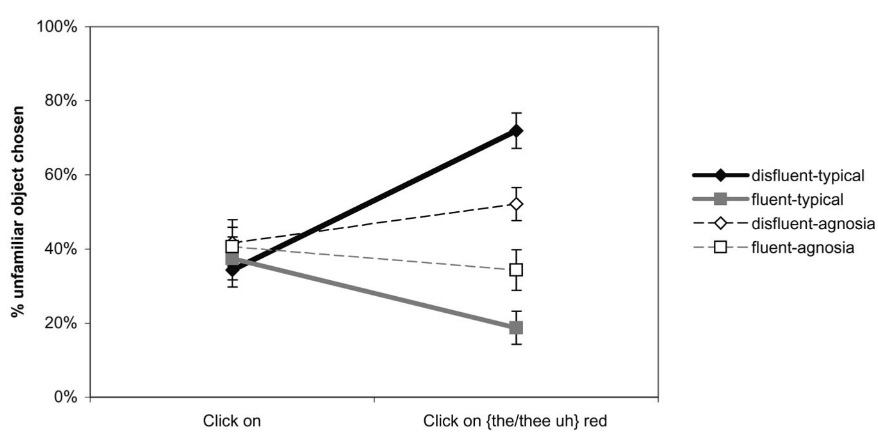

Figure 5. Results for Experiment 2 (gating task): mean percentage of unfamiliar objects chosen in a gating task following short (*Click on* . . .) or long (*Click on the/thee uh red* . . .) fragments. Each item had four objects, of which two were unfamiliar and two familiar. Error bars represent the standard error of the mean.

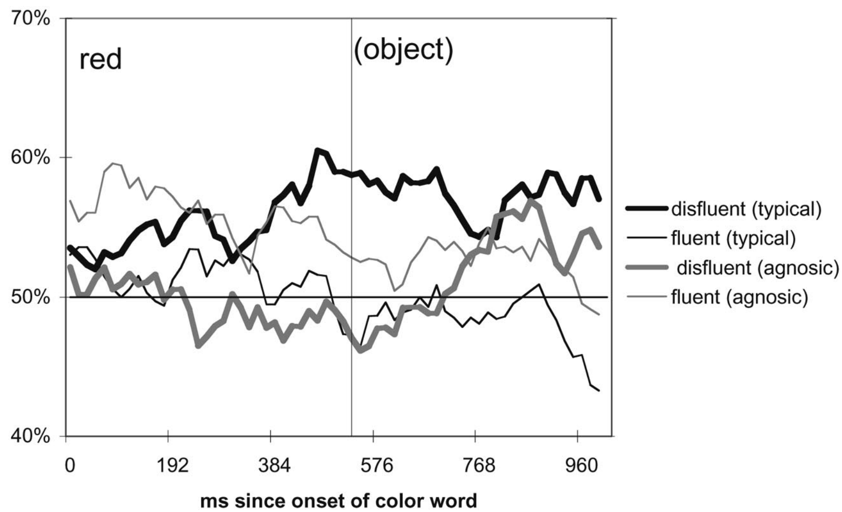

Figure 6. Results for Experiment 2 (eye-tracking task). Typical speaker condition (black lines) versus agnosic speaker condition (gray lines): proportion looks to the unfamiliar color-matched object, from 0 to 1,000 ms after the onset of the color word. Proportion looks are calculated out of all looks to both familiar and unfamiliar color-matched objects. A horizontal line at 50% represents equal looks to familiar and unfamiliar color-matched objects. The vertical line represents the average onset of the target object description.

For technical reasons, we analyzed the data from this and the following experiment using a 16-ms window for each data point.

Discussion

These results reveal that the disfluent bias toward particular kinds of hard-to-describe objects is not reflexive. Instead, listeners can modulate their assumptions about which objects are hard for a particular speaker and use these modified assumptions to guide on-line expectations about referring expressions. When listening to a typical adult speaking, unfamiliar objects are considered the most likely referent following a disfluent instruction. When listening to a speaker with object agnosia, participants look equally at both familiar and unfamiliar objects, for both fluent and disfluent stimuli.

We interpret this as evidence that listeners are sensitive to the fact that typical speakers are more likely to have difficulty with unfamiliar than familiar objects, but that speakers with object agnosia are likely to have difficulty with all objects. As such, this experiment provides evidence that the disfluency bias involves drawing an inference about the abilities of a particular speaker. This suggests that the unfamiliarity effect is not limited to a simple association between unfamiliar hard-to-name objects and disfluent speech. In terms of the hypothetical inferences in Figure 1, this establishes that the black arrows involve an inference rather than a simple association.

This finding is consistent with a parsimonious account of the two established disfluency biases, where disfluency increases the expectancy of both discourse–new (Arnold et al., 2004) and unfamiliar objects. If these had been reflexive associations, listeners would need to separately calculate the co-occurrence of disfluency with each kind of difficult-to-describe object. By contrast, if listeners infer the most likely cause of disfluency in a particular situation, one would expect biases toward many different kinds of difficult-to-describe objects, including new objects, unfamiliar objects, and—for agnosic speakers—simple familiar objects as well.

Although it is clear that comprehension involves some amount of inferencing, the data from Experiments 1 and 2 do not establish the grain at which listeners make inferences about the source of

Table 5
Results for Experiment 2 (Eye-Tracking Task): Average Looks to the Familiar and Unfamiliar Color-Matched objects, Collapsing Across Target Condition (Unfamiliar vs. Familiar), and Inferential Statistics for the Disfluency × Agnosia Interaction

Variable	Familiar looks: Disfluent	Familiar looks: Fluent	Unfamiliar looks: Disfluent	Unfamiliar looks: Fluent
Typical speaker (%)	28 (1.2)	36 (1.4)	36 (1.7)	34 (1.2)
Agnosic speaker (%)	31 (3.5)	31 (1.7)	31 (1.5)	36 (1.7)
Disfluency × Speaker	F_1 (1, 46) = 4.64, $p < .05$		F_1 (1, 46) = 7.7, $p < .01$	
	F_2 (1, 23) = 8.98, $p < .01$		F_2 (1, 23) = 5.06, $p < .05$	

Note. Standard errors are in parentheses.

disfluency. Listeners in Experiments 1 and 2 heard instructions from only one kind of speaker and from only one kind of situation (i.e., one with unfamiliar and familiar objects). It is therefore possible that participants calculated the ease of referring to familiar and unfamiliar objects once for the particular speaker in that particular situation, rather than repeatedly making that judgment on-line as the speech unfolded. It is also not clear whether the link between disfluency and the act of describing something difficult is limited to cases in which there is no competing cause for disfluency, such as the presence of a distracting noise.

In Experiment 3, we examined what happens when naming difficulty is not the only or even the most plausible cause of disfluency. For example, speakers may experience difficulty for reasons not related to the planning and production process itself (e.g., if they are distracted by a noise). If listeners make local inferences about the cause of disfluency (e.g., distraction vs. naming difficulty), then evidence of distraction during an utterance might eliminate the expectation that the speaker is referring to something difficult to describe.

Experiment 3

Experiment 3 investigated whether the unfamiliarity bias would disappear in the presence of plausible alternative cause for disfluency: speaker distraction. Stimuli like those in Experiments 1 and 2 were compared with versions that contained audible evidence of speaker distraction—beeps and construction noises.[4] We used the same cover story, telling participants that the instructions were recorded in an earlier phase of the study, where the speaker gave instructions to a partner. We modified the story enough to indicate that the noises were likely to distract the speaker as well.

We manipulated distraction in two ways. In Experiment 3a, we added beeps to half the stimuli. We told participants that the speaker had to do the secondary task of pressing a button when they heard their computer beep—a task that should be perceived as a source of distraction. In Experiment 3b, we cross-spliced construction noises into half the stimuli. In this case, we told participants that the instructions were recorded while the participants heard prerecorded construction noises, in an effort to study communication in noisy settings. The construction noise manipulation builds on the fact that everyone has experienced conversation in noisy settings, as occurs when there is construction nearby. The distracting effect of construction noises needs no explanation and thus provides an ecological test of speaker distraction. The two manipulations produced the same results and so are presented together.

If listeners perceived the beeps and construction noises as evidence of speaker distraction, they may have attributed the disfluency to the distraction and not to the difficulty of describing an unfamiliar word. Thus, a reduction in the unfamiliarity bias would indicate a highly flexible and on-line attribution about the cause of disfluency. For stimuli that had no distraction, the established unfamiliarity bias should be replicated. Note that the distracting noises occurred equally for fluent and disfluent stimuli and fillers, so that any observed distraction effects could not be the effect of an experiment-specific learned contingency but rather would reflect an inference based on the knowledge of the usual effects of distraction.

Experiment 3 additionally provided an important test of the characteristics of disfluency that led to the bias observed in Experiments 1 and 2. The instructions in those experiments were recorded to achieve a naturalistic impression of disfluency, which meant that the pronunciation of *Click on* was longer and had a higher pitch in the disfluent conditions, compared with the fluent ones. Given findings that accenting can impact speech comprehension (cf. Dahan et al., 2002), it was important to establish that the disfluency results were not merely a side effect of the lengthening and higher pitch on *Click on*. Experiment 3 thus used stimuli that cross-spliced *Click on* across experimental conditions.

Method

Participants: Gating Study

Sixteen students from the University of North Carolina at Chapel Hill participated in Experiment 3a (beeps), and 16 participated in Experiment 3b (construction noises) in exchange for course credit.

Participants: Eye-Tracking Study

A total of 46 students from the University of North Carolina at Chapel Hill participated in Experiment 3a (beeps) as partial fulfillment of a requirement for an introductory psychology class. Data from 6 participants were not included in the analysis—2 because of a programming error, 1 because of interruptions during the experiment, and 3 additional participants were chosen at random to be excluded because the list rotation was not complete. Thirty-nine different participants had also participated in an initial version of the experiment that was discovered to have an error in the sound files and was thus not analyzed.

Thirty-six people from the community at the University of Rochester participated in exchange for pay in Experiment 3b (construction noises). Data from 4 participants were not analyzed—1 because of excessive track loss and 3 because they said afterward that they thought the construction noises had been spliced into the instructions after they had been produced.

Design and Materials

The same objects and descriptions were used as for Experiments 1 and 2. As for Experiment 2, the pairings between unfamiliar and familiar items were changed to match the first phoneme of the description for the unfamiliar and familiar objects in a particular item to facilitate cross-splicing. In some cases the description of the unfamiliar object was changed slightly or the familiar object was replaced with a different object to create a better match between initial phonemes.

The auditory stimuli were nearly identical to those in Experiments 1 and 2 in terms of the words used. However, all instructions were re-recorded to limit the disfluency manipulation to the region beginning at *the/thee uh*. The *Click on* portion was initially recorded to be very similar across the disfluent and fluent conditions of each item. Otherwise the stimuli had the same characteristics as

[4] Another version of Experiment 3a (beeps) was also conducted, with essentially the same results.

in Experiments 1 and 2. Each item was cross-spliced so that the *Click on* segment was identical across all versions of each item, *the* was the same for both fluent conditions, and *thee uh* was the same for both disfluent conditions. The pause between *Click on* and *thee uh* in all disfluent conditions was later lengthened by splicing in 230 ms of background noise to provide adequate time for the disfluency to sound like it was in reaction to the distraction noise. This served as the basic sound-file set.

The basic set was then modified to create the distraction conditions. Experiment 3a used a beep, and Experiment 3b used a loud clang that sounded like a metal object being struck or dropped. The onset of the beep fell at the end of *Click* or beginning of *on*, and the onset of the clang (which was longer) fell earlier, during *Click*. We hypothesized that the disfluency was most likely to sound like a reaction to the distraction if the disfluency occurred immediately after the distracting noise. The stimuli therefore started out with a fluent-sounding *Click on*, during which time a distracting noise was heard on half the stimuli. In the disfluent conditions, this was followed by a 230-ms pause and the disfluent *thee uh*. To support the impression that the disfluency was a reaction to distraction, the first author recorded the instructions with distraction in mind during the disfluent conditions.

Half of the fillers had beeps and construction noises (of various types) at different places in the instruction. Distraction noises were distributed equally across the fillers with familiar and unfamiliar targets. For Experiment 3b, half of the experimental stimuli had another construction noise (a drill) spliced into the later part of the instruction after the critical region.

Procedure and Apparatus

Both gating and on-line experiments were conducted, following the same procedures as in Experiments 1 and 2. We told participants that the instructions were recorded by a naïve speaker who either had to press a button at the beep (Experiment 3a) or heard occasional construction noises while giving instructions (Experiment 3b).

Results

Postexperiment Questionnaire Results

The postexperiment questionnaire was used to assess whether the distraction manipulations were successful in their effect on comprehenders. What is at stake here is not whether speakers are usually disfluent in these conditions but rather what listeners perceive when they hear disfluency following a distraction. We analyzed responses to three of the postexperiment questions: (a) *Did the {beeps/construction noises} sound like they might have distracted the speaker?*, (b) . . . *did they distract you?*, and (c) *Did the {beeps/construction noise} sound like they might have caused the disfluency?* The questionnaire was orally administered, and the experimenter wrote down the answers. Table 6 shows the percentage of participants who said that the noises were a possible or probable cause of distraction and disfluency.

The responses revealed that the majority of participants perceived the beeps or construction noises as distracting to the speaker. The fact that less than half reported the noises as distracting to themselves reveals that this perception is a judgment from the speaker's perspective and not just a general assessment about how distracting the noises are. An important finding is that the majority of participants also said that the disfluency occurred at least some of the time as a result of the distraction. In fact, some responses to the third question suggested that participants perceived that disfluency was more likely in the distraction conditions (e.g., *yes—usually she paused when it beeped*, or *Possibly, they did sync-up very well*). This occurred even though disfluency was equally distributed across unfamiliarity and distraction conditions in both experimental stimuli and fillers. It must be noted that this tendency was not universal, and some participants made comments (such as *uh . . . not so much. Just trying to think of what to say, rather than being affected by the noise*). Nonetheless, the general tendency was to see the disfluency as a likely reaction to the distracting noise.

Table 6
Responses to the Postexperiment Questionnaire for Experiment 3: Percentage of Participants Who Said Yes or Possibly/Probably for Each Question

Question	Gating procedure (32 participants)	Eyetracking procedure (76 participants)
Did beeps/construction noises distract the speaker?	93	82
Did beeps/construction noises distract you?	48	43
Did the beeps/construction noises cause disfluency?	75	71

Note. Values indicate percentages.

Gating Task

The gating task (see Figure 7) revealed that participants had a preference for choosing one of the unfamiliar objects following a disfluent fragment, whether or not there was a distracting noise present. This effect only occurred with the longer fragments. Recall that in this experiment the short fragment had been cross-spliced and was identical across conditions, thus we would not expect a difference for the short fragments. The unfamiliarity bias with disfluent stimuli emerged in analyses of variance as an interaction between disfluency and fragment length, $F_1(1, 30) = 27.51$, $p < .001$; $F_2(1, 23) = 15.55$, $p < .005$.

There were no effects or interactions with distraction. Neither were there any effects or interactions with experiment (Experiment 3a vs. Experiment 3b), except for a trend toward choosing the unfamiliar object more often in Experiment 3b (with the construction manipulation), which was significant in the items analysis only, $F_1(1, 30) = 1.07$, $p = .31$; $F_2(1, 23) = 4.70$, $p < .05$. It is important to note that the same unfamiliarity preference for disfluent instructions occurred with both the beep and construction manipulations.

Eye-Tracking

The unfamiliarity bias with disfluent stimuli also guided on-line hypotheses about the referent of the unfolding input (see Figure 8). Whether the distraction was present (gray lines) or not (black lines), the disfluent instructions (thick lines) resulted in more looks to the

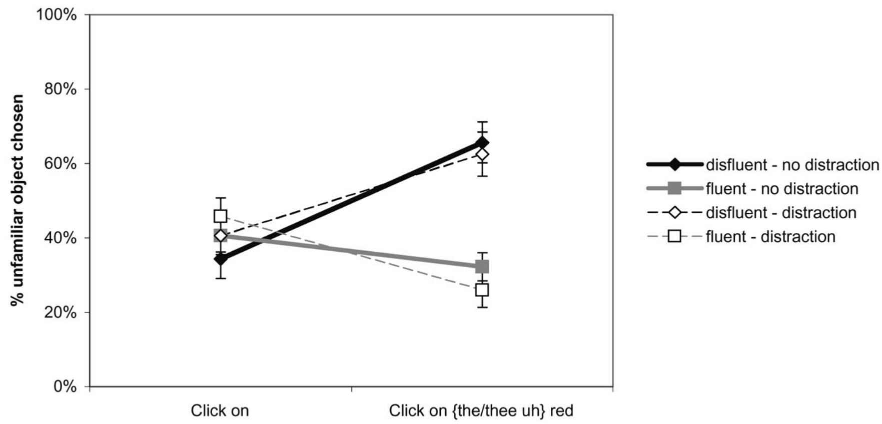

Figure 7. Results for Experiment 3 (gating task): mean percentage of unfamiliar objects chosen in a gating task following short (*Click on . . .*) or long (*Click on the/thee uh red . . .*) fragments. Each item had four objects, of which two were unfamiliar and two familiar. Error bars represent the standard error of the mean.

unfamiliar objects than the familiar objects, shown here by the tendency for about 60% looks to the unfamiliar object during the critical region. By comparison, the fluent instructions (thin lines) led to about equal looks to both unfamiliar and familiar objects, so that the proportion of looks to the unfamiliar object in Figure 8 falls around the 50% line. This was reflected in a significant interaction between disfluency and unfamiliarity, with both competitor looks and target looks as the dependent variable (see Table 4). There were no effects or interactions with distraction.

The interaction was robust even if we considered the subset of participants (71% of the total) who indicated on the postexperiment questionnaire that they thought the disfluency could have been caused by the distraction: target looks, $F_1(1, 50) = 13$; competitor looks, $F_1(1, 50) = 29$ (item analyses were not possible because of missing cells). Even for this set of participants, there were no effects or interactions with distraction.

Discussion

For both beeps and construction noises, the bias toward unfamiliar objects with disfluent instructions was robust against the presence of distraction. This occurred in both the gating task and the on-line experiment, despite the fact that listeners tended to think that the distraction was a likely cause of disfluency when asked explicitly.

It is notable that the disfluency effect occurred in all experiments, despite the fact that disfluent utterances referred equally as often to familiar and unfamiliar objects. This means that

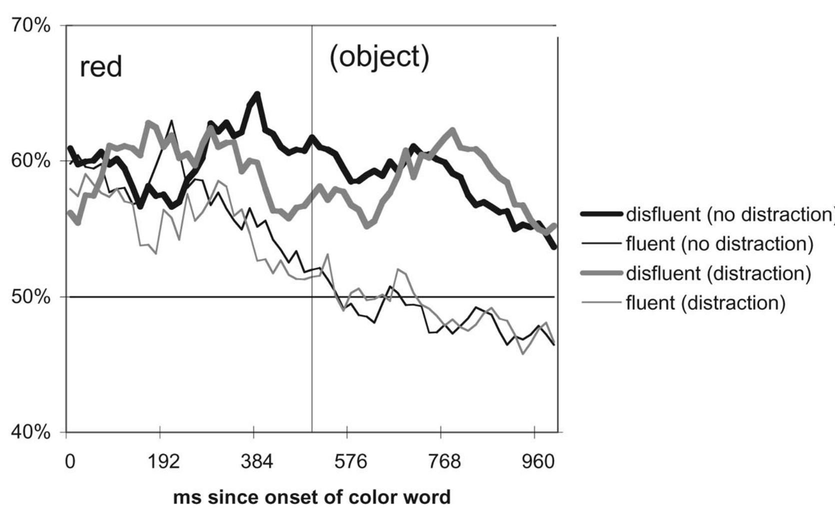

Figure 8. Results for Experiment 3 (eye-tracking task). No-distraction condition (black lines) versus distraction condition (gray lines): proportion looks to the unfamiliar color-matched object, from 0 to 1,000 ms after the onset of the color word. Proportion looks are calculated out of all looks to both familiar and unfamiliar color-matched objects. A horizontal line at 50% represents equal looks to familiar and unfamiliar color-matched objects. The vertical line represents the average onset of the target object description.

listeners could not simply have learned the co-occurrence pattern in the experimental input. By contrast, we observed no effect of the distracting noises, which also occurred equally for disfluent/fluent utterances and familiar/unfamiliar referents. One question, which we leave to future research, is whether listeners require a contingency between distraction and disfluency to draw the inference of a causal relationship between the two. If so, this would differ from the bias toward unfamiliar objects with disfluent stimuli, which occurs in the absense of such contingency.

These findings thus suggest that there are limits to the kinds of inferences that listeners typically draw, but there are several potential interpretations of what this limitation is. One possibility is that listeners do attribute the disfluency to distraction, but that attributions to one cause do not preclude the simultaneous attribution of disfluency to the difficulty associated with describing difficult objects. That is, perhaps the attribution system is generous, in essence "covering all bases." In terms of Figure 1, this would suggest that the inferences represented by gray arrows are not mutually exclusive. Another possibility is that listeners can make the inference that disfluency results from cognitive load, but distraction occurs too rapidly to infer that it is the cause of cognitive load. However, this possibility is unlikely because distraction did nothing to modulate the off-line judgments in the gating task either. Finally, it may be that listeners are more likely to attribute disfluency to production–internal problems, such as naming difficulty, than to attribute it to outside distraction. This might be the case if planning difficulty is a more frequent cause for disfluency.

Although there are several potential explanations for why distraction did not modulate the unfamiliarity bias, Experiment 3 provides strong evidence for the existence of the unfamiliarity bias with disfluent stimuli.

General Discussion

The principal finding, which emerged in all experiments here, is that disfluency in speech directs listeners' attention to unfamiliar objects. This builds on the finding that disfluency causes a bias toward previously unmentioned (new) objects (Arnold et al., 2004). Disfluency biases toward both unfamiliar and new objects have rapid, on-line effects, influencing listeners' hypotheses about the unfolding referential expression. Unfamiliar objects and new objects are also generally considered more difficult to describe and are more likely to result in disfluency, compared with familiar objects (see Arnold & Tanenhaus, 2007, and the naming experiment described here).

We used the disfluency bias toward unfamiliar objects as a vehicle for investigating the degree to which language comprehension is affected by inferences about why the speaker said something in a particular way. Experiment 2 found that the unfamiliarity bias with disfluent stimuli was flexible and depended on what was perceived as difficult for a particular speaker. When comprehenders thought that the speaker had object agnosia, they appeared to make the inference that both familiar and unfamiliar objects should be difficult to describe. This resulted in an elimination of the unfamiliarity bias, even though the same stimuli yielded an unfamiliarity bias for listeners who thought the instructions were spoken by a typical speaker. These results suggest that speakers do indeed make inferences about which objects are hard for a particular speaker to describe and perhaps even for particular situations.

The results of Experiment 2 are relevant to an ongoing debate about the extent to which listeners can take into account speaker perspective and knowledge in real-time language processing (Barr & Keysar, 2006). Some researchers have suggested that real-time reference resolution operates egocentrically, without taking into account the knowledge or perspective of the speaker (e.g., Keysar, Barr, Balin, & Brauner, 2000; Keysar, Barr; Balin, & Paek, 1998; see also Pickering & Garrod, 2004). Our results demonstrate that real-time reference resolution can be modulated by explicit top–down knowledge about the speaker's cognitive state and the circumstances likely to cause that particular speaker to be disfluent. Thus they are inconsistent with the strongest form of the egocentrism hypothesis.

Experiment 3, on the other hand, established some limits to the flexibility of the unfamiliarity bias. Listeners heard instructions that either included evidence that the speaker was distracted, or did not. We predicted that if listeners could dynamically choose the most likely cause of disfluency, the unfamiliarity bias (which depends on an attribution of naming difficulty) should disappear in the distraction condition. However, it did not. In terms of the hypothetical inferences in Figure 1, these findings indicate that there are limits to making inferences about the cause of disfluency, as represented by the gray arrows. It may be that inferences at this level are blind to the presence of an alternative plausible cause, or there may be a preference to seek a production–internal interpretation.

The results of these experiments together delimit the possible inferential or automatic processes that may be involved during reference comprehension. The strongest evidence here supports an inferential account. It is clear that reference resolution benefits from inferences about what kinds of things might be difficult to describe for a particular speaker. Thus, despite the fact that the link between disfluency and the act of describing new or unfamiliar objects could be learned from experience, people appear to be flexible about the kinds of objects they associate with the act of describing something difficult.

However, there is one way that we could still salvage an associationist account of the unfamiliar and new biases. Perhaps the system has a learned association between disfluency and a *class* of objects, for example, "those objects that are difficult to name." However, if so, the system would still need to make inferences about which objects fall in this class for any given situation. Although unfamiliar objects may be considered difficult for most people, they are no more difficult than familiar objects for speakers with object agnosia. Thus, it could be that the gray arrow linking disfluency with "describing something difficult" could be prelearned, whereas the black arrows require a situation-specific inference. If so, this would explain the tenacity of the unfamiliarity bias in the face of distraction.

Although an association between disfluency and a class of difficult-to-describe objects is consistent with the data, it blurs the distinction between an inference attributing disfluency to the difficulty of naming a particular object, and using a prelearned link between disfluency and hard-to-describe objects. Even a prelearned association would require listeners to weigh the evidence for a particular situation in deciding which objects are the most difficult to name. How different is this from inferring that the disfluency occurred because of difficulty naming that object?

Indeed, the association may not be with a category of objects per se but instead with the state of having difficulty. Thus, the association account and the inference account both involve a kind of mindreading—that is, modeling the speaker's cognitive processes and judging what is difficult for that speaker.

The approaching similarity of the inferential and automatic explanations of the disfluency effect highlights the complex array of mechanisms that may be involved in the use of disfluency or any other cue during language processing. Inferences and associations are not mutually exclusive; for example, attribution may be built upon prelearned associations between disfluency and naming difficulty.

Here we have considered the mechanisms involved in normal language comprehension by examining how disfluency affects reference comprehension. Disfluency leads to on-line biases toward complex, unfamiliar objects, which we have suggested is the result of the assumption that those objects are difficult to name and therefore cause difficulty for the speaker. The robustness of the effects suggests some limits to the flexibility of inferring that the speaker is having production difficulty. However, we found strong evidence that listeners flexibly identify objects that are difficult for the speaker, revealing an adaptable language system.

References

Allopenna, P. D., Magnuson, J. S., & Tanenhaus, M. K. (1998). Tracking the time course of spoken word recognition using eye movements: Evidence for continuous mapping models. *Journal of Memory and Language, 38,* 419–439.

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73,* 247–264.

Arnold, J. E., & Tanenhaus, M. K. (2007). Disfluency effects in comprehension: How new information can become accessible. In E. Gibson & N. Perlmutter (Eds.), *The processing and acquisition of reference.* Cambridge, MA: MIT Press.

Arnold, J. E., Tanenhaus, M. K., Altmann, R., & Fagnano, M. (2004). The old and thee, uh, new. *Psychological Science, 15,* 578–581.

Arnold, J. E., Wasow, T., Losongco, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language, 76,* 28–55.

Bailey, K. G. D., & Ferreira, F. (2003). Disfluencies affect the parsing of garden-path sentences. *Journal of Memory and Language, 49,* 183–200.

Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In C. Cavé, I. Guaïtella, & S. Santi (Eds.), *Oralité et gestualité: Interactions et comportements multimodaux dans la communication* [Orality and gestuality: Multimodal interaction and behavior in communication] (pp. 597–600). Paris: L'Harmattan.

Barr, D. J., & Keysar, B. (2006). Perspective taking and the coordination of meaning in language use. In M. J. Traxler & M. A. Gernsbacher (Eds.), *Handbook of Psycholinguistics*. Amsterdam: Elsevier.

Beattie, G. W. (1979). Planning units in spontaneous speech: Some evidence from hesitation in speech and speaker gaze direction in conversation. *Linguistics, 17,* 61–78.

Beattie, G. W., & Butterworth, B. L. (1979). Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech, 22,* 201–211.

Bell, A., Jurafsky, D., Fosler-Lussier, E., Girand, C., Gregory, M., & Gildea, D. (2003). Effects of disfluencies, predictability, and utterance position on word form variation in English conversation. *Journal of the Acoustical Society of America, 113,* 1001–1024.

Boersma, P., & Weenink, D. (2003). Praat: Doing phonetics by computer (Version 4.1.13) [Computer program]. Amsterdam: Institute of Phonetic Sciences.

Boersma, P., & Weenink, D. (2006). Praat: Doing phonetics by computer (Version 4.4.24) [Computer program]. Amsterdam: Institute of Phonetic Sciences.

Bortfield, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. *Language and Speech, 32,* 229–259.

Brennan, S. E., & Williams, M. (1995). The feeling of another's knowing: Prosody and filled pauses as cues to listeners about the metacognitive states of speakers. *Journal of Memory and Language, 34,* 383–398.

Chafe, W. L. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–56). New York: Academic Press.

Chambers, C. G., Tanenhaus, M. K., Eberhard, K. M., Carlson, G. N., & Filip, H. (2002). Circumscribing referential domains during real-time language comprehension. *Journal of Memory and Language, 47,* 30–49.

Clark, H. H., & Fox Tree, J. E. (2002). Using *uh* and *um* in spontaneous speaking. *Cognition, 84,* 73–111.

Clark, H. H., & Wasow, T. (1998). Repeating words in spontaneous speech. *Cognitive Psychology, 37,* 201–242.

Cohen, J. D., MacWhinney, B., Flatt, M. R., & Provost, J. (1993). PsyScope: A new graphic interactive environment for designing psychology experiments. *Behavioral Research Methods, Instruments, & Computers, 25,* 257–271.

Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology, 6,* 84–107.

Dahan, D., Magnuson, J. S., & Tanenhaus, M. K. (2001). Time course of frequency effects in spoken word recognition: Evidence from eye movements. *Cognitive Psychology, 42,* 317–367.

Dahan, D., & Tanenhaus, M. K. (2005). Looking at the rope when looking for the snake: Conceptually mediated eye movements during spoken-word recognition. *Psychonomic Bulletin & Review, 12,* 453–459.

Dahan, D., Tanenhaus, M. K., & Chambers, C. G. (2002). Accent and reference resolution in spoken language comprehension. *Journal of Memory and Language, 47,* 292–314.

Dahan, D., Tanenhaus, M. K., & Salverda, A. P. (2007). The influence of visual processing on phonetically-driven saccades in the "visual world" paradigm. In R. P. G. Van Gompel, M. H. Fischer, W. S. Murray, & R. L. Hill (Eds.), *Eye movements: A window on mind and brain* (pp. 471–486). Oxford, England: Elsevier.

Eberhard, K. M., Spivey-Knowlton, M. J., Sedivy, J. C., & Tanenhaus, M. K. (1995). Eye movements as a window into real-time spoken language comprehension in natural contexts. *Journal of Psycholinguistic Research, 24,* 409–436.

Farah, M. J. (1990). *Visual agnosia: Disorders of object recognition and what they tell us about normal vision.* Cambridge, MA: MIT press.

Ferreira, F., & Swets, B. (2002). How incremental is language production? Evidence from the production of utterances requiring the computation of arithmetic sums. *Journal of Memory and Language, 46,* 57–84.

Fox Tree, J. E. (2002). Interpreting pauses and *ums* at turn exchanges. *Discourse Processes, 34,* 37–55.

Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing "the" as "thee" to signal problems in speaking. *Cognition, 62,* 151–167.

Gennari, S., & MacDonald, M. C. (2006). Acquisition of negation and quantification: Insights from adult production and comprehension. *Language Acquisition, 13,* 125–168.

Goldman-Eisler, F. (1968). *Psycholinguistics: Experiments in spontaneous speech.* London: Academic Press.

Grodner, D. J., & Sedivy, J. C. (2007). The effect of speaker-specific information on pragmatic inferences. In E. Gibson & N. Perlmutter

(Eds.), *The processing and acquisition of reference.* Cambridge, MA: MIT Press.

Gundel, J. (1988). Universals of topic-comment structure. In E. M. Hammond et al. (Eds.), *Studies in syntactic typology* (pp. 209–239). Philadelphia: John Benjamins.

Keysar, B., Barr, D. J., Balin, J. A., & Brauner, J. S. (2000). Taking perspective in conversation: The role of mutual knowledge in comprehension. *Psychological Science, 11,* 32–38.

Keysar, B., Barr, D. J., Balin, J. A., & Paek, T. S. (1998). Definite reference and mutual knowledge: Process models of common ground in comprehension. *Journal of Memory and Language, 39,* 1–20.

Krahmer, E., & Swerts, M. (2005). How children and adults produce and perceive uncertainty in audiovisual speech. *Language and Speech, 48,* 29–53.

Longhurst, E. (2006). ExBuilder [Computer program]. Rochester, NY: University of Rochester.

MacDonald, M. C. (1999). Distributional information in language comprehension, production, and acquisition: Three puzzles and a moral. In B. MacWhinney (Ed.), *The emergence of language* (pp. 177–196). Mahwah, NJ: Erlbaum.

Marslen-Wilson, W. (1973, August 24). Linguistic structure and speech shadowing at very short latencies. *Nature, 244,* 522–523.

Marslen-Wilson, W. (1975, July 18). Sentence perception as an interactive parallel process. *Science, 189,* 226–228.

Marslen-Wilson, W. (1987). Functional parallelism in spoken word recognition. *Cognition, 25,* 71–102.

Oviatt, S. (1995). Predicting spoken disfluencies during human–computer interaction. *Computer Speech and Language, 9,* 19–35.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27,* 169–226.

Prince, E. (1992). The ZPG letter: Subjects, definiteness, and information-status. In S. Thompson & W. Mann (Eds.), *Discourse description: Diverse analyses of a fund raising text* (pp. 295–325). Philadelphia: John Benjamins.

Race, D. S., & MacDonald, M. C. (2003). The use of "that" in the production and comprehension of object relative clauses. In *Proceedings of the 25th Annual Meeting of the Cognitive Science Society* (pp. 946–951). Mahwah, NJ: Erlbaum.

Sacks, O. (1985). *The man who mistook his wife for a hat and other clinical tales.* New York: Harper & Row.

Schachter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology, 60,* 362–367.

Schachter, S., Rauscher, F. H., Christenfeld, N., & Tyson Crone, K. (1994). The vocabularies of academia. *Psychological Science, 5,* 37–41.

Siegman, A. W. (1979). Cognition and hesitation in speech. In A. W. Siegman & S. Feldstein (Eds.), *Of speech and time: Temporal speech patterns in interpersonal contexts* (pp. 151–178). Hillsdale, NJ: Erlbaum.

Smith, V. L., & Clark, H. H. (1993). On the course of answering questions. *Journal of Memory and Language, 32,* 25–38.

Snodgrass, J. G., & Vanderwart, M. (1980). A standardized set of 260 pictures: Norms for name agreement, image agreement, familiarity, and visual complexity. *Journal of Experimental Psychology: Human Learning & Memory, 6,* 174–215.

Swerts, M. (1998). Filled pauses as markers of discourse structure. *Journal of Pragmatics, 30,* 485–496.

Swerts, M., & Geluykens, R. (1994). Prosody as a marker of information flow in spoken discourse. *Language and Speech, 37,* 21–43.

Swerts, M., & Krahmer, E. (2005). Audiovisual prosody and feeling of knowing. *Journal of Memory and Language, 53,* 81–94.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1995, June 16). Integration of visual and linguistic information in spoken language comprehension. *Science, 268,* 1632–1634.

Tanenhaus, M. K., Spivey-Knowlton, M. J., Eberhard, K. M., & Sedivy, J. C. (1996). Using eye movements to study spoken language comprehension: Evidence for visually mediated incremental interpretation. In T. Inui & J. L. McClelland (Eds.), *Attention and performance XVI: Information integration in perception and communication* (pp. 457–478). Cambridge, MA: MIT Publishing.

Received July 6, 2006
Revision received May 17, 2007
Accepted May 23, 2007 ■