I Forget What That's Called! Children's Online Processing of Disfluencies Depends on Speaker Knowledge

Adriel John Orena
McGill University

Katherine S. White
University of Waterloo

Speech disfluencies can convey information to listeners: Adults and children predict that filled pauses (e.g., *uhh*) will be followed by referents that are difficult to describe or are new to the discourse. In adults, this is driven partly by an understanding that disfluencies reflect processing difficulties. This experiment examined whether 3½-year-olds' use of disfluencies similarly involves inferences about processing difficulty. Forty children were introduced to either a knowledgeable or a forgetful speaker, who then produced fluent and disfluent utterances. Children exposed to the knowledgeable speaker looked preferentially at novel, discourse-new objects during disfluent utterances. However, children who heard the forgetful speaker did not. These results suggest that, like adults, children modify their expectations about the informativeness of disfluencies on a speaker-specific basis.

Language comprehension is highly efficient because of its predictive nature: Listeners predict what a speaker is referring to before the end of a word (Allopenna, Magnuson, & Tanenhaus, 1998; Fernald, Swingley, & Pinto, 2001) and, in certain contexts, before the word is uttered (Altmann & Kamide, 1999; Borovsky, Elman, & Fernald, 2012). One of the cues that listeners can use to predict a speaker's intended referent is the presence of speech disfluencies. Filled-pause disfluencies (e.g., *umms* and *uhhs*) tend to occur in predictable locations—usually preceding difficult words (i.e., words that are novel or infrequent) or discourse-new words (i.e., words that have not been previously mentioned in the conversation; Barr, 2001)—and the presence of filled pauses leads listeners to expect these types of referents. Arnold, Hudson Kam, and Tanenhaus (2007) demonstrated that adult listeners were more likely to fixate on novel objects when instruction phrases contained filled pauses (e.g., "Click on *thee, uhh,* . . .") than when they were fluent. Kidd, White, and Aslin (2011a) found that 2½-year-olds also made use

of such disfluencies to guide their predictions about a speaker's intended referent. Thus, this "disfluency effect" emerges at a young age.

In the present study, we ask what type of processing underlies children's predictive use of filled-pause disfluencies. One possibility is that children's predictions are based on the low-level distributional patterns exhibited by filled pauses. That is, since filled pauses are likely to precede less familiar or discourse-new words, children may automatically search for novel, discourse-new referents whenever they hear filled pauses. It is certainly plausible that children have learned the co-occurrence between filled pauses and particular types of referents. Toddlers can discriminate between fluent and disfluent adult-directed speech, even when segmental content is removed (Soderstrom & Morgan, 2007). Moreover, filled pauses, in addition to being prosodically marked, often involve articles with marked phonetic forms (e.g., *thee* instead of *thuh*), and by 13 months, infants are sensitive to the typical phonetic forms of common articles (Shi, Werker, & Cutler, 2006). Therefore, it is quite likely that young learners register the presence of disfluencies in their speech input. Given that even infants are capable of learning statistical relations between elements at a number of levels (Gómez, 2002; Gómez & Gerken, 1999; Saffran, Aslin, & Newport, 1996; Yu & Smith,

2007), it is also likely that young learners have the statistical tools to uncover and use the relations between filled-pause disfluencies and particular types of referents.

Another possibility is that children's predictions are based on higher level inferences about the source of disfluencies. Many filled pauses are caused by the demands of real-time speech production (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001) and adult listeners are aware that production difficulties can lead to disfluency (Clark & Fox Tree, 2002). On this view, listeners infer that a speaker is disfluent because of a retrieval difficulty, and that the reason for the difficulty is that the speaker is searching for a less familiar or less recently accessed word. Consistent with this, Arnold et al. (2007) showed that adults' use of disfluencies can be modulated by speaker-specific knowledge: Participants were less likely to treat disfluencies as predictive of novel referents when they were told that the speaker suffered from object agnosia (a condition in which even familiar object recognition is difficult). When listeners had reason to believe that the speaker had unusual patterns of retrieval difficulty, the expectation that novel words selectively cause retrieval difficulty—and therefore disfluencies—was canceled. Bosker, Quené, Sanders, and de Jong (2014) similarly found that adults disregard disfluency as a cue to upcoming referents when listening to non-native speakers (whose patterns of disfluency production are more irregular). Adults also appear to take a speaker's background knowledge into account, using disfluencies as a predictive cue only when referents are discourse-new for that particular speaker (Barr & Seyfeddinipur, 2010). Collectively, these studies suggest that situation-specific inferences about the likely source of disfluencies modulate adults' use of disfluencies during spoken language processing.

It is possible that children's use of filled-pause disfluencies is likewise mediated by speaker-specific inferences, as even young children have demonstrated a surprising amount of pragmatic competence. For example, children's use of eye gaze to determine reference can be overridden when the linguistic or pragmatic context makes it uninformative (Nurmsoo & Bloom, 2008). Children also use a conversation partner's perspective or background knowledge to guide their referential interpretations (Grassmann, Stracke, & Tomasello, 2009; Southgate, Chevallier, & Csibra, 2010). Finally, they use the perceived knowledge state of a speaker to decide whether the speaker is a reliable information source (Birch, Vauthier, & Bloom, 2008; Koenig & Harris,

2005; Sabbagh & Baldwin, 2001), but cancel inferences about a speaker's competence in some contexts—for example, if an inaccurate speaker did not have access to relevant visual information (Brosseau-Liard & Birch, 2011; Nurmsoo & Robinson, 2009). These findings suggest that children's pragmatic understanding may be sufficiently developed that they process disfluencies in the same way as adults.

The present study examined whether children make use of speaker-specific inferences when processing filled-pause disfluencies. Analogous to the Arnold et al. (2007) experiment in which adult participants were told that a speaker had object agnosia, we gave our child participants additional information about the speaker. If children's predictive use of disfluencies is driven solely by the statistical relations in the environment between disfluencies and classes of referents, we expect children to exhibit the disfluency effect regardless of knowledge about specific speakers. However, if children—like adults—engage in the type of inference process described above, they might show more flexible, speaker-specific, interpretation of disfluencies.

Method

Three-and-a-half-year-old children were tested using a preferential looking-while-listening paradigm. In the history phase, children were introduced to a puppet speaker who was either knowledgeable or forgetful. In the test phase, this puppet then produced fluent and disfluent instructions. If children modulate their use of disfluencies based on attributions of speaker difficulty, they should look preferentially at novel, discourse-new objects when the knowledgeable speaker produces disfluencies, but suspend this bias when the forgetful speaker produces disfluencies.

Participants

Forty children were randomly assigned to two conditions (20 per condition): The knowledgeable speaker condition (13 males, M_{age} = 42.1 months, range = 40.3–43.7 months) and the forgetful speaker condition (10 males, M_{age} = 41.7 months, range = 40.1–43.7 months). The group of participants was predominantly Caucasian, and all were monolingual English-speaking children recruited from the Kitchener-Waterloo region in Canada. Data collection occurred from July 2013 to January 2015. One additional child participated but was replaced

because he left the room in the middle of the experiment. All participants were full-term and had no hearing or visual problems. Children received either a T-shirt or a book as compensation.

Stimuli

History phase stimuli consisted of videotaped puppet shows. Two puppets discussed six familiar objects (apple, balloon, chair, cow, flower, and truck), one object at a time. These objects are among the earliest acquired words listed in the MacArthur–Bates Communicative Development Inventories (Dale & Fenson, 1996). Two native speakers of English (one male and one female) produced the audio for the puppets in a child-directed manner. The male speaker voiced the role of the narrator puppet and the female speaker voiced the role of the target puppet. Audio was recorded in a soundproofed room using Praat (Boersma & Weenink, 2009), with a sampling rate of 44.1 kHz. Audio files were edited to match average intensity ($M = 60$ dB). Each scene of the puppet show was filmed separately, and edited using iMovie. The movie files were dubbed with the separately recorded audio stimuli.

For the Test phase, the still pictures of objects were taken from the Kidd et al. (2011a) stimulus set, which includes 32 colored pictures of familiar and novel objects (see Kidd et al., 2011a, for more information). Each trial contained one unique familiar–novel object pair, for a total of 16 trials. None of the familiar objects from the history phase appeared in the test phase. The audio stimuli for the test trials were produced by the target speaker (the female speaker from the history phase).

Procedure

The experiment was conducted in a dimly lit, soundproofed room. Children were seated on their parent's lap in front of a 42-in. plasma screen. Parents wore headphones playing masking music and were asked to not communicate with their child or influence their behavior in any way. A camera mounted below the plasma screen recorded the children's eye movements for later offline coding. We used PsyScope X Software (Cohen, MacWhinney, Flatt, & Provost, 1993) from a Macintosh computer in an adjacent room to control stimulus presentation.

Children saw an alternation of history and test blocks (History 1 → Test 1 → History 2 → Test 2). Each history block introduced three unique familiar objects, and each test block contained eight unique test trials. The history and test blocks alternated to remind children of the speaker's characteristics halfway through the experiment.

The critical difference between the two conditions was the history phase. The first history block began with the target puppet being introduced by the narrator puppet. In the knowledgeable speaker condition, the target puppet was simply introduced as a friend, Sally; in the forgetful speaker condition, she was introduced as Sally, a forgetful friend who did not remember the names of objects. In order to show evidence of Sally's knowledge (or lack thereof), the narrator then asked her to give the names of three familiar objects. In the knowledgeable speaker condition, Sally answered correctly for all three objects (e.g., "I know, that's an apple!"). However, in the forgetful speaker condition, Sally claimed that she had forgotten the names of these objects (e.g., "I forget what that's called!"). At the end of each history block, the narrator introduced the test block by stating that Sally was going to talk about some pictures.

The test blocks were the same for the two conditions. See Figure 1 for the schematic of a single trial in the test phase. Each test trial began with a central attention-grabbing stimulus. Once the child looked to the center, the trial was initiated and a pair of objects (one familiar and one novel) was presented side-by-side, three times in succession. The first two presentations of each object pair were accompanied by the audio stimuli "Ooh! Look at that" and "Ooh! How nice." While the audio was playing, a video of the target puppet appeared centrally and directed attention to the familiar object by turning toward it. Note that the speaker always looked at the familiar object during the first two presentations, establishing the familiar object as the discourse-old object. Labeling of the familiar object was not used, as fluent labeling would have conflicted with the information from the history phase that the speaker in the forgetful speaker condition was forgetful. This video also reinforced the identity of the test speaker as the target puppet from the history phase. The third, critical, utterance was either a fluent or disfluent instruction to look at one of the objects in the pair ("Look! Look at the X!" or "Look! Look at *thee, uhh*, X!"). There was no visual of the speaker during this presentation. Half of the test trials were fluent and half were disfluent. Fluent and disfluent trials were equally likely to occur with familiar and novel object targets. Objects remained on screen for 3.5 s after the onset of the target word label during the third presentation. The assignment of object pairs to trial type (fluent or
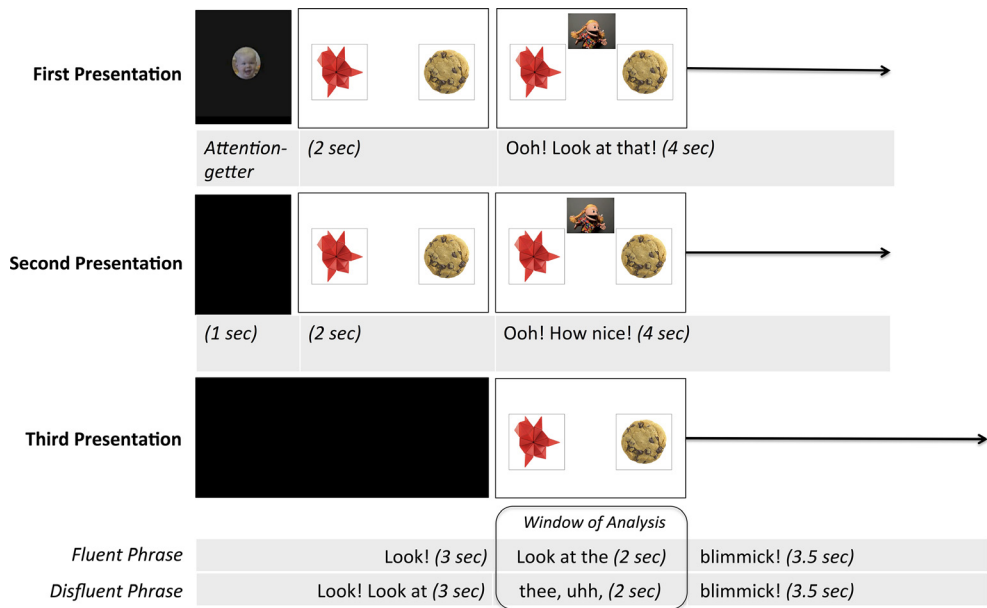
Figure 1. Schematic of a single trial in the test phase.

disfluent; familiar or novel target) and the order of these pairs were counterbalanced across participants in each condition (the two conditions had identical counterbalancing).

Coding

Sessions were coded offline by a coder blind to the stimuli, frame by frame (33 ms/frame) using customized coding software. Each frame was coded as a look to the left, right, speaker, or other. To check for coding reliability, a second coder coded a random sample of participants (25% of participants; $n = 10$). The two coders achieved an average reliability of 91% across all frames of all trials.

In order to ensure that children knew where the novel and familiar objects would be on the screen during the third presentation of the objects, trials in which children did not look at both objects during the first two presentations of the objects were discarded. For participants in the knowledgeable speaker condition, there were 14 such trials, accounting for 4.38% of all trials; for participants in the forgetful speaker condition, there were 18 such trials, accounting for 5.63% of all trials.

Results

First Two Presentations of Objects

We first asked whether children were able to follow the visual cues (i.e., turning and pointing at the object) during the first two presentations of the objects. These cues established the discourse-old status of the familiar object. To determine whether this cueing was effective, we calculated the proportion of looking time to the *familiar* object precueing (2 s) and postcueing (4 s) by dividing the looking time to the familiar object by looking time to the novel and familiar objects. A mixed three-way analysis of variance (ANOVA) with condition (knowledgeable speaker vs. forgetful speaker), presentation (first vs. second), and cueing (precueing vs. postcueing) showed a main effect of presentation, $F(1, 38) = 11.57$, $p = .002$, with children looking more at the familiar object during the second presentation ($M = 0.63$, $SE = 0.02$) than the first ($M = 0.57$, $SE = 0.01$), as well as a main effect of cueing, $F(1, 38) = 134.77$, $p < .001$, with children looking more at the familiar object after cueing. There was also an interaction between presentation and cueing, $F(1, 38) = 4.42$, $p = .042$. During the first presentation, the proportion of looking to the familiar object significantly increased from 0.43 (*prior to* cueing) to 0.71 (*after* the start of cueing), $t(39) = 10.95$, $p < .001$. The same pattern was found during the second presentation, albeit slightly weaker: The proportion looking to the familiar object significantly increased from 0.53 to 0.73, $t(39) = 6.92$, $p < .001$. There was no main effect of condition ($p = .565$), but there was a marginally significant interaction between condition and presentation, $F(1, 38) = 3.29$, $p = .077$. No other interactions were significant ($p > .114$ in all cases). Thus, children in both conditions were able to follow the

puppet's visual cues equally well, rendering the familiar object discourse-old.

Looking to Target Object Following Naming

Next, we confirmed that children reliably identified the target object after it was named during the third presentation. As in previous research, it was estimated that young children require ~270 ms to fixate a visual stimulus in response to an auditory stimulus (Canfield, Smith, Bresnyak, & Snow, 1997). Accordingly, we shifted the analysis forward by eight frames (264 ms) posttarget onset to account for this stimulus-response latency. We calculated the proportion of looking time to the target object as looking time to the *target* object divided by looking time to both objects. During the 2-s period following naming, children reliably looked at the target object: On familiar-target trials, children looked more toward the familiar object (mean proportion familiar = 0.80, SE = 0.02). Likewise, on novel-target trials, children looked more toward the novel object (mean proportion novel = 0.60, SE = 0.02). One-sample t tests against .5 showed that both were significantly above chance, $t(39)$ = 17.76, p < .001, and $t(39)$ = 4.09, p < .001, respectively. Children in the two conditions did not differ in their proportion of looking to the familiar target object, $t(38)$ = 0.29, p = .769, or to the novel target object, $t(38)$ = 0.30, p = .766. Therefore, children in both conditions were successful in mapping the familiar labels to their respective objects, and in using disambiguation to associate the novel labels with the novel objects.

Critical Window

Finally, we turned to the critical window to examine whether children's looking behavior differed as a function of fluency and speaker. Following Kidd et al. (2011a), the critical window was defined as the 2-s period *before* the target object was labeled, the period corresponding to the period of disfluency (and the material "Look, look at" in fluent trials; Figure 1). As above, we shifted the analysis window forward by 264 ms.

To test whether children showed a disfluency effect, we calculated the proportion of looking time to the *novel* object during the critical window as the looking time to the novel object divided by looking time to both objects. Figure 2 displays the proportion looking to the novel object during this window for fluent versus disfluent trials (collapsed across target object type).

The mean proportions were submitted to a two-way mixed ANOVA, with condition (knowledgeable speaker vs. forgetful speaker) as a between-subjects factor and trial type (fluent vs. disfluent) as a within-subjects factor. There were no main effects of condition, $F(1, 38)$ = 0.22, p = .656, or trial type, $F(1, 38)$ = 2.92, p = .095. However, as hypothesized, there was a significant interaction between the two factors, $F(1, 38)$ = 4.48, p = .041. To explore this interaction, t tests compared looking proportions in disfluent and fluent trials for each condition separately. Children in the knowledgeable speaker condition had a significantly higher proportion of looking to the novel object during disfluent trials than fluent trials, $t(19)$ = 3.61, p = .002. However, for the forgetful speaker condition, there was no difference in proportion looking to the novel object in the two types of trials, $t(19)$ = 0.24, p = .814. Because the relevant baseline for determining whether there is a disfluency effect is the child's behavior during fluent trials (and not chance level), these comparisons reveal a disfluency effect in the knowledgeable speaker condition only: Only in this condition did the presence of disfluencies significantly alter children's looking behavior from baseline.

However, comparisons against chance (0.5) reveal another interesting finding: In the knowledgeable speaker condition, the proportion of looking to the novel object was significantly different from chance during fluent trials (M = 0.42, SE = 0.03), $t(19)$ = −2.93, p = .009, but not during disfluent trials (M = 0.51, SE = 0.04), $t(19)$ = 0.21, p = .837, suggesting that children expected the speaker to continue talking about the familiar object
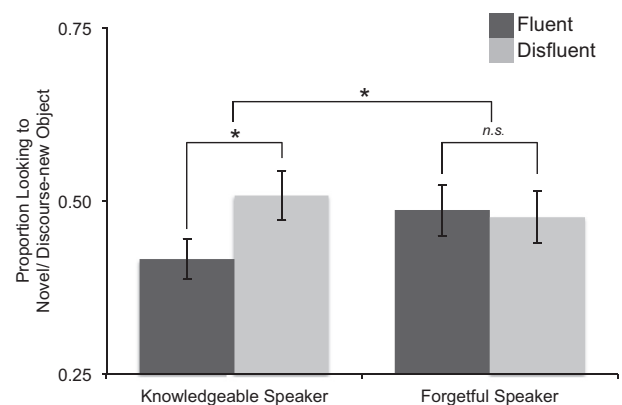

Figure 2. Proportion looking to the novel, discourse-new object during fluent and disfluent trials for the knowledgeable speaker and forgetful speaker conditions. Error bars are standard error.
*p = .05.

during fluent trials, but not during disfluent trials (see Arnold, Fagnano, & Tanenhaus, 2003, for a similar pattern). In contrast, this bias was not observed in the forgetful speaker condition: Neither proportion was significantly different from chance: disfluent trials ($M = 0.48$, $SE = 0.04$), $t(19) = -0.64$, $p = .532$, and fluent trials ($M = 0.49$, $SE = 0.04$),

$t(19) = -0.38$, $p = .705$. Taken together, children in the knowledgeable speaker condition selectively increased their attention to novel objects when they heard disfluencies, but children in the forgetful speaker condition did not.

For visualization, the time course of looking over the critical window is depicted in Figure 3. This figure
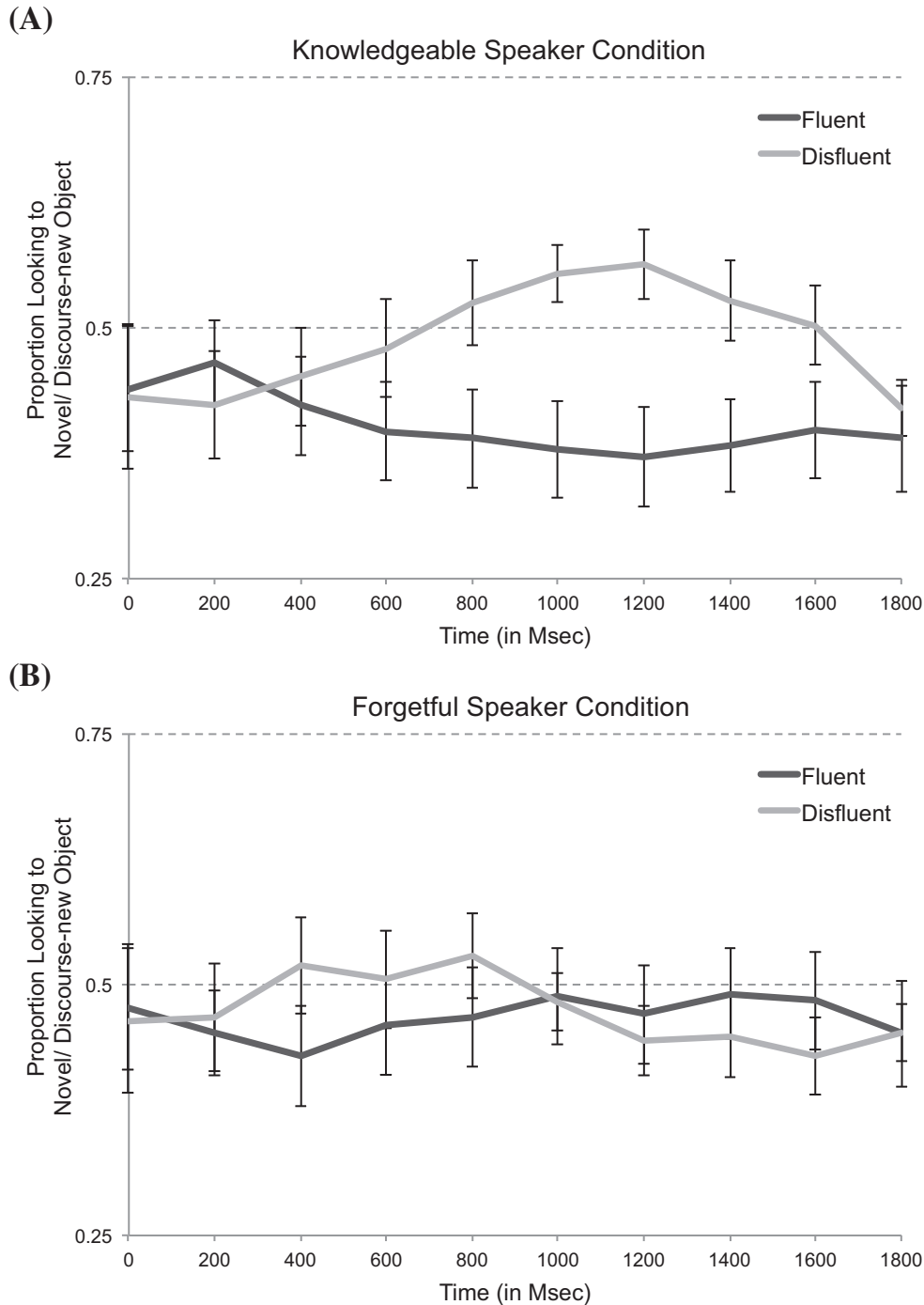
(A)

(B)

Figure 3. Time course of proportion looking to the novel, discourse-new object during the 2-s critical window for children in the (A) knowledgeable speaker and (B) forgetful speaker conditions. Error bars are standard error.

demonstrates that for children in the knowledgeable speaker condition, attention to the novel object during disfluent and fluent trials diverged around 600 ms after the onset of the disfluency. In contrast, for children in the forgetful speaker condition, looking for the two trial types did not differ throughout the 2-s critical window.

Discussion

Prior work has shown that children use disfluencies predictively (Kidd et al., 2011a), but has not addressed the mechanisms by which disfluencies influence their expectations. Here, we asked whether young children's predictive use of disfluencies reflects only learned statistical associations between disfluencies and particular types of referents, or whether it can be modulated by speaker-specific inferences. Our results show that, like adults, children did not use disfluency as a cue to the speaker's referential intent automatically. Instead, they appeared to consider the speaker's knowledge state. When children heard the knowledgeable speaker, they used disfluencies predictively: Although they appeared to assume that the speaker would continue referring to the familiar, discourse-old object during fluent trials, the presence of disfluencies significantly altered their behavior, causing them to increase their looking to the novel, discourse-new object. In contrast, when children heard the forgetful speaker, they did not use disfluency as a cue for the speaker's referential intent: There was no difference in their looking behavior for fluent and disfluent trials. It is important to note that the children in the forgetful speaker condition did not hear the speaker produce disfluencies with familiar referents during the History phase. Therefore, they had no direct evidence of disfluencies preceding familiar referents. Instead, their failure to use the disfluency cue must reflect an inference that the speaker's (lack of) knowledge could lead to unpredictable patterns of disfluency.

In addition to elucidating how children make predictive use of disfluencies, these findings have two important implications. First, they show that having information about a speaker's knowledge state can affect children's online language processing. Previous studies on speaker-specific inferences have shown how children's inferences about a speaker's knowledge (based on information provided by a speaker) affect their decision about

whether to believe or seek subsequent information from that speaker (Birch et al., 2008; Koenig & Woodward, 2010; Sobel & Corriveau, 2010). Our findings demonstrate that children's perception of a speaker's knowledge state can also have effects in a very different domain—in particular, in how they treat speech cues during online language processing. Interestingly, the effects on processing were quite specific: Children in both conditions looked at the target objects (familiar and novel) after hearing the target labeled. Thus, our speaker manipulation did not cause children in the forgetful speaker condition to disregard the speaker's labeling or weaken their use of disambiguation for her novel labels. However, it is interesting to note that, during fluent trials, children in the forgetful speaker condition (unlike the knowledgeable speaker condition) did not appear to assume that the speaker would continue talking about the discourse-old object. Thus, children may have ignored the discourse cues presented by that speaker.

Second, our results clarify exactly what types of referents are considered in children's disfluency-based predictions. In adults, filled pauses cause looking to novel objects and to discourse-new objects individually (Arnold et al., 2003; Arnold et al., 2007). In the first study to demonstrate a disfluency effect in toddlers, Kidd et al. (2011a) manipulated *both* object novelty and discourse status simultaneously to maximize the possibility of obtaining an effect. More recently, Owen and Graham (2013) found that 3½-year-olds show a disfluency effect when discourse status alone is manipulated, but may not when object novelty alone is manipulated (Owen & Graham, 2014). If it is true that children make disfluency-based predictions based on discourse status alone, our results demonstrate that explicit verbal discourse is unnecessary. In our test trials, the familiar object was highlighted during the first two presentations by the speaker's visual attention and a referentially ambiguous utterance ("Ooh, look at that!"). In the knowledgeable speaker condition, children nonetheless increased their proportion looking to the other object during a subsequent disfluency. Thus, overt production of a word is not required to give it discourse-old status. This indicates that the relevant inference about retrieval difficulty concerns whether the speaker has had the referent in mind, and not the act of production.

Of course, the question of precisely how children learn that disfluencies reflect speaker difficulty remains unanswered. One possibility is that this

understanding comes from their observation of how others produce disfluencies. Kidd, White, and Aslin (2011b) estimated, based on CHILDES (Child Language Data Exchange System) transcripts, that filled pauses occur in 1 of every 1,000 words in speech from adults to children. Although this rate is lower than speech between adults (Fox Tree, 1995), this number likely underestimates the true value, as disfluencies are often not transcribed. Moreover, children's learning of disfluencies may also occur from the presence of disfluencies in speech between adults, as toddlers can discriminate fluent and disfluent adult-directed speech (Soderstrom & Morgan, 2007). Thus, children may have had ample opportunity to observe the relations between production difficulty and the presence of disfluencies.

An alternative possibility is that children's understanding that disfluencies reflect difficulty is driven by their own production difficulties. In other words, the contexts in which children produce filled pauses may provide them with insights into the types of linguistic material that are difficult to produce. Kidd et al. (2011b) estimated that at age 2, the rate of disfluencies in children's own speech was roughly 1 in every 230 words—again they argued that this may be an underestimation of the true disfluency rate. Thus, it is possible that the 3½-year-old children in our study learned about the relations between production difficulties and filled pauses from their own prior experiences with production difficulty.

Note that the uses of statistical associations and higher order inferences are not mutually exclusive. Children may detect the distributional properties of filled-pause disfluencies and also, through either self-observation or observation of others, understand that it is the difficulty of producing words that causes disfluencies. However, our results suggest that by the age of 3½ years, children's processing of disfluencies cannot be solely statistical in nature. Further support for this is the fact that the size of the disfluency effect did not diminish over the course of the experiment (despite disfluencies preceding familiar targets half of the time). This suggests that children's processing was more influenced by their inferences about speaker difficulty than by the statistical distribution of disfluencies during the experiment.

To conclude, our findings suggest that 3½-year-olds use evaluations of a speaker's knowledge state to make speaker-specific inferences about the source of speech disfluencies, and that these inferences modulate their use of disfluencies during online language processing. Thus, from an early age, children show flexible, speaker-specific use of speech cues during language processing.

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. doi:10.1006/jmla.1997.2558

Altmann, G. T. M., & Kamide, Y. (1999). Incremental interpretation at verbs: Restricting the domain of subsequent reference. *Cognition, 73,* 247–264. doi:10.1016/S0010-0277(99)00059-1

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of Psycholinguistic Research, 32,* 25–36. doi:10.1023/A:1021980931292

Arnold, J. E., Hudson Kam, C. L., & Tanenhaus, M. K. (2007). If you say thee uh you are describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning Memory and Cognition, 33,* 914–930. doi:10.1037/0278-7393.33.5.914

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 gestualiét: interactions et comportements multimodaux dans la communication* [Voice and gesture: Multimodal interactions and behaviors during communication]. (pp. 597–600). Paris, France: L'Harmattan.

Barr, D. J., & Seyfeddinipur, M. (2010). The role of fillers in listener attributions for speaker disfluency. *Language and Cognitive Processes, 25,* 441–455. doi:10.1080/01690960903047122

Birch, S. A. J., Vauthier, S. A., & Bloom, P. (2008). Three- and four-year-olds spontaneously use others' past performance to guide their learning. *Cognition, 107,* 1018–1034. doi:10.1016/j.cognition.2007.12.008

Boersma, P., & Weenink, D. (2009). Praat: Doing phonetics by computer (version 5.1.05) [Computer software]. Retrieved from http://www.praat.org/

Borovsky, A., Elman, J. L., & Fernald, A. (2012). Knowing a lot for one's age: Vocabulary skill and not age is associated with anticipatory incremental sentence interpretation in children and adults. *Journal of Experimental Child Psychology, 112,* 417–436. doi:10.1016/j.jecp.2012.01.005

Bortfeld, 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, 44,* 123–147. doi:10.1177/00238309010440020101

Bosker, H. R., Quené, H., Sanders, T., & de Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language, 75,* 104–116. doi:10.1016/j.jml.2014.05.004

Brosseau-Liard, P. E., & Birch, S. A. J. (2011). Epistemic states and traits: Preschoolers appreciate the differential informativeness of situation-specific and person-specific cues to knowledge. *Child Development*, *82*, 1788–1796. doi:10.1111/j.1467-8624.2011.01662.x

Canfield, R. L., Smith, E. G., Brezsnyak, M. P., & Snow, K. (1997). Information processing through the first year of life: A longitudinal study using the visual expectation paradigm. *Monographs of the Society for Research in Child Development*, *62*(2), 1–145.

Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, *84*, 73–111. doi:10.1016/S0010-0277(02)00017-3

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

Dale, P. S., & Fenson, L. (1996). Lexical development norms for young children. *Behavior Research Methods, Instruments, and Computers*, *28*, 125–127. doi:10.3758/BF03203646

Fernald, A., Swingley, D., & Pinto, J. P. (2001). When half a word is enough: Infants can recognize spoken words using partial phonetic information. *Child Development*, *72*, 1003–1015. doi:10.1111/1467-8624.00331

Fox Tree, J. E. (1995). The effects of false starts and repetitions on the processing of subsequent words in spontaneous speech. *Journal of Memory and Language*, *34*, 709–738. doi:10.1006/jmla.1995.1032

Gómez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science*, *13*, 431–436. doi:10.1111/1467-9280.00476

Gómez, R. L., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*, 109–135. doi:10.1016/S0010-0277(99)00003-7

Grassmann, S., Stracke, M., & Tomasello, M. (2009). Two-year-olds exclude novel objects as potential referents of novel words based on pragmatics. *Cognition*, *112*, 488–493. doi:10.1016/j.cognition.2009.06.010

Kidd, C., White, K. S., & Aslin, R. N. (2011a). Toddlers use speech disfluencies to predict speakers' referential intentions. *Developmental Science*, *14*, 925–934. doi:10.1111/j.1467-7687.2011.01049.x

Kidd, C., White, K. S., & Aslin, R. N. (2011b). Learning the meaning of "um:" Toddlers' developing use of speech disfluencies as cues to speakers' referential intentions. In I. Arnon & E. V. Clark (Eds.), *Experience, variation, and generalization: Learning a first language* (pp. 91–106). Amsterdam, Netherlands: Benjamins.

Koenig, M. A., & Harris, P. L. (2005). The role of social cognition in early trust. *Trends in Cognitive Sciences*, *9*, 457–459. doi:10.1016/j.tics.2005.08.006

Koenig, M. A., & Woodward, A. L. (2010). Sensitivity of 24-month-olds to the prior inaccuracy of the source: Possible mechanisms. *Developmental Psychology*, *46*, 815–826. doi:10.1037/a0019664

Nurmsoo, E., & Bloom, P. (2008). Preschoolers' perspective taking in word learning: Do they blindly follow eye gaze? *Psychological Science*, *19*, 211–215. doi:10.1111/j.1467-9280.2008.02069.x

Nurmsoo, E., & Robinson, E. J. (2009). Children's trust in previously inaccurate informants who were well or poorly informed: When past errors can be excused. *Child Development*, *80*, 23–27. doi:10.1111/j.1467-8624.2008.01243.x

Owen, S. J., & Graham, S. A. (2013). *Thee, uhh, role of discourse status in three-year-olds' understanding of disfluent utterances*. Poster presented at the annual CUNY Conference on Human Sentence Processing, Columbia, SC.

Owen, S. J., & Graham, S. A. (2014). *The disfluency effect in children: A cue to discourse status but not object familiarity*. Poster presented at the biennial International Conference on Infant Studies, Berlin, Germany.

Sabbagh, M. A., & Baldwin, D. A. (2001). Learning words from knowledgeable versus ignorant speakers: Links between preschoolers' theory of mind and semantic development. *Child Development*, *72*, 1054–1070. doi:10.1111/1467-8624.00334

Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928. doi:10.1126/science.274.5294.1926

Shi, R., Werker, J. F., & Cutler, A. (2006). Recognition and representation of function words in English-learning infants. *Infancy*, *10*, 187–198. doi:10.1207/s15327078in1002_5

Sobel, D. M., & Corriveau, K. H. (2010). Children monitor individuals' expertise for word learning. *Child Development*, *81*, 669–679. doi:10.1111/j.1467-8624.2009.01422.x

Soderstrom, M., & Morgan, J. L. (2007). Twenty-two-month-olds discriminate fluent from disfluent adult-directed speech. *Developmental Science*, *10*, 641–653. doi:10.1111/j.1467-7687.2007.00605.x

Southgate, V., Chevallier, C., & Csibra, G. (2010). Seventeen-month-olds appeal to false beliefs to interpret others' referential communication. *Developmental Science*, *13*, 907–912. doi:10.1111/j.1467-7687.2009.00946.x

Yu, C., & Smith, L. B. (2007). Rapid word learning under uncertainty via cross-situational statistics. *Psychological Science*, *18*, 414–420. doi:10.1111/j.1467-9280.2007.01915.x