Children’s attribution of disfluency to different sources

Si On Yoon (sion-yoon@uiowa.edu)
Department of Communication Sciences and Disorders, University of Iowa
Iowa City, IA 52242 USA

Cynthia L. Fisher (clfishe@illinois.edu)
Department of Psychology, University of Illinois at Urbana-Champaign
Champaign, IL 61820 USA

Abstract

Disfluency in speech leads listeners, even two-year-old children, to expect the speaker to refer to novel and discourse-new objects. Previous evidence suggests this link between disfluency and discourse novelty is not driven simply by tracking of co-occurrence statistics connecting disfluency with reference to a new object, but also by integrating extra-linguistic information about the speaker’s perspective. We asked whether children can attribute a speaker’s disfluency to different sources – language planning difficulty vs. distraction from the conversation. We tested children’s processing of disfluency when interacting with an engaged versus a distracted speaker. When the engaged speaker was disfluent, children looked more at a novel and discourse-new image than at a familiar and just-named image, consistent with the existing literature. This disfluency effect was attenuated when the speaker was distracted, suggesting that four-year-old children can flexibly attribute a speaker’s disfluency to different sources in online interpretation of disfluent speech.

Keywords: Speech disfluency; Eye-tracking; Pragmatic inference; Attention; Source of disfluency

Introduction

Disfluency is common in everyday conversation (Brennan & Schober, 2001; Bortfeld, et al., 2001). Adult speakers often become hesitant or disfluent, producing filled pauses (e.g., “Look at thee… uh…”), repeating words, or restarting a phrase. Filled-pause disfluencies are regularly observed in speech to children, though less commonly than in speech to adults due to the predominance of short and simple sentences in child-directed speech (1 every 1000 words in speech to 2-year-olds vs. 6 every 100 words in speech to adults; Kidd et al., 2011; Fox Tree, 1995). As children grow older, filled pauses become more common.

Though disfluency can stem from multiple causes, it often reflects difficulty in planning utterances or retrieving lexical items (Clark & Fox Tree, 2002; Clark & Wasow, 1998; Ferreira, 1991; Fraundorf & Watson, 2013; Smith & Clark, 1993). Accordingly, although filled pause disfluencies (e.g., “um” or “uh”) do not carry linguistic meaning, adult listeners tend to interpret disfluency as a sign of planning difficulty, and therefore to predict aspects of how the utterance will unfold after the disfluency (Arnold, Hudson Kam, & Tanenhaus, 2007; Arnold, et al., 2004; Barr & Seyfeddinipur, 2010). For example, adults expect disfluent descriptions to refer to discourse-new entities or entities that are hard to describe. This disfluency effect has been observed in children as young as two years old (Kidd et al., 2011); two-year-old children looked more at a novel and discourse-new image (e.g., a pretzel-shaped object) than at a familiar and just-named image (e.g., an apple) following a speaker’s disfluency. Such findings suggest that both children and adults attribute speaker disfluency to planning difficulty and predict upcoming referents accordingly.

However, disfluency does not always reflect linguistic planning difficulty. Other factors can interrupt speakers’ utterances, including attempts to carry out two tasks at once, distracting the speaker’s attention from the conversation (e.g., checking messages on a cellphone while talking with another person). Less explored is whether listeners are able to interpret disfluency flexibly based on the inferred source of the disfluency (e.g., planning difficulty vs. distraction). Here, we explored this question by examining how four-year-old children processed disfluency when they interacted with an engaged versus a distracted speaker.

Previous evidence establishes that, in adulthood, the link between disfluency and discourse novelty is not a fixed one that might be based on simple tracking of co-occurrence statistics linking disfluency and reference to a new object, but also flexibly integrates information about the speaker. Adult listeners’ interpretation of disfluency as a cue to discourse novelty can be tailored to a particular speaker’s knowledge (Barr & Seyfeddinipur, 2010, Arnold et al., 2007, Yoon & Brown-Schmidt, 2014). For example, the typical “disfluency=new reference” prediction is reduced when adults interact with a speaker who has difficulty naming familiar objects (Arnold et al., 2007). Such findings suggest that adult listeners make situation-specific inferences and process disfluency accordingly (see also Heller et al., 2014).

Do preschoolers, like adults, interpret disfluency flexibly? Or do they rigidly treat disfluency as a cue to novelty (e.g., discourse newness and/or object novelty), regardless of who is speaking? Previous studies of children’s processing of disfluency have shown that children interpret disfluency differently depending on particular speakers’ characteristics or preferences (Orena & White, 2015; Thacker, Chambers, & Graham, 2018). For example, Orena and White (2015) found that 3.5-year-old children interpreted disfluency as a predictor of reference to a novel and discourse-new object when listening to a knowledgeable speaker who named objects properly, but canceled this expectation when listening to a forgetful speaker who often forgot the names of ordinary objects. Further, in a live conversation, four-year-old children learned about two different partners' knowledge states and later used this information appropriately when interpreting a partner's disfluency (Jin et al., in preparation). In Jin et al. (in preparation), children first established distinct shared knowledge with two partners consecutively in a referential communication game. One partner shared labels of animal tangrams, and the other partner shared the labels of vehicle tangrams, with the children. In a later test block, in trials containing a disfluency, children looked more at a tangram that was unfamiliar from the speaker's perspective, though not from the child's own perspective. These results suggest that instead of using simple statistics, children make inferences about the source of planning difficulty based on a particular speaker's knowledge.

We built on these findings to assess children's sensitivity to the differing sources of disfluency. As noted earlier, disfluency can result from difficulty in naming an unfamiliar and discourse-new object, but could also result from being distracted by another task. Even a knowledgeable speaker should be more disfluent when dividing attention between speaking and another task (e.g., Barch & Berenbaum, 1994; Oomen & Postma, 2001). Thus, if children can see that the speaker is distracted, and can reason about disfluency as caused by multiple kinds of difficulty, they might suspend their ordinary expectation that disfluency signals novelty.

Previous work provides reasons to expect preschoolers might succeed in this task. For example, when learning object names, children are sensitive to the direction of the speaker's visual attention (Baldwin, 1991; 1993) or aspects of the speaker's behavior including distraction (e.g., Jaswal & Malone, 2007). We examined how children process disfluency in live conversation, when the speaker is engaged vs. visibly distracted by another task. If children flexibly infer the likely source of disfluency based on the speaker's behavior (engaged vs. distracted), children interacting with an engaged speaker should expect disfluent utterances to refer to an item that is novel and discourse-new, but children interacting with a distracted speaker should reduce or suspend this expectation.

Experiment

Participants

Thirty-two 4-year-old children (48.6 – 59.2 months; M=53.6; 16 girls) participated in the experiment. All were acquiring English as their native language. Each child's parent gave written informed consent.

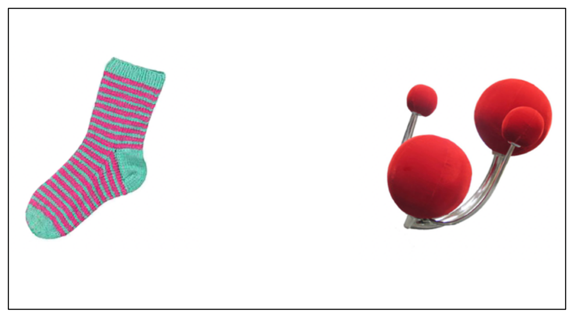
Materials and apparatus

Children sat at a table in front of a computer monitor. In each trial of the task, two images were presented on the screen – one familiar and nameable image and one novel, unnameable image (Figure 1). The images were adapted from Kidd et al. (2011); some images were modified. In Kidd et al. (2011), 2-year-old children looked more at the image that was both novel and discourse-new than the image that was familiar and just-named when the speaker was disfluent. However, this disfluency effect was not replicated when the image was novel, but not discourse-new (Owen, Thacker, & Graham, 2017). For this reason, we retained the same reference types tested in Kidd et al. (2011): novel and discourse-new vs. familiar and just-named.

The child faced the monitor, and an experimenter (the speaker) stood behind the child and to the left, to reduce looks to the speaker during the task. A camera centered above the monitor recorded children's eye movements.

Procedure

The child and the experimenter performed a referential communication task (Krauss & Weinheimer, 1966). The experimenter introduced a 'secret-card game' (Jin et al., in preparation) in which the experimenter had a deck of cards, and in each trial described one card to the child; the child's task was to point to the matching picture on the monitor. The experimenter introduced the game as follows: "Today we're going to play a game together. I have these secret cards and you'll have two pictures on your screen. I will tell you what I have on my secret card, and you will point to it."



Fluent instruction: Look at the sock. Now point to the sock/bleet.
Disfluent instruction: Look at the sock. Now Point to thee… uh… sock/bleet.

Figure 1. Example test display in Experiment 1.

To create a premise for the distraction manipulation, before the experimenter started the task, she looked at a clock on the wall and said "Oh, the clock is broken. Let me fix this first. Can you wait for me? I'll fix it really quickly." The experimenter fiddled with the clock and let the child help. After a while (approximately a minute later), in the engaged speaker condition, the experimenter fixed the clock and continued the task ("Oh, I fixed it (changing a battery). It's working now. Let's start the game."). In contrast, in the distracted speaker condition, the experimenter said that she could not fix the clock and decided to continue the task and fix the clock simultaneously ("Oh, I don't think I can fix this. Let's start the game. I'll keep trying to fix it. I think I can play the game and fix the clock at the same time."). During the secret-card game, the experimenter in the distracted speaker condition kept fiddling with the clock while she gave instructions to the child. The experimenter's role (distracted vs. engaged) was manipulated between subjects. In both conditions, the speaker was cooperative and attentive to the secret-card game, but in the distracted condition only, the speaker also undertook another task.

The experiment started with two practice trials (containing no disfluency) followed by 16 critical trials. In each trial, the child saw one familiar and one novel image on the screen (Figure 1), and heard 'secret card' instructions consisting of two sentences. The first sentence always referred to the familiar image, and the second could refer either to the familiar image again or to the novel image (e.g., "Look at the sock. Now point to the sock/bleet"). Thus, in the second sentence, reference type was manipulated within subjects: in half of the trials the second sentence named the familiar and just-named image, and in the other half it named the novel and discourse-new image (cf. Kidd et al., 2011; see also Owen, Thacker, & Graham, 2017). The first sentence of each instruction was always produced fluently, and we manipulated disfluency within subjects in the second sentence: the second instruction was fluent for half of the trials (e.g., "Point to the sock/bleet.") and disfluent for the other half (e.g., "Point to thee… uh… sock/bleet"). While the experimenter described her secret card following a script, she read the script as naturally as possible. The location of the target image was counterbalanced across children.

Following the secret-card game, children completed the Peabody Picture Vocabulary Test (PPVT; Dunn & Dunn, 2007).

Coding

The onset of the instructions (e.g., "Point to…") and the critical noun (e.g., "sock") were identified from the video recording of each session, because all critical instructions were produced live. The critical noun was produced on average 530ms (range 303ms to 878ms) after the onset of "Point" in fluent trials and 2169ms (range 1545ms to 3151ms) after the onset of "Point" in disfluent trials (Figures 2 and 4).

Children's eye fixations during the critical instructions (left or right image on the screen, or away from the screen) were coded frame by frame from silent video. Frames in which the child's eyes were not visible (e.g., eyes closed or turned to look at the experimenter) were coded as missing. In each analysis time window (defined below), trials with more than 2/3 frames missing were excluded from the analysis (4.8% in the Distracted speaker condition and 4.1% in the Engaged speaker condition). Missing frames typically reflected the child's own distraction, signaled by looks away from the monitor. The proportion of excluded trials did not differ across conditions; this suggests that children in the two conditions were about equally engaged in the task.

Reliability was assessed for 20% of the participants. Inter-coder reliability was high, with the two coders agreeing on 97.7% of coded video frames.

Predictions

In fluent trials, we predicted that children would readily identify the target image during the second sentence, and that their gaze patterns would not differ between the distracted and engaged speaker conditions in both the pre-noun and noun windows.

In disfluent trials, if children process disfluency differently based on speaker engagement, then during a period of disfluency (e.g., "thee… uh….", before noun onset) they should look more at the novel and discourse-new image than at the familiar and just-named image in the engaged speaker condition than in the distracted speaker condition. In other words, the expectation that "disfluency=new reference" should be attenuated in the distracted speaker condition.

Alternatively, if children do not take the speaker's engagement into consideration in processing disfluency, their gaze should not differ between the two conditions.

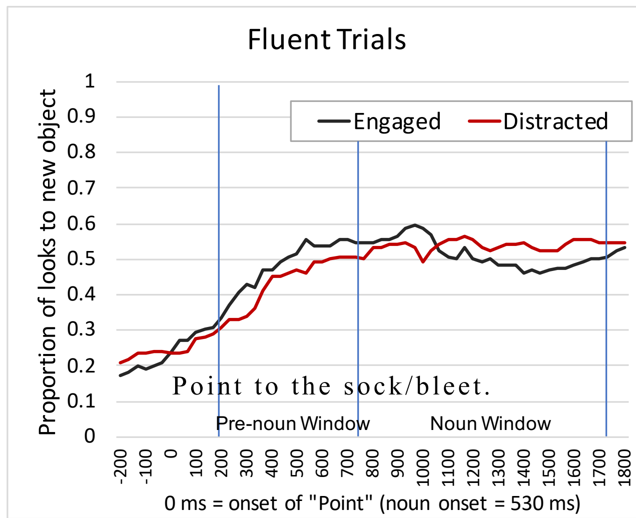


Figure 2. The proportion of fixations to the new object following the onset of "Point" for fluent trials in Experiment 1.

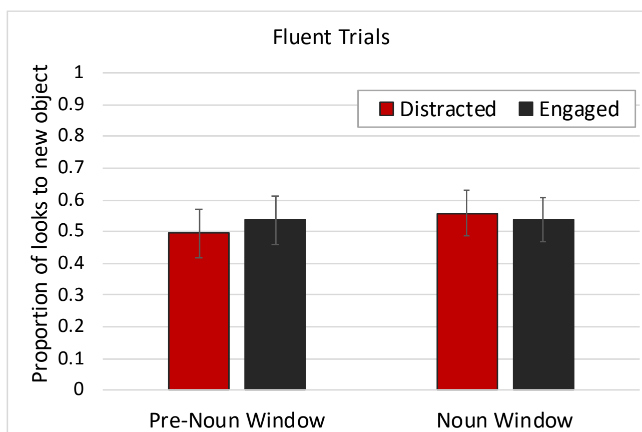


Figure 3. The Proportion of fixations to the new object in each window for fluent trials.

Results

Test trials were analyzed for fluent and disfluent trials separately, as the latency between the onset of "Point" and the critical noun was significantly longer in disfluent trials than in fluent trials, making them difficult to directly compare (see Yoon & Brown-Schmidt, 2014; Jin et al., in preparation; Arizpe et al., 2019).

Fluent Trials Figure 2 shows the proportion of looks to the new object in fluent trials, measured from the onset of 'Point,' the first word of the second instruction. We analyzed children's eye movements in two a-priori time-windows following the analysis in our prior work (Yoon & Brown-Schmidt 2014; Jin et al., in preparation) (Figure 2 and 3): a pre-noun window from 200ms after the onset of "Point" to 200ms after noun onset, and a noun window from 200ms to 1200ms after noun onset. The pre-noun window reflected the processing of referential descriptions before hearing the critical noun, and the noun window reflected the processing of the critical noun. Predictive processing is typically reflected in the pre-noun window, but can also emerge or linger in a later time-window (e.g., Experiment 1 in Jin et al., in preparation; Lew-Williams & Fernald, 2007). Both windows were offset by 200ms to consider the time to program and launch an eye movement (Hallet, 1986).

As Figure 2 shows, looks to the new object started low, because the first instruction had just referred to the familiar object, but began to increase approximately 200ms after the onset of "Point". After the critical noun, fixations to the new object hovered around 50% in both conditions, because the speaker referred equally often to the familiar/just-named image and to the novel/discourse-new image.

The proportion of looks to the novel object was analyzed in a mixed-effects model with a Gaussian link function with subjects and item as random effects. In all analyses, models were fit using the lmer package in R, with the maximal random effects structure for subjects and items. In cases where the maximal model did not converge, a backwards-fitting procedure was used to identify the model with the largest random effects structure that would converge (see Barr., et al., 2013). The model included speaker (distracted vs. engaged) and time-window (pre-noun vs. noun window) as fixed effects. The dependent measure was the proportion of looks to the novel object. The model (Table 1) revealed no main effect of time-window ($t=0.71$, $p>.05$.) or speaker condition ($t=0.69$, $p>.05$.), and no interaction between time-window and speaker ($t=-0.77$, $p>.05$).

Consistent with prior findings (Jin et al., in preparation), in fluent trials children quickly identified the target after hearing the critical noun. Speaker engagement did not affect children's interpretation of referential expressions in fluent trials.

Disfluent Trials Figure 4 shows the proportion of looks to the new object during the second instruction in disfluent trials. Children looked more at the new object before the onset of the critical noun in the engaged than in the distracted speaker condition, suggesting that speaker engagement affected children's processing of disfluency.

In disfluent trials, the onset of the critical noun was on average 2,169ms after the onset of "Point". Because the latency between the onset of "Point" and the critical noun was longer for disfluent trials than fluent trials, we analyzed children's eye movements in three time-windows that were determined a-priori based on the prior work (Figure 4 and 5): two pre-noun windows that equally divided the time from 200ms after the onset of "Point" to 200ms after noun onset (pre-noun window 1: 200ms – on average 1,285 ms after the onset of "Point"; pre-noun window 2: on average 1,285ms – 2,369ms), and a noun window from 200ms to 1200ms after noun onset. The two pre-noun windows assessed children's processing of the disfluency prior to hearing the critical noun, and the noun window reflected the processing of the critical noun.

Table 1: Proportion of looks to novel objects in fluent trials. Mixed effect model with speaker (distracted (-0.5) vs. engaged (0.5)) and time window (pre-noun vs. noun) as fixed effects. The pre-noun window is treated as baseline.

	Estimate	SE	t-value	p-value
(intercept)	0.48	0.09	5.25	<.001
speaker	0.12	0.17	0.69	0.49
window	0.03	0.05	0.71	0.48
speaker*window	-0.07	0.09	-0.77	0.45
Random effects	Variance	SD		
subject				
(intercept)	0.14	0.37		
window	0.03	0.16		
item				
(intercept)	0.02	0.13		
speaker	0.003	0.05		
window	0.004	0.06		
Residual	0.15	0.39		

A mixed-effect model that included speaker (distracted vs. engaged) and time-window as fixed effects was used to examine children's processing of disfluent expressions (Table 2). The dependent measure was the proportion of looks to the novel image. The model revealed a significant main effect of speaker ($t=2.09$, $p<.05$) and a significant interaction of speaker and time-window (-2.10, $p<.05$). The main effect of window was not significant ($t=-0.92$, $p>.05$).

Separate planned analyses explored the significant interaction between speaker and time-window. During the first pre-noun window, a significant effect of speaker ($t=2.08$, $p<.05$) showed that children looked at the novel image more in the engaged speaker condition than in the distracted speaker condition. This speaker effect was not significant in the second pre-noun window ($t=0.82$, $p>.05$) or the noun window ($t=-0.31$, $p>.05$). The early disfluency effect that emerged in the first half of the disfluency is consistent with our prior finding in Jin et al. (in preparation).

This result suggests that children interpret disfluency differently depending on speaker engagement. When the speaker was fully engaged in the conversation, children attributed the speaker's disfluency to speech planning difficulty, and expected reference to an image that was difficult to name (novel and discourse-new). In contrast, when the speaker was distracted, children attributed the speaker's disfluency to a different source, their distraction, and as a result the disfluency effect (disfluency=new image) was attenuated. This speaker engagement effect emerged in the first half of the disfluency and disappeared in the noun window.

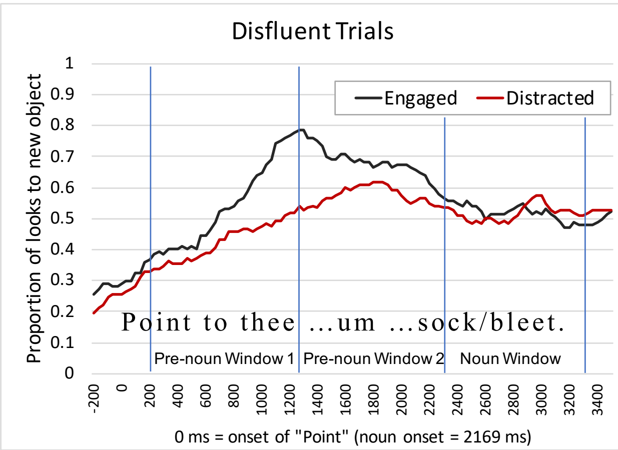


Figure 4. The proportion of fixations to the new object following the onset of "Point" for disfluent trials in Experiment 1.

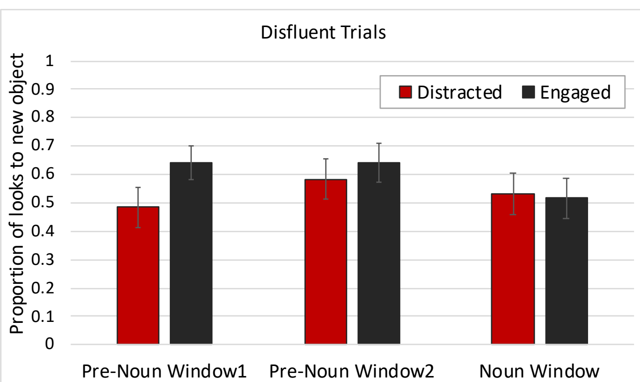


Figure 5. The Proportion of fixations to the new object in each window for disfluent trials.

Discussion

The results show that different sources of a speaker's disfluency modulate children's online processing of disfluency. When the speaker was disfluent, 4-year-old children looked at a novel and discourse-new object more than a familiar object if the speaker was fully engaged in the task, but their gaze was more equally distributed between the two objects when the speaker was distracted. Evidence from early eye-gaze revealed that children expected a novel referent as soon as they heard disfluent expressions in the engaged speaker condition, but this expectation was attenuated in the distracted speaker condition. This finding suggests that 4-year-old children understand that disfluencies in speech can reflect multiple sources of difficulty.

An alternative explanation for our findings might be that children paid less attention to the task in the distracted speaker condition. However, children's engagement in the task was not different between the engaged and distracted speaker conditions, reflected in the proportion of looks away (1.6% vs. 1.7%). This finding weighs against the possibility that children might pay less attention to the task when the experimenter was distracted versus engaged, and corroborates the claim that children flexibly attribute disfluency to situation-specific sources and process it accordingly.

Table 2: Proportion of looks to novel objects in disfluent trials. Mixed effect model with speaker (distracted (-0.5) vs. engaged (0.5)) and time window (pre-noun1 vs. pre-noun2 vs. noun) as fixed effects. The pre-noun window1 is treated as baseline. Values in bold indicate significant results.

	Estimate	SE	t-value	p-value
(intercept)	0.61	0.06	9.91	<.001
speaker	0.25	0.12	**2.09**	**0.04**
window	-0.02	0.02	-0.92	0.37
speaker*window	-0.09	0.04	**-2.13**	**0.04**
Random effects	Variance	SD		
subject				
(intercept)	0.07	0.26		
window	0.01	0.07		
item				
(intercept)	0.005	0.07		
speaker	<.001	0.02		
window	<.001	0.03		
Residual	0.13	0.37		

This striking finding suggests that children make situation-specific inferences about the likely source of disfluencies and use these inferences in online language processing. In our study, a critical feature of the distracted-speaker manipulation was that the speaker was still cooperative in the task. A few studies have examined how attention modulates the establishment of shared knowledge in adults' conversation (Craycraft & Brown-Schmidt; 2018; Rosa et al., 2015). These studies have shown that partners' attention modulates conversational language processing, such that speakers tend not to assume they have successfully established shared knowledge when the listener was inattentive. In these studies, an inattentive conversational partner repeatedly checked their cell phone or looked around the room, appearing significantly disengaged from the conversation. In the present study, we presented a cooperative but distracted speaker who did not violate Grice's cooperative principle (Grice, 1975). Even in the distracted speaker condition, the speaker was attentive to the secret-card communication task, though she also kept busy with a concurrent task. Thus, in our study, children were encouraged to attribute the speaker's disfluency to a situation-specific source even when interacting with a distracted speaker.

An open question is whether the disfluency effect (disfluency=novel objects) shown in the engaged speaker condition was driven by image novelty or discourse status. In our study, the novel object on the screen was both discourse-new and a hard-to-name object (Kidd et al., 2012) and it is less clear which factor (discourse-new vs. novelty) drove the effect (see also Owen, Thacker, Graham, 2017). Another open question is whether the live interaction we adopted in our study played an important role in children's sophisticated inferences. Adult listeners are more likely to take their conversational partners' perspective in an interactive conversational setting than in a non-interactive setting (e.g., talking/listening in front of a computer screen, presenting a photo of a conversational partner) (Brown-Schmidt, 2009; Horton & Spieler, 2007; Yoon & Stine-Morrow, 2019). It is known that interlocutors can utilize more available social cues, such as a partner's prompt feedback (e.g., providing backchannels, nodding), during interactive conversation. In future work we hope to examine whether children, like adults, can benefit from interactive social interactions in their online language processing.

In conclusion, we have shown that children are sensitive to the source of a speaker's disfluency and flexibly adjust how they process disfluency accordingly. Rather than processing disfluency based on simple associations between disfluency and novel referents, our findings point to a high degree of sophisticated inferences children can make during the online processing of disfluency. Different sources of disfluency modulate children's predictions about upcoming referents; when the speaker was engaged in conversation, they predicted novel referents following the speaker's disfluency whereas this prediction was attenuated when the speaker was distracted. Our results expand previous findings and show that children can appreciate different sources of disfluency and flexibly process it in live conversation.

References

Arizpe, J. M., Noles, D. L., Tsao, J. W., & Chan, A. W. (2019). Eye movement dynamics differ between encoding and recognition of faces. Vision (Basel, Switzerland). DOI:10.3390/vision3010009.

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.

Arnold, J. E., Tanenhaus, M. K., Altmann, R., & Fagnano, M. (2004). The old and thee, uh, new. *Psychological Science, 15,* 578–582.

Baldwin, D. A. (1991). Infants' contribution to the achievement of joint reference. *Child Development, 62*, 875-890.

Baldwin, D. A. (1993). Early referential understanding: Infants' ability to recognize referential acts for what they are. *Developmental Psychology, 29*(5), 832–843.

Barch, D., & Berenbaum, H. (1994). The relationship between information processing and language production. *Journal of Abnormal Psychology, 103,* 241–250.

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

Bortfeld, H., Leon, S. D., Bloom, J. E., Schober, M. F., & Brennan, S. E. (2001). Disfluency rates in spontaneous speech: Effects of age, relationship, topic, role, and gender. *Language and Speech, 44,* 123-147.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44,* 274-296.

Brown-Schmidt, S. (2009). Partner-specific interpretation of maintained referential precedents during interactive dialog. *Journal of Memory and Language, 61,* 171-190.

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.

Craycraft, N. N., & Brown-Schmidt, S. (2018). Compensating for an inattentive audience. *Cognitive Science, 42*, 1504-1528.

Dunn, L. M., & Dunn, D. M. (2007). *Peabody picture vocabulary test* (4th ed.). Circle Pines, MN: American Guidance Service.

Ferreira, F. (1991). Effects of length and syntactic complexity on initiation times for prepared utterances. *Journal of Memory and Language, 30,* 210–233.

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.

Fraundorf, S. H., & Watson, D. G. (2013). Alice's adventures in um-derland: Psycholinguistic sources of variation in disfluency production. *Language and Cognitive Processes. 29,* 1083-1096.

Grice, Paul (1975). "Logic and conversation". In Cole, P.; Morgan, J. (eds.). *Syntax and semantics*. 3: Speech acts. New York: Academic Press. pp. 41–58.

Hallett, P. E. (1986). Eye movements. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 10.1–10.112). New York, NY: Wiley.

Heller, D., Arnold, J. E., Klein, N., & Tanenhaus, M. K. (2014). Inferring difficulty: Flexibility in the real-time processing of disfluency. *Language and Speech, 58,* 190-203.

Horton, W. S., & Spieler, D. H. (2007). Age-related differences in communication and audience design. *Psychology and Aging, 22,* 281-290.

Jaswal, V. K., & Malone, L. S. (2007). Turning believers into skeptics: 3-year-olds' sensitivity to cues to speaker credibility. *Journal of Cognition and Development, 8,* 263-283.

Jin, K., Yoon, S. O., Brown-Schmidt, S., & Fisher, C. L. (in preparation). What's new to you? Preschoolers' online attribution of disfluency.

Kidd, C., Katherine, S. W., & Aslin, R. N. (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. Developmental Science, 14, 925-934.

Krauss, R. M., & Weinheimer, S. (1966). Concurrent feedback, confirmation, and the encoding of referents in verbal communication. *Journal of Personality and Social Psychology, 4,* 343–346.

Lew-Williams, C., & Fernald, A. (2007). Young children learning Spanish make rapid use of grammatical gender in spoken word recognition. *Psychological Science, 18,* 193-198.

Oomen, C.C.E. & Postma, A. (2001). Effects of divided attention on the production of filled pauses and repetitions. *Journal of Speech, Language, and Hearing Research, 44,* 99701994.

Orena, A. J., & White, K. S. (2015). I forget what that's called! Children's online processing of disfluencies depends on speaker knowledge. *Child Development, 86,* 1701-1709.

Owens, S. J., & Graham, S. A. (2016). Thee, uhh disfluency effect in preschoolers: A cue to discourse status. *British Journal of Developmental Psychology, 34,* 388-401.

Owens, S. J., Thacker, J. M., & Graham, S. A. (2018). Disfluencies signal reference to novel objects for adults but not children. *Journal of Child Language, 45,* 581-609.

Rosa, E. C., Finch, K. H., Bergeson, M., & Arnold, J. E. (2015). The effects of addressee attention onprosodic prominence. *Language, Cognition and Neuroscience, 30,* 48–56.

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

Thacker, M. T., Chambers, C. G., & Graham, S. A. (2018). When it is apt to adapt: Flexible reasoning guides children's use of talker identity and disfluency cues. *Journal of Experimental Child Psychology, 167,* 314-327.

Yoon, S. O., & Brown-Schmidt, S. (2014). Adjusting conceptual pacts in three-party conversation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 40,* 919-937.

Yoon, S. O., & Stine-Morrow, E. A. L. (2019). Evidence of preserved audience design with aging in interactive conversation. *Psychology and Aging, 34,* 613-623.