

That's thee, uuh blicket! How does disfluency affect children's word learning?

First Language

1–18

© The Author(s) 2019

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/0142723719873499

journals.sagepub.com/home/fla

Katherine S. White , Elizabeth S. Nilsen,
Taylor Deglint and Janel Silva

University of Waterloo, Canada

Abstract

Disfluencies, such as 'um' or 'uh', can cause adults to attribute uncertainty to speakers, but may also facilitate speech processing. To understand how these different functions affect children's learning, we asked whether (dis)fluency affects children's decision to select information from speakers (an explicit behavior) and their learning of specific words (an implicit behavior). In Experiment 1a, 31 3- to 4-year-olds heard two puppets provide fluent or disfluent descriptions of familiar objects. Each puppet then labeled a different novel object with the same novel word (again, fluently or disfluently). Children more frequently endorsed the object referred to by the fluent speaker. We replicated this finding with a separate group of 4-year-olds in Experiment 1b ($N = 31$) and a modified design. In Experiment 2, 62 3- to 4-year-olds were trained on new words, produced following a disfluency or not, and were subsequently tested on their recognition of the words. Children were equally accurate for the two types of words. These results suggest that while children may prefer information from fluent speakers, they learn words equally well regardless of fluency, at least in some contexts.

Keywords

Disfluency, language processing, selective learning, speaker reliability, word learning

Introduction

One of the most challenging aspects of language learning and processing is that a single element can have multiple functions. For example, a word or a morpheme can have more than one meaning (e.g., the word 'bank'; the plural noun ending, 3rd person verb

Corresponding author:

Katherine S. White, Psychology Department, University of Waterloo, 200 University Avenue West, Waterloo, ON N2L 3G1, Canada.

Email: white@uwaterloo.ca

ending, and possessive 's'). Similarly, prosody can convey information about emotion, information structure, syntactic structure, and word identity (in lexical stress and tonal languages). One speech cue that has a wide range of functions is disfluency (e.g., filled pauses such as 'um' or 'uh'), which occurs on average 6 out of every 100 words in speech between adults (Fox Tree, 1995; Shriberg, 2001). Although speech directed to children is more fluent (Newport, Gleitman, & Gleitman, 1977), disfluencies still occur (Kidd, White, & Aslin, 2011b). Disfluencies convey information about a speaker's certainty in the information being expressed (Brennan & Williams, 1995), are involved in the coordination of turn-taking behavior (e.g., by holding a turn while information is being retrieved (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Clark & Fox Tree, 2002), and reflect planning or retrieval difficulty (Bortfeld et al., 2001; Fox Tree & Clark, 1997; Schacter, Christenfeld, Ravina, & Bilous, 1991). Thus, the processing of disfluencies taps into multiple competencies underlying successful language processing and communication.

Adult listeners are sensitive to the various functions of disfluencies. First, disfluencies appear to signal processing difficulty, leading listeners to make predictions about what a person will say following a disfluency (i.e., something that would be difficult in that situation). For example, it is generally more difficult to retrieve object names that are less familiar or have not been mentioned in recent discourse. Adults look more towards these sorts of referents when they hear a disfluent instruction relative to when they hear a fluent instruction (Arnold, Fagnano, & Tanenhaus, 2003; Arnold, Hudson Kam, & Tanenhaus, 2007; Arnold, Tanenhaus, Altmann, & Fagnano, 2004). Disfluencies also heighten adults' attention and memory for information. For instance, they are faster at identifying a target word when it is preceded by a disfluency (at least for some disfluencies, such as 'uh'; Fox Tree, 2001). This may be for a variety of reasons, such as the presence of an additional segmentation cue, heightened attention resulting from the disfluency itself, or the extra time afforded to processing (Corley & Hartsuiker, 2011). Finally, adults are better at remembering material preceded by disfluencies in recognition tasks (Corley, Macgregor, & Donaldson, 2007; Corley & Stewart, 2008; Fraundorf & Watson, 2011).

While the above studies identify the ways disfluency can benefit adults' word processing, listeners make (generally) negative inferences about speakers on the basis of disfluency. Among other social attributions, highly disfluent speakers are rated as less articulate, less emotionally and communicatively competent, and less intelligent (Amick, Chang, Wade, & McAuley, 2017; Christenfeld, 1995; Von Tiling, 2011). Furthermore, adult listeners judge speakers as less confident and less likely to know an answer when their utterances are preceded by filled pauses (Brennan & Williams, 1995; Christenfeld, 1995). Therefore, although disfluencies cause increased attention and memory for linguistic material, they can also cause negative attributions about a speaker and lower a listener's confidence in the speaker's knowledge about specific material.

Previous research into the influence of disfluency has tended to separate online processing effects from disfluency-based attributions about speakers or the information they are conveying. To our knowledge, no previous research has explicitly tested listeners' learning of, or memory for, *new* information that is preceded by a disfluency. (In past studies investigating memory for words, adults heard previously known words;

Corley et al., 2007; Corley & Stewart, 2008). It is unclear what might happen in learning situations where disfluency is potentially playing various roles (e.g., signaling uncertainty, increasing attention). In the present study, we explore this question in preschool-age children.

Children have shown sensitivity to some of the functions of disfluencies. For example, children as young as 2 years show heightened attention and look at more difficult to access referents in the presence of a disfluency, prior to hearing the referent label (Kidd, White, & Aslin, 2011a; Morin-Lessard & Byers-Heinlein, 2019; Orena & White, 2015; Owens & Graham, 2016). Like adults, they also modulate their predictive use of disfluencies based on a speaker's characteristics or their pattern of disfluency use (Orena & White, 2015; Thacker, Chambers, & Graham, 2018; but see Thacker, Chambers, & Graham, 2018b for limitations on this ability). However, the types of implicit predictions children make during disfluency may be narrower than for adults, in that they are less likely than adults to anticipate a novel target on the basis of a disfluency (Owens, Thacker, & Graham, 2018). Although disfluencies have been shown to affect children's predictions about upcoming referents, there is no research yet addressing the effects of disfluencies on children's word learning. In particular, it is not clear whether the presence of disfluency in a learning situation would promote better retention of a label, as a result of increased attention.

At the level of attributions about speakers, while adults clearly view disfluencies as indicative of an unreliable source, the literature on children's perception of disfluencies is sparse. Giolas and Williams (1958) found that young children's choice of a potential teacher was negatively affected by the presence of disfluencies during the teacher's story reading. In another study, children older, but not younger, than 4 years judged that a disfluent speaker did not speak well and showed a preference for playing with the fluent speaker (Ezrati-Vinacour, Platzky, & Yairi, 2001). Thus, by the age of 4, children show social preferences for fluent speakers. However, it is not known whether they view these speakers as more reliable informational sources and preferentially learn from them over more disfluent speakers.

A growing body of research has revealed that children are sensitive to a number of static (e.g., age, expertise, language background) and dynamic (e.g., expressions of ignorance, certainty) characteristics of their conversation partners and that these characteristics affect children's evaluation of these individuals as information sources. One particularly strong contributor to children's selective learning is a speaker's history of accuracy (e.g., Birch, Vauthier, & Bloom, 2008; Koenig, Clément, & Harris, 2004; Koenig & Harris, 2005; Koenig & Sabbagh, 2013). They also consider a potential informant's degree of certainty, as expressed either verbally or nonverbally. For example, 3- to 4-year-old children are more likely to endorse the information provided by a speaker who says, 'I know it's in the red box' than a speaker who says, 'I think/guess it's in the blue box' (Jaswal & Malone, 2007; Sabbagh & Baldwin, 2001; Tenney, Small, Kondrad, Jaswal, & Spellman, 2011). In addition, children as young as 2 years imitate novel actions performed by a visibly confident actor (e.g., upright posture) over those of an uncertain actor (e.g., shrugged shoulders; Birch, Akmal, & Frampton, 2010; Brosseau-Liard & Poulin-Dubois, 2014). Thus, children track the accuracy and certainty of particular individuals over time, and they show preferential learning from those who have

been accurate and confident. However, the cues investigated thus far involve nonverbal signals and explicit verbal signals, typically in combination (such as incorrect answers, declarations of ignorance ['I don't know'], shrugged shoulders, and facial expressions of uncertainty), as opposed to speech disfluency in isolation. Although children demonstrate sensitivity to other highly salient features of speech (showing strong preferences for learning from native over non-native accented speakers; Kinzler, Corriveau, & Harris, 2011), it is not clear whether children use disfluency alone as a cue to reliability.

Present study

The overarching question addressed by the present study is how disfluencies affect children's word learning. We examined this in two ways, using two different methodologies. First, in Experiment 1, we asked whether fluency affects who children choose to learn new words from. If children consider disfluent speakers to be uncertain and thus less reliable, they should choose to learn new words from fluent speakers. Second, in Experiment 2, we asked whether the fluency with which a single speaker presents information affects children's learning. If disfluency causes heightened attention and better memory, words produced disfluently could be learned better than words produced fluently. However, if disfluency is interpreted as speaker uncertainty with that specific label, these words could be learned less well.

Experiment 1a: speaker choice

Participants

The data from 31 children aged 3–4 years (19 female; mean age = 47.7 months; range 36.4–59.5 months) from the Kitchener-Waterloo region of Canada were included in this study (16 3-year-olds aged 36 to 47.9 months; 16 4-year-olds aged 48 to 59.9 months). An additional four participants were tested, but their data were not included due to non-completion (3) and performance on a standardized measure that indicated insufficient receptive language skills for understanding the familiar labels used in the task (1) (Wechsler Preschool and Primary Scale of Intelligence – Third Edition; Wechsler, 2002). Children had no known auditory or visual conditions; their parents gave informed consent for their participation. Children were given a small toy, t-shirt, or book as a token of appreciation. This experiment, as well as the two following experiments, were approved by a Research Ethics Committee at the University of Waterloo.

Materials

Familiarization audio. Two female native English speakers recorded audio sequences in which three familiar objects were named. There were four sets of familiar objects: plane, spoon, cow; flower, book, ball; bear, car, banana; and clock, dog, balloon. Each speaker produced both fluent and disfluent utterances for these four sets of objects (e.g., 'This is a plane; this is a spoon; this is a cow'; 'This is a uhhh plane; this is a uhhh spoon; this is a uhhh cow').



Figure 1. Still shot of test video for Experiment 1.

Test audio. For the test stimuli, the same two speakers produced utterances in which a novel object was labeled. Again, both speakers produced a fluent and disfluent version of each utterance (e.g., ‘This is a mido/This is a uhhh mido’). There were four novel labels: mido, blicket, dosa, taki. A single disfluency from each speaker was spliced in for all of that speaker’s disfluent utterances (familiarization and test). The length of the disfluencies for the two speakers were approximately 1300 ms and 1500 ms.

Familiarization video. Video recordings were made of four puppets, two with brown hair and two with blond hair (the puppets with the same hair color had different hair styles and clothing styles). Each video contained one of the puppets moving behind a table and pointing to the three familiar objects while mouthing ‘This is a (uhhh) X’. Recordings were made of each of the four puppets presenting two different sets of familiar objects, with both fluently and disfluently mouthed utterances (different videos), for a total of 16 videos. These 16 videos were then combined such that each trial contained a fluent speaker and a disfluent speaker labeling the same objects, with four fixed pairings of speakers (AB, CD, AD, BC).

Test video. Each test video contained the two puppets for that trial (either A and B, C and D, A and D, or B and C), each standing behind a different novel object (across the four trials, there were eight unique novel objects). One puppet pointed to the object in front of them and labeled it; the second puppet then pointed to the other object and labeled it with the same label (see Figure 1). For each of the puppet pairings, eight videos were recorded. These videos differed in the novel object labeled by each puppet (each puppet pair always appeared with the same two objects, but the object assigned to each puppet differed), the puppets’ fluency, and which puppet spoke first. Children were assigned to one of these counterbalanced conditions.

The recorded audio was overlaid onto these familiarization and test videos. (The audio was recorded first, and the puppet recordings were done in synch with the audio to ensure that the timing was as natural as possible.)

Design and procedure

The stimuli were presented on a laptop. Each child completed four trials (puppet combinations AB, CD, AD, BC, in this order). Each puppet's fluency remained consistent in the two trials in which they appeared.

After a warm-up trial in which children watched a video of a familiar object (i.e., a frog) being labeled fluently by a male puppet and then had to point to the object in an array of four familiar objects, children were simply told they would watch a video. No additional information was provided. For each trial, children first viewed the familiarization videos in which one puppet labeled familiar objects (disfluently or fluently) and then another puppet labeled the same objects (disfluently or fluently). Children then viewed the test video where the two puppets were side by side and each provided the same label for a different novel object. The video was then paused by the experimenter (with the puppets and test objects still on the screen). Children were asked to point to the object that was the referent of the novel label used on that trial (e.g., 'Which one is the mido?').

Results

Data were subjected to a $2(\text{age group}) \times 2(\text{trial pair})$ ANOVA. Age group was included as a between-group factor in the analyses as past work has demonstrated differences in judgments of speaker reliability between 3- and 4-year-olds (Koenig & Harris, 2005). Further, as each puppet was viewed twice, analyses included trial pair as a check to see if repeated exposure had an influence on children's choices. The dependent variable was the speaker's choice (0 for a disfluent speaker, 1 for a fluent speaker for each trial, for a total possible score of 2 for each trial pair). Means are displayed in Figure 2, left columns.

There was no significant effect of age ($p = .721$) and no interaction between age and trial pair ($p = .399$). However, there was a significant effect of trial pair, wherein fluent speakers were chosen more often in the first trial pair ($F(1,29) = 5.74, p = .023$). One-sample t -tests against chance demonstrated that children showed a preference for the fluent speaker in the first two trials ($M = 1.26, SD = .58; t(30) = 2.49, p = .018$), but not in the second two trials ($M = .84, SD = .52; t(30) = -1.72, p = .096$).¹ In fact, individual trial means (where chance = .5) showed a distinct change in performance between trials two and three (.63, .63, .41, .44). Thus, in the first two trials, children demonstrated sensitivity to fluency (choosing labels from the fluent speakers), but there was a reversal in their choices once they saw the same puppets again (albeit non-significant). It may have been the case that seeing the same puppets again led them to change their decisions, potentially as an attempt to maximize their chances of being accurate. Thus, we consider their performance on the first two trials, when they saw each puppet for the first time, to be the most representative of their preferences.

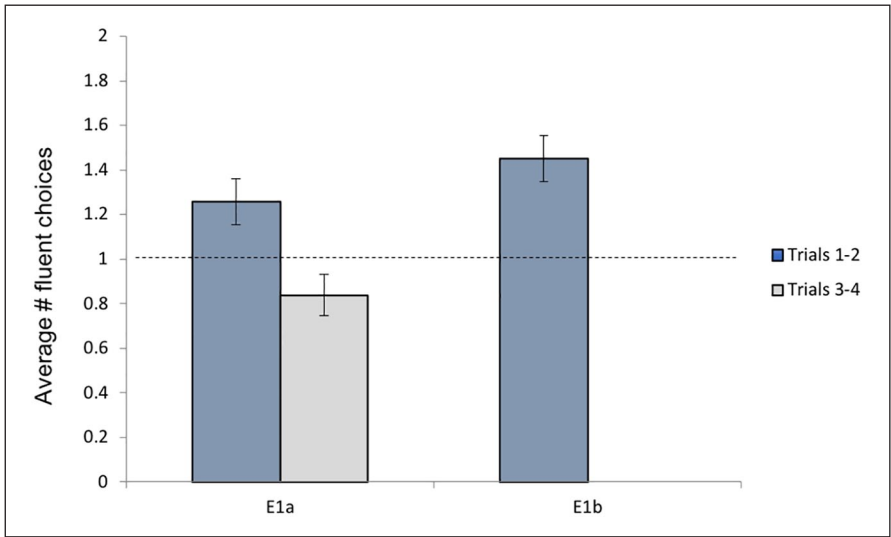


Figure 2. Choices in Experiment 1, out of 2. A higher value reflects more choice of the fluent speaker (the dotted line at 1 represents chance). Error bars represent standard error.

Experiment 1b

The results of Experiment 1a suggest that 3- to 4-year-old children are sensitive to a speaker’s fluency and use this to evaluate them as an information source. However, this conclusion was based on an analysis that included trial pair. To determine whether this finding is robust, we conducted a replication of this experiment.

Participants

Thirty-one children aged 4 years (14 female; mean age = 52.8 months; range = 48.2–59.7 months) from the Kitchener-Waterloo community were tested in this study. Given the lack of age difference in Experiment 1a, we tested only a single age group in this experiment. As in Experiment 1a, children had no known auditory or visual conditions; their parents gave informed consent for their participation. Children were given a small toy, t-shirt, or book as a token of appreciation.

Materials

Given the change in behavior between the first and second pair of trials in Experiment 1a, we presented children with only two trials (this way, there was no repetition of individual speakers). This meant that only half of the familiar and novel objects used in Experiment 1a were presented. In this experiment, participants always saw puppet A paired with puppet B, and puppet C paired with puppet D. However, as in Experiment 1a, the fluency of the puppets during familiarization, the order in which the puppets spoke, and the novel object assigned to each puppet were counterbalanced across participants.

Design and procedure

Participants took part in the same warm-up trial as Experiment 1a prior to the familiarization trials. We made two modifications to the familiarization presentation compared to Experiment 1a, to make our procedures consistent with past work on speaker reliability when speech cues have been manipulated (e.g., Sobel & Macris, 2013). First, the two novel objects for each trial were presented in a static image before the puppets appeared with the familiar objects. During this static presentation, the children were provided with more detailed instructions than in 1a: *'Here are two new toys. We have to figure out what one of these toys is called. First, we're going to listen to two girls tell us about some things, and then you can pick one of them to help us figure out the name of the new toy! One of these girls knows a lot and the other one doesn't. You can only ask one of the girls, so let's pay attention and listen carefully!'* Following this, the labeling of familiar and novel objects proceeded as in Experiment 1a. Second, to place the emphasis on the knowledge of the speakers, a small modification was made to the test question in that children were asked *'So, which girl do you think was right about the X?'*

Additionally, so as to explore whether children noticed the disfluencies presented, children were asked if there was anything funny about the way the girls were talking. This question was only asked after the second trial so that children would not be explicitly cued to speaker differences when making their choices.

Results

See Figure 2, rightmost column, for mean performance. We conducted a one-sample *t*-test against chance (1). This test revealed that children showed a preference for the fluent speaker ($t(30) = 4.43, p < .001$). Children who explicitly pointed out the disfluency (i.e., mentioned *'uhhh'* or *'umm'*) when asked if there was anything funny about the way the girls were talking, showed a greater preference for fluent speakers ($M = 1.73, SD = .47$) than those who did not report anything about the disfluency ($M = 1.3, SD = .57; t(29) = 2.12, p = .042$). Moreover, three of the 10 participants who explicitly noted the disfluency stated that this meant the girl *'didn't know as much'*.

Experiment 2: word learning

Experiment 1 shows that 3- to 4-year-olds choose to obtain information about novel words from fluent speakers. Therefore, when given a choice between informants, children prefer to learn from fluent speakers. However, disfluencies are a relatively common feature of speech, and all speakers produce them (Fox Tree, 1995; Shriberg, 2001). How do these disfluencies affect learning? In Experiment 2, we asked whether fluency affects the learning of specific information (e.g., whether disfluently produced words are privileged in learning or memory). To do this, we provided children with brief training on novel words (produced either fluently or disfluently) in an interactive task and subsequently tested their recognition of these words in a preferential looking task.

Method

Participants. The data from 62 children aged 3–4 years (37 female; mean age = 47.4 months; range = 37.2–59.8 months; 32 3-year-olds) from the Kitchener-Waterloo community were analyzed in this study. An additional 17 participants were dropped due to: the experimenter's failure to produce disfluencies and training stimuli according to the familiarization scheme (3), children's failure to follow instructions during the familiarization task (4), experimenter failure to record the test session (1), hyperactivity and lack of interest during the test session (3), or failure to look at both objects during the baseline portion for more than four of the test trials (6). Children had no known auditory or visual conditions and their primary language was English (i.e., no more than 20% of another language spoken at home). Parents gave informed consent for their children's participation and all children received a book or t-shirt as a token of appreciation.

Materials

Familiarization objects. Ten small objects were used: two objects for the initial practice trial (a block and a toy banana), four familiar objects (an infant sippy cup, a spoon, a toy car, and a toy apple), and four unfamiliar objects that were each given novel names (dax, blicket, tanzer, glark). Still images of these objects were used during the word recognition test.

Procedure

Familiarization. The familiarization phase consisted of an interactive sorting game. Children sat across from the experimenter. Two empty bins were placed in front of the child on each side of the table. The toys used in the sorting game were placed in a third bin under the table beside the experimenter out of view of the child. Children were told they were going to play a sorting game and the experimenter would tell them which of the two toys on the table needed to go in each box. They first completed a practice trial with the block and banana. After the practice trial, the experimenter placed all eight remaining objects in a row, said they were her favorite toys, and proceeded to label each one fluently, thereby providing at least one clear labeling event per object. The objects were then placed in front of the child one pair at a time. Each pair included one familiar and one unfamiliar object (pairings were fixed across children). The experimenter then asked the child to select the novel object: 'Look at the [object name]. Can you put the [object name] in the box?' Children uniformly selected the novel object, as expected given their bias to map novel labels to novel objects (Markman, Wasow, & Hansen, 2003; Merriman & Bowman, 1989).

During this sorting game, instructions for two of the four novel objects were produced disfluently ('Look at thee, uh [object name]. Can you put thee, uh [object name] in the box?'). The audio from the experimenter was recorded to ensure that disfluencies were produced in the intended positions and that each word was repeated exactly twice during the familiarization. The average length of the disfluencies produced was 810 ms, which is consistent with natural speech (Shriberg, 2001). The labeling of the familiar objects was always fluent and always followed the novel object in the pair. The name given to a particular novel object was fixed across participants. However, the assignment of each novel object to fluency condition (fluent or disfluent) was counterbalanced across two

groups of participants. In addition, object pairs were presented in two fixed orders during familiarization (either Fluent-Disfluent-Fluent-Disfluent or Disfluent-Fluent-Disfluent-Fluent). Therefore, there were four counterbalancing groups.

After the objects were sorted, the child, parent, and experimenter immediately moved down the hallway from the playroom to the test room to complete the looking task.

Test. Participants' word recognition was assessed using the intermodal preferential looking procedure. Children sat on their parents' lap in front of a visual display (36 in × 21 in) approximately 1.5 ft in front of them. Parents listened to music over noise-reducing headphones. Children's looking was recorded with a video camera, located under the display, which was linked to an external monitor and video recorder in an adjacent room for experimenter-viewing purposes and later off-line coding.

Each trial began with the presentation of two objects on the screen (in silence) for 3 seconds. With the objects still on the screen, an audio recording instructed children to look at one of the objects. Instructions were recorded by a female speaker (not the experimenter) in a child-friendly speech style and presented at approximately 65 dB. For familiar target trials, children were asked, 'Do you see a [familiar object]?'; for novel target trials, children were asked to 'Look at the [unfamiliar object]' or 'Find the [unfamiliar object]'. All labels were produced in fluent sentences, and different sentences were used to maintain children's interest. During the looking task, novel objects were paired with novel objects and familiar objects with familiar objects. These pairings prevented participants from determining the referents of the novel words during the test phase (by mapping novel labels to novel objects). For the novel objects, each object appeared with one object of the same fluency type twice (i.e., both had been labeled fluently during familiarization or both had been labeled disfluently) and one object of a different fluency type twice (i.e., one had been labeled fluently during familiarization and one labeled disfluently), for a total of eight trials. These eight trials were arranged in two blocks. Each pairing occurred once per block, and for each pairing the target object switched between the two blocks. The familiar object pairs served as fillers to maintain children's attention in the task. The familiar objects had fixed pairings that were repeated twice (for a total of four familiar trials). Again, the target object in these pairings switched between blocks. For all participants, targets appeared half of the time on the left and half of the time on the right side of the screen. Children were assigned to one of four counterbalancing lists (differing in the side of the objects and trial order). Between each trial, a brief attention-getter (cartoon) played until the experimenter advanced to the next trial.

A timeline of individual test trial events is depicted in Figure 3. Presentation of the stimuli was programmed in PsyScope X (Cohen, MacWhinney, Flatt, & Provost, 1993). Given that each trial lasted 8.5 seconds, the total test was approximately 2–2.5 minutes.

Video coding. Children's looking during test trials was coded frame-by-frame (33 ms frames) for each trial using customized software by trained coders blind to trial type. Each test trial had two 3000 ms components: (1) the 3000 ms silent period of time prior to the onset of the instruction (baseline phase) and (2) the 3000 ms period following the onset of the target word in the instruction (naming phase). This naming phase started 230 ms after the onset of the target word to account for motor planning, as is convention in both child

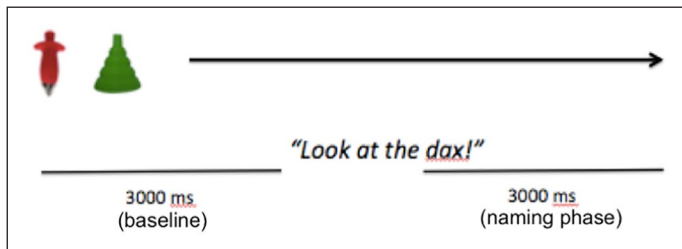


Figure 3. Timeline for test trials in Experiment 2. Novel labels appeared in two sentence frames: ‘Look at the X’ and ‘Find the X’. The average duration of ‘Look at the’ (prior to the target word) was 880 ms (range 850–920 ms); the average duration of the full sentence was 1660 ms (1590–1740 ms). The average duration of ‘Find the’ (prior to the target word) was 858 ms (830–900ms); the average duration of the full sentence was 1595 ms (1490–1700 ms).

and adult word recognition studies (e.g., Swingley & Aslin, 2000). The dependent measure assessing word recognition was the difference between children’s looking proportions to the target object during these two periods, otherwise known as the naming effect (a significant increase following the label indicates recognition; see Singh, Goh, & Wewalaarachchi, 2015; White & Morgan, 2008). Reliability coding of a random selection of 15 participants (~24% of the total) found that coders were consistent on 95% of the coded frames.

Results

Baseline phase. During the baseline phase (prior to object labeling), children looked at the target objects an average of 50.2% ($SD = .06$) of the time. When familiar and novel target trials were considered separately, these averages were 49.9% ($SD = .09$) and 50.4% ($SD = .06$), respectively.

Familiar words. While our main focus was on children’s behavior on novel target trials, the naming effect was also examined in the familiar word trials as a check to see whether children were engaged in the looking task. Difference scores (the naming effect) are presented in Figure 4. A paired t -test revealed that there was no difference between the 3- and 4-year-olds in the size of the naming effect for familiar words, $t(60) = 1.14, p = .26$. Collapsed across age, the average naming effect was .28. A one-sample t -test against chance (0) was significant, $t(61) = 13.46, p < .001$. Thus, children recognized the familiar words and showed engagement during test.

Novel words. To address our main aim, namely to determine if children’s learning of novel words was influenced by fluency, we conducted a repeated measures ANOVA including Age and familiarization Fluency. There were no significant effects of either age ($p = .87$) or fluency ($p = .837$), and no interaction ($p = .808$). One-sample t -tests demonstrated that children showed recognition of novel words presented both fluently and disfluently during familiarization at greater than chance levels (i.e., 0), $t(61) = 3.06$,

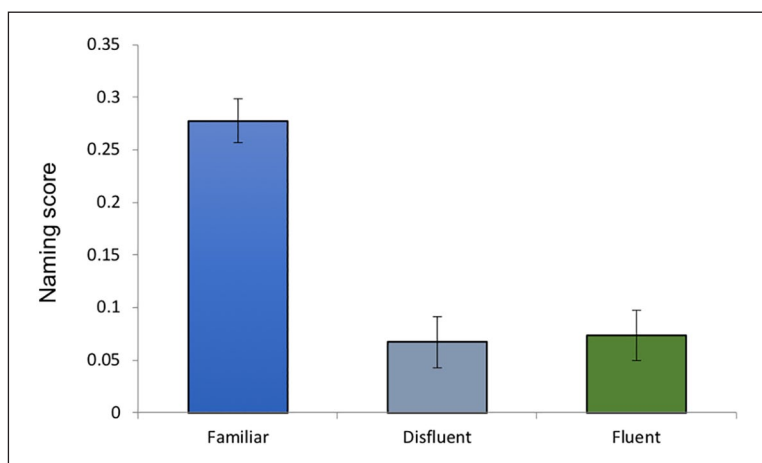


Figure 4. Naming scores for Experiment 2 (proportion looking to the target during naming – proportion looking to the target during baseline). Chance is 0. Error bars represent standard error.

$p = .003$ and $t(61) = 2.75$, $p = .008$, respectively. The pattern of results was consistent when the naming phase proportion alone was considered.

General discussion

The present study asks how disfluencies affect children's word learning. In Experiment 1, we found that 3- to 4-year-old children preferentially endorsed a novel object label from a fluent (as opposed to a disfluent) speaker. In Experiment 2, we asked whether the presence of a disfluency has effects on children's learning of a subsequent word. Fluently and disfluently produced labels were learned equally well. Together, these results suggest that while children prefer information from fluent speakers, they may learn information equally well whether it is presented in a fluent or disfluent manner. We consider each of these findings in turn.

Previous research has demonstrated that children consider a variety of speaker characteristics in their determination of who to get information from, including a speaker's history of accuracy (e.g., Corriveau, Meints, & Harris, 2009; Koenig et al., 2004; Koenig & Sabbagh, 2013), their degree of certainty (Birch et al., 2010; Sabbagh & Baldwin, 2001; Tenney et al., 2011), and one feature of their speech, accent (Kinzler et al., 2011). The present study is the first to show that children also use disfluency in isolation (in the absence of any other linguistic or nonverbal cues to certainty). The preference for fluency was particularly strong in children who reported noticing the disfluency of the speakers. Moreover, some of these children explicitly stated that the presence of a disfluency meant that the speaker had less knowledge. Thus, it appears that when disfluency is explicitly detected it gets used during the evaluation of information sources and that at least some children make an explicit link between disfluency and uncertainty. This is a reasonable strategy, because in many cases disfluency does signal uncertainty (Brennan & Williams, 1995).

Future research could explore whether disfluency has broader effects on children's learning beyond word learning (for instance, in children's imitation of behavior; Birch et al., 2010). In addition, it would be useful to determine whether contextual factors influence children's interpretation of disfluencies. For instance, as disfluency often signals new information, are children more likely to accept information from speakers who demonstrate disfluency in the context of discussing newly acquired information (where it would be expected; Heller, Arnold, Klein, & Tanenhaus, 2014) versus from speakers who are disfluent when presenting familiar information (where it would be less expected)? Moreover, do children adjust their reliance on (dis)fluent speakers according to the rate at which speakers are producing the disfluencies (as they do with relative accuracy; Ronfard & Lane, 2018)?

Another open question is whether children take a speaker's *history* of fluency into account, or whether children considered the speaker's fluency only during the delivery of the critical (novel label) information. In the case of speaker accuracy, children do consider information presented prior to the learning phase (Corriveau et al., 2009). However, when evaluating accuracy, there is no local cue presented during the learning phase itself (i.e., it is impossible to know whether a novel label presented by a speaker is correct or not). This is not the case with fluency, where the information to be learned can be presented either fluently or not. Therefore, it may be that children will prioritize local fluency (the fluency with which the critical information is delivered). However, there is some evidence that 2-year-olds continue to show more imitation of a previously confident actor than a previously unconfident actor (as expressed by both verbal and nonverbal cues), even for a subsequent trial in which both actors are neutral (Brosseau-Liard & Poulin-Dubois, 2014).

In Experiment 2, we asked about more local effects of disfluency on children's word learning. Disfluencies are a relatively common feature of speech, occurring, on average, every 6 per 100 words in conversations between adults (Fox Tree, 1995). Therefore, children will hear disfluencies even from speakers they consider reliable (e.g., their parents). How do these disfluencies affect learning? On the one hand, since disfluency may be indicative of uncertainty about upcoming material (Brennan & Williams, 1995), children may be more cautious about learning that material. However, in both adults and children, disfluency has been shown to cause heightened attention to certain referents during word processing, and in adults, disfluency also leads to better memory for subsequent material (Corley et al., 2007; Corley & Stewart, 2008; Fraundorf & Watson, 2011). In Experiment 2, children were presented with novel words, two of which followed disfluencies, and two of which did not. We found that children learned both sets of words, those presented fluently and disfluently, and learned them equally well.

Why did we find neither a disadvantage nor advantage for disfluently presented material? One possibility is that features of our design attenuated children's treatment of disfluency as a cue to certainty. The experimenter in Experiment 2 indicated that she knew the names of the unfamiliar objects, naming them all fluently prior to the sorting task. Perhaps children would have failed to learn the disfluently produced labels if we had not included this initial naming round. However, to the extent that the presence of disfluency might serve as a more automatic cue to a speaker's certainty, the disfluently produced words should still have been at a disadvantage compared to those presented a total of three times fluently. Therefore, although children appear to take disfluency into account when they are asked to

explicitly obtain information from speakers, any reduced confidence that they might have in disfluently presented information does not appear to affect their learning of that information. There is evidence to suggest that while children show a preference for learning from speakers deemed to be credible, they also learn from less credible or less confident speakers in some contexts (Kim, Paulus, & Kalish, 2017; Krogh-Jespersen & Echols, 2012). This seems to be especially evident when there is no conflicting evidence from more credible informants (Vanderbilt, Heyman, & Liu, 2014), as was the case in Experiment 2. Thus, children may learn the disfluent labels despite the fact that they perceive disfluency to be indicative of less certainty. However, learning from less credible informants may not be as robust, in that children are less likely to generalize the information to new speakers or exemplars (Koenig & Woodward, 2010; Sabbagh & Shafman, 2009).

We also observed no facilitative effect of disfluencies in Experiment 2, in contrast to some previous work with adults (Corley et al., 2007; Corley & Stewart, 2008). For example, when words are preceded by disfluencies, adult listeners are faster to identify whether they match a previously presented cue word and to select an associated image (Corley & Hartsuiker, 2011; Fox Tree, 2001). Adults also have better memory for sentence-final words preceded by disfluencies in unexpected recognition tasks, and their ERP responses show that semantic integration of acoustically modified words (presented with amplitude distortions of certain frequencies) into a sentence context is facilitated by the presence of disfluency (Collard, Corley, Macgregor, & Donaldson, 2008; Corley et al., 2007; Corley & Stewart, 2008). It is possible that disfluency leads to heightened attention in adults precisely because its distributional properties signal to listeners that new or difficult information is likely to follow (Arnold et al., 2007; Arnold et al., 2004; Barr, 2001). Children also make some predictions about what will follow a disfluency (Kidd et al., 2011a; Orena & White, 2015; Owens & Graham, 2016). However, they may not predict the same range of possible referents that adults do. In particular, they do not appear to predict that a less familiar referent will follow a disfluency (Owens et al., 2018). Therefore, children in our task, presented with familiar and novel objects during training, may not have selectively predicted the novel objects during disfluency, attenuating any potential learning advantage for those items. Alternatively, it is possible that the nature of our learning procedure enhanced learning of novel words, regardless of fluency. First, novel words were presented during training in the context of familiar words, which may boost word learning compared to isolated learning, as it allows children to map the novel word to novel object through disambiguation (Zosh, Brinster, & Halberda, 2013; though see Bion, Borovsky, & Fernald, 2013). Second, the ostensive nature of the context may have led children to be maximally engaged and open to receiving information (Csibra & Gergely, 2009), such that there was no added benefit of disfluency.

Of course, because disfluencies can both signal uncertainty and enhance online speech processing, it is possible that both functions were in play, but they cancelled one another out during learning (leading to neither an advantage nor disadvantage for material presented disfluently). In addition, it is also possible that children evaluate the function of disfluencies differently across experimental tasks/situations. For instance, children are very sensitive to nonverbal expressions of uncertainty (e.g., Birch et al., 2010), which did not accompany the disfluencies here. Therefore, disfluencies presented in combination

with such information may have had led to a different outcome in this type of learning situation. Additionally, the type of material to be processed may matter. Previous work demonstrating beneficial effects of disfluency on adults' recognition memory has not involved learning of novel linguistic material, but the recognition of known words (Corley et al., 2007; Corley & Stewart, 2008). Therefore, it would be interesting to test what happens when these two potential functions of disfluencies compete during word learning in adults, as well as whether children's memory for familiar material is enhanced by the presence of disfluency.

In sum, disfluency conveys a wide range of information in human communication. The present results demonstrate that 3- to 4-year-old children use fluency to select an information source, suggesting that, like adults, they view disfluency unfavorably at an explicit level. However, they did not demonstrate more implicit effects of disfluency on their learning of specific information (in terms of either enhanced or diminished learning). Future research should explore how children's sensitivity to and use of disfluency in communication and processing develop and, in particular, how the multiple functions of disfluency interact in real-world communicative situations.

Acknowledgements

The authors would like to thank Shaneene Heupel, Megan Smith, Anisha Varghese, Alexandra Rett, Erin Kim, Jennifer Smith, Robin Rosentreter, Shaquille Sealy, and Joel Le Forestier for their assistance in creating the stimuli, participant scheduling, and data collection. We thank the participating children and families for their time and support of this research. Portions of these data were presented at the Society for Research in Child Development 2017 meeting.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship and/or publication of this article: This work was supported by a Social Sciences and Humanities Research Council (SSHRC) Insight Development Grant awarded to the first two authors.

ORCID iD

Katherine S. White  <https://orcid.org/0000-0003-2631-7711>

Note

1. Separated by age groups, performance over trial pairs was as follows: 3-year-olds: $M_s = 1.19$, $.94$ $SD_s = .66$, $.57$; 4-year-olds: $M_s = 1.33$, $.73$, $SD_s = .49$, $.46$.

References

- Amick, L. J., Chang, S.-E., Wade, J., & McAuley, J. D. (2017). Social and cognitive impressions of adults who do and do not stutter based on listeners' perceptions of read-speech samples. *Frontiers in Psychology*, 8, 1148.
- Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal thee, um, new information. *Journal of Psycholinguistic Research*, 32, 25–36.
- 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, K. K., Altmann, R. J., & Fagnano, M. (2004). The old and thee, uh, new: Disfluency and reference resolution. *Psychological Science*, 15, 578–582.
- Barr, D. J. (2001). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In C. Cave, I. Guatella, & 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, France: L'Harmattan.
- Bion, R. A. H., Borovsky, A., & Fernald, A. (2013). Fast mapping, slow learning: Disambiguation of novel word-object mappings in relation to vocabulary learning at 18, 24, and 30 months. *Cognition*, 126, 39–53.
- Birch, S. A. J., Akmal, N., & Frampton, K. L. (2010). Two-year-olds are vigilant of others' non-verbal cues to credibility. *Developmental Science*, 13, 363–369.
- 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.
- 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.
- 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.
- Brousseau-Liard, P. E., & Poulin-Dubois, D. (2014). Sensitivity to confidence cues increases during the second year of life. *Infancy*, 19, 461–475.
- Christenfeld, N. (1995). Does it hurt to say um? *Journal of Nonverbal Behavior*, 19, 171–186.
- Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speaking. *Cognition*, 84, 73–111.
- 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. Available from <http://psy.cns.sissa.it>
- Collard, P., Corley, M., Macgregor, L., & Donaldson, D. I. (2008). Attention orienting effects of hesitations in speech: Evidence from ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 34, 696–702.
- Corley, M., & Hartsuiker, R. J. (2011). Why um helps auditory word recognition: The temporal delay hypothesis. *PLoS ONE*, 6, e19792.
- Corley, M., Macgregor, L. J., & Donaldson, D. I. (2007). It's the way you, er, say it: Hesitations in speech affect language comprehension. *Cognition*, 105, 658–668.
- Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass*, 2, 589–602.
- Corriveau, K. H., Meints, K., & Harris, P. L. (2009). Early tracking of informant accuracy and inaccuracy. *British Journal of Developmental Psychology*, 27, 331–342.
- Csibra, G., & Gergely, G. (2009). Natural pedagogy. *Trends in Cognitive Sciences*, 13, 148–153.
- Ezrati-Vinacour, R., Platzky, R., & Yairi, E. (2001). The young child's awareness of stuttering-like disfluency. *Journal of Speech, Language and Hearing Research*, 44, 368–380.
- 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.
- Fox Tree, J. E. (2001). Listeners' uses of um and uh in speech comprehension. *Memory & Cognition*, 29, 320–326.
- Fox Tree, J. E., & Clark, H. H. (1997). Pronouncing 'the' as 'thee' to signal problems in speaker. *Cognition*, 62, 151–167.
- Fraundorf, S. H., & Watson, D. G. (2011). The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language*, 65, 161–175.

- Giolas, T. G., & Williams, D. E. (1958). Children's reactions to nonfluencies in adult speech. *Journal of Speech, Language, and Hearing Research*, 1, 86–93.
- 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.
- Jaswal, V. K., & Malone, L. S. (2007). Turning believers into skeptics: 3-year-olds' sensitivity to cues to speaker reliability. *Journal of Cognition and Development*, 8, 263–283.
- Kidd, C., White, K. S., & Aslin, R. N. (2011a). Toddlers use speech disfluencies to predict speakers' referential intention. *Developmental Science*, 14, 925–934.
- 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* (Trends in Language Acquisition Research, pp. 91–106). Amsterdam, The Netherlands: John Benjamins.
- Kim, S., Paulus, M., & Kalish, C. (2017). Young children's reliance on information from inaccurate informants. *Cognitive Science*, 41, 601–621.
- Kinzler, K. D., Corriveau, K. H., & Harris, P. L. (2011). Children's selective trust in native-accented speakers. *Developmental Science*, 14, 106–111.
- Koenig, M. A., Clément, F., & Harris, P. L. (2004). Trust in testimony: Children's use of true and false statements. *Psychological Science*, 15, 694–698.
- Koenig, M. A., & Harris, P. (2005). Preschoolers mistrust ignorant and inaccurate speakers. *Child Development*, 76, 1261–1277.
- Koenig, M. A., & Sabbagh, M. A. (2013). Selective social learning: New perspectives on learning from others. *Developmental Psychology*, 49, 399–403.
- 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.
- Krogh-Jespersen, S., & Echols, C. H. (2012). The influence of speaker reliability on first versus second label learning. *Child Development*, 83, 581–590.
- Markman, E. M., Wasow, J. L., & Hansen, M. B. (2003). Use of the mutual exclusivity assumption by young word learners. *Cognitive Psychology*, 47, 241–275.
- Merriman, W. E., & Bowman, L. L. (1989). The mutual exclusivity bias in children's word learning. With commentary by Brian MacWhinney. *Monographs of the Society for Research in Child Development*, 54(3–4, Serial No. 220), 1–129.
- Morin-Lessard, E., & Byers-Heinlein, K. (2019). *Uh* and *euh* signal novelty for monolinguals and bilinguals: Evidence from children and adults. *Journal of Child Language*, 46, 522–545.
- Newport, E., Gleitman, H., & Gleitman, L. (1977). Mother, I'd rather do it myself: Some effects and non-effects of maternal speech style. In C. E. Snow & C. A. Ferguson (Eds.), *Talking to children* (pp. 109–149). Cambridge, UK: Cambridge University Press.
- 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.
- Ronfard, S., & Lane, J. D. (2018). Preschoolers continually adjust their epistemic trust based on an informant's ongoing accuracy. *Child Development*, 89, 414–429.
- 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.

- Sabbagh, M. A., & Shafman, D. (2009). How children block learning from ignorant speakers. *Cognition*, 112, 415–422.
- Schacter, S., Christenfeld, N., Ravina, B., & Bilous, F. (1991). Speech disfluency and the structure of knowledge. *Journal of Personality and Social Psychology*, 60, 362–367.
- Shriberg, E. (2001). To ‘errrr’ is human: Ecology and acoustics of speech disfluencies. *Journal of the International Phonetic Association*, 31, 153–169.
- Singh, L., Goh, H. H., & Wewalaarachchi, T. D. (2015). Spoken word recognition in early childhood: Comparative effects of vowel, consonant and lexical tone variation. *Cognition*, 142, 1–11.
- Sobel, D. M., & Macris, D. M. (2013). Children’s understanding of speaker reliability between lexical and syntactic knowledge. *Developmental Psychology*, 49, 523–532.
- Swingle, D., & Aslin, R. N. (2000). Spoken word recognition and lexical representation in very young children. *Cognition*, 76, 147–166.
- Tenney, E. R. A., Small, J. E., Kondrad, R. L., Jaswal, V. K., & Spellman, B. (2011). Accuracy, confidence, and calibration: How young children and adults assess credibility. *Developmental Psychology*, 47, 1065–1077.
- Thacker, J. M., Chambers, C. G., & Graham, S. A. (2018a). Five-year-olds’ and adults’ use of paralinguistic cues to overcome referential uncertainty. *Frontiers in Psychology*, 9, 143.
- Thacker, J. M., Chambers, C. G., & Graham, S. A. (2018b). 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.
- Vanderbilt, K. E., Heyman, G. D., & Liu, D. (2014). In the absence of conflicting testimony young children trust inaccurate informants. *Developmental Science*, 17, 443–451.
- Von Tiling, J. (2011). Listener perceptions of stuttering, prolonged speech, and verbal avoidance behaviors. *Journal of Communication Disorders*, 44, 161–172.
- Wechsler, D. (2002). *Wechsler preschool and primary scale of intelligence* (3rd ed.). San Antonio, TX: The Psychological Corp.
- White, K. S., & Morgan, J. L. (2008). Sub-segmental detail in early lexical representations. *Journal of Memory and Language*, 59, 114–132.
- Zosh, J. M., Brinster, M., & Halberda, J. (2013). Optimal contrast: Competition between two referents improves word learning. *Applied Developmental Science*, 17, 20–28.