HHS Public Access

Author manuscript

Lang Cogn Neurosci. Author manuscript; available in PMC 2021 January 01.

Published in final edited form as:

Lang Cogn Neurosci. 2020 ; 35(1): 93–105. doi:10.1080/23273798.2019.1628284.

Processing of Self-Repairs in Stuttered and Non-Stuttered Speech

Matthew W. Lowder[1], **Nathan D. Maxfield**[2], **Fernanda Ferreira**[3]

[1.]Department of Psychology, University of Richmond

[2.]Department of Communication Sciences and Disorders, University of South Florida

[3.]Department of Psychology, University of California, Davis

Abstract

Previous research suggests that listeners can use the presence of speech disfluencies to predict upcoming linguistic input. But how is the processing of typical disfluencies affected when the speaker also produces atypical disfluencies, as in the case of stuttering? We addressed this question in a visual-world eye-tracking experiment in which participants heard self-repair disfluencies while viewing displays that contained a predictable target entity. Half the participants heard the sentences spoken by a speaker who stuttered, and half heard the sentences spoken by the same speaker who produced the sentences without stuttering. Results replicated previous work in demonstrating that listeners engage in robust predictive processing when hearing self-repair disfluencies. Crucially, the magnitude of the prediction effect was reduced when the speaker stuttered compared to when the speaker did not stutter. Overall, the results suggest that listeners' ability to model the production system of a speaker is disrupted when the speaker stutters.

Keywords

disfluencies; repairs; stuttering; prediction; eye movements

The speech we encounter in the real world is rarely spoken with complete fluency. Rather, everyday speech regularly contains hesitations, false starts, self-repairs, repetitions, filled pauses such as *uh* and *um*, and a variety of other types of *disfluency*, affecting approximately six to ten percent of the words we hear (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Fox Tree, 1995). Acknowledging that disfluency is a common feature of everyday language, the field of psycholinguistics has slowly shifted away from research paradigms that once focused exclusively on the processing of idealized, perfectly fluent utterances, and has begun to systematically study the effects of speech disfluency on language processing. Indeed, research over the past decade or so has yielded strong evidence that both the online processing and offline interpretation of language is affected by the

Address correspondence to: Matthew W. Lowder, University of Richmond, 114 UR Drive, Richmond, VA 23173, mlowder@richmond.edu.

Data Availability Statement

The data and analysis scripts that support the findings of this study are openly available via Open Science Framework and can be accessed at https://osf.io/3xfpv/.

presence of disfluencies (e.g., Arnold, Fagnano, & Tanenhaus, 2003; Arnold, Hudson Kam, & Tanenhaus, 2007; Arnold, Tanenhaus, Altmann, & Fagnano, 2004; Bailey & Ferreira, 2003, 2007; Barr & Seyfeddinipur, 2010; Brennan & Schober, 2001; Corley, 2010; Corley, MacGregor, & Donaldson, 2007; Corley & Hartsuiker, 2011; Ferreira & Bailey, 2004; Ferreira, Lau, & Bailey, 2004; Fraundorf & Watson, 2011; Kidd, White, & Aslin, 2011; Lau & Ferreira, 2005; Lowder & Ferreira, 2016a, 2016b, in press; Maxfield, Lyon, & Silliman, 2009; Maxfield & Ferreira, 2019). A major theme of this work is that disfluencies provide the listener with important information about the internal mental state of the speaker. For example, in the case of self-repairs (e.g., *At the intersection turn left, uh I mean right*), the speaker's disfluency indicates to the listener that a portion of the preceding utterance (the *reparandum*; e.g., *left*) was spoken in error and should be replaced by new material (the *repair*; e.g., *right*).

Crucially, however, disfluencies do not always indicate that the speaker is having a typical problem with higher-level language planning. Instead, disfluencies sometimes signal a very different and atypical type of speech production problem, as in the case of stuttering. Stuttering involves the chronic production of within-word disfluencies that are characterized by the repetition, audible prolongation, or silent blocking of speech segments (e.g., *At the intersection t-t-t-turn right*) (Wingate, 1964). Struggle behaviors and signs of emotional distress might also accompany instances of stuttering (Wingate, 1964). In adults, stuttering is more likely to occur on content words, and on words that are phonetically longer, lower in frequency, and less predictable from context (Bloodstein & Bernstein Ratner, 2008). In addition, the frequency of stuttering tends to be higher in sentences with more complex grammatical structures (Tsiamtsiouris & Cairns, 2013). Although some of these cues may be informative to listeners (e.g., listeners may detect the speaker's emotional state, and also may come to associate instances of stuttering with less predictable information and/or with grammatically more complex sentences), stuttering-like disfluencies do not typically reveal anything else about the current mental state of the speaker. Nevertheless, it still may be the case that the listener's processing of a self-repair disfluency is impaired to some extent when it comes from a speaker who stutters as opposed to a speaker who does not stutter, particularly in light of recent theoretical advancements suggesting that basic sentence-processing operations are affected by differences in speaker-specific characteristics. The purpose of the current study is to directly investigate this possibility.

Noisy Channel models of language comprehension (e.g., Gibson, Bergen, & Piantadosi, 2013; Gibson, Piantadosi, Brink, Bergen, Lim, & Saxe, 2013) propose that listeners routinely combine linguistic input with relevant knowledge in order to recover a speaker's intended meaning, particularly when the signal is distorted by speaker, listener, or environmental noise. Applied to the domain of speech disfluencies, this framework predicts that listeners will actively anticipate forthcoming material when the speaker signals that he or she has made an error or will even anticipate that the speaker has made an error when a given word or phrase seems implausible with the preceding sentence context. Indeed, we have recently provided direct evidence for these ideas. Lowder and Ferreira (2016b) showed that when listeners encountered self-repair disfluencies (e.g., *The woman went to the animal shelter and brought home a dog uh I mean a rabbit…*), they used semantic information about the reparandum (e.g., *dog*) to generate predictions about what they expected the upcoming repair would be (e.g., *cat*). This tendency to predict the upcoming repair was stronger than the tendency to predict the second conjunct in a coordination control condition (e.g., *… brought home a dog and also a rabbit…*). Further, Lowder and Ferreira (in press) have demonstrated that listeners are sensitive to cues of contextual plausibility and speaker certainty and use these cues to rapidly anticipate and implicitly correct speaker error, even before receiving an explicit error signal from the speaker.

Importantly, the Noisy Channel account also predicts that the listener's ability to recover the speaker's intended meaning depends crucially on the ease with which the listener is able to model the speaker's production system. This suggests that the listener's ability to predict the identity of an upcoming self-repair might be disrupted when the speaker stutters because listeners may have trouble determining whether disfluencies in the utterance signal a self-repair or a more general speech production problem. Consistent with this idea, there is previous work demonstrating that speaker-specific features can systematically affect listeners' online language processing. For example, Grodner and Sedivy (2011) showed that although listeners are typically able to draw rapid contrastive inferences (e.g., upon hearing *Pick up the tall…* they quickly look to the taller of two glasses in a display), this effect is reduced when participants are told that the speaker suffers from an impairment that affects language and social skills. In addition, Van Berkum, van den Brink, Tesink, Kos, & Hagoort (2008) conducted an event-related potential (ERP) study to examine whether listeners could use perceptual properties of the speaker to access stereotype-driven inferences (based on age, gender, or social status) and then rapidly integrate this information with the linguistic content. They found that perceived mismatches between the speaker and the linguistic content (e.g., hearing a man say *I might be pregnant…*) elicited a larger N400 effect compared to when the speaker and the linguistic content matched. Other work has focused on how listeners process speech coming from a native versus non-native speaker. For example, using ERPs, Hanulíková, van Alphen, van Goch, and Weber (2012) showed that sentences with grammatical violations spoken by a native speaker elicited a classic P600 effect, whereas the same violations spoken by a non-native speaker who had a clear foreign accent elicited no P600 effect. These results suggest that listeners can use the foreign accent as a cue that the speaker may not have complete mastery of the language, leading listeners to adjust their model of the speaker's production system to allow more room for error. Finally, Bosker, Quené, Sanders, and de Jong (2014) investigated whether listeners would engage in different patterns of predictive processing when hearing a filled pause (e.g., *um*) from a native versus a non-native speaker. When a native speaker produced a filled pause, listeners anticipated that the speaker was about to refer to a low-frequency as compared to a high-frequency object. In contrast, when a non-native speaker produced a filled pause, listeners did not show this bias to anticipate a low-frequency object. Again, the results suggest that listeners use information about the speaker to adjust their expectations: listeners tended to attribute native-speaker disfluency to difficulty with lexical selection, whereas listeners tended to attribute non-native-speaker disfluency to a wider range of possible causes.

Although there have been efforts to understand how stuttering-like disfluencies affect listener comprehension and listener perceptions, it remains to be determined how stuttering affects specific sentence processing mechanisms such as prediction. There have, however, been some studies investigating how listeners' memory for linguistic content is affected by stuttering. Cyprus, Hezel, Rossi, and Adams (1984) used a between-subjects design to determine the extent to which simulated stuttering affected recall of an audio-recorded message, using mildly versus severely stuttered speech together with a no-stuttering control. In addition, disfluencies varied in whether they appeared on words of high versus low information value (essentially, on content words or function words). College students recalled less material when severe stuttering was present on words of high information value; all other conditions were statistically equivalent. The authors concluded that severe stuttering makes the speech of an adult who stutters difficult to remember, particularly on words that are semantically informative. More recently, Panico and Healey (2009) observed that free and cued recall of even mildly stuttered speech was impaired compared to fluent controls, and listeners indicated that more mental effort was required to comprehend stuttered speech.

Additional work has examined how stuttering affects listeners' perceptions and judgments of the speaker. Susca and Healey (2001) modified a single stuttered sample using digital signal processing methods to create versions that varied in stuttering severity. They observed that the more stuttering in the speech samples, the more negative the terms used to describe the speech. Interestingly, naïve listeners apparently could distinguish normal speech from stuttered speech with all disfluencies and pauses removed, suggesting the presence of other features differentiating normal from stuttered speech. Susca and Healey (2002) used the same speech stimuli and found that more severe stuttering was associated with perceptions that the speaker seemed flustered as well as less intelligent and educated, and that the listener was boring and hard to understand. Interestingly, listeners noted that the speech was "hard to follow to predict next words", and that earlier portions of the speech were hard to remember (see also Panico, Healey, Brouwer, & Susca, 2005). This suggests, at least informally, that the presence of stuttering-like disfluencies might impact specific sentence processing mechanisms such as those associated with prediction and revision.

In a recently published study, Maxfield and Ferreira (2019) investigated effects of stuttering-like disfluencies on the processing and interpretation of Garden Path (GP) sentences. One finding was that disambiguating verbs in GP sentences elicited a P600 ERP effect in the absence of any disfluencies but not when stuttering-like disfluencies were present in the sentence preambles. In addition, listeners hearing GP sentences containing stuttering had more accurate interpretations than listeners who heard GP sentences without stuttering. This combination of findings suggests that listeners hearing sentences containing stuttering abandon prior expectations, possibly because they lack a speaker model defined by the presence of stuttering. In the presence of stuttering, listeners appeared to focus on the input itself instead of using prior knowledge to form early (erroneous) interpretations of GP sentences, reducing the likelihood of experiencing Garden Paths (thus attenuating P600 activation to disambiguating verbs), and increasing accuracy of sentence interpretations. This discovery opens the question of how other aspects of sentence processing may be affected by the presence of stuttering-like disfluencies.

In the current visual-world eye-tracking experiment, we investigate whether listeners are sensitive to the speaker-specific feature of stuttering during the online processing of self-repairs. This general issue has received surprisingly little attention given both its theoretical and practical importance. The theoretical significance arises from the key role of cognitive mechanisms which allow comprehenders to integrate information about the speaker with the linguistic content so as to build a richer representation of the sentence or utterance. In terms of practical significance, stuttering is a common speech disorder, affecting as many as 8% of speakers at some point in their lifetime, although a much smaller percentage of adults stutter chronically (slightly less than 1%) (Yairi & Seery, 2015). A better understanding of how stuttering affects the ability of typically-fluent speakers to understand spoken utterances is potentially useful for directing treatment to the aspects of stuttering that are most communicatively disruptive.

Additionally, investigating the processing of speech repairs allows us to assess how prediction mechanisms that are used to efficiently process the speech of typical speakers change when listeners hear speech from a person who stutters. Current approaches to prediction in language comprehension assume that the ability to predict is based in part on listeners having built a model of the speaker based on prior experiences (Kuperberg & Jaeger, 2016). Given that most listeners will have had limited experience interacting with people who stutter, they may have weak priors to integrate with the input, and those weak priors will discourage prediction and encourage a heavier weighting of the input.

Participants listened to sentences as in Example 1, which were adapted from Lowder and Ferreira (2016b, Experiment 1). We employed a between-subjects design such that half of the participants heard the sentences spoken by a speaker who stuttered and half of the participants heard the sentences spoken by a speaker who does not stutter (in fact the same speaker recorded all sentences; see Method section). Participants listened to these sentences while they viewed images on the computer screen. The array of images was identical across the four conditions and always consisted of an image representing the first noun phrase (NP1; e.g., *a dog*), an image representing the second noun phrase (NP2; e.g., *a rabbit*), an image representing a random distractor (e.g., *a plant*), and an image representing a critical unnamed distractor (e.g., *a cat*), which was the most predictable continuation of the sentence but was never spoken in any of the conditions.

1a. *The woman next door went to the animal shelter and brought home a dog uh I mean a rabbit, even though her apartment doesn't allow pets.* (Self-Repair)

1b. *The woman next door went to the animal shelter and brought home a dog and also a rabbit, even though her apartment doesn't allow pets.* (Coordination)

1c. *The woman next door went to the animal shelter and brought home a dog, even though her apartment doesn't allow pets*. (NP1 Only)

1d. *The woman next door went to the animal shelter and brought home a rabbit, even though her apartment doesn't allow pets.* (NP2 Only)

This design thus allowed us to pursue several research goals. First, we aimed to replicate the basic findings reported in Lowder and Ferreira (2016b) relating to prediction in the processing of self-repair disfluencies. Specifically, we hypothesized that listeners would show a greater tendency to fixate the critical distractor item during the *uh I mean* portion of self-repair disfluencies (1a) compared to the *and also* portion of coordination structures (1b). Second, the between-subjects manipulation of Speaker allowed us to investigate how the magnitude of this prediction effect might be modulated when the sentences are spoken by a person who stutters versus a person who does not stutter. Finally, we also examined how the Speaker manipulation would affect listeners' processing of the NP1-Only and NP2-Only conditions (1c and 1d). Given that these two conditions involve the mention of only one target item, they provide a straightforward test of how the Speaker manipulation might influence the speed with which listeners direct their attention to the corresponding image, as well as how long listeners tend to remain fixated on the image before shifting their attention elsewhere.

Method

Participants

Sixty-four undergraduate students at the University of South Carolina participated in this experiment in exchange for course credit in Introductory Psychology. They were all native English speakers and reported normal or corrected-to-normal vision. The experimental protocols were approved by the University of South Carolina Institutional Review Board, and all participants provided written informed consent.

Materials and Design

The same 40 experimental items used in Lowder and Ferreira (2016b, Experiment 1) were used in the current study (see Example 1). Construction of the experimental items began with pairs of highly semantically related words (e.g., *dog-cat*, *salt-pepper*, *bread-butter*). The Self-Repair and Coordination conditions (1a and 1b) were constructed such that the first noun from the pair (e.g., *dog*) served as NP1 in the sentence (i.e., the reparandum or the first conjunct), whereas the second noun from the pair (e.g., *cat*) did not appear in the sentence. Instead, the noun that served as NP2 (e.g., *rabbit*) was a plausible but less predictable continuation of the sentence (see Lowder & Ferreira for details on norming of stimuli). The only difference between the Self-Repair and Coordination conditions was the words that intervened between NP1 and NP2 (*uh I mean* versus *and also*). The NP1-Only and NP2-Only conditions (1c and 1d) were constructed by inserting only one of the NPs into the sentence while holding the rest of the sentence constant. Each set of experimental items was associated with a visual display that consisted of four color images representing NP1 (e.g., *a dog*), NP2 (e.g., *a rabbit*), the critical distractor (e.g., *a cat*), and a random distractor (e.g., *a plant*). The full set of experimental stimuli appears in the Appendix of Lowder and Ferreira (2016b).

The materials also included a set of 45 filler items representing a variety of different sentence types, each with a corresponding visual array of four images. Ten of these filler items contained a repair disfluency in which the actual repair was highly predictable based on the reparandum. The purpose of these items was to discourage participants from developing strategies based on a realization that the most predictable repair in the critical items was never actually spoken. More information on the creation of these 10 predictable filler items are available in Lowder and Ferreira (2016b).

A 23-year-old male Caucasian with a clinical diagnosis of stuttering was recruited to produce the verbal stimuli. He spoke Standard American English as a native and only language. As a result of speech therapy, the speaker was able to reliably produce non-stuttered speech by employing subtle fluency control with acceptable naturalness. Fluency control involved (a) initiating the first word of each speech group with gentle voicing and slightly exaggerated articulatory movements, and (b) maintaining fluency in each speech group by sustaining air flow and using soft articulatory contacts. The speaker was also able to produce speech with fluency controls off, during which his speech contained somewhat frequent stuttering and some associated struggle behavior. With fluency controls off, his stuttering was mild to moderate in severity.

The speaker recorded all sentences twice: once without fluency controls (Stutter condition) and once with fluency controls (No-Stutter condition). Sentences spoken in the Stutter condition always contained one or at most two instances of stuttering per sentence (including filler sentences), which is considered relatively mild in terms of stuttering severity (Healey, 2010). For the experimental sentences, the stuttering always occurred in the sentence preamble, with no instances of stuttering from the first critical word (e.g., *dog*) through the rest of the sentence. The type of stuttering in each sentence was always a sound repetition, a sound prolongation, or a combination of these. Analysis of stuttering frequency for sentences in the Stutter condition revealed the following mean rates of stuttered syllables per sentence for each sentence type: Self-Repair – 5.73%; Coordination – 5.78%; NP1-Only – 5.53%; NP2-Only – 5.84%; Fillers – 5.95%. The full set of spoken materials is available at https://osf.io/3xfpv/.

Speaker was manipulated between-subjects such that 32 participants were exposed to the Stutter condition and 32 were exposed to the No-Stutter condition. The 40 sets of experimental sentences and their corresponding visual displays were counterbalanced within-subjects such that each participant was presented with only one version of each item. Assignment of participants to stimulus lists was random.

A brief post-experiment questionnaire was created that asked participants to rate the extent to which they agreed with four statements on a scale from 1 (strongly disagree) to 7 (strongly agree). The statements were: “This person is a competent speaker”, “This person is a fluent speaker”, “This person read the sentences easily”, and “I felt comfortable listening to this person.”

Procedure

Eye movements were recorded with an EyeLink 1000 tower system (SR Research). The tracker sampled the location of the eye at a rate of 1,000 Hz and automatically parsed the samples into fixations and saccades. The tracker was calibrated at the beginning of each session and was recalibrated throughout the session as needed, at the discretion of the experimenter. Forehead and chinrests were used to minimize head movements. Participants were told that they would view images on the computer screen while also listening to sentences. They were explicitly told that “some of the sentences might contain errors,” but were instructed to just try to understand each sentence. At the start of each trial, a fixation point was presented in the center of the screen, which served as an opportunity for the experimenter to check for signs of calibration drift. If the experimenter judged the participant's gaze to have drifted from this fixation cross, the tracker was recalibrated. Once the participant's gaze was steady on this fixation point, the experimenter pressed a button that presented the visual display for the trial. After a 3,000 ms delay, the corresponding sentence was presented via headphones. After the sentence finished playing, the images disappeared and the fixation point for the next trial appeared.

Participants were first presented with five of the filler sentences. After this warm-up block, the remaining sentences were presented pseudorandomly under the constraint that no more than two trials from the same within-subjects condition could be presented consecutively. The locations of the four images within a visual display were randomized on each trial.

After the eye-tracking session, the participant completed the post-experiment questionnaire. The entire session lasted approximately 40 minutes.

Analysis

Statistical analyses on the eye-tracking data were performed using the lme4 software package in R (Bates, Maechler, & Dai, 2012). The dependent variable was whether a particular image was fixated within a given time window (coded as "1" if a fixation was made and "0" if it was not). The analyses were conducted in two phases. First, we investigated whether listeners' processing of the Self-Repair versus Coordination conditions was modulated by the Speaker manipulation. We fit mixed effects models (Baayen, Davidson, & Bates, 2008; Jaeger, 2008) that included Sentence (Self-Repair vs. Coordination), Speaker (Stutter vs. No-Stutter), and their interaction as fixed effects as well as subjects and items as crossed random effects. The random effects structure included the maximally appropriate random intercepts as well as by-subject and by-item random slopes. In cases where the maximal model failed to converge, the random effects structure was sequentially simplified until convergence was achieved (Barr, Levy, Scheepers, & Tily, 2013). The second phase of our analysis investigated whether the Speaker manipulation affected listeners' gaze patterns in the NP1-Only and NP2-Only conditions. For the NP1-Only condition, we analyzed the effect of Speaker on looks to the picture corresponding to NP1, whereas for the NP2-Only condition, we analyzed the effect of Speaker on looks to the picture corresponding to NP2. For these analyses, Speaker was entered as a fixed effect, and subjects and items were entered as crossed random effects. Again, the maximally appropriate random effects structure was employed.

Within each sound file, we marked the onsets of target words that were then used to create time windows for analysis. For all conditions, we marked the onset of the first critical noun as zero on the time line. For the Self-Repair and Coordination conditions, we also marked the onsets of the immediately following word (*uh* vs. *and*), as well as the onset of the second critical noun. Three time windows were then constructed around these critical word onsets. Window 1 measured from the onset of the first critical noun to the onset of *uh* or *and*. Window 2 measured from the onset of *uh* or *and* to the onset of the second critical noun. Finally, Window 3 measured from the onset of the second critical noun until 1,000 ms had elapsed. For the NP1-Only and NP2-Only conditions, we also constructed three time windows: Window 1 measured from the onset of the critical noun until 1,000 ms had elapsed, Window 2 encompassed the next 1,000 ms, and Window 3 encompassed the next 1,000 ms. Given that it takes approximately 200 ms to program and launch a saccade (Matin, Shao, & Boff, 1993), these analysis windows were offset by a constant latency of 200 ms, as is common practice in visual-world eyetracking experiments (see, e.g., Barr, 2008).

As the utterances were recorded to sound as natural as possible, this led to one systematic difference across conditions in the duration of the time windows. Specifically, Window 2 was substantially longer for the Self-Repair condition (1,181 ms) than for the Coordination condition (619 ms). Crucially, however, there were no significant Speaker-related differences in the duration of any time windows, nor were there any Sentence-by-Speaker interactions. Thus, any significant effects of Speaker or Sentence-by-Speaker interactions observed in fixations to images while listening to the sentences cannot be attributed to differences in duration of the utterances.

Results

Self-Repair and Coordination

Figure 1 displays the probability of fixating each of the four picture types for the Self-Repair and Coordination conditions as a function of Speaker. Table 1 shows the results of the mixed-effects analyses. Visual inspection of Window 1 suggests similar fixation patterns across all conditions, with participants shifting their gaze toward the named entity (e.g., *dog*). We did observe a marginally significant main effect of Sentence in Window 1, with slightly more looks to NP1 in the Self-Repair condition than in the Coordination condition. This difference was unexpected, given that the two conditions were identical up to this point. However, the difference was only marginally significant, and we did not observe such an effect in our previous work (Lowder & Ferreira, 2016b); we therefore do not discuss it further. No other main effects or interactions emerged in Window 1.

Within Window 2, we observed significant main effects of Sentence in looks to all three pictures of interest. Participants were more likely to fixate NP1, NP2, and the critical distractor in the Self-Repair condition versus the Coordination condition. Although this pattern replicates the pattern reported in Lowder and Ferreira (2016b), the effects observed in the current experiment are likely due, at least in part, to basic differences in the duration of this window, with listeners having more time to fixate items in the Self-Repair condition than in the Coordination condition. Crucially, however, there was a significant Sentence-by-Speaker interaction in looks to the critical distractor, such that the prediction effect (i.e., the greater tendency to look at the critical distractor in the Self-Repair versus Coordination condition) was smaller in the Stutter condition than in the No-Stutter condition.[1] There was also a marginally significant main effect of Speaker in looks to NP2, with listeners being somewhat more likely to fixate this picture in the No-Stutter condition than the Stutter condition. Importantly, these Speaker-related effects cannot be explained by differences in duration of the time windows.

[1]We also conducted a supplementary analysis using proportion of fixations per participant per trial as our dependent measure. The results were identical to those presented here. Most notably, there was a significant Sentence-by-Speaker interaction, $p < .005$, reflecting the smaller prediction effect in the Stutter condition compared to the No-Stutter condition.

Within Window 3, we again observed significant main effects of Sentence in looks to all three pictures, although the pattern was different than that observed in Window 2. Participants were more likely to fixate NP1 in the Coordination condition versus the Self-Repair condition, but were more likely to fixate NP2 and the critical distractor in the Self-Repair condition versus the Coordination condition. Again, there was a significant Sentence-by-Speaker interaction in looks to the critical distractor, such that the prediction effect was smaller in the Stutter condition than in the No-Stutter condition.

NP1-Only and NP2-Only

Figure 2 displays the probability of fixating each of the four picture types for the NP1-Only and NP2-Only conditions as a function of Speaker. Table 2 shows the results of the mixed-effects analyses. Visual inspection of the fixation plots suggests that in both the NP1-Only and NP2-Only conditions, listeners in both Speaker conditions initially directed their gaze to the named entity. However, whereas listeners in the No-Stutter condition appear to hold their gaze on this picture, listeners in the Stutter condition appear more likely to slowly drift away. The results of the statistical analyses are consistent with this observation: There were no effects of Speaker in Window 1 or Window 2 (i.e., the first 2,000 ms following onset of the target word); however, both the NP1-Only and NP2-Only conditions showed a significant effect of Speaker in Window 3 (i.e., from 2,000 ms to 3,000 ms following onset of the target word), such that listeners in the Stutter condition were less likely to fixate the target picture during this time window than listeners in the No-Stutter condition.

Post-experiment questionnaire

There were significant effects of Speaker in participants' ratings of all four items on the questionnaire. Specifically, participants in the Stutter condition rated the speaker as less competent in his speech (3.56 vs. 5.13, $t(62) = -4.08$, $p < .001$), less fluent (3.72 vs. 5.22, $t(62) = -3.48$, $p < .005$), harder to understand (2.84 vs. 4.34, $t(62) = -3.89$, $p < .001$), and rated themselves as less comfortable (4.50 vs. 5.34, $t(62) = -2.00$, $p < .05$) compared to participants in the No-Stutter condition. These findings confirm that the presence of stuttering affected listener perceptions of the speaker and of their understanding of the speaker despite the fact that the same person produced the Stutter and No-Stutter stimuli.

Cross-experiment analysis

Given the theoretical importance of the Sentence-by-Speaker interaction we observed in looks to the critical distractor, we conducted an exploratory cross-experiment analysis in which we compared the key results from the current experiment to results from Experiment 1 of Lowder and Ferreira (2016). The experiments are similar in that they contain identical within-subjects manipulations of Sentence; they are different in that the current experiment has a between-subjects manipulation of Speaker (i.e., Stutter vs. No-Stutter), whereas there was no between-subjects manipulation in Lowder and Ferreira, making it akin to the No-Stutter condition presented here. Further, the two experiments differ in the identity of the speaker, which naturally leads to systematic differences in prosodic characteristics of the utterances. This is important because it allows us to test the hypothesis that the effects observed in the current experiment are due to stuttering *per se*, as opposed to a manipulation of the identity of the speaker. An analysis comparing the No-Stutter condition in the current experiment with Experiment 1 of Lowder and Ferreira tested for effects of Sentence (Self-Repair vs. Coordination) on probability of fixating the critical distractor in Window 2 (i.e., the onset of *uh* versus *and*). The results revealed a robust main effect of Sentence (estimate = 0.68, $SE = 0.07$, $z = 10.00$, $p < .001$), demonstrating the greater tendency to fixate the critical distractor during this region in the Self-Repair condition versus the Coordination condition. Crucially, there was no significant effect of Experiment (estimate = −0.12, $SE = 0.11$, $z = -1.08$), nor a Sentence-by-Experiment interaction (estimate = −0.05, $SE = 0.07$, $z = -0.83$). We also conducted a similar analysis comparing the Stutter condition in the current experiment with Experiment 1 of Lowder and Ferreira. The results again revealed a robust main effect of Sentence (estimate = 0.55, $SE = 0.07$, $z = 8.08$, $p < .001$) and no main effect of Experiment (estimate = −0.01, $SE = 0.13$, $z = -0.03$). The Sentence-by-Experiment interaction was not significant (estimate = 0.10, $SE = 0.07$, $z = 1.51$, $p = 0.13$), but the pattern was suggestive of the pattern observed in the current experiment – that is, the magnitude of the prediction effect in Lowder and Ferreira was somewhat larger compared to the magnitude of the effect in the Stutter condition in the current experiment.[2] Note, however, that analyses comparing the Stutter condition from the current experiment with Experiment 1 of Lowder and Ferreira should be interpreted with caution, given that the two speakers were different individuals whose speech differed in many ways besides simply the presence or absence of stuttering. Nevertheless, we believe that the patterns from this cross-experiment analysis, combined with the significant Sentence-by-Speaker interaction observed in the current experiment, further supports the conclusion that the presence versus absence of the stutter in the current experiment was the key factor that contributed to the different patterns of prediction we observed between the Self-Repair and Coordination conditions.

Discussion

Although we encounter self-repair disfluencies on a daily basis as part of everyday language use, our ability to process these utterances may depend on the context in which we hear them; specifically, the ease with which we are able to anticipate the upcoming repair may be modulated by relevant characteristics of the speaker. We tested this possibility in the current visual-world eye-tracking experiment, in which participants heard self-repair disfluencies and control sentences spoken by the same speaker who either stuttered in the sentence preamble (i.e., by turning off fluency controls) or did not stutter (i.e., by using fluency controls while producing the sentences). In these utterances, the entity that constituted the most predictable word was represented on the screen pictorially but was never actually spoken. Our results yielded three main findings. First, we replicated previous work (Lowder & Ferreira, 2016b) showing that listeners engage in robust predictive processing when hearing self-repair disfluencies. That is, listeners were more likely to fixate the critical distractor picture (e.g., *cat*) when listening to self-repairs (e.g., *…brought home a dog uh I mean…*) versus coordination structures (e.g., *…brought home a dog and also…*). Second, the magnitude of this prediction effect was significantly reduced when the sentences contained stuttering compared with no stuttering. Finally, when the speaker named only one

[2]When the analysis was conducted using proportion of fixations per participant per trial as the dependent measure, the Sentence-by-Experiment interaction was fully significant, $p < .01$.

critical entity (e.g., *…brought home a dog…* or *…brought home a rabbit…*), listeners quickly shifted their gaze to the named entity, regardless of speaker condition; however, when the utterance came from a speaker who stuttered, listeners were quicker to move their gaze away from the critical picture compared to listeners who heard a speaker who did not stutter. In the remainder of this section, we discuss each of these findings in turn, focusing on the implications this work has for models of predictive language processing in general and processing of stuttered speech more specifically.

The finding that listeners actively predict an upcoming repair during the processing of self-repair disfluencies is consistent with a broader trend in the literature that casts prediction as an inherent feature of language processing (e.g., Kuperberg & Jaeger, 2016) and cognition in general (e.g., Clark, 2013). Although current debates have begun to raise serious questions regarding the extent to which linguistic prediction occurs under “normal” circumstances—considering, for example, the incredibly rapid speeds at which we typically perceive fluent speech or move our eyes across a page of written text (see, e.g., Ferreira & Lowder, 2016; Huettig & Mani, 2016; Luke & Christianson, 2016)—the domain of speech disfluencies seems to represent a language context in which predictive processing is especially robust. That is, relatively long breaks in the flow of fluent speech in which the speaker signals production difficulty (e.g., *thee uhh uh*) or in which the speaker signals that an error has occurred (e.g., *uh I mean*) seem to serve as explicit cues to the listener to actively anticipate what the speaker will say next. As we have argued previously (Ferreira & Lowder, 2016; Lowder & Ferreira, 2016a, 2016b, in press), the predictive processing that occurs during the comprehension of self-repairs seems to involve the generation of a set of weighted candidates that stand in semantic contrast to the reparandum. For example, hearing an utterance like *…the woman brought home a dog uh I mean…* is immediately interpreted to mean that the woman did *not* bring home a dog, but rather she brought home something else instead. The listener then generates a set of likely repairs (e.g., *cat, rabbit, gerbil*), in much the same way as has been proposed in the processing of contrastive focus (Rooth, 1992).

Importantly, however, the current work demonstrated speaker-specific modulation of listeners’ predictive processing: the robust prediction effect we observed in listeners who heard a speaker without a perceptible stutter was significantly reduced when the same utterances were spoken by a speaker who does stutter. Because the same speaker recorded all sentences, it is difficult to attribute the differences in fixation patterns to anything other than the presence versus absence of stuttering. Our findings are consistent with Noisy Channel models of language comprehension (Gibson et al., 2013), according to which listeners cope with the noise in everyday language by actively trying to uncover the speaker’s intended meaning. This requires the listener to combine the input with relevant linguistic and contextual knowledge, and to then use this knowledge to adopt an internal model of the speaker’s production system. Our results suggest that listeners had difficulty forming an accurate model of the speaker’s production system when the speaker produced stuttering-like disfluencies, which led to less efficient predictive processing compared to listeners who modeled the production system of the same speaker when he did not stutter. In other words, it was difficult for these listeners to be sure whether a disfluency that signaled an upcoming self-repair was an informative cue about the mental state of the speaker or was instead a stuttering-related disfluency.

Another speaker-related difference we observed was in fixation patterns in the NP1-Only and NP2-Only conditions, such that listeners in the Stutter condition were quicker to move their gaze away from the named picture compared to listeners in the No-Stutter condition. Although we cannot say with certainty why this difference emerged, one possibility is related to the speaker-related differences we observed in participants' responses to items on the post-experiment questionnaire. Not only did listeners in the Stutter condition rate the individual as being a less competent speaker, less fluent, and less easy to understand than listeners in the No-Stutter condition, but they also indicated that they felt less comfortable listening to him. This finding is consistent with previous work showing that listeners tend to form negative judgments of speakers who produce stuttered speech (e.g., Panico et al., 2005; Susca & Healey, 2001, 2002), and may suggest that listeners in the Stutter condition in the current study became impatient with the speaker, and so were quicker to move away from the target picture in anticipation of the next trial. A related possibility is that listeners in the Stutter condition focused more on the stuttering rather than the linguistic content of the utterance, or perhaps even disengaged from the speaker entirely and began mind-wandering because they found the speaker uninteresting. These interpretations are supported by evidence that fluent listeners hearing stuttered speech are physiologically aroused (evident in a significant increase in skin conductance response relative to observing a fluent condition) and appear to maintain attention to the presence of stuttering (evident in a significant decrease in heart rate relative to observing a fluent condition) (Guntupalli, Kalinowski, Nanjundeswaran, Saltuklaroglu & Everhart, 2006). These same effects have also been observed in fluent listeners watching and hearing stuttered speech (Guntupalli, Everhart, Kalinowski, Nanjundeswaran & Saltuklaroglu, 2007; see also Bowers, Crawcour, Saltuklaroglu & Kalinowski, 2010) including graduate students in speech-language pathology (Guntupalli, Nanjundeswaran, Dayalu & Kalinowski, 2012), as well as in people who stutter watching and hearing stuttered speech (Zhang, Kalinowski, Saltuklaroglu & Hudock, 2010). The implication is that hearing, as well as watching and hearing, stuttered speech can elicit a visceral reaction in listeners which might, ultimately, distract from ongoing sentence processing.

Overall, our findings are consistent with previous research showing that the listener's ability to process and comprehend spoken language depends on characteristics of the speaker (e.g., Bosker et al., 2014; Grodner & Sedivy, 2011; Hanulíková et al., 2012; Van Berkum et al., 2008). Whereas this previous work has tended to focus on whether the speaker is perceived as a native versus a non-native speaker or whether the content of the speech matches listeners' perceptions of the speaker, the current work demonstrates that the speaker's tendency to stutter also affects basic sentence processing mechanisms. As discussed in the Introduction, previous work has shown that stuttering can impair listeners' memory for the content of utterances (Cyprus et al., 1984; Panico & Healey, 2009). The current work extends these findings by showing that the presence of stuttering affects mechanisms of linguistic prediction that are detectable during the moment-to-moment online processing of language.

Further, this work serves to highlight the importance of considering stuttering and other fluency disorders in the mainstream psycholinguistic literature. As noted in the Introduction, the past several years have seen a shift toward research that has begun to more carefully consider the effects of speech disfluencies on sentence processing and language comprehension. Crucially, however, this work has focused almost exclusively on what we might call “normal” speech disfluencies such as filled pauses (e.g., Arnold et al., 2003, 2004, 2007; Bailey & Ferreira, 2003, 2007; Barr & Seyfeddinipur, 2010; Corley et al., 2007; Corley & Hartsuiker, 2011; Fraundorf & Watson, 2011; Kidd et al., 2011; Maxfield et al., 2009) and self-repairs (e.g., Corley, 2010; Ferreira et al., 2004; Lau & Ferreira, 2005; Lowder & Ferreira, 2016a, 2016b, in press). Although this body of work has been fundamental in demonstrating that listeners can use disfluencies such as these to make inferences about the speaker’s state of mind and predict upcoming linguistic input, the work we have presented here clearly demonstrates that this is only part of the story of the role of speech disfluencies in sentence processing. Indeed, a complete theory of language comprehension must provide a mechanistic account of not only the circumstances under which the presence of speech disfluencies makes language processing more efficient, but also when and how the presence of speech disfluencies hinders sentence processing and language comprehension.

Finally, we believe that a more complete theory of this sort would have implications for treatment of fluency disorders, perhaps leading to the development of therapies specifically targeted to the aspects of stuttering most disruptive to communicative success. For example, there is some evidence that repetition- and prolongation-type stuttering on non-keywords reduces sentence recall accuracy, while prolongation-type stuttering on keywords yields similar sentence recall accuracy as hearing fluent speech (Hulit, 1976). It was also shown that relatively infrequent stuttering on keywords yielded similar sentence recall accuracy as hearing fluent speech (Cyprus, Hezel, Rossi & Adams, 1984; although see Panico & Healey, 2009). From this perspective, working with a person who stutters to modify how they stutter on keywords, and/or to reduce the frequency of stuttering on keywords, might reduce impacts of stuttering on sentence processing in listeners. In addition, it has been shown that self-acknowledgement of stuttering can decrease negative listener reactions to stuttered speech (Lee & Manning, 2010). More generally, being instructed to ignore versus attend to the presence of stuttering has been shown to reduce negative impacts of stuttering on sentence recall (Sander, 1965). A particularly interesting domain for future research will be to determine whether these various speaker strategies reduce negative impacts of stuttering on the processing and interpretation of sentences in listeners hearing stuttered speech.

Acknowledgements

We thank Alissa Belmont for assistance in preparing the stimuli and John Lim for assistance in conducting the experiment. We also thank the speaker for volunteering his time to record the utterances.

Disclosure of Interest

The authors report no conflict of interest. This research was supported in part by grants from NICHD (F32 HD084100) awarded to M.W.L. and NIDCD (R56 DC013545) awarded to N.D.M. and F.F.

References

Arnold JE, Fagnano M, & Tanenhaus MK (2003). Disfluencies signal theee, um, new information. Journal of Psycholinguistic Research, 32, 25–36. doi: 10.1023/A:1021980931292 [PubMed: 12647561]

Arnold JE, Hudson Kam CL, & Tanenhaus MK (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

Arnold JE, Tanenhaus MK, Altmann R, & Fagnano M (2004). The old and thee, uh, new. Psychological Science, 15, 578–581. doi: 10.1111/j.0956-7976.2004.00723.x [PubMed: 15327627]

Baayen RH, Davidson DH, & Bates DM (2008). Mixed-effects modeling with crossed random effects for subjects and items. Journal of Memory and Language, 59, 390–412. doi: 10.1016/j.jml.2007.12.005

Bailey KGD, & Ferreira F (2003). Disfluencies affect the parsing of garden-path sentences. Journal of Memory & Language, 49, 183–200. doi: 10.1016/S0749-596X(03)00027-5

Bailey KGD, & Ferreira F (2007). The processing of filled pause disfluencies in the visual world In van Gompel RPG, Fischer MH, Murray WS, & Hill RL (Eds.), Eye movements: A window on mind and brain (pp. 485–500). Oxford, UK: Elsevier.

Barr DJ (2008). Analyzing 'visual world' eyetracking data using multilevel logistic regression. Journal of Memory and Language, 59, 457–474. doi: 10.1016/j.jml.2007.09.002

Barr DJ, Levy R, Scheepers C, & Tily HJ (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. Journal of Memory and Language, 68, 255–278. doi: 10.1016/j.jml.2012.11.001

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

Bates D, Maechler M, & Dai B (2012). lme4: Linear mixed-effects models using S4 classes. Available at http://lme4.r-forge.r-project/org.

Bloodstein O, & Bernstein Ratner N (2008). A handbook on stuttering. New York: Thomson Delmar Learning.

Bortfeld H, Leon S, Bloom J, Schober M, & Brennan S (2001). Disfluency rates in conversation: Effects of age, relationship, topic, role, and gender. Language and Speech, 44, 123–147. doi: 10.1177/00238309010440020101 [PubMed: 11575901]

Bosker HR, Quené H, Sanders T, & de Jong NH (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

Bowers AL, Crawcour SC, Saltuklaroglu T, & Kalinowski J (2010). Gaze aversion to stuttered speech: a pilot study investigating differential visual attention to stuttered and fluent speech. International Journal of Language & Communication Disorders, 45, 133–144. doi: 10.3109/13682820902763951 [PubMed: 22748027]

Brennan SE, & Schober MF (2001). How listeners compensate for disfluencies in spontaneous speech. Journal of Memory and Language, 44, 274–296. doi: 10.1006/jmla.2000.2753

Clark A (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. Behavioral and Brain Sciences, 36, 181–204. doi: 10.1017/S0140525X12000477 [PubMed: 23663408]

Corley M (2010). Making predictions from speech with repairs: Evidence from eye movements. Language and Cognitive Processes, 25, 706–727. doi: 10.1080/01690960903512489

Corley M, & Hartsuiker RJ (2011). Why *um* helps auditory word recognition: The temporal delay hypothesis. PLoS One, 6, e19792. doi: 10.1371/journal.pone.0019792 [PubMed: 21611164]

Corley M, MacGregor LJ, & Donaldson DI (2007). It's the way that you, er, say it: Hesitations in speech affect language comprehension. Cognition, 105, 658–668. doi: 10.1016/j.cognition.2006.10.010 [PubMed: 17173887]

Cyprus S, Hezel RT, Rossi D, & Adams MR (1984). Effects of simulated stuttering on listener recall. Journal of Fluency Disorders, 9, 191–197. doi: 10.1016/0094-730X(84)90012-3

Ferreira F, & Bailey KGD (2004). Disfluencies and human language comprehension. Trends in Cognitive Sciences, 8, 231–237. doi: 10.1016/j.tics.2004.03.011 [PubMed: 15120682]

Ferreira F, Lau EF, & Bailey KGD (2004). Disfluencies, language comprehension, and tree adjoining grammars. Cognitive Science, 28, 721–749. doi: 10.1207/s15516709cog2805_5

Ferreira F, & Lowder MW (2016). Prediction, information structure, and good-enough language processing. Psychology of Learning and Motivation, 65, 217–247. doi: 10.1016/bs.plm.2016.04.002

Fox Tree JE (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

Fraundorf SH, & Watson DG (2011). The disfluent discourse: Effects of filled pauses on recall. Journal of Memory and Language, 65, 161–175. doi: 10.1016/j.jml.2010.06.004 [PubMed: 21765590]

Gibson E, Bergen L & Piantadosi ST (2013). Rational integration of noisy evidence and prior semantic expectations in sentence interpretation. Proceedings of the National Academy of Sciences, 110, 8051–8056. doi:10.1073/pnas.1216438110

Gibson E, Piantadosi ST, Brink K, Bergen L, Lim E, & Saxe R (2013). A noisy-channel account of crosslinguistic word-order variation. Psychological Science, 24, 1079–1088. doi: 10.1177/0956797612463705 [PubMed: 23649563]

Grodner DJ, & Sedivy J (2011). The effect of speaker-specific information on pragmatic inferences In Gibson E and Pearlmutter N (Eds.), The processing and acquisition of reference (pp. 239–272). Cambridge, MA: MIT Press.

Guntupalli VK, Everhart DE, Kalinowski J, Nanjundeswaran C, & Saltuklaroglu T (2007). Emotional and physiological responses of fluent listeners while watching the speech of adults who stutter. International Journal of Language & Communication Disorders, 42, 113–129. doi: 10.1080/10610270600850036 [PubMed: 17365090]

Guntupalli VK, Kalinowski J, Nanjundeswaran C, Saltuklaroglu T, & Everhart DE (2006). Psychophysiological responses of adults who do not stutter while listening to stuttering. International Journal of Psychophysiology, 62, 1–8. doi: 10.1016/j.ijpsycho.2005.11.001 [PubMed: 16414137]

Guntupalli VK, Nanjundeswaran C, Dayalu VN, & Kalinowski J (2012). Autonomic and emotional responses of graduate student clinicians in speech–language pathology to stuttered speech. International Journal of Language & Communication Disorders, 47, 603–608. doi: 10.1111/j.1460-6984.2012.00162.x [PubMed: 22938070]

Hanulíková A, van Alphen PM, van Goch MM, & Weber A (2012). When one person's mistake is another's standard usage: The effect of foreign accent on syntactic processing. Journal of Cognitive Neuroscience, 24, 878–887. doi: 10.1162/jocn_a_00103 [PubMed: 21812565]

Huettig F, & Mani N (2016). Is prediction necessary to understand language? Probably not. Language, Cognition and Neuroscience, 31, 19–31. doi:10.1080/23273798.2015.1072223

Hulit LM (1976). Effects of nonfluencies on comprehension. Perceptual and Motor Skills, 42, 1119–1122. doi: 10.2466/pms.1976.42.3c.1119

Jaeger TF (2008). Categorical data analysis: Away from ANOVAs (transformation or not) and towards logit mixed models. Journal of Memory and Language, 59, 434–446. doi: 10.1016/j.jml.2007.11.007 [PubMed: 19884961]

Kidd C, White KS, & Aslin RN (2011). Toddlers use speech disfluencies to predict speakers' referential intentions. Developmental Science, 14, 925–934. doi: 10.1111/j.1467-7687.2011.01049.x [PubMed: 21676111]

Kuperberg GR, & Jaeger TF (2016). What do we mean by prediction in language comprehension? Language, Cognition and Neuroscience, 31, 32–59. doi: 10.1080/23273798.2015.1102299

Lau EF, & Ferreira F (2005). Lingering effects of disfluent material on comprehension of garden path sentences. Language and Cognitive Processes, 20, 633–666. doi: 10.1080/01690960444000142

Lee K, & Manning WH (2010). Listener responses according to stuttering self-acknowledgment and modification. Journal of Fluency Disorders, 35, 110–122. doi: 10.1016/j.jfludis.2010.04.001 [PubMed: 20609332]

Lowder MW, & Ferreira F (2016a). Prediction in the processing of repair disfluencies. Language, Cognition and Neuroscience, 31, 73–79. doi: 10.1080/23273798.2015.1036089

Lowder MW, & Ferreira F (2016b). Prediction in the processing of repair disfluencies: Evidence from the visual-world paradigm. Journal of Experimental Psychology: Learning, Memory, and Cognition, 42, 1400–1416. doi: 10.1037/xlm0000256

Lowder MW, & Ferreira F (in press). I see what you meant to say: Anticipating speech errors during online language processing. Journal of Experimental Psychology: General. Advance online publication. doi: 10.1037/xge0000544

Luke SG, & Christianson K (2016). Limits on lexical prediction during reading. Cognitive Psychology, 88, 22–60. doi: 10.1016/j.cogpsych.2016.06.002 [PubMed: 27376659]

Matin E, Shao KC, & Boff KR (1993). Saccadic overhead: Information processing time with and without saccades. Perception & Psychophysics, 53, 372–380. doi: 10.3758/BF03206780 [PubMed: 8483701]

Maxfield ND, & Ferreira F (2019). Backward-looking sentence processing in typically disfluent versus stuttered speech: ERP evidence. Language, Cognition and Neuroscience, 34, 561–579. doi: 10.1080/23273798.2018.1533140

Maxfield ND, Lyon JM, & Silliman ER (2009). Disfluencies along the garden path: Brain electrophysiological evidence of disrupted sentence processing. Brain & Language, 111, 86–100. doi: 10.1016/j.bandl.2009.08.003 [PubMed: 19765813]

Panico J, & Healey EC (2009). Influence of text type, topic familiarity, and stuttering frequency on listener recall, comprehension, and mental effort. Journal of Speech, Language, and Hearing Research, 52, 534–546. doi: 10.1044/1092-4388(2008/07-0238)

Panico J, Healey EC, Brouwer K, & Susca M (2005). Listener perceptions of stuttering across two presentation modes: A quantitative and qualitative approach. Journal of Fluency Disorders, 30, 65–85. doi: 10.1016/j.jfludis.2005.01.003 [PubMed: 15769499]

Rooth M (1992). A theory of focus interpretation. Natural Language Semantics, 1, 75–116. doi: 10.1007/BF02342617

Sander EK (1965). Comments on investigating listener reaction to speech disfluency. Journal of Speech & Hearing Disorders, 30, 159–165. doi: 10.1044/jshd.3002.159 [PubMed: 14281428]

Susca M, & Healey EC (2001). Perceptions of simulated stuttering and fluency. Journal of Speech, Language, and Hearing Research, 44, 61–72. doi: 10.1044/1092-4388(2001/006)

Susca M, & Healey EC (2002). Listener perceptions along a fluency-disfluency continuum: A phenomenological analysis. Journal of Fluency Disorders, 27, 135–161. doi: 10.1016/S0094-730X(02)00126-2 [PubMed: 12145983]

Tsiamtsiouris J, & Cairns HS (2013). Effects of sentence-structure complexity on speech initiation time and disfluency. Journal of fluency disorders, 38, 30–44. doi: 10.1016/j.jfludis.2012.12.002 [PubMed: 23540911]

Van Berkum JJA, van den Brink D, Tesink CMJY, Kos M, & Hagoort P (2008). The neural integration of speaker and message. Journal of Cognitive Neuroscience, 20, 580–591. doi: 10.1162/jocn.2008.20054 [PubMed: 18052777]

Wingate ME (1964). A standard definition of stuttering. Journal of Speech and Hearing Disorders, 29, 484–489. doi: 10.1044/jshd.2904.484 [PubMed: 14257050]

Yairi E, & Seery CH (2015). Stuttering: Foundations and clinical applications (Second Edition). Upper Saddle River, NJ: Pearson.

Zhang J, Kalinowski J, Saltuklaroglu T, & Hudock D (2010). Stuttered and fluent speakers' heart rate and skin conductance in response to fluent and stuttered speech. International Journal of Language & Communication Disorders, 45, 670–680. doi: 10.3109/13682820903391385 [PubMed: 20017588]

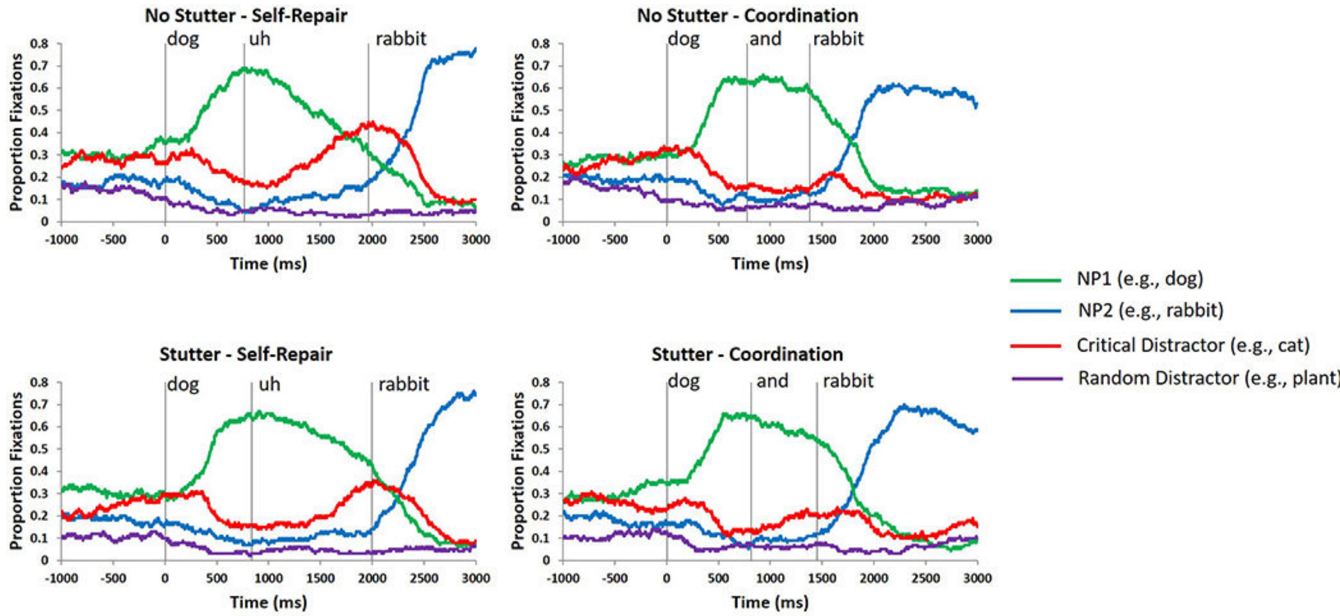

Figure 1:
Fixation plots for Self-Repair and Coordination conditions. Vertical lines represent mean onset times for critical words.

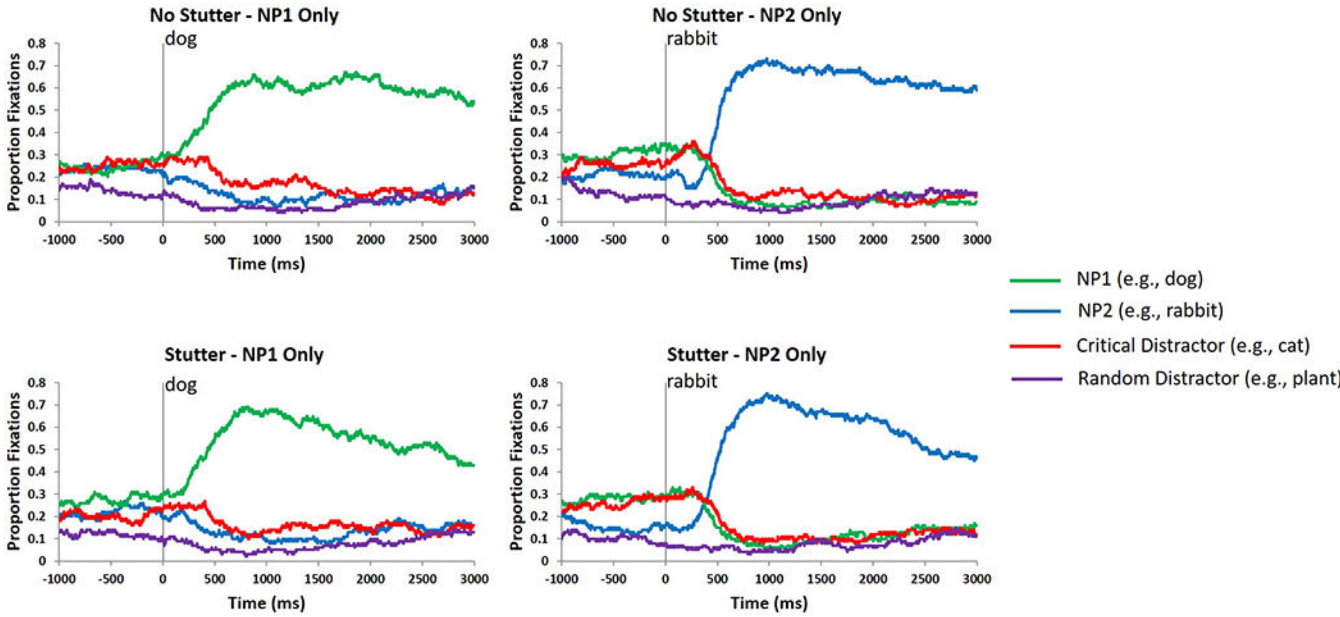

Figure 2:

Fixation plots for NP1-Only and NP2-Only conditions.

Table 1

Results of mixed effects analyses for Self-Repair and Coordination conditions.

	NP1				NP2				Critical Distractor			
Model parameters	**Estimate**	***SE***	***z***	***P***	**Estimate**	***SE***	***z***	***P***	**Estimate**	***SE***	***z***	***P***
Window 1												
Intercept	1.57	.14	11.26	< .001	−1.00	.11	−9.87	< .001	−.18	.10	−1.77	< .10
Sentence	.13	.07	1.82	< .10	−.06	.06	−1.01	*n.s.*	.00	.06	.02	*n.s.*
Speaker	.03	.13	.20	*n.s.*	−.05	.08	−.70	*n.s.*	−.09	.06	−1.56	*n.s.*
Sentence*Speaker	−.01	.07	−.19	*n.s.*	.02	.06	.33	*n.s.*	.01	.06	.10	*n.s.*
Window 2												
Intercept	1.44	.16	9.14	< .001	−1.15	.15	−7.82	< .001	−.24	.13	−1.89	< .10
Sentence	.21	.07	2.93	< .005	.28	.07	4.09	< .001	.58	.06	9.15	< .001
Speaker	−.08	.14	−.55	*n.s.*	−.17	.10	−1.70	< .10	−.11	.10	−1.05	*n.s.*
Sentence*Speaker	.06	.07	.86	*n.s.*	−.07	.07	−1.08	*n.s.*	−.16	.06	−2.51	< .02
Window 3												
Intercept	−.28	.12	−2.36	< .02	2.51	.24	10.40	< .001	−.23	.09	−2.49	< .02
Sentence	−.44	.06	−7.09	< .001	.24	.09	2.63	< .01	.19	.06	3.27	< .005
Speaker	.03	.10	.28	*n.s.*	.13	.21	.64	*n.s.*	−.06	.07	−.85	*n.s.*
Sentence*Speaker	.02	.06	.33	*n.s.*	−.13	.09	−1.35	*n.s.*	−.14	.06	−2.35	< .02

Table 2

Results of mixed effects analyses for NP1-Only and NP2-Only conditions.

	NP1-Only (looks to NP1)				NP2-Only (looks to NP2)			
Model parameters	**Estimate**	***SE***	***z***	***P***	**Estimate**	***SE***	***z***	***P***
Window 1								
Intercept	1.89	.22	8.70	< .001	2.10	.25	8.55	< .001
Speaker	.06	.20	.32	*n.s.*	.06	.19	.34	*n.s.*
Window 2								
Intercept	1.56	.19	8.40	< .001	2.27	.28	8.24	< .001
Speaker	−.08	.16	−.53	*n.s.*	.06	.26	.24	*n.s.*
Window 3								
Intercept	.99	.20	4.94	< .001	1.13	.17	6.61	< .001
Speaker	−.33	.16	−2.05	< .05	−.34	.17	−2.04	< .05