BOTH NATIVE AND NON-NATIVE DISFLUENCIES TRIGGER LISTENERS' ATTENTION

Hans Rutger Bosker[1], Jade Tjiong[2], Hugo Quené[2], Ted Sanders[2], and Nivja H. de Jong[2]

[1]Max Planck Institute for Psycholinguistics, Nijmegen; [2]UiL OTS, Utrecht University
HansRutger.Bosker@mpi.nl

ABSTRACT

Disfluencies, such as *uh* and *uhm*, are known to help the listener in speech comprehension. For instance, disfluencies may elicit prediction of less accessible referents and may trigger listeners' attention to the following word. However, recent work suggests differential processing of disfluencies in native and non-native speech. The current study investigated whether the beneficial effects of disfluencies on listeners' attention are modulated by the (non-)native identity of the speaker. Using the Change Detection Paradigm, we investigated listeners' recall accuracy for words presented in disfluent and fluent contexts, in native and non-native speech. We observed beneficial effects of both native and non-native disfluencies on listeners' recall accuracy, suggesting that native and non-native disfluencies trigger listeners' attention in a similar fashion.

Keywords: disfluencies, attention, non-native speech, Change Detection Paradigm.

1. INTRODUCTION

Disfluencies are "phenomena that interrupt the flow of speech and do not add propositional content to an utterance" [13], such as silent pauses, filled pauses (e.g., *uh* and *uhm*), slow speech, corrections, and repetitions. Despite their negative effects on listeners' impressions of the speaker's fluency level [5], disfluencies may have beneficial effects on the cognitive processes involved in speech comprehension, such as prediction and attention. For example, because disfluencies tend to occur before less accessible lexical items [19, 3, 18, 15], listeners may use disfluencies to predict more complex content to follow [1].

However, recent work suggests that there are differences in the way native and non-native fluency characteristics are processed (e.g., [7]). For instance, the beneficial effect of disfluencies on prediction may be attenuated when listening to a non-native speaker. That is, where native disfluencies may elicit prediction of low-frequency referents, disfluencies in non-native speech do not [6]. This attenuation has been argued to be due to variation in the distribution of disfluencies in native and non-native speech. Non-native speakers produce more disfluencies than native speakers and with a more irregular distribution [12, 22, 16]. As such, they are worse predictors of the word to follow, attenuating listeners' predictive strategies.

Apart from prediction effects, disfluencies are also known to trigger listeners' attention to the following word [9] as evidenced by higher recall accuracy of words heard in a disfluent context compared to a fluent context [11, 8, 14]. However, it is unknown whether the beneficial effects of disfluencies on listeners' attention are also modulated when listening to non-native speech. Using the Change Detection Paradigm [8], the current study compared how native and non-native disfluencies affect listeners' retention of words that were heard either in a fluent or a disfluent context (i.e., following a filled pause).

Given the finding that disfluency effects on prediction are attenuated in non-native speech [6], one may similarly hypothesize that effects of non-native disfluencies on attention will also be attenuated. Because native disfluencies introduce less accessible information, listeners may benefit from raising their attention as a precautionary measure to ensure timely comprehension of the unexpected information. However, non-native disfluencies follow a more irregular distribution, and, therefore, raised attention levels in response to non-native disfluencies may not prove advantageous to the listener. Therefore, the attentional effects of non-native disfluencies may be different from native disfluencies (e.g., attenuated).

Alternatively, the effects of disfluencies on attention have also been interpreted in terms of automatic cognitive consequences of temporal delay. The Temporal Delay Hypothesis [10] argues that temporal delay – inherent to disfluency – facilitates listeners' comprehension of the following content (e.g., better retention) by allowing more time to orient to the upcoming information. Following this hypothesis, native and non-native disfluencies would have similar effects on listeners' attention since they both delay the onset of the following word. The current study was designed to compare these two hypotheses.

2. METHOD

2.1. Participants

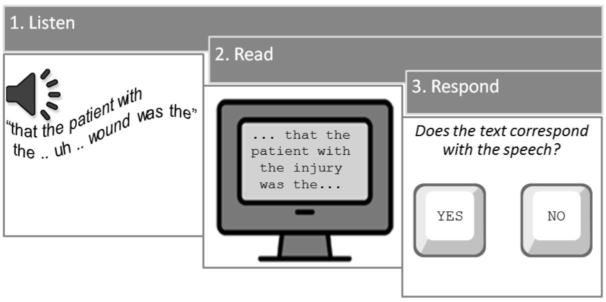
A sample of 80 native Dutch participants with normal hearing took part with implicit informed consent in accordance with local and national guidelines (M_{age}=23.3, SD_{age}=5.8, 11M/69F). Participants were randomly assigned to the native or non-native speaker condition.

2.2. Materials

A native speaker of Dutch was recorded (male, age=25) producing disfluent versions of 36 experimental story passages. These story passages were adopted from Collard [8] and consisted of three sentences. The passages were fluent except for a single filled pause (*uh*) preceding a target word. The speaker was instructed to speak as clearly as possible and to make the disfluencies sound as natural as possible. A highly proficient non-native speaker of Dutch (male, age=43, L1 Hebrew, LoR=13 years), reporting adequate knowledge of Dutch (self-reported CEFR level C1) and extensive experience with using Dutch in daily life, listened to the native recordings and subsequently imitated the native speech. Thus, matching native and non-native speech materials were obtained.

Fluent versions of the story passages were created by excising the filled pause from the disfluent version (at positive-going zero-crossings, using Praat [4]). If removing the disfluency led to an unnatural result, we instead inserted a disfluency into a fluent sentence, which was required for three native passages. Using this splicing method, fluent and disfluent versions of story passages were acoustically identical except for a filled pause appearing before a particular target word.

Figure 1: Schematical representation of the Change Detection Paradigm. Example of the Close Change condition.



2.3. Procedure

In our Change Detection Paradigm (schematically represented in Figure 1, adopted from [8]), participants listened to the fluent and disfluent passages, containing a particular target word, and then saw a written transcript of the passage. Their task was to indicate whether the transcript matched the spoken passage or not. The passages appeared in three conditions:

1. **No Change condition**:
   the transcript is identical to the spoken passage.
   (e.g., target word *wound* → *wound*)
2. **Distant Change condition**:
   the transcript contains one substitution involving a semantically unrelated noun.
   (e.g., *wound* → *handkerchief*)
3. **Close Change condition**:
   the transcript contains one substitution involving a semantically related noun.
   (e.g., *wound* → *injury*)

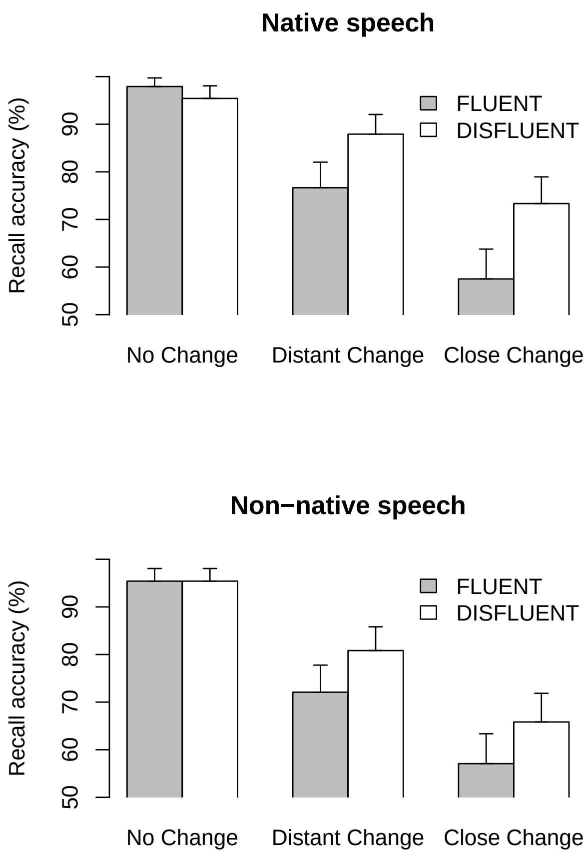
Target words from the three change conditions were matched in the log-transformed frequency of occurrence per million words, obtained from SUBTLEX-NL [17]. They were also matched in the number of characters. Target words always appeared halfway through the passage in a prepositional phrase that was out of focus.

To avoid the participants becoming accustomed to the co-occurrence of target words and disfluencies, 18 filler passages were included in the experiment which contained disfluencies without subsequent substitutions in other parts of the spoken passages. Trials were presented in pseudo-randomized order using a Latin-square design, such that all participants listened to all conditions without repeating passages. If the participant detected a substitution, he/she was asked to report the word from the audio passage that had been replaced. Finally, global accent ratings of both speakers were collected from participants using scales ranging from 1 (no accent) to 9 (very strong accent).

3. RESULTS

The accent ratings revealed a clear distinction between native and non-native speech (M_N=1.23, M_{NN}=8.08), indicating that participants clearly perceived a foreign accent in the non-native speech materials. For the comparison of recall accuracy for native vs. non-native speech, trials in which participants noticed a substitution but failed to provide the correct target word were coded as 'incorrect'. Overall recall accuracy is given in Figure 2 and was analyzed using a Generalized Linear Mixed Model (GLMM; [20]) as implemented in the `lme4` library [2] in R [21], with crossed random effects of Participants and Items. This GLMM included fixed effects of Nativeness (intercept: native speech), Disfluency (intercept: fluent speech), Change Condition (intercept: Close Change), and their interactions. This model revealed (1) an effect of Disfluency, showing a beneficial effect of disfluency on recall accuracy ($\beta = 0.86$, $z = 4.21$, $p < 0.001$), (2) effects of the different Change Conditions (No Change > Distant Change > Close Change), and (3) an interaction between Disfluency and No Change ($\beta = -1.81$, $z = -2.94$, $p = 0.003$), showing a smaller disfluency effect in the No Change condition, most likely due to a ceiling effect. However, no interactions were found between the factor Nativeness and any other predictor. The lack of interactions with the factor Nativeness indicates similar recall accuracy across native and non-native speech.

Figure 2: Mean recall accuracy in percentages. Error bars enclose 1.96 x SE, 95% CIs.



4. DISCUSSION

Results revealed that disfluencies have a beneficial effect on participants' recall accuracy. When our participants were presented with a transcript of an earlier spoken passage, they were more accurate in detecting a change in this text when the target word in the spoken passage had been preceded by a disfluency. This beneficial effect of disfluency was found for both native and non-native disfluencies, suggesting that both native and non-native disfluencies induce heightened attention to the following content.

These findings are in line with the Temporal Delay Hypothesis [10] arguing that the delay inherent to both native and non-native disfluencies allows listeners more time to orient to the upcoming information. However, one may consider an alternative explanation related to the perceived proficiency of our non-native speaker. Since our speech materials consisted of a variety of story passages with high lexical diversity and perfect grammatical accuracy, our non-native speaker produced relatively proficient Dutch speech. This may have indicated, to our listeners, a relatively high L2 proficiency. This, in turn, may have led listeners to treat non-native speech as similar to native speech. Future studies, manipulating perceived proficiency, may investigate how different (perceived) levels of L2 proficiency can affect the way non-native disfluencies are processed.

Alternatively, the absence of modulation of the disfluency effect for non-native speech may have been a result of our particular speech materials. Because we wanted to match the native and non-native speech as closely as possible, we used scripted passages (adopted from [8]). Listeners may have been aware that our speakers 'acted out' the story passages, thus preventing them from interpreting the non-native disfluencies as authentically different from the native disfluencies. Future experiments, involving spontaneous non-native speech materials and matched native counterparts, may shed light on the generalizability of the present findings.

Despite the fact that we cannot draw definitive conclusions about how non-native disfluencies affect listeners' perceptual mechanisms, our results, nonetheless, emphasize the role of attention in an account of disfluency processing. Disfluencies trigger listeners' attention with consequences for the retention of words following the disfluency.

5. REFERENCES

[1] Arnold, J. E., Hudson Kam, C. L., Tanenhaus, M. K. 2007. If you say -thee uh- you're describing something hard: The on-line attribution of disfluency during reference comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33(5), 914–930.

[2] Bates, D., Maechler, M., Bolker, B. 2012. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-39.

[3] Beattie, G. W., Butterworth, B. L. 1979. Contextual probability and word frequency as determinants of pauses and errors in spontaneous speech. *Language and Speech* 22(3), 201–211.

[4] Boersma, P., Weenink, D. 2012. Praat: doing phonetics by computer [computer program]. Version 5.3.18.

[5] Bosker, H. R., Quené, H., Sanders, T. J. M., De Jong, N. H. 2014. The perception of fluency in native and non-native speech. *Language Learning* 64, 579–614.

[6] Bosker, H. R., Quené, H., Sanders, T. J. M., De-Jong, N. H. 2014. Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of Memory and Language* 75, 104–116.

[7] Bosker, H. R., Reinisch, E. 2015. Normalization for speechrate in native and nonnative speech. *Proceedings of the 18th International Congress of Phonetic Sciences 2015 [ICPhS XVIII], Glasgow*.

[8] Collard, P. 2009. *Disfluency and listeners' attention: An investigation of the immediate and lasting effects of hesitations in speech*. PhD thesis The University of Edinburgh.

[9] Collard, P., Corley, M., MacGregor, L. J., Donaldson, D. I. 2008. Attention orienting effects of hesitations in speech: Evidence from ERPs. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 34(3), 696–702.

[10] Corley, M., Hartsuiker, R. J. 2011. Why um helps auditory word recognition: The temporal delay hypothesis. *PLoS One* 6(5), e19792.

[11] Corley, M., MacGregor, L. J., Donaldson, D. I. 2007. It's the way that you, er, say it: Hesitations in speech affect language comprehension. *Cognition* 105(3), 658–668.

[12] Davies, A. 2003. *The native speaker: Myth and reality*. Clevedon: Multilingual Matters.

[13] 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(6), 709–738.

[14] Fraundorf, S. H., Watson, D. G. 2011. The disfluent discourse: Effects of filled pauses on recall. *Journal of Memory and Language* 65(2), 161–175.

[15] Hartsuiker, R. J., Notebaert, L. 2010. Lexical access problems lead to disfluencies in speech. *Experimental Psychology* 57(3), 169–177.

[16] Kahng, J. 2014. Exploring utterance and cognitive fluency of L1 and L2 English speakers: Temporal measures and stimulated recall. *Language Learning* 64(4), 809–854.

[17] Keuleers, E., Brysbaert, M., New, B. 2010. SUBTLEX-NL: A new measure for Dutch word frequency based on film subtitles. *Behavior Research Methods* 42(3), 643–650.

[18] Levelt, W. J. M. 1983. Monitoring and self-repair in speech. *Cognition* 14, 41–104.

[19] Maclay, H., Osgood, C. E. 1959. Hesitation phenomena in spontaneous English speech. *Word* 15, 19–44.

[20] Quené, H., Van den Bergh, H. 2008. Examples of mixed-effects modeling with crossed random effects and with binomial data. *Journal of Memory and Language* 59(4), 413–425.

[21] R Development Core Team, 2012. R: A language and environment for statistical computing. ISBN 3-900051-07-0.

[22] Skehan, P. 2009. Modelling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics* 30(4), 510–532.