Perception of disfluencies in non-native speech

Rajalakshmi Satarai Madhavan (University of Göttingen) and Martin Corley (University of Edinburgh)

Background:

Disfluencies are usually defined as the false starts, hesitations, and filled pauses that occur in speech (Corley & Stewart, 2008). They are common in spontaneous speech, and can occur due to difficulties in lexical access (Arnold, Losongco, Wasow, & Ginstrom, 2000). Listeners use disfluencies to predict upcoming words: For example, listeners look more towards a low frequency object (LFO: objects with names that occur rarely during speech) when preceded by disfluency (Arnold, Fagnano & Tanenhaus, 2003; Arnold, Kam & Tanenhaus, 2007). Research about perception of disfluencies in non-native speech shows that disfluencies do not influence native listeners' predictions in the same way (Bosker, Quené, Sanders & De Jong, 2014). However, there has been no investigation to date into whether these differences stem from difficulties in comprehending non-standard accents, or from taking the speaker's perspective and attributing any disfluencies to general difficulties in formulation. The aim of this study was to distinguish these two views.

Methodology and procedure:

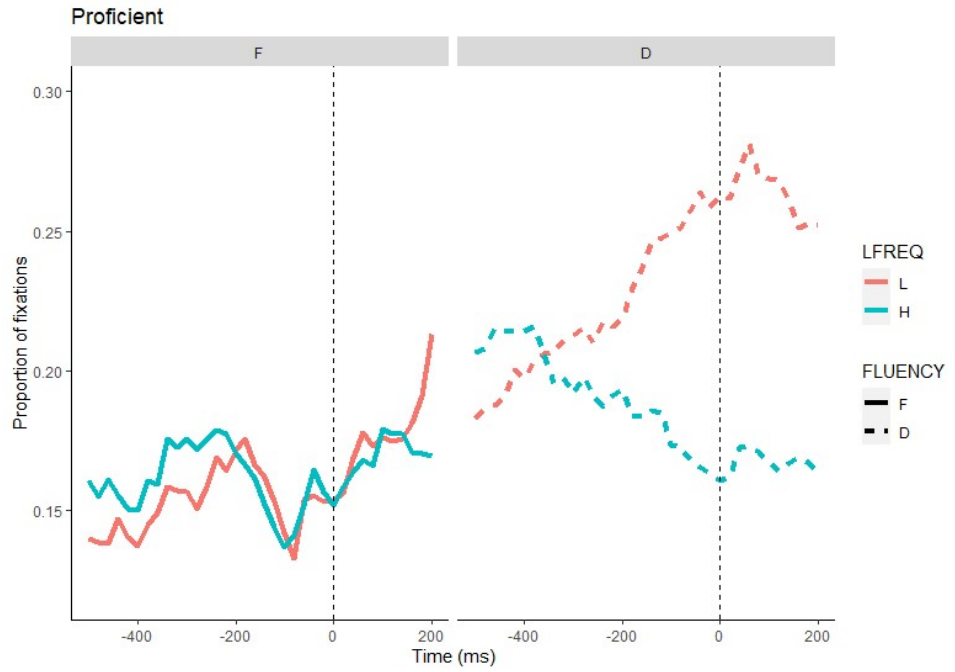
Sixty participants performed an eye-tracking study where they were randomly assigned to either a 'proficient' or 'non-proficient' non-native speaker (30 participants per condition). In both conditions the same speaker of Indian English introduced himself differently, in a brief audio recording, so as to appear 'proficient' ("I enjoy reading historical fiction") or 'nonproficient' ("I only learnt English for a short amount of time"). The only difference between conditions was the content of the introductory stories; the accent of the speaker stayed the same, and identical recordings of experimental items, using the accent and pronunciation of Indian English, were used in both conditions. Participants were presented with pictures of a high-frequency (e.g., egg) and a low-frequency (e.g., wheelbarrow) object. The speaker then gave either fluent [*Click on the…*] or disfluent [*Click on thee uh…*] instructions to click on one of the objects in the visual array. After the task was complete, participants were given a questionnaire to assess their exposure to non-native accents.

Results:

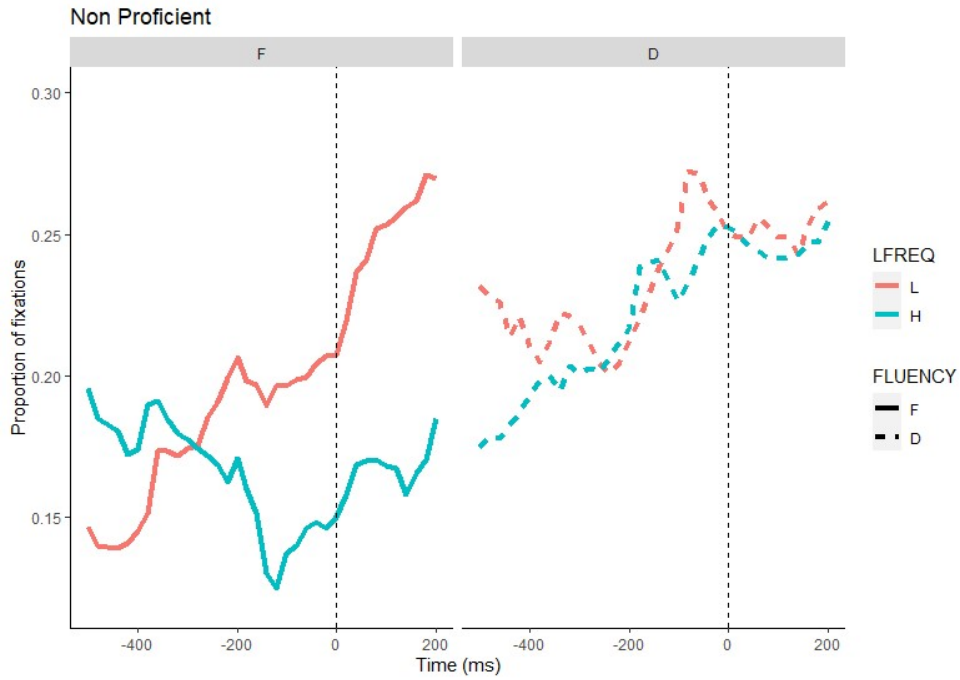
Two linear mixed models were run, one for Fluent and one for Disfluent trials, predicting the proportions of looks toward the LFO in the visual array with predictors of proficiency and linear and quadratic time terms. Following Bosker et al. (2014) , the time window for analysis ran from the start of the sentence to the onset of the target word. There were no effects of any predictors in the fluent trials. However, for disfluent trials, both time terms, and the interactions between proficiency condition and time terms were significant. Participants in the proficient speaker condition looked more towards the LFOs in the disfluent trials. The analysis of the questionnaire answers showed no differences between participants in exposure to non-native speech in daily life.

Discussion:

When they encounter non-native-sounding speech, listeners engage in perspective taking. In the present study, they anticipated the low-frequency referent when the supposedly 'proficient' speaker was disfluent, while there was no bias toward the low frequency referent when they were listening to the supposedly 'nonproficient' speaker. Thus, it appears that listeners are able modulate their assumptions about non-native speakers' disfluencies, perhaps inferring that a less proficient speaker is more likely to be disfluent for reasons other than retrieving a low-frequency object name.



(a)



(b)

Figure 1 shows the proportion of looks toward both high and low frequency objects in the (a) proficient and (b) non-proficient condition. The vertical black dashed lines show the time of the target onset. The graphs with the full lines show fluent trials, and those with the dashed lines show disfluent trials. The pink lines and blue show proportion of fixations toward high frequency objects and low frequency objects respectively.

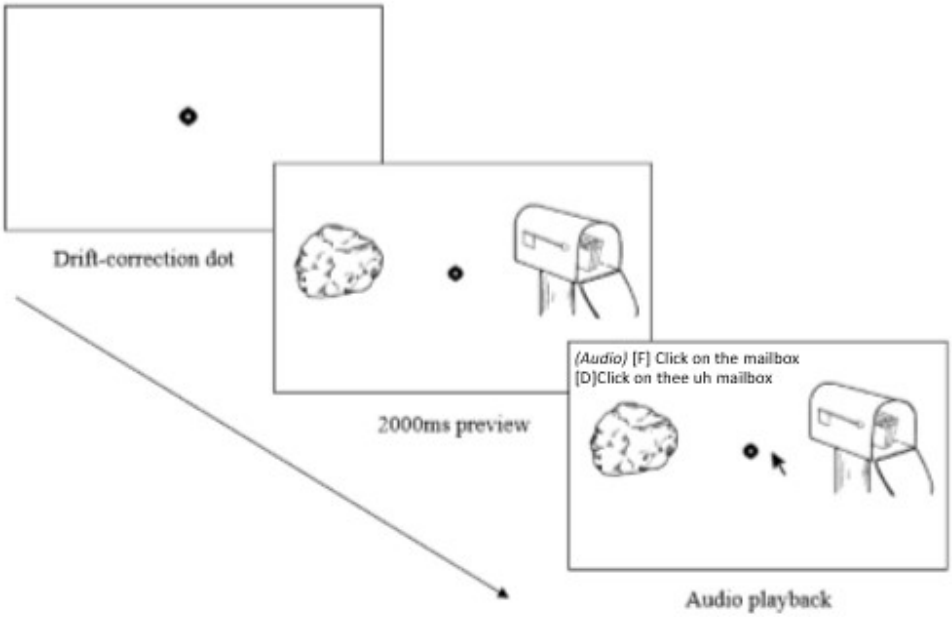


Figure 2 shows the timeline of one fluent or disfluent trial.

Key References

Arnold, J. E., Losongco, A., Wasow, T., & Ginstrom, R. (2000). Heaviness vs. newness: The effects of structural complexity and discourse status on constituent ordering. *Language, 76*(1), 28-55.

Arnold, J. E., Fagnano, M., & Tanenhaus, M. K. (2003). Disfluencies signal theee, um, new information. *Journal of psycholinguistic research, 32*(1), 25-36.

Arnold, J. E., Kam, C. L. H., & 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*(5), 914.

Bosker, H. R., Quené, H., Sanders, T., & De Jong, N. H. (2014). Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not. *Journal of memory and language, 75*, 104-116.

Corley, M., & Stewart, O. W. (2008). Hesitation disfluencies in spontaneous speech: The meaning of um. *Language and Linguistics Compass, 2*(4), 589-602.