Journal of Memory and Language 75 (2014) 104–116

Contents lists available at ScienceDirect

Journal of Memory and Language

journal homepage: www.elsevier.com/locate/jml

Native 'um's elicit prediction of low-frequency referents, but non-native 'um's do not

Hans Rutger Bosker[a,*], Hugo Quené[b], Ted Sanders[b], Nivja H. de Jong[b]

[a] Max Planck Institute for Psycholinguistics, PO Box 310, 6500 AH Nijmegen, The Netherlands
[b] Utrecht Institute of Linguistics OTS, Utrecht University, Trans 10, 3512 JK Utrecht, The Netherlands

ARTICLE INFO

Article history:
Received 25 July 2013
revision received 16 May 2014
Available online 22 June 2014

Keywords:
Prediction
Disfluency bias
Disfluency
Hesitation
Non-native speech
Speech comprehension

ABSTRACT

Speech comprehension involves extensive use of prediction. Linguistic prediction may be guided by the semantics or syntax, but also by the performance characteristics of the speech signal, such as disfluency. Previous studies have shown that listeners, when presented with the filler *uh*, exhibit a *disfluency bias* for discourse-new or unknown referents, drawing inferences about the source of the disfluency. The goal of the present study is to study the contrast between native and non-native disfluencies in speech comprehension. Experiment 1 presented listeners with pictures of high-frequency (e.g., a hand) and low-frequency objects (e.g., a sewing machine) and with fluent and disfluent instructions. Listeners were found to anticipate reference to low-frequency objects when encountering disfluency, thus attributing disfluency to speaker trouble in lexical retrieval. Experiment 2 showed that, when participants listened to disfluent non-native speech, no anticipation of low-frequency referents was observed. We conclude that listeners can adapt their predictive strategies to the (non-native) speaker at hand, extending our understanding of the role of speaker identity in speech comprehension.

Prediction in human communication lies at the core of language production and comprehension (see Kutas, DeLong, & Smith, 2011; Pickering & Garrod, 2007, for reviews). Most research into language-mediated prediction has focused on prediction elicited by semantic (e.g., Altmann & Kamide, 1999), syntactic (e.g., Van Berkum, Brown, Zwitserlood, Kooijman, & Hagoort, 2005; Wicha, Moreno, & Kutas, 2004) or phonological properties (e.g., DeLong, Urbach, & Kutas, 2005) of the linguistic input. But listeners form linguistic predictions not only based on *what* is said, but also on *how* it is said. That is, performance aspects of the speech signal also affect prediction, such as prosodic characteristics (Dahan, Tanenhaus, & Chambers, 2002; Weber, Grice, & Crocker, 2006) and disfluencies (Arnold, Tanenhaus, Altmann, & Fagnano, 2004; Arnold, Hudson Kam, & Tanenhaus, 2007). This study corroborates that disfluency may indeed guide prediction by showing that listeners—upon encountering an *uh*—anticipate reference to a low-frequency referent, but only when listening to a native speaker.

Disfluencies are "phenomena that interrupt the flow of speech and do not add propositional content to an utterance" (Fox Tree, 1995, p. 709), such as silent pauses, filled pauses (e.g., *uh*'s and *uhm*'s), corrections, and repetitions. Disfluency is a common feature of spontaneous speech: it is estimated that six in every hundred words are affected by disfluency (Bortfeld, Leon, Bloom, Schober, & Brennan, 2001; Fox Tree, 1995). Disfluencies follow a non-arbitrary distribution: they tend to occur before dispreferred or more complex content, such as open-class words (Maclay & Osgood, 1959), unpredictable lexical items (Beattie &

* Corresponding author.
E-mail address: HansRutger.Bosker@mpi.nl (H.R. Bosker).

http://dx.doi.org/10.1016/j.jml.2014.05.004

Butterworth, 1979), color names of low-frequency (Levelt, 1983), or names of low-codability images (Hartsuiker & Notebaert, 2010).

Traditionally, the disfluent character of spontaneous speech was thought to disrupt the mechanisms involved in speech perception (Martin & Strange, 1968). It was assumed to pose a continuation problem for listeners (Levelt, 1989), who were thought to be required to edit out disfluencies in order to process the remaining linguistic input. Thus, disfluencies would uniformly present obstacles to comprehension and need to be excluded in order to study speech comprehension in its 'purest' form (cf. Brennan & Schober, 2001).

Experimental evidence has shown, however, that disfluencies may help rather than hinder the listener. Disfluencies may aid comprehenders to avoid erroneous syntactic parsing (Brennan & Schober, 2001; Fox Tree, 2001), to attenuate context-driven expectations about upcoming words (Corley, MacGregor, & Donaldson, 2007; MacGregor, Corley, & Donaldson, 2010), to speed up word recognition (Corley & Hartsuiker, 2011), and to improve recognition memory (Collard, Corley, MacGregor, & Donaldson, 2008; Corley et al., 2007; Fraundorf & Watson, 2011; MacGregor et al., 2010). Moreover, they may guide prediction of the following linguistic content.

Arnold, Fagnano, and Tanenhaus (2003) and Arnold et al. (2004) investigated whether listeners use the increased likelihood of speakers to be disfluent (e.g., saying 'thee *uh* candle' instead of 'the candle') when speakers refer to new as compared to given information (Arnold, Wasow, Losongco, & Ginstrom, 2000). In eye-tracking experiments using the Visual World Paradigm, participants' eye fixations revealed that, prior to target onset, listeners were biased to look at a discourse-new referent when presented with a disfluent utterance: a *disfluency bias* toward discourse-new referents (Arnold et al., 2003, 2004). Subsequently, Arnold et al. (2007) extended the disfluency bias to the reference resolution of known vs. unknown objects (cf. Watanabe, Hirose, Den, & Minematsu, 2008). Upon presentation of a disfluent sentence such as 'Click on thee *uh* red [target]', listeners were found to look more at an unknown object (an unidentifiable abstract symbol) prior to target onset as compared to a known object (e.g., an ice-cream cone).

Additional experiments in Arnold et al. (2007) and Barr and Seyfeddinipur (2010) demonstrated that the mechanism responsible for the disfluency bias is a perspective-taking process. In the second experiment reported in Arnold et al. (2007), the authors tested whether (1) listeners 'simply' associated unknown or discourse-new referents with disfluency, or that (2) listeners actively made rapid inferences about the source of the disfluency (e.g., when the speaker is perceived to have trouble in speech production, the most probable source of difficulty is the unfamiliarity of the unknown referent). This second experiment was identical to their first experiment, except that participants were now told that the speaker suffered from object agnosia (a condition involving difficulty recognizing simple objects). Results revealed that the preference for unknown referents following a disfluency, observed in the first experiment, disappeared in the second experiment. This suggests that listeners draw inferences about the speaker's cognitive state (e.g., having equal difficulty naming known and unknown objects) which modulates the extent to which disfluency guides prediction.

This raises the question how disfluency affects prediction in a much more common situation, namely when listeners are confronted with disfluencies in non-native speech. Non-native speech is all the more vulnerable to disfluency due to, for instance, incomplete mastery of the second language (L2) or a lack of automaticity in L2 speech production (De Bot, 1992; Segalowitz, 2010). These factors lead to a higher incidence of disfluencies in non-native speech, and it causes a different distribution of disfluencies (relative to the regularities in native disfluency production; Davies, 2003; Kahng, 2013; Riazantseva, 2001; Skehan & Foster, 2007; Skehan, 2009; Tavakoli, 2011). As a consequence, the distribution of disfluencies in non-native speech may be argued to be, from the native listeners point of view, more variable than the disfluency distribution in native speech.

This different distribution of non-native disfluencies might affect listeners' predictive strategies in two possible ways: first, the disfluency bias may be attenuated when listening to non-native speech—similar to what was found for speech from an object-agnosic patient (Arnold et al., 2007). Because of their higher incidence and wider distribution, non-native disfluencies are, to the listener, worse predictors of the word to follow (as compared to native disfluencies). Thus, native listeners may refrain from using non-native disfluencies for prediction, leading to a reduction or elimination of the preference for more complex referents upon hearing non-native disfluent speech.

This first hypothesis is supported by the observed attenuation of comprehension processes when listening to non-native speech. For instance, Hanulíková, Van Alphen, Van Goch, and Weber (2012) report a classical P600 effect for grammatical gender violations in native speech. In contrast, when the same violations were produced by a non-native speaker with a foreign accent, no P600 effect was observed. The foreign accent in the non-native speech presumably served as a cue for listeners to adjust their comprehension strategies of grammatically ill-formed sentences. Hanulíková et al. (2012) argue that prior experience with non-native speakers producing syntactic errors lies at the core of this cognitive modulation. Similarly, prior experience with the different distribution of non-native disfluencies may attenuate listeners' predictive strategies.

Alternatively, the disfluency bias for dispreferred or more complex referents may be enhanced when listening to non-native speech. Naming more complex objects is (even) more cognitively demanding for a non-native speaker than it is for a native speaker. As a consequence, the likelihood of a disfluency preceding a more complex word may be argued to be higher in non-native speech than in native speech. If native listeners take this into account, their anticipation of more complex information following non-native disfluency may be enhanced (relative to native speech). As such, the disfluency bias may have an even stronger presence in the comprehension of non-native speech.

This second hypothesis is supported by studies of lexical retrieval of monolinguals and bilinguals. For instance,