Downstep in Boston Radio News Corpus

Tae-Jin Yoon
University of Illinois at Urbana-Champaign
oral or poster

This paper presents acoustic evidence for a categorial distinction between downstepped and non-downstepped high-toned pitch accents (H* vs. !H*) in Boston Radio News Corpus and offers an explanation for the contradictory findings in two prior acoustic studies of downstep. In the ToBI labeling system a high-toned pitch accent is labeled as downstepped if (i) it follows a downstep trigger and (ii) it is realized in a compressed pitch range relative to the trigger. A recurrent question with regard to downstep in English is whether the pitch accent in the compressed pitch range is categorically different from the pitch accent in the uncompressed or normal pitch range.

Previous studies investigate downstepped and non-downstepped pitch accents in English with data collected from laboratory experiments (Liberman and Pierrehumbert 1984) and with data from the Radio News corpus (Dainora 2001), but with contradictory results. Liberman and Pierrehumbert (1984) report that downstepped pitch accent differs categorically from non-downstepped pitch accents, but Dainora (2001) shows that downstepped and non-downstepped pitch accents represent different ends of a single distribution. She reports on the measure of pitch drop from a high pitch accent to a following high-toned pitch accent - both downstepped and non-downstepped. She shows that the pitch drop measure defines a unimodal distribution where downstepped and non-downstepped pitch accents are considered as a single group, and this finding contradicts the claim that downstepped pitch accents form a distinct category.

In this paper, I provide evidence in support of the claim that the downstepped pitch accent is categorically distinct from the non-downstepped pitch accent. My analysis is based on data from the Boston Radio News corpus, the same data used in Dainora (2001). Pitch values were measured from pitch peaks in the second pitch accent of 421 H* !H* sequences, where the second pitch accent is downstepped, and of 434 H* H* sequences, where the second pitch accent is not downstepped. Note that the preceding pitch accent is identical in both types of sequences. I argue that Dainora's study fails to consider the possibility that the pitch drop measure is sensitive to the pitch peak (maximal F0) of the first H* in the studied sequences. To investigate this hypothesis, I plotted the pitch peak values from the first H* against the pitch peak values from the second pitch accent (H* or !H*) in the pitch accent sequences, and applied a separate linear regression analysis for the downstepped and non-downstepped data. The results show that pitch peaks from non-downstepped H* and downstepped !H* are both highly correlated with the pitch peak of the preceding H*, but that the two accents define different clouds in the data plot, and the regression line for the downstepped !H* has a significantly lower slope (0.51 for !H* versus 1.0 for H*). This finding provides evidence for downstep as a distinct category, in support of the finding of Liberman and Pierrehumbert (1984).

Reference

Dainora, A. 2001. *An Empirically based probabilistic model of intonation in American English*. PhD. dissertation, University of Chicago.

Liberman, M. and J. Pierrehumbert. 1984. Intonational invariance under changes in pitch and length. In M. Aronoff and R. Oehrle (eds.), *Language Sound Structure*, 157-234. Cambridge, MA: MIT Press.