

A Perception Study in Black and White: Examining the Effects of Intonational Variables on Judgments of Ethnicity

American English listeners are fairly adept at making rapid judgments of speaker ethnicity, and research has shown that these judgments affect the ways in which speakers and listeners interact (Purnell et. al 1999, Foreman 2000, Thomas and Reaser 2004). Unfortunately, the social and linguistic factors that influence these judgments are still not well understood. This study tests the effect of different intonational contours and filtering conditions to better illuminate how listeners make racialized judgments about black and white American voices.

Forty American English native listeners (balanced for self-identified white/black race and male/female gender) heard sound clips that had been uttered by American English speakers with one black parent and one white parent. For intonational consistency, each clip contained two of the same type of pitch accents (either two H* or two L+H*) and had an L-L% boundary tone, and were labelled for broad or narrow focus. These contours were selected because of earlier work indicating potential differences in use between some communities of black and white speakers (Loman 1975, McLarty 2011, Holliday 2016). Listeners rated the voices in two different tasks, using a computer-based laboratory survey. In Task 1, listeners heard two repetitions of 48 phrases and were asked to judge whether the phrases had been uttered by a black or white speaker. In this task, phrases were low-pass filtered to eliminate frequencies above 400 Hz to reduce judgments based on phrase content, as well as to limit the listener's ability to rely primarily on segmental information. In Task 2, the same listeners were asked to rate the same clips, but in this condition, the clips were presented in their original form, with segmental and content information preserved.

Results of logistic regression models in R indicate that listeners appear to employ opposing strategies to determine speaker ethnicity, depending on the information available to them. In the low-pass filtered condition, tokens with the L+H* contour were significantly more likely to be rated as "black" than tokens with the H* contour, but only in the broad focus condition ($p < .001$). In the unfiltered condition, however, tokens with the L+H* contour were less likely to be rated as black than those with the H* contour, with no difference between focus conditions. These results indicate that listener perception of intonational contours as well as their interactions with focus may depend upon other available linguistic information. This result is especially of interest due to the fact that in realistic situations, listeners often hear speech filtered through background noise, distance or other voices, and most previous studies have not addressed these types of differences between filtered and unfiltered speech.

These findings are important for future research on ethnic identification because they begin to pinpoint the ways in which listeners make judgments when they have access to different types of linguistic information. The results also bring us closer to understanding the cues that listeners rely on when making ethnicity judgments, which in turn has an impact for issues such as social evaluation and linguistic profiling, and discrimination.