The Influence of Perceived Race on Ratings of Children’s Speech: Accuracy Ratings
Karen Evans1,2, Benjamin Munson1, and Jan Edwards3

(1) Department of Speech-Language-Hearing Sciences, University of Minnesota, Minneapolis; (2) Hopkins Public Schools, Hopkins, Minnesota; (3) Department of Communication Sciences and Disorders, University of Wisconsin, Madison

http://www.learningtotalk.org

Introduction

- ASHA emphasizes knowledge of the “influences of one’s own beliefs and biases in providing effective services” as essential for Speech-Language Pathologists (SLPs) serving culturally and linguistically diverse populations (ASHA, 2004). SLPs use real-time binary accuracy judgments to score standardized articulation assessments, from which they make eligibility and treatment decisions. Research is only beginning to unpack the effects of listener bias on speech perception.

- African-American English (AAE) is a distinct English dialect characterized by phonological and morpho-syntactic features including glottalization of final /l/, the reduction of final consonant clusters (including ones that mark morphology), and the vocalization or deletion of final /l/ (Thomas, 2007). African-American children, who may or may not be speakers of AAE, are consistently over-represented in Special Education (US Department of Education, 2001).

- Perceptual judgments are known to be affected by assumed age, gender, race, and region of origin of the talker (interrater variability).

- Stein Casasanto (2008) previously found that listeners were quicker to process AAE-like cluster reductions when shown a picture of an African-American speaker versus a European-American speaker.

Research Questions

- Does imputed race of child speakers affect listeners’ judgments of speech accuracy? That is, can we replicate imputed speaker-race effects found elsewhere with adult talkers?
- Do speaker-race effects differ for speech patterns that comply with AAE phonological patterns?
- Do clinical experience, knowledge of AAE features, and implicitly or explicitly held attitudes towards AAE affect these accuracy judgments?

Methods

Participants

Inexperienced Listeners: 40 adults (20 for baseline rating task and 20 for audio-visual rating task) without training in Speech-Language Pathology

Experienced Listeners for Audio-Visual Task: 12 pediatric Speech-Language Pathologists and 8 advanced graduate students in Speech-Language Pathology

Stimuli

- Elicited as part of a larger study on dialect variation (see Edwards, Gross, MacDonald, & Seidenberg, 2010). Talkers were African-American children between 4 and 9 years old from the Madison, WI area.

- We selected 160 speech samples, 10 each of 16 different words from 4 coda groups (final /t/, final /l/, final stop + plural /s/, and final stop + plural /lds/).

- Reductions or omissions are allowed in AAE for all of these types of codas. Final consonants in tokens for each word ranged from completely present to completely reduced.

- Tokens were classified as “accurate” (i.e. fully produced codas) or “mismatched” (i.e. reduced or omitted codas) by both (1) skilled transcription and (2) the results of the baseline rating task, below.

Acknowledgments

This work was supported by a Wisconsin Institutes for Discovery Seed grant to Mark Seidenberg, by NSF grant BCS-0729140 to Jan Edwards, by NIH grant R01 DC02532 to Jan Edwards, Mary E. Beckman, and Benjamin Munson, and by NICHD P50 HD03352 grant to the Waisman Center at the University of Wisconsin. We are very grateful to all of the adult participants as well as to the children who contributed the speech samples and their families. We also thank Mary E. Beckman for useful input on this work, and both Mandi Proue and Carol-June Leonard for valuable assistance in participant recruitment and testing.

Figure 1 (left). Analog Scale used in the baseline rating task.

Figure 2 (right). The Visual Analogue Scale used in the audio-visual rating task (Continued)

Figure 3 (left). Probability density functions of the implicit and explicit attitudes test scores for the participants in the current study and the participants in the study by Babel (2012) previously used by Babel (2012) to measure racial bias.

Figure 4 (left). Scatterplot showing the relationship between the implicit and explicit race attitude scores on the implicit racial bias test and the accuracy ratings for /ld/ words.

Figure 5 (left). Scatterplot showing the relationship between the implicit and explicit race attitude scores on the implicit racial bias test and the accuracy ratings for /s/ words.

Figure 6. Scatterplots showing the relationship between the implicit and explicit race attitude scores on the implicit racial bias test and the accuracy ratings for /lds/ words.

Methods (Continued)

Baseline Rating Task: Participants listened to each token and rated the token’s accuracy by clicking along a visual analog scale (Figure 1).

Audio-Visual Rating Task: Similar to baseline task, but speech samples were paired with photos of African-American (AA) or European-American (EA) children. Listeners were told that the visual images would be used to evaluate the participants’ implicit attitudes towards AAE, and that this knowledge could affect their accuracy judgments. The following attitude and knowledge measures:

- Implicit Attitude Measure: Implicit Association Test (Greenwald, McGhee, & Schwartz, 1998) uses differential reaction time to measure implicitly held biases. We implemented a version previously used by Babel (2012) to measure racial bias.

- Explicit Attitude Measure: Participants scored on 7-point Likert-type scale. Higher scores indicated more anti-AAE attitudes while lower scores indicated pro-AAE attitudes.

- Cross-linguistic implicit bias measure: A dictionary of AAE words is associated with positive attitudes while a dictionary of EA words is associated with negative attitudes. Participants are asked to rate the association of each dictionary to AAE and EA and compute their scores by subtracting the EA dictionary score from the AAE dictionary score, resulting in a single score that can range from −5 to +5, with higher scores indicating more anti-African American attitudes.

Results

- Both groups of listeners rated speech as less accurate overall when paired with African-American children’s faces (Figure 2). Interestingly, the ratings were more categorical in the audio-visual rating task than in the baseline rating task.

- By and large, the effects of race were qualitatively similar for the listeners with and without clinical training.

- Imputed race affected the accuracy ratings for final /ld/ final stop-plural /s/ words, and final /l/ words most strongly (Figures 3, 4, and 5).

- Knowledge of AAE did not mediate the effect of imputed race on accuracy ratings. Implicit and explicit attitudes toward African Americans affected the ratings of /l/-final words (Figure 6). Individuals with negative explicit attitudes toward African American English rated words to be more accurate when paired with African American children’s faces, perhaps reflecting stereotypic suppression. Individuals with positive implicit attitudes toward African Americans rated words as more accurate when paired with African American children’s faces.

Implications

- Equivalent forms may be rated differently in clinical settings when produced by African American versus European American children.

- Clinicians need to be mindful of the ways that their perceptual abilities may be affected by characteristics of the populations they serve.

- Additional research should continue to tease apart the listener- and speaker-related factors that may bias speech perception, in order to ultimately prepare clinicians to provide more effective services.

References


- Edwards, J., Gross, M., MacDonald, M., Brown, M., & Seidenberg, M. (2010). Examining aspects of code switching ability in chil


