Influences of Perceived Racial Identity on Human Talker Identification

Tyler K. Perrachione1,2,3, Joan Y. Chiao4,5 and Patrick C.M. Wong6,7

1Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge Massachusetts, U.S.A.
2Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, Department of Psychology, Northwestern University Interdepartmental Neuroscience Program, Cognitive Science Program, Department of Otolaryngology – Head & Neck Surgery, Northwestern University, Evanston, Illinois, USA
tkp@mit.edu pwong@northwestern.edu

Abstract

Talker identification, the process by which human listeners recognize individuals by their voice, is one of the most poorly understood abilities of the human auditory system. Current psychological models of talker identification rely on strict analogies to face perception and the visual system, despite differences in the objects of perception in these two modalities. Here we investigate the existence of an own-race bias in voice perception - a phenomenon which has been pivotal in the study of face perception.

Our results demonstrate an own-race bias in talker identification: Listeners of different ethnic backgrounds show an advantage for identifying individual voices of the same race as themselves. However, unlike in vision, the own-race bias in talker identification manifests based on the perceived, but not actual, race of a talker. The influence of perceived race suggests physical (voice structural) cues do not give rise to this effect. Instead, the own-race bias in talker identification is a result of listeners' asymmetric exposure to talkers' socially-acquired manners of expression (i.e. the dynamic features of voice and speech). Such manners of expression may be stereotypically associated with a particular ethnic group, although not actually exhibited by all members of that group.

Background

- AA listeners were more conservative (less likely to respond “AA”) than EA listeners.

Results:

- 10 sentences

Procedure

- Listeners were normal hearing.

Individual Talker Identification

Stimuli:

- AA-Actual: 5 most categorizable AA voices; EA-Perceived: 5 least categorizable AA voices; EA-Actual: 5 most categorizable EA voices.

Listeners:

- 11 AA (10 female, M = 19.8 yrs); 12 EA (9 female, M = 19.5 yrs) listeners; all normal hearing.

Procedure:

- Learned to recognize voices from 5 training sentences with feedback; tested on identification abilities with 5 novel sentences with no feedback. Heard voices from each group separately, counterbalanced order.

Results:

- Listener Race effect (F(2,20) = 7.494, p < 0.002) (own-race bias)
- Main effect of Talker Group (F(2,20) = 5.914, p < 0.006)
- No effect of Listener Race (p = 0.818)
- Subjects treated EA-Perceived voices more like EA-Actual than AA-Actual (F(5,10) = 6.87, p < 0.005)

Discussion

- Our results support the Perceived-Race Hypothesis, and reject the Actual-Race Hypothesis, as well as structure-only models of voice perception / talker identification.
- Listeners use more than structural cues to categorize voices by race.
- The voices of certain AA individuals are more often categorized as EA.
- Listeners exhibit an own-race bias in talker identification based not on the actual race of the talker, but whether the talker and listener share manners of expression
- AA listeners are more accurate at identifying voices that “sound” AA, EA listeners are more accurate at identifying voices that “sound” EA, regardless of their actual race.
- Talker identification involves not only distinguishing vocal structure, but also individual, culturally acquired manners of expression.

Fig. 1

An integrated model of person identification based on Perrachione & Wong (2007). Current psychological models of talker identification (e.g. Belin et al., 2004) posit the use of only vocal structure information (green), which is strictly analogous to the perception of faces (blue). The current study highlights critical contributions of speech perception mechanisms in talker identification (orange).

Fig. 2

Voice-Race Categorization

Stimuli: 12 African-American (AA) & 12 European-American (EA) males read 10 sentences; Matched for height [p = 0.743], weight [p = 0.110], or age [p = 0.950] (Lass & Brown, 1978)

Listeners: 10 AA (all female, M = 19.8yrs); 10 EA (7 female, M = 20.2yrs) listeners; all normal hearing

Procedure: Indicated perceived race (AA/EA) of talker on 480 trials (12 voices x 2 races x 10 sentences)

Results:

- AA voices more accurately categorized than AA voices [F(1,18) = 58.147, p < 0.001]
- Accuracy did not differ by listener race [p = 0.991], nor was there an interaction [p = 0.728]
- AA listeners were more sensitive to racial voice information than EA listeners [mean d'AA = 2.339, mean d'EA = 1.931, F(1,20) = 2.105, p < 0.05]
- AA listeners were more conservative (less likely to respond “AA”) than EA listeners [mean d'EA = 1.584, mean d'EA = 0.725, F(1,20) = 2.284, p < 0.035]

Fig. 3

Individual Talker Identification

Stimuli: AA-Actual: 5 most categorizable AA voices; EA-Perceived: 5 least categorizable AA voices; EA-Actual: 5 most categorizable EA voices.

Listeners: 11 AA (10 female, M = 21.1yrs); 12 EA (9 female, M = 19.5yrs) listeners; all normal hearing

Procedure: Learned to recognize voices from 5 training sentences with feedback; tested on identification abilities with 5 novel sentences with no feedback. Heard voices from each group separately, counterbalanced order.

Results:

- Listener Race effect (F(2,20) = 7.494, p < 0.002) (own-race bias)
- Main effect of Talker Group (F(2,20) = 5.914, p < 0.006)
- No effect of Listener Race (p = 0.818)
- Subjects treated EA-Perceived voices more like EA-Actual than AA-Actual (F(5,10) = 6.87, p < 0.005)