

# Influences of Perceived Racial Identity on Human Talker Identification



NORTHWESTERN UNIVERSITY

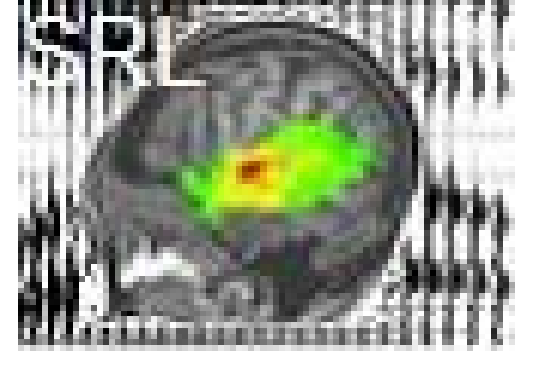
Tyler. K. Perrachione<sup>1,2,5</sup>, Joan. Y. Chiao<sup>3,5</sup> and Patrick. C.M. Wong<sup>2,4,6</sup>

<sup>1</sup>Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge Massachusetts, U.S.A.

<sup>2</sup>Roxelyn and Richard Pepper Department of Communication Sciences and Disorders, <sup>3</sup>Department of Psychology, <sup>4</sup>Northwestern University Interdepartmental Neuroscience Program, <sup>5</sup>Cognitive Science Program, <sup>6</sup>Department of Otolaryngology – Head & Neck Surgery, Northwestern University, Evanston, Illinois, USA

tkp@mit.edu

pwong@northwestern.edu



Speech Research Laboratory  
http://www.soc.northwestern.edu/wong/

## 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 specifically 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

- Is there an own-race bias in talker identification, and what is its perceptual basis?
- Face perception distinguishes invariant structural features of faces (Bruce & Young, 1984)
  - Asymmetric cultural experiences with racial differences in structural features and configurations leads to an own-race bias in face perception (Meisner & Brigham, 2001)
- Current models of talker identification are strictly analogized to face perception: Vocal identity is computed only from invariant structural features of the voice (Belin et al., 2004)
  - Actual-Race Hypothesis: Structure-only models predict an own-race bias in talker identification would result from asymmetric cultural experiences with racial differences in the structural features of voices
- However, there are reasons to believe any own-race bias in talker identification would be independent of vocal structure:
  - Voice structural features (e.g. oral or pharyngeal cavities' volumes or lengths, mean F1, F2 or F3 of /a/) do not differ significantly between races (Xue et al., 2006)
  - Phonetic information is sufficient for talker identification (Remez et al., 1997)
  - Talker identification draws on speech perception abilities (Perrachione & Wong, 2007)
  - Individuals of the same race differ in manners of speech (Evans & Iverson, 2004) and members of different races often share one dialect (Thomas & Reaser, 2004)
  - Perceived-Race Hypothesis: An own-race bias in voice perception results from listeners' asymmetric exposure to different manners of talking. Listeners will show an advantage for identifying voices not only of their own race, but also voices of another race that share dialectal features of the listener

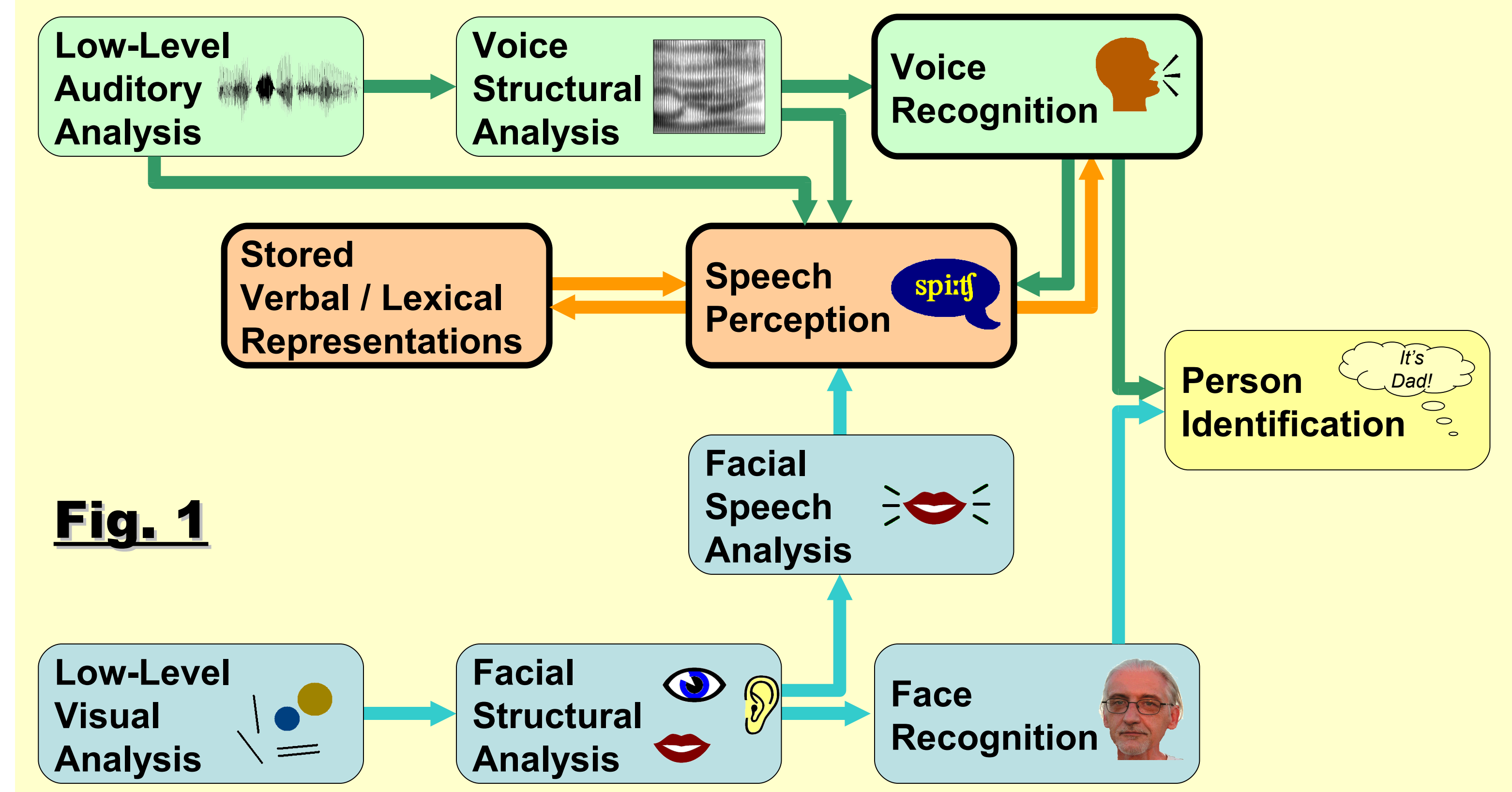


Fig. 1

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).

## 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 [ $p = 0.905$ ] (Lass & Brown, 1978)

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

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

### Results:

- EA 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  $d'_{EA} = 1.931, t(18) = 2.105, p < 0.05$ ]
- AA listeners were more conservative (less likely to respond "AA") than EA listeners [mean  $\ln \beta_{AA} = 1.584$ , mean  $\ln \beta_{EA} = 0.725, t(18) = 2.284, p < 0.035$ ]

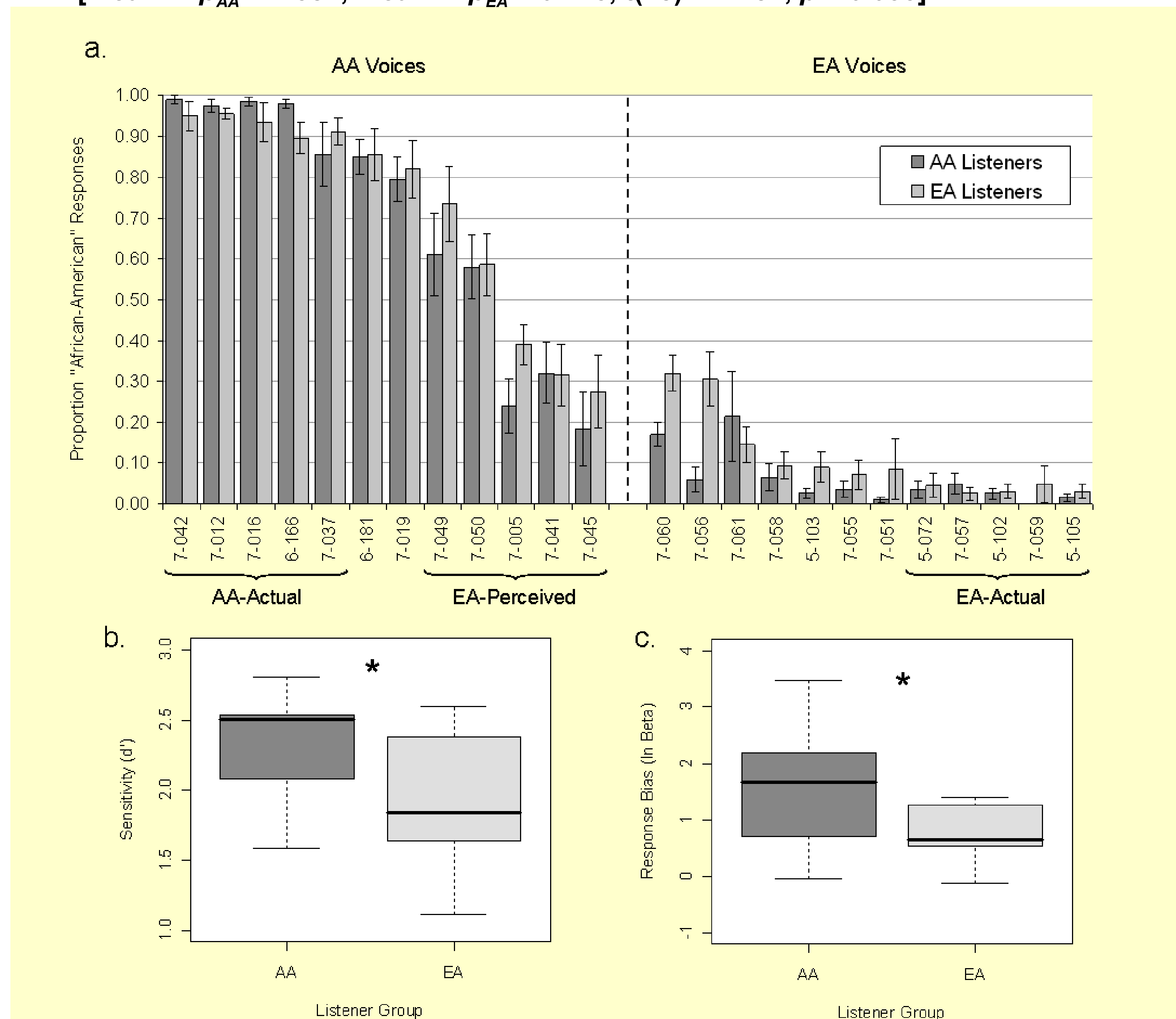


Fig. 2

## 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,  $M = 21.1$  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  $\times$  Talker Group interaction [ $F(2,20) = 7.494, p < 0.002$ ] (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 [ $z_{diff} = 1.687, p < 0.05$ ]

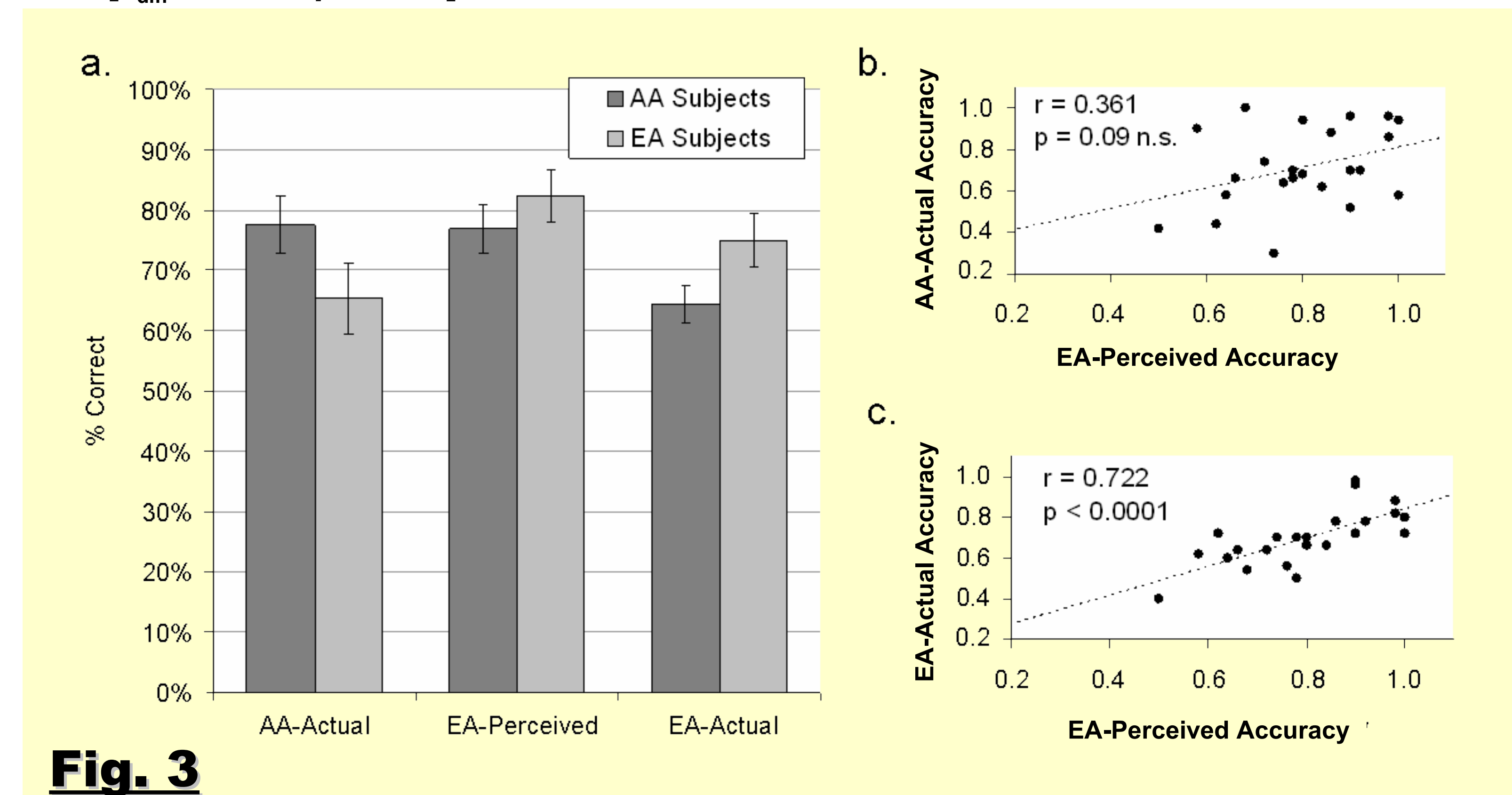


Fig. 3

## 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.

Belin, P., Fecteau, S. & Bédard, C. (2004) *Trends. Cogn. Sci.*, **3**, 129-135.  
 Bruce, V. & Young, A. (1986) *Brit. J. Psychol.*, **77**, 305-327.  
 Evans, B.G. & Iverson, P. (2004) *J. Acoust. Soc. Am.*, **115**, 352-361.  
 Lass, N.J. & Brown, W.S. (1978) *J. Acoust. Soc. Am.*, **63**, 1218-1220.  
 Meissner, C.A. & Brigham, J.C. (2001) *Psychol. Public Pol. L.*, **7**, 3-35.  
 Perrachione, T.K. & Wong, P.C.M. (2007) *Neuropsychologia*, **45**, 1899-1910.

Remez, R.E., Fellowes, J.M., Rubin, P.E. (1997) *J. Exp. Psychol. Human*, **23**, 651-666.  
 Thomas, E.R. & Reaser, J. (2004) *J. Socioling.*, **8**, 54-87.  
 Xue, S.A., Hao, G.J.P. & Mayo, R. (2006) *Clin Linguist Phonet*, **20**, 691-702.

We thank Geshri Gunasekara, Johnston Chen, Louisa Ha, Peter Hsieh, and Tasha Dees for their assistance. This work was supported by the National Institutes of Health (USA) grants R03HD051827 and R21DC007468 to P.W.