Increased Left-Hemisphere Contribution To Native- Versus Foreign-Language Talker Identification Revealed by Dichotic Listening

Tyler K. Perrachione[1,3] & Patrick C.M. Wong[2,4,5]

[1]Department of Linguistics, [2]Roxelyn & Richard Pepper Department of Communication Sciences and Disorders, [3]Cognitive Science Program, [4]Northwestern University Interdepartmental Neuroscience Program, [5]Program in Computational Biology and Bioinformatics
Northwestern University, Evanston, Illinois, U.S.A.
pwong@northwestern.edu

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

Abstract

Previous studies of human listeners' ability to identify speakers by voice have revealed a reliable language-familiarity effect: Listeners are better at identifying voices when they can understand the language being spoken. It has been claimed that talker identification is facilitated in a familiar language because of functional integration between the cognitive systems underlying speech and voice perception. However, prior studies have not provided specific evidence demonstrating neural integration between these two systems.

Using dichotic listening as a means to assess the role of each hemisphere in talker identification, we show that listeners' right-, but not left-, ear (left-hemisphere) performance better predicts overall accuracy in their native than non-native language. By demonstrating functional integration of speech perception regions (classical left-hemisphere language areas) in a talker identification task, we provide evidence for a neurologic basis underlying the language-familiarity effect.

Introduction

Behavioral evidence suggests a functional integration between the cognitive systems responsible for the perception of speech and voice (talker identity):

- **Language-Familiarity Effect: Listeners who understand talkers' speech are better at identifying those voices than listeners who are unfamiliar with the language being spoken (Perrachione & Wong, 2007; Goggin et al., 1991).**
- **Talker-Normalization and Talker-Specific Learning: Variability due to voice can affect memory for spoken words (Bradlow, Nygaard, & Pisoni, 1999), speed and accuracy of word recognition (Mullenix & Pisoni, 1990), and even perception of vowel quality (Johnson, 1990).**

Specific neural systems underlie these complex auditory perception abilities:

- **Speech Perception is supported primarily by a network of left-hemisphere cortical structures, including especially the superior temporal gyrus (Wernicke's area) and anterior temporal lobe.**
- **Voice Perception appears to be supported by the superior temporal sulcus, especially in the right hemisphere (e.g. Belin et al., 2000).**

We employ a behavioral measure (dichotic listening) of neurologic function to demonstrate the biological integration of speech- and voice-perception abilities, likely underlying the language-familiarity effect.

Methods

Subjects:
- **English L1 Group: 12 individuals (10 female), age 18-29yrs (M = 22.2)**
- **Mandarin L1 Group: 13 individuals (9 female) age 18-31yrs (M = 23.6)**
- **All were right-handed, no known auditory or neurologic impairment**

Stimuli:
- **10 English sentences read by 5 male L1 talkers (19-26yrs, M = 21.6)**
- **10 Mandarin sentences read by 5 male L1 talkers (21-26yrs, M = 22.6)**
- **Digitally recorded at 22.05kHz, normalized to 70dB SPL, 2.3sec duration**
- **5 sentences were designated "practice" the other 5 as "test" sentences**

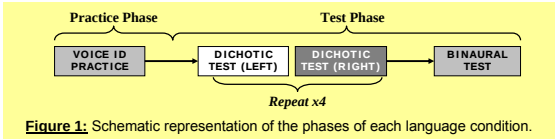
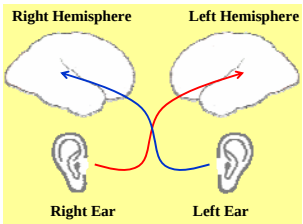


Figure 1: Schematic representation of the phases of each language condition.

Procedure:
- **Primary task was to identify the talker. "Which voice do you hear?"**
- **Subjects participated in both language conditions, English & Mandarin.**

Practice Phase:
- **Subjects were introduced to the voices and practiced recognizing them with feedback.**

Dichotic Test Phase:
- **Directed alternately to attend to the left or the right ear for blocks of 25 stimuli**
- **200 stimulus presentations (5 target voices × 4 distracter voices × 5 sentences × 2 ears)**
- **Used as a measure of "ear advantage" or "hemispheric involvement"**



Binaural Test Phase:
- **Same voice & sentence in each ear**
- **25 stimulus presentations (5 targets × 5 sentences)**
- **Used as a measure of "overall talker identification accuracy"**

Results

Language Familiarity Effect

- **Significant Group × Condition interaction [$F(1,23) = 50.024$, $p < 0.001$]**
  - **Subjects more accurate at talker identification in their native language**
- **Marginally higher performance by Mandarin group [$F(1,23) = 3.372$, $p = 0.079$]**

Lateralization / Ear Advantage

- **Significant Condition × Ear interaction [$F(1,23) = 6.58$, $p < 0.02$]**
  - **Left-Ear Advantage for all subjects when identifying English talkers**
  - **No such advantage for either group when identifying Mandarin talkers**
- **Marginal main effect of Ear ($p = 0.07$) likely driven by English condition LEA**

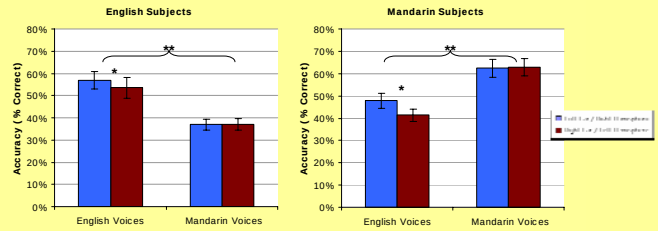


Figure 3: Talker identification accuracy. Error bars = SEM, $p < 0.02$ = * and $p < 0.001$ = **.

Hemispheric Contribution Analyses

- **Right-Ear (Left-Hemisphere) accuracy significantly better predicts overall accuracy in both English [$z = 2.121$, $p < 0.02$] and Mandarin [$z = -1.961$, $p < 0.03$]**

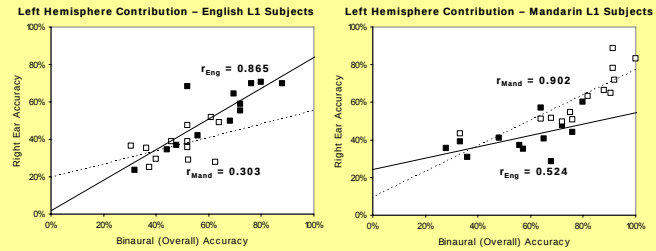


Figure 4: Predictive capacity of Right-Ear accuracy on overall accuracy by Group and Condition. Filled squares (■) and a solid line represent the English condition; Open squares (□) and a dashed line represent the Mandarin condition.

		Condition		
		Spearman's Rho		
Subject Group	*Ear*	*English*	*Mandarin*	*Difference (z)*
English L1	*Left*	0.882	0.599	1.470 (*n.s.*)
	Right	**0.865**	**0.303**	**2.121** ($p < 0.02$)
Mandarin L1	*Left*	0.545	0.812	-1.135 (*n.s.*)
	Right	**0.524**	**0.902**	**-1.961** ($p < 0.03$)

Table 1: Correlation between accuracy in each ear and overall accuracy by Group and Condition

Discussion

This study reveals three important features of the cognitive and neural systems responsible for talker identification abilities:

- **We further confirm the Language Familiarity Effect in talker identification**
- **There is a significant Left-Ear Advantage for identifying voices speaking in English, but not in Mandarin, regardless of listeners' language background.**
  - **Possibly related to different temporal properties of prosody in English and Mandarin interacting with hemispheric differences in temporal processing**
- **The predictive capacity of Right Ear / Left Hemisphere performance is a significantly better predictor of overall accuracy in one's native versus non-native language only**
  - **Increased left-hemisphere involvement in native language talker is a likely neurologic basis for the Language Familiarity Effect**

References

Belin, P., Zatorre, R., Lafaille, P., Ahad, P., Pike, B. (2000) *Nature*, *403*, 309-312.
Bradlow, A., Nygaard, L., Pisoni, D. (1999) *Percept. Psychophys.* *61*, 206-219.
Goggin, J., Thompson, C., Strube, G., Simental, L. (1991) Memory & Cognition 19, 448-458.
Johnson, K. (1990) J. Account. Soc. Am. 88, 642-654.
Mullenix, J., Pisoni, D. (1990) *Percept. Psychophys.* *47*, 379-390.
Perrachione, T., Wong, P. (2007) Neuropsychologia 45, 1899-1910.

Thanks to Geshri Gunasekera, Ajith Kumar Uppunda, Tasha Dees, and Janet Pierrehumbert for help on this project.
This work is supported by the National Institutes of Health (U.S.A.) grants HD051827 & DC007468 awarded to P.W.