A language familiarity effect for talker identification in forward but not time-reversed speech

Sara C. Dougherty, Deirdre E. McLaughlin & Tyler K. Perrachione

Department of Speech, Language, & Hearing Sciences, Boston University

BOSTON UNIVERSITY

Background

- Previous research has demonstrated that the ability to accurately identify human voices depends on listeners' familiarity with the language being spoken – a phenomenon known as the *Language Familiarity Effect* (LFE). (Goggin *et al.,* 1991; Perrachione & Wong, 2007)

- Previous studies report that linguistic processes involved in speech perception and language comprehension facilitate native-language talker identification. (Perrachione *et al.,* 2011)

- However, a recent study exploring subjective judgments of talker dissimilarity alleges that talker identification may depend on familiarity with general auditory statistics of speech, not language-specific speech processing. (Fleming *et al.,* 2014)

- The purpose of this study was to assess these two competing hypotheses:
    - **(1)** The LFE depends on linguistic processing (speech comprehension), or
    - **(2)** The LFE results from familiar, but ultimately non-linguistic, properties of speech.

Methods

Participants:

- Native English-speaking adults (N=16) with no knowledge of Mandarin.
- Native Mandarin-speaking adults (N=14) in the United States for < 4 years.
- All participants had no known or suspected speech, language or hearing impairments.

Stimuli:

- We parametrically manipulated talker identification across two levels of two conditions: **Language** (English / Mandarin) and **Intelligibility** (Forward / Reversed).
- 10 sentences recorded in each language, plus time-reversed versions of the same.
    - *English stimuli:* List #13 of the "Harvard Sentences." (IEEE, 1969)
    - *Mandarin stimuli:* List #1 of the "Mandarin Speech Perception Test." (Fu et al., 2011)
- 10 female native speakers of American English recorded the English sentences.
- 10 female native speakers of standard Mandarin recorded the Mandarin sentences.
- Talkers in each language had regionally homogeneous accents.

English Sentences
Type out three lists of orders.
The harder he tried the less he got done.
The cup cracked and spilled its contents.

Mandarin Sentences
今天的阳光真好
jīn tiān de yáng guāng zhēn hǎo
It's a nice sunny day.
晚上一块去跳舞
wǎn shàng yī kuài qù tiào wǔ
Let's go dancing together tonight.
对面有两所高中
duì miàn yǒu liǎng suǒ gāo zhōng
There are two high schools across the street.

Listening Task:

I. Familiarization Phase: Listeners passively attend each voice speaking. (10 trials × 5 blocks)

II. Active Practice: Listeners are quizzed on each familiarization sentence; they automatically receive corrective feedback. (10 trials × 5 blocks)

III. Test Phase: Listeners are tested on their ability to identify talkers when listening to the sentences they were trained on and to generalize talker identity to novel sentences. (50 trials)
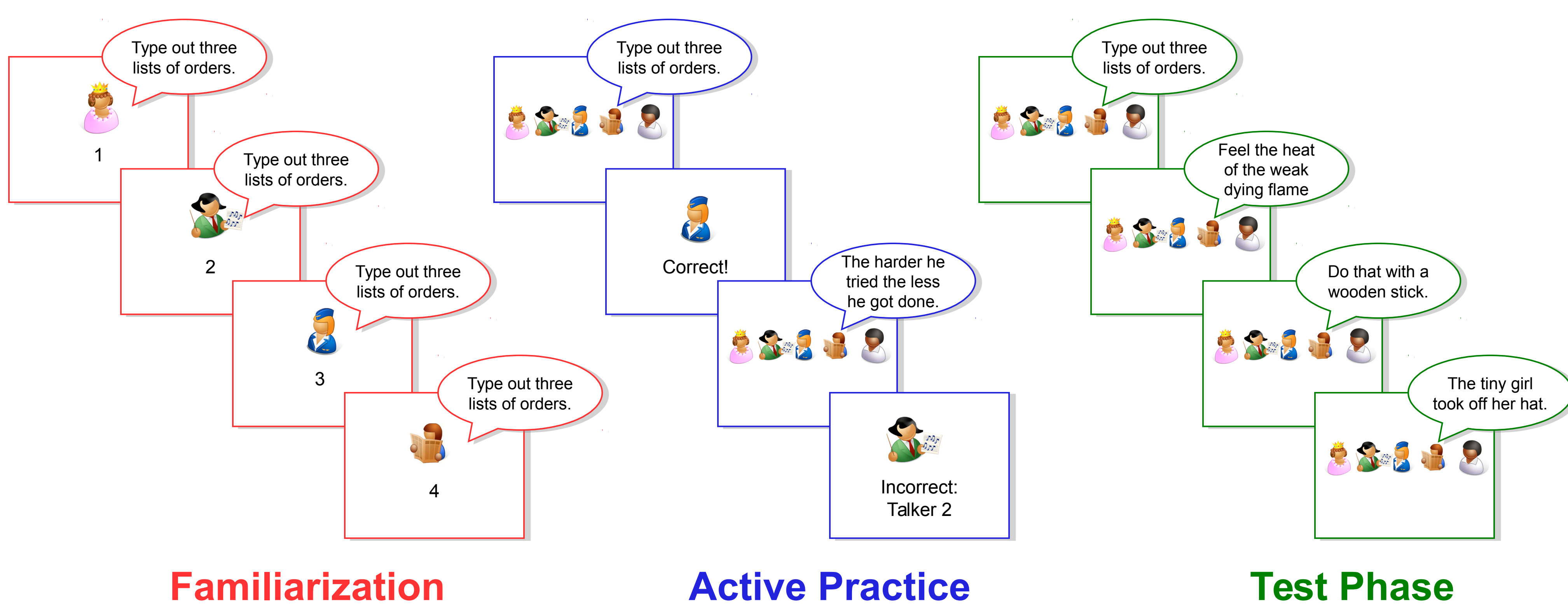
**Familiarization**   **Active Practice**   **Test Phase**

Results

English Listeners:

- More accurate talker identification for forward English-speaking voices than time-reversed ones.
(Forward: 71% > Reversed: 39%); $z = 10.06$, $p < 2\times10^{-16}$)

- Equally accurate talker identification for Mandarin-speaking voices regardless of whether they were forward or time-reversed.
(Forward: 47% = Reversed: 45%); $z = 0.88$, $p = 0.38$)

- Equally accurate talker identification for time-reversed voices regardless of language being spoken.
(English: 39% = Mandarin: 45%); $z = 1.22$, $p = 0.22$).

Mandarin Listeners:

- More accurate talker identification for forward Mandarin-speaking voices than time-reversed ones.
(Forward: 81% > Reversed: 48%); $z = 9.57$, $p < 2\times10^{-16}$)

- Equally accurate talker identification for time-reversed voices regardless of language being spoken
(Mandarin: 48% = English: 46%); $z = 0.33$, $p = 0.74$)

- Equal overall ability to identify voices from time-reversed speech as English speakers ($z = 1.10$, $p = 0.27$)

- More accurate talker identification overall than English listeners for forward speech ($z = 2.06$, $p < 0.04$) (cf. Xie & Myers, 2015)
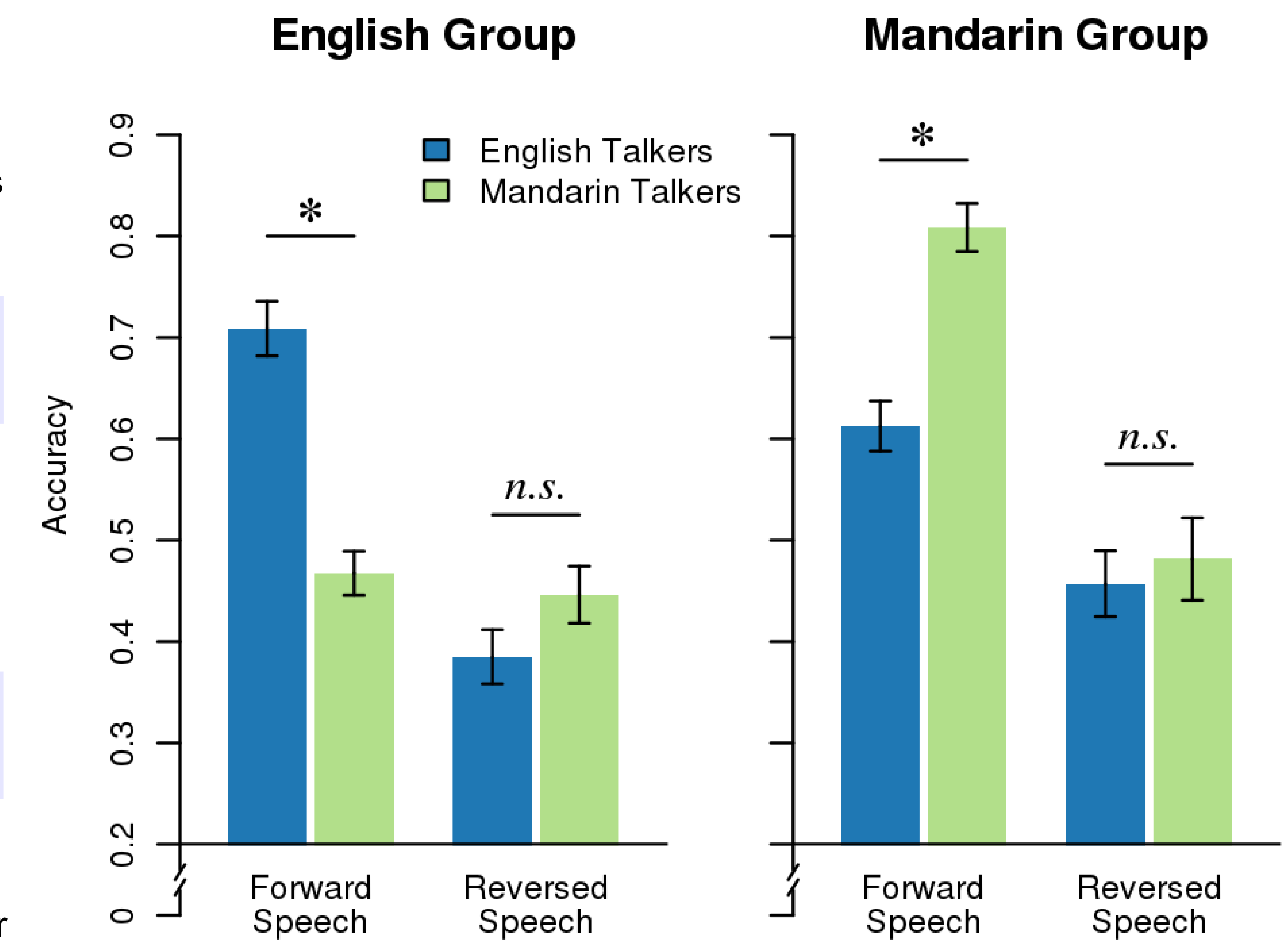
There is no language-familiarity effect for talker identification from time-reversed (incomprehensible) speech.

Conclusions

- Our results **support** the hypothesis that the **linguistic processes** involved in speech perception and language comprehension are responsible for the LFE (Perrachione *et al.* 2011; Creel and Bregman, 2014).
    - Both English and Mandarin listeners exhibited the LFE (better native language talker identification) for unaltered, comprehensible, forward speech

- Our findings **do not support** the predictions of a recent experiment alleging that the LFE is independent of speech comprehension, based on subjective judgments of voice dissimilarity (Fleming *et al.,* 2014)
    - Neither the English listeners nor the Mandarin listeners exhibited the LFE when identifying talkers from time-reversed speech

- Listeners' subjective judgments about voice attributes (*e.g.*, similarity) may not correspond to their actual abilities to perceptually identify those same voices

- Experiments testing mechanisms or processes underlying perception and memory for human voices should utilize designs that reflect ecological voice recognition / talker identification behaviors

References

Barr *et al.* (2013). *J. Mem. Lang.* 68, 255-278.
Creel & Bregman (2014). *Cognition*. 130, 85.
Fleming *et al.* (2014). *P. Natl. Acad. Sci.*111, 13795-13798.
Goggin *et al.* (1991). *Mem. Cogn.* 19, 448.
Fu *et al.* (2011). *J. Acoust. Soc. Am.* 129, EL267-EL273.
IEEE. (1969). *Audio Electroacoust.* 17, 225-246.
Perrachione & Wong (2007). *Neuropsychologia*. 45, 1899+.
Perrachione *et al.* (2011). *Science*. 333, 595.
Thompson (1987). *Appl. Cognitive Psych.* 1, 121-131.
Xie & Myers (2015). *J. Acoust. Soc. Am.* 137, 419-432.

Acknowledgments

BU Communication Neuroscience Research Laboratory

http://sites.bu.edu/cnrlab/
cnrlab@bu.edu