Orthogonal interference of indexical information occurs even when phonetic contrasts are unambiguous across talkers

Ja Young Choi, Elly R. Hu, & Tyler K. Perrachione

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

BOSTON UNIVERSITY

Summary

- Talker normalization facilitates speech processing by reducing the degrees of freedom for the many-to-many mapping between encountered speech and phonemic representations.
- Listeners rapidly identified pairs of words with varying acoustic-phonetic similarity, spoken by a single talker or mixed talkers.
- Listeners were always slower at identifying words in the mixed-talker condition than in the single-talker condition.
- Orthogonal interference of indexical variability was greatest in the hard phonetic contrast condition.
- The effect of indexical variability was significant even in the easy phonetic contrast condition.
- Acoustic distinctiveness of inividual tokens did not have a significant effect on rapid auditory categorization.
- Talker normalization appears to be an obligatory component of speech processing.
- The facilitatory effect of talker normalization varies as a function of potential ambiguity of acoustic-to-phonetic mapping of the given speech sounds.

Methods

Participants

- Native English-speaking adults (N=24) with no speech, language or hearing disorder
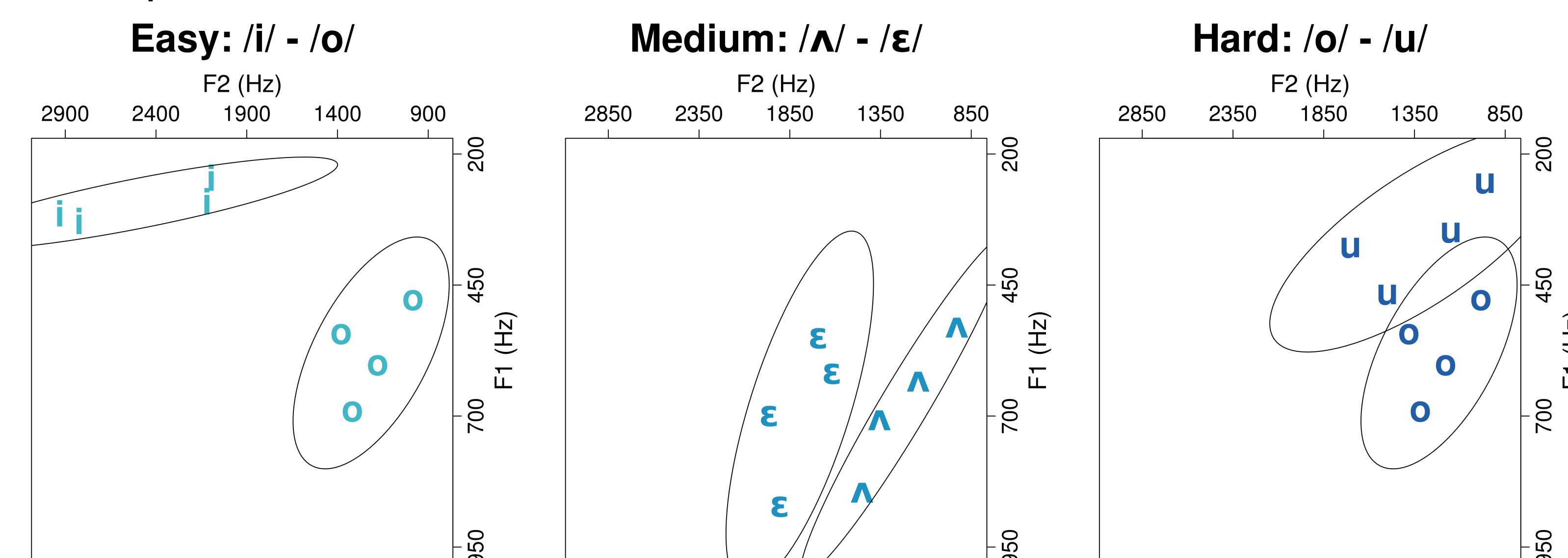
Design & Procedure

- We used a speeded classification task, parametrically manipulating **indexical variability** and difficulty of **phonetic contrast**:
  - Indexical variability: **single talker** vs. **mixed talkers**
  - Phonetic contrast: **easy** vs. **medium** vs. **hard**
- Participants identified spoken words as quickly and accurately as possible.
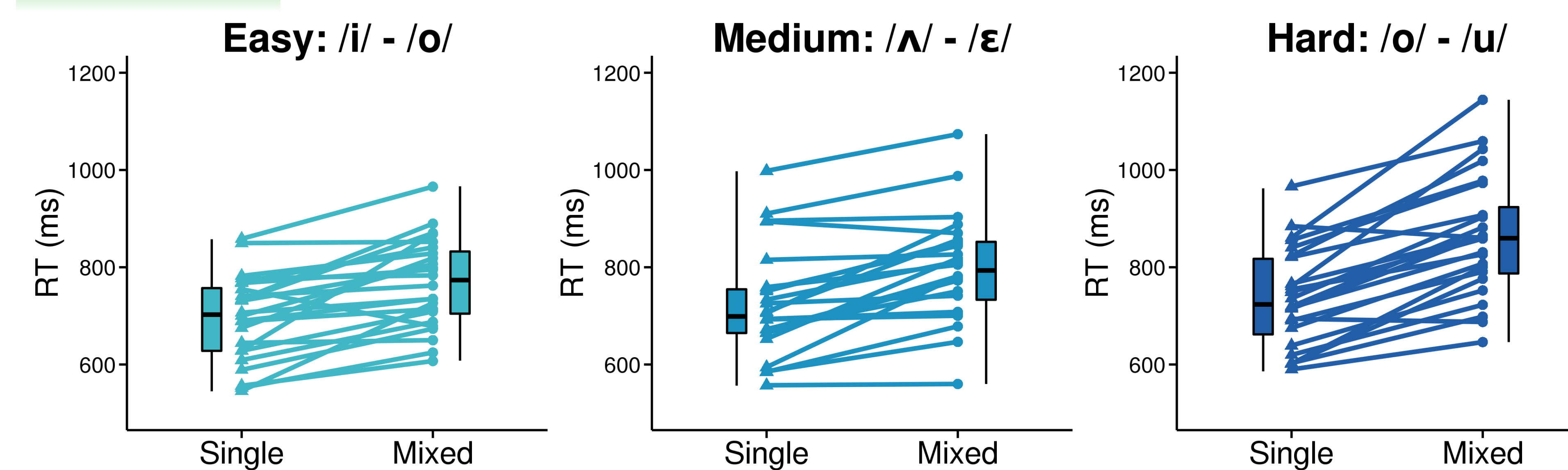- 48 trials in each condition, with 2000ms interval between each trial

A. Single talker   **B. Mixed talkers**

Vowels

Stimuli

Five naturally spoken English words with the same onset (/b/) and coda (/t/) consonants but different vowel nuclei (/i/, /ɛ/, /ʌ/, /o/, /u/):
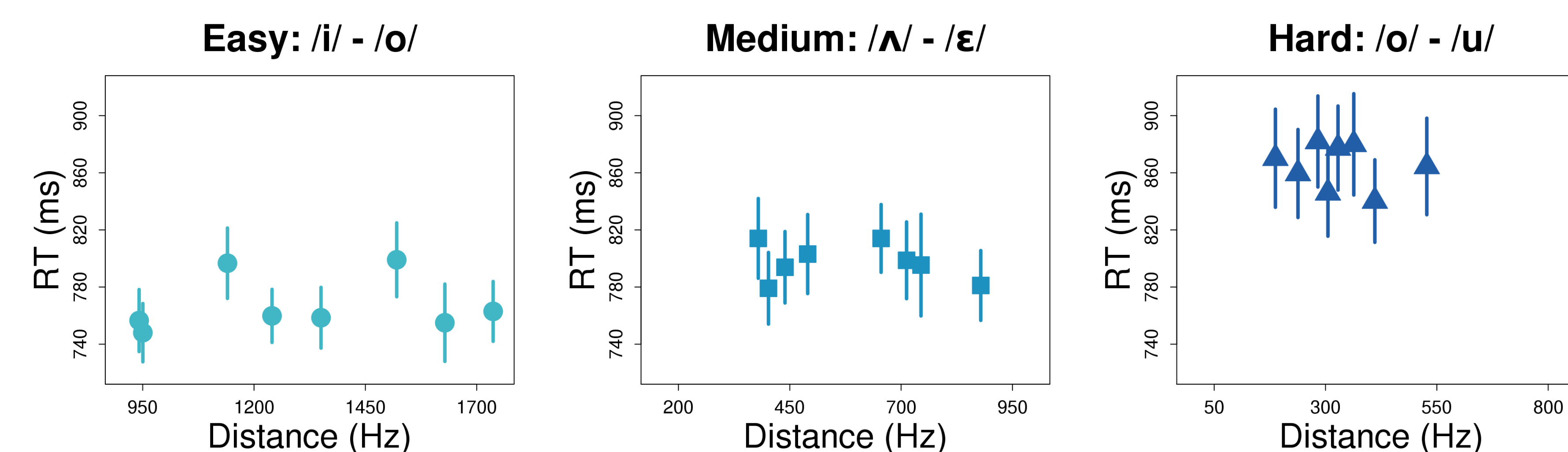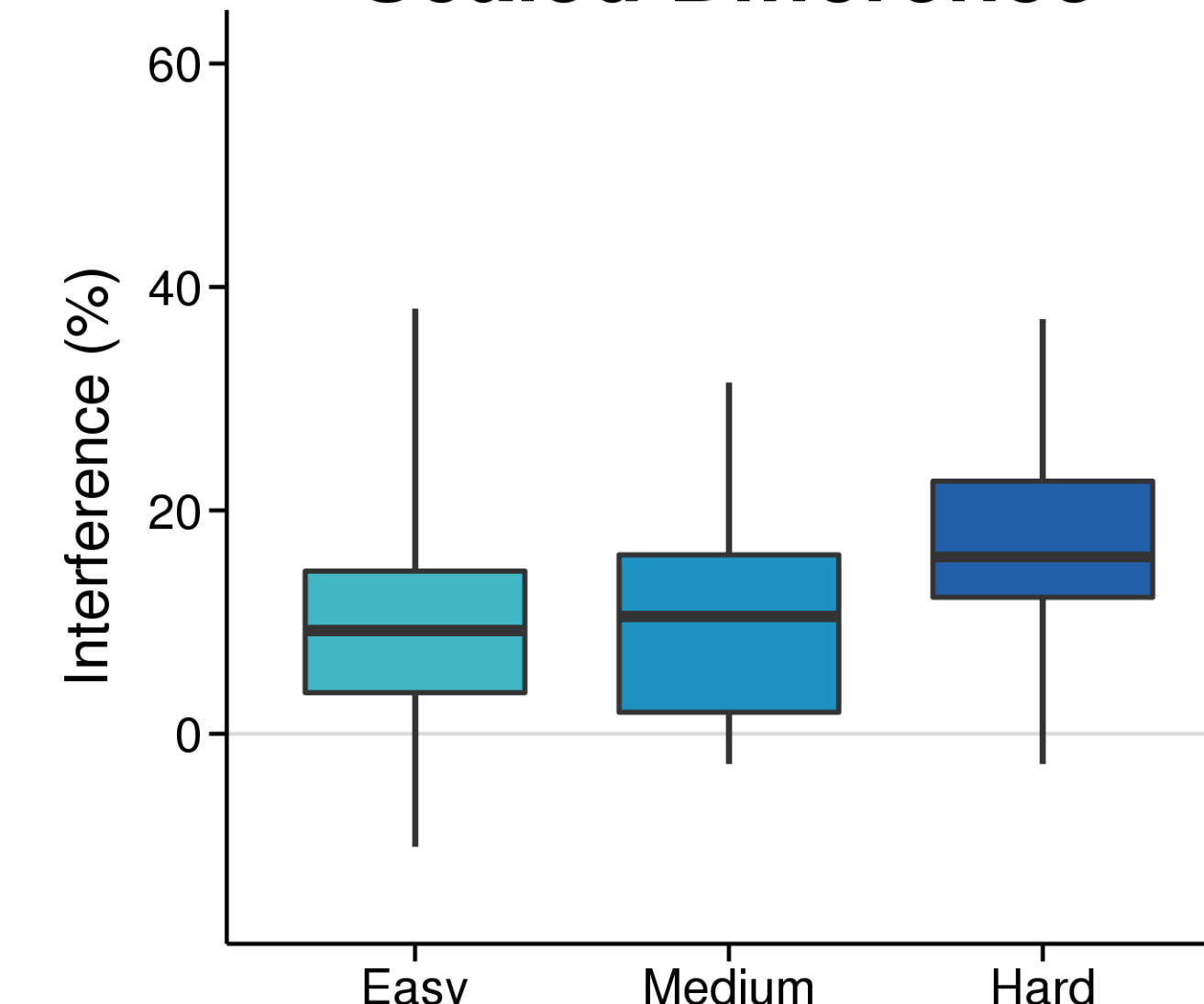**"beet", "bet", "but", "boat", "boot"**
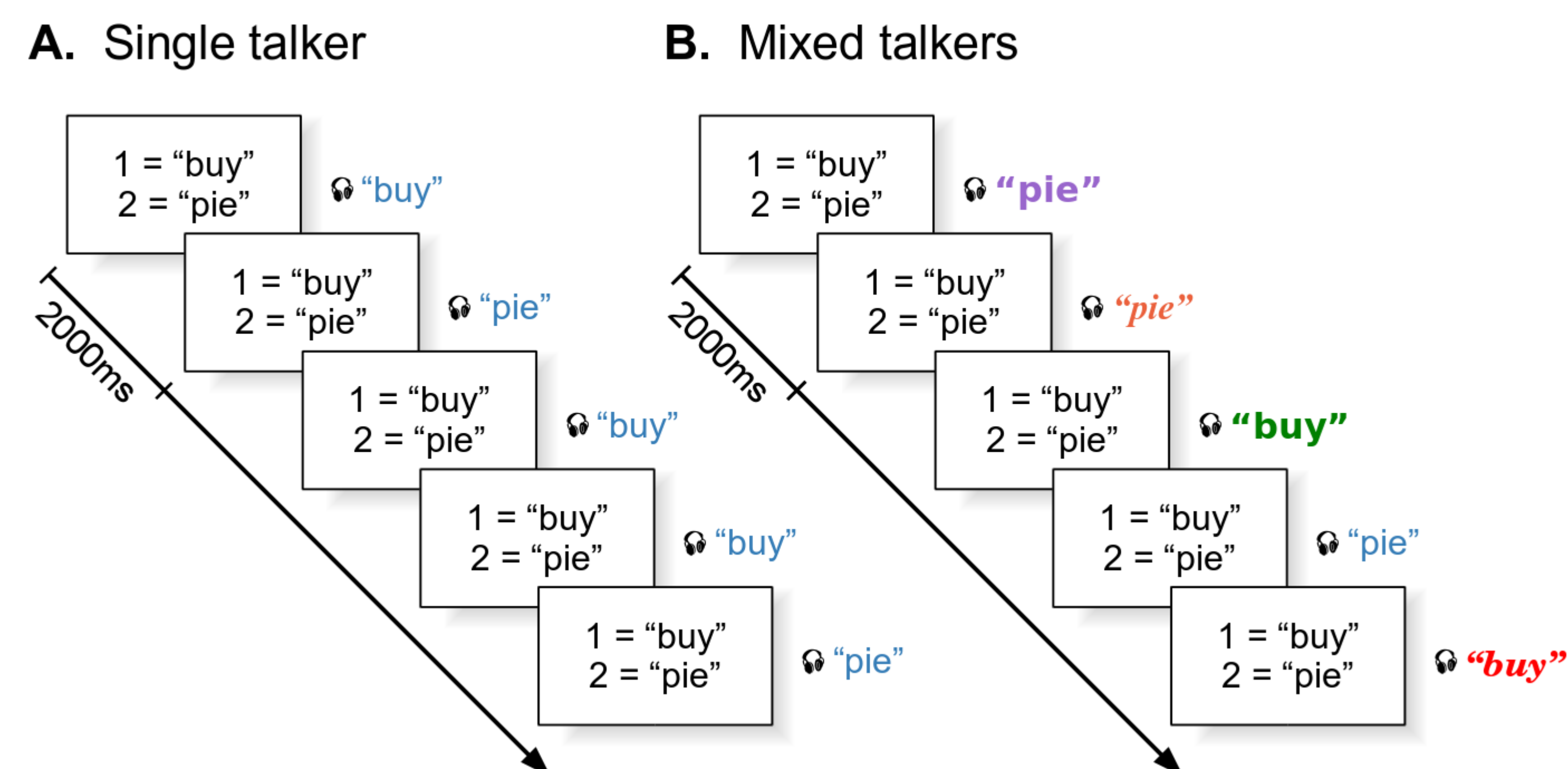
- All stimuli were recorded by four native speakers of American English (2 male, 2 female)
- Phonetic similarity was manipulated based on the distance between the two vowels in F1×F2 vowel space.

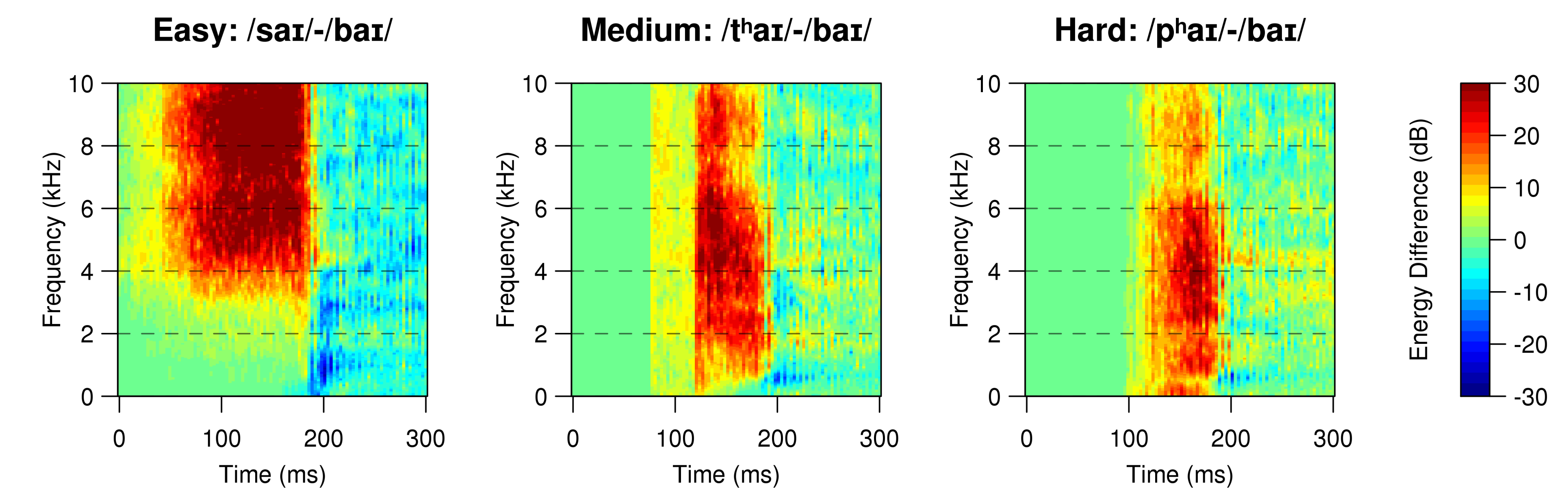### Results

- Word identification accuracy was at ceiling (99% ± 2%).

Talker normalization facilitates speech processing even when the phonetic contrast is acoustically unambiguous.

- Indexical variability introduced a significant delay in word identification even in the easy phonetic contrast condition ($p < 1.6 \times 10^{-5}$).

The effect of talker normalization varies as a function of phonetic dissimilarity of target speech sounds.

- Orthogonal interference of indexical variability in hard condition was significantly larger than easy ($p < 4.4 \times 10^{-5}$) and medium ($p < 6.5 \times 10^{-5}$) conditions.


- Acoustic distinctiveness was defined as euclidean distance between the each vowel token and the centroid of the other vowel in F1×F2 space.
- Acoustic distinctiveness of each token did not have a significant effect on response times.

Consonants

Stimuli

Four naturally spoken English words, sharing the same vowel nucleus (/aɪ/) but starting with different onset consonants: **"buy", "sigh", "tie", "pie"**
- All stimuli were recorded by the same four speakers from Experiment 1.

- We manipulated phonetic dissimilarity between the two target consonants based on the number of shared phonetic features (voicing, manner, place of articulation).

Results
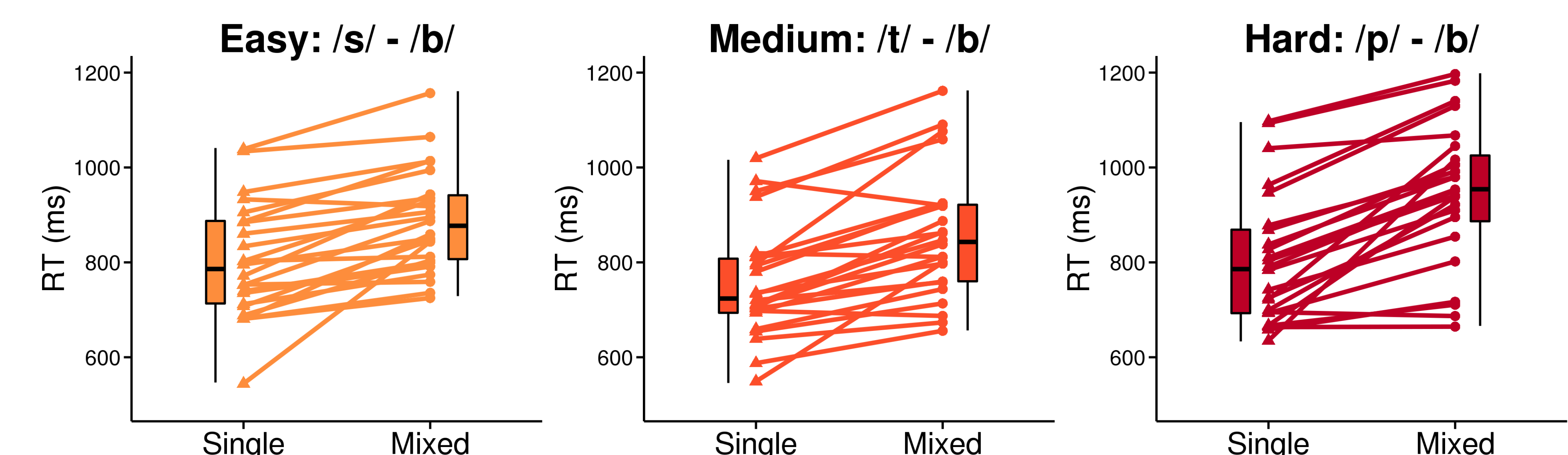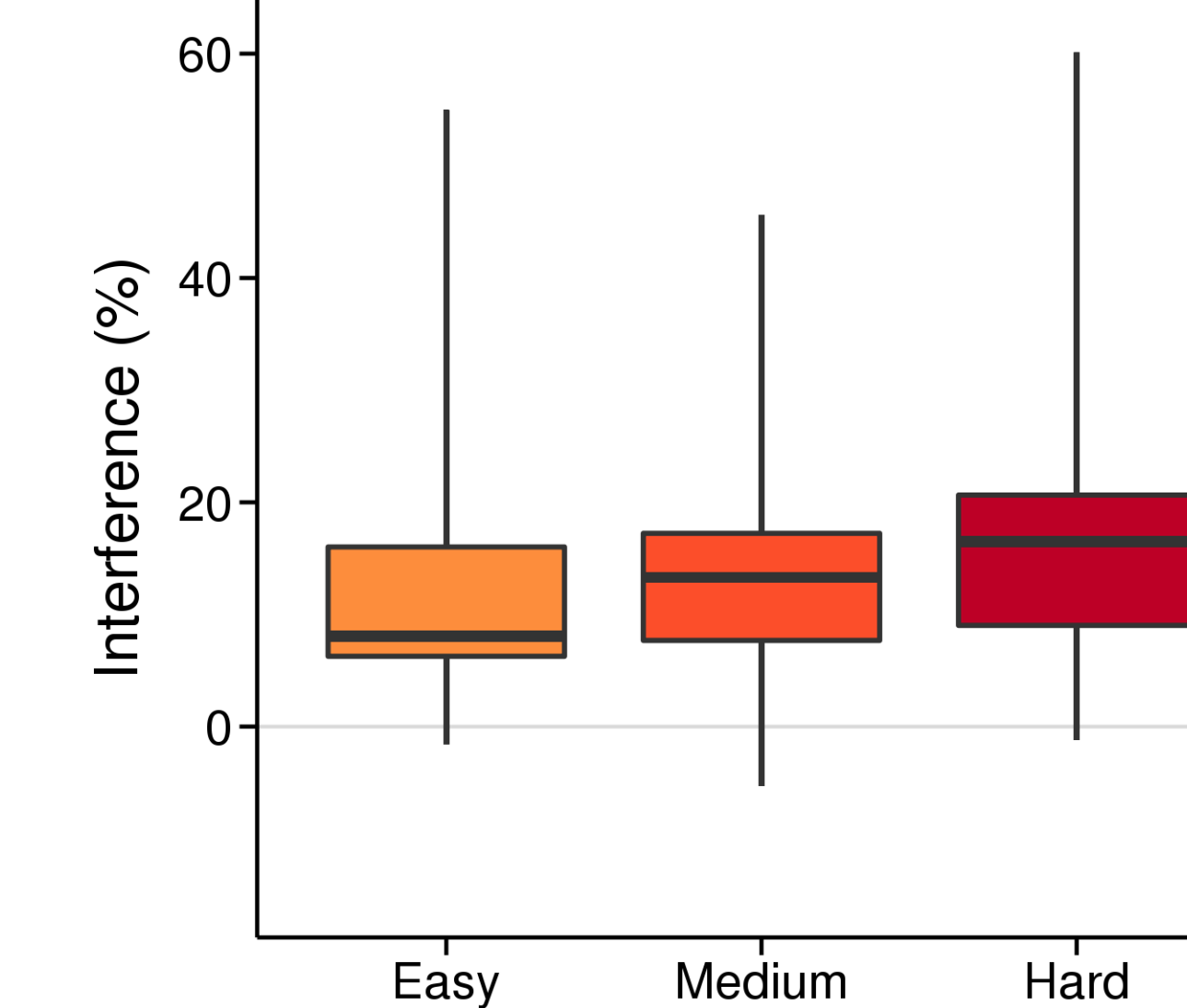
- Word identification accuracy was at ceiling (99% ± 1%).

- Indexical variability introduced a significant delay in response times in all three conditions, even in the easy phonetic contrast condition where the acoustic-phonetic mapping was unambiguous ($p < 1.3 \times 10^{-5}$).

- Interference of indexical variability was significantly greater in hard condition than in easy ($p < 0.0002$) and medium ($p < 0.004$) conditions.

The results from the experiment with vowels were replicated, showing significant facilitatory effect of talker normalization even when the phonetic contrast was acoustically unambiguous and the effect being greater in the hard contrast condition than in the easy or medium ones.

References

Barr et al. (2013) J. Mem. Lang. 68, 255-278.
Garner (1974). The processing of information and astructure. Potomac, MD: Erlbaum.
Green et al. (1997). Perception & Psychophysics, 59, 675-692.
Hillenbrand et al. (1995). J. Acoust. Soc. Am. 97, 3099-3111.
Mullennix & Pisoni (1990). Perception & Psychophysics, 47, 379-390.
Tomiak et al. (1991). J. Acoust. Soc. Am. 90, 2363.
Volaitis & Miller(1992). J. Acoust. Soc. Am., 92, 723-735.

Acknowledgments

BU Communication Neuroscience Research Laboratory

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