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

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 individual tokens did not have a significant effect on response times.
- Phonetic similarity was manipulated based on the distance between the two vowels in F1×F2 vowel space.
- 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

- Easy: /l - /o/
- Medium: /a - /e/
- Hard: /o - /i/

Stimuli

- Five naturally spoken English words with the same onset (/b/) and coda (/t/) consonants but different vowel nuclei (/i/, /ɪ/, /ʌ/, /a/, /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 space.

Results

- Acoustic distinctiveness of individual tokens did not have a significant effect on response times.
- Acoustic distinctiveness was defined as euclidean distance between 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% ± 2%).
- Indexical variability introduced a significant delay in word identification even in the easy phonetic contrast condition (p < 1.6 × 10^-14).
- 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 greater than in easy (p < 6.5 × 10^-2) and medium (p < 8.5 × 10^-3) conditions.

References

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