Perceptual learning in speech reflects rapid adjustment of global mappings from acoustic cues to phonological categories

Samantha Chiu[1], Cheyenne Munson Toscano[2], Joe Toscano[2], Bob McMurray[1]

[1]Department of Psychological and Brain Sciences, University of Iowa and [2]Department of Psychological and Brain Sciences, Villanova University

Introduction

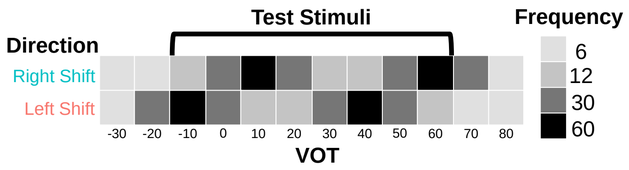
- Listeners need to account for variability due to coarticulation, speaking rate[1], gender[2], accent[3], and individual talker[4,5] to categorize speech sounds.
- Because talker differences are unique to each talker, talker differences are consistently novel and must be learned.
- What do listeners do when encountering multiple novel talkers?
  - ***Generalize*** categories learned from prior talkers[4].
  - Learn new ***talker-specific*** categories for each talker[5].

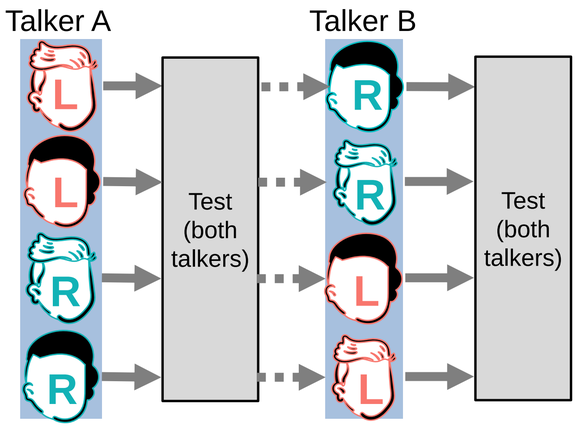
General Methods

- All participants were tested online and recruited through Prolific
- Participants must pass a headphone screener to ensure reasonable audio quality.

Experiment 1: Distributional Learning of Multiple Talkers (n=160)

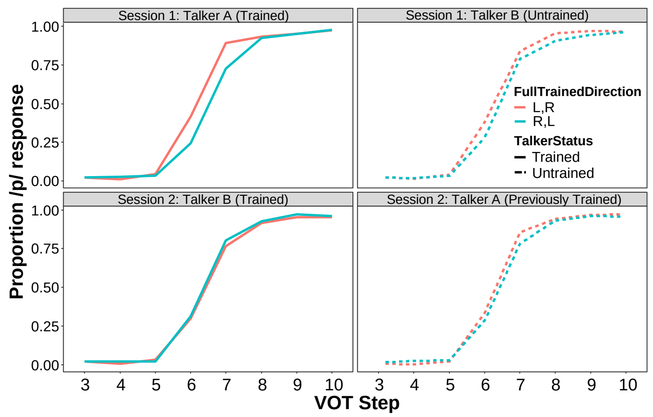
- Session 1: Learn novel talker (A) with unique VOT categories.
- Session 2: Learn novel talker (B) with categories shifted in the other direction.
- Training block (30 minutes) followed by test block on both Talker A & B.





Experiment 1 Research Questions:

- Session 1: Do listeners generalize categories from a learned talker to a novel talker?
- Session 2: Do listeners learn talker-specific categories between two learned talkers?

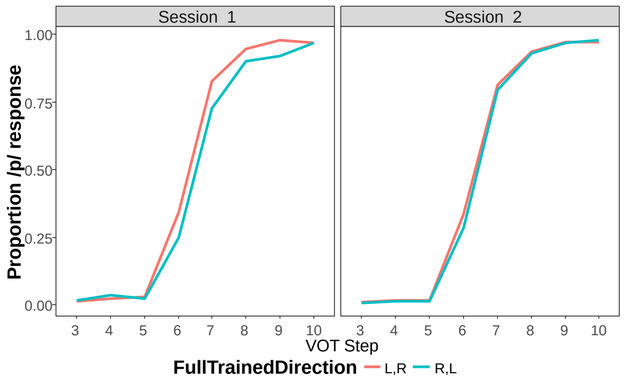


Experiment 1 Conclusions:

- Listeners can learn talker-specific VOTs from distributional learning (Session 1: Talker A)
- Listeners generalize VOT boundaries to a novel talker (Session 1: Talker B)
- Listeners do not retain VOT boundaries of multiple talkers (Session 2)
  - Talker A: Interference from Talker B or decay of learning over time?
  - Talker B: Need stronger learning cue?

Experiment 2: Interference or Decay? (n=160)

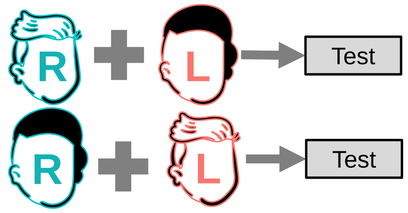
Without interference of learning Talker B, will listeners retain Talker A?



Experiment 2 Conclusions:

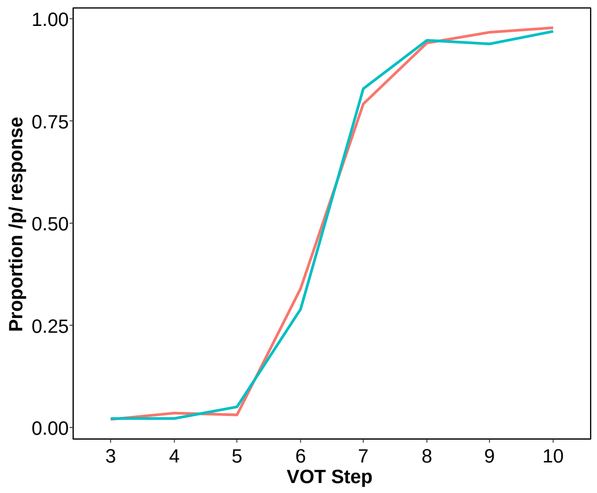
- Listeners can learn talker-specific categories (Session 1).
  - Replication of Experiment 1, Session 1.
- Listeners do not retain learning; learning decays over time.

Experiments 3 and 4: Learning Multiple Talkers Simultaneously



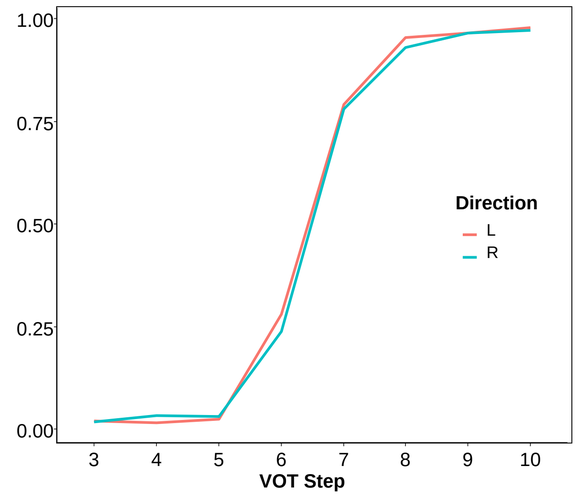
Experiment 3: Interleaved (n = 64)

Will talker-specific learning occur when both talkers are interleaved?



Experiment 4: Feedback (n = 64)

Will talker-specific learning occur with both talkers interleaved and with feedback?



Listeners do not appear to learn two distinct talkers at all.

Conclusions & Future Directions

- During rapid adaptation, listeners continually adjust a single VOT boundary to match the current talker. Though imperfect, "good-enough adaptation" may be useful.
- Distributional learning can support learning native phoneme categories[6] and adapting to talker-specific categories; these mechanisms may be one and the same.

References

1. Miller, J.L., Volaitis, L.E. Effect of speaking rate on the perceptual structure of a phonetic category. Perception & Psychophysics 46, 505–512 (1989). https://doi.org/10.3758/BF03208147
2. Strand, E. A., & Johnson, K. (1996). Gradient and Visual Speaker Normalization in the Perception of Fricatives. Natural Language Processing and Speech Technology, 14–26. https://doi.org/10.1515/9783110821895-003
3. Bradlow, A. R., & Bent, T. (2008). Perceptual adaptation to non-native speech. *Cognition*, *106*(2), 707–729. https://doi.org/10.1016/j.cognition.2007.04.005
4. Kraljic, T., & Samuel, A. G. (2006). Generalization in perceptual learning for speech. *Psychonomic Bulletin & Review*, *13*(2), 262–268. https://doi.org/10.3758/BF03193841
5. Eisner, F., & McQueen, J. M. (2005). The specificity of perceptual learning in speech processing. *Perception and Psychophysics*, *67*(2), 224–238. https://doi.org/10.3758/BF03206487
6. Maye, J., Werker, J. F., & Gerken, L. (2002). Infant sensitivity to distributional information can affect phonetic discrimination. *Cognition*, *82*(3), B101-B111. https://doi.org/10.1016/S0010-0277(01)00157-3

Acknowledgements

This research was supported by NSF 1945069 and NSF 1945994.