Auditory, Visual, and Speech Category Learning in the Same Individuals

Casey L. Roark and Bharath Chandrasekaran

University of Pittsburgh Department of Communication Science & Disorders
and the Center for the Neural Basis of Cognition

Introduction

Perceptual category learning

- Category learning spans the senses yet has typically been studied separately in vision and audition.
- Recent theories have expanded theories of visual category learning to audition (Francis & Nusbaum, 2002; Goudbeek, et al., 2009; Maddox, et al., 2013; Yi et al., 2014).
- It is not yet known whether auditory and visual category learning rely on domain-general or modality-specific mechanisms. We leverage a within-subjects approach to examine learning of auditory and visual categories in the same individuals.

Artificial and natural category learning

- Learning of artificial perceptual categories is typical in laboratory studies with implications often applied to learning of natural categories, such as speech or object categories. Direct comparisons are rare.
- We compare how the same individuals learn artificial (auditory and visual) categories and natural non-native speech categories.

Methods

- 30 Pittsburgh community members ages 18-32 learned both nonspeech auditory and visual categories (Figure 1A; order counterbalanced across participants) in the same session.
- A subset of participants ($N = 22$) also completed a natural speech categorization task (Figure 1B). Stimuli were natural speech stimuli produced by native Mandarin speakers.
- Auditory and visual tasks: 300 trials across 6 blocks with feedback + 58-trial generalization test (novel exemplars with no feedback).
- Mandarin speech task: 208 trials across 4 blocks + 40 trial generalization test.

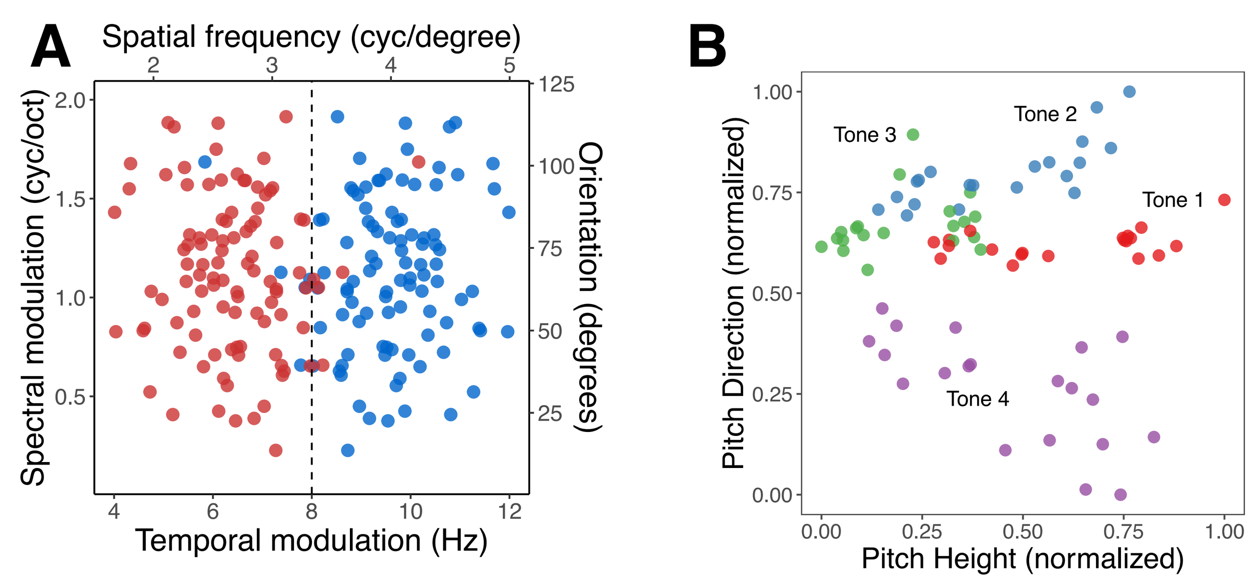
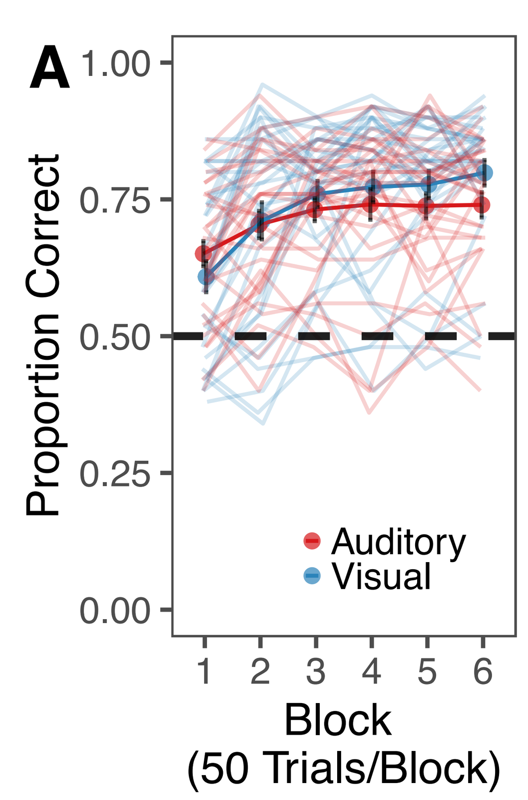


Figure 1. A. Stimulus distributions for the nonspeech and visual tasks. Auditory stimuli were defined by temporal modulation and spectral modulation dimensions. Visual stimuli were defined by spatial frequency and orientation dimensions. Categories were defined based on one dimension. B. Mandarin tone categories defined by pitch height and pitch direction dimensions.

Auditory and Visual Category Learning

- Participants learned the nonspeech auditory and visual categories well with substantial individual variability among participants (Figure 2A).
- Participants learned the nonspeech and visual categories equally well ($F(1,29) = 1.33$, $p = 0.26$, $\eta_p^2 = .04$), with visual categories having a slightly steeper slope across blocks ($F(5,145) = 2.83$, $p = 0.018$, $\eta_p^2 = .09$).
- Accuracies in the final blocks of the two tasks were significantly positively correlated ($r(28) = 0.61$, $p = 0.00034$)



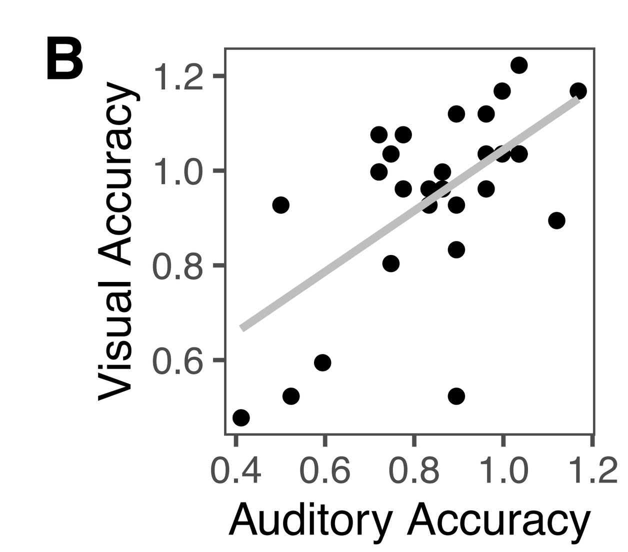


Figure 2. A. Average accuracy across blocks in the auditory and visual tasks. B. Correlation between final block accuracy for the nonspeech auditory and visual tasks (logit transform).

Auditory and Visual Category Generalization

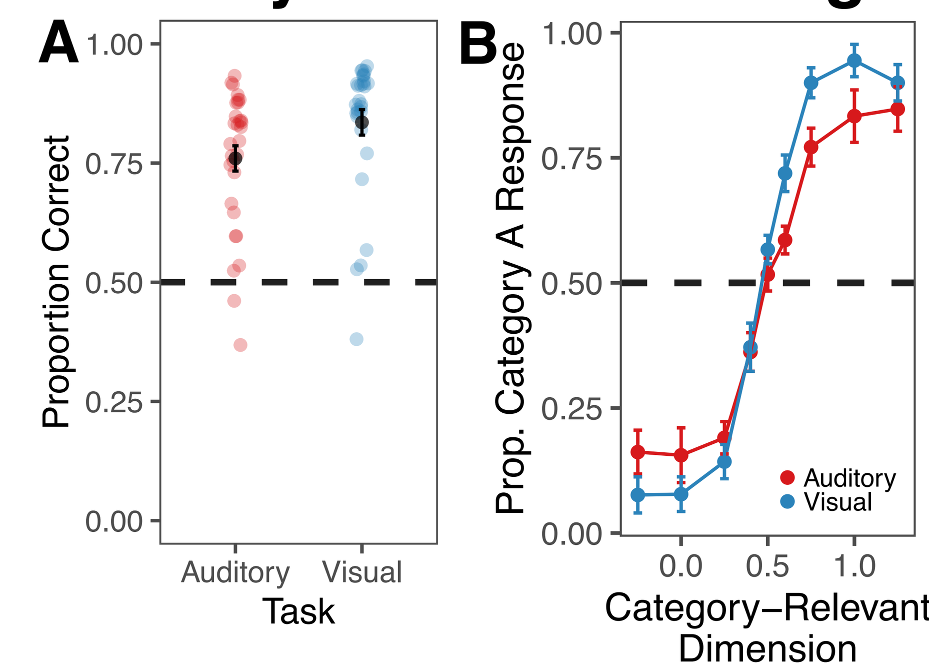


Figure 3. A. Average accuracy in generalization test for the two tasks. B. Average accuracy in generalization test based on the category-relevant dimension (on x-axis, 0.5 reflects the category boundary).

- Participants successfully generalized to novel exemplars and were more accurate for visual categories than auditory (Figure 3A, $M_{diff} = 7.57\%$, $t(29) = 2.86$, $p = 0.0078$, $d = 0.52$).
- Participants had more 'categorical' representations for visual categories relative to auditory categories, indicated by a steeper categorization curve (Figure 3B, $F(8,232) = 3.51$, $p = 0.001$ $\eta_p^2 = .11$)
- Participants had difficultly selectively attending to the category-relevant dimension in the auditory task relative to visual task, evidenced by differences in the decision boundaries participants used to separate the categories (Figure 4).

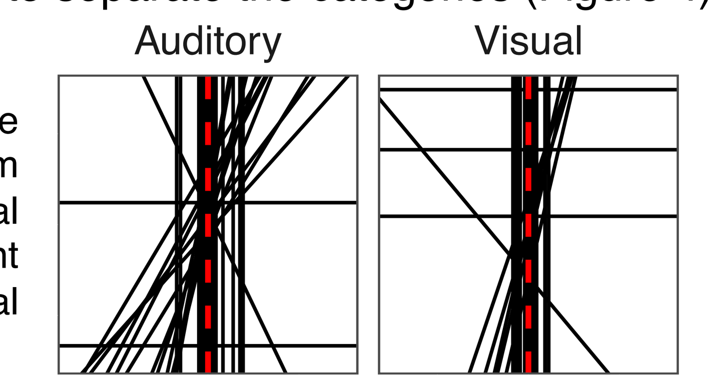
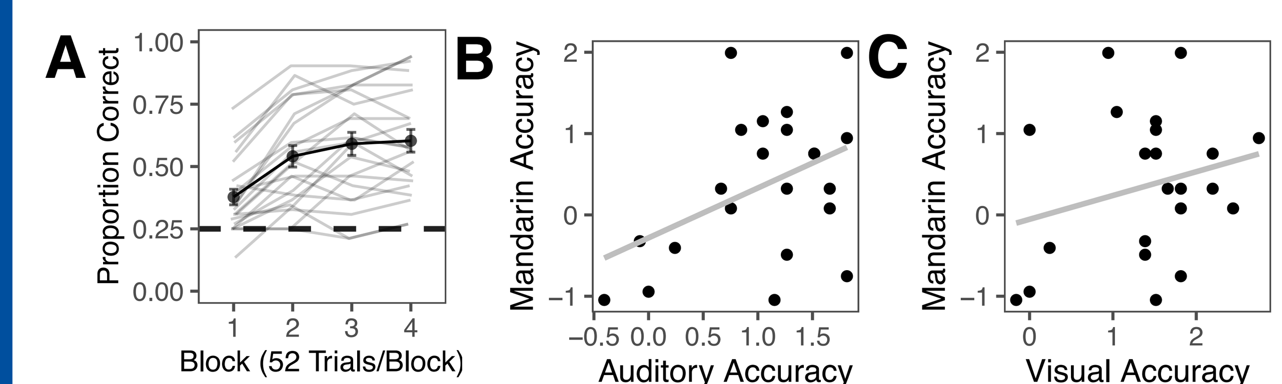


Figure 4. Decision bounds in the generalization test derived from decision bound computational models fit for each participant (black lines) relative to optimal boundary (red dashed line).

Nonspeech/Visual and Natural Speech Category Learning



- Substantial individual variability in ability to learn Mandarin speech categories (Figure 5A).
- Significant positive correlation between auditory and Mandarin accuracies ($r(20) = 0.43$, $p = 0.044$), but not between visual and Mandarin accuracy ($r(20) = 0.25$, $p = 0.26$). Need more power to know whether there are differences.

Figure 5. A. Average accuracy across blocks in the Mandarin speech task. B. Correlation between final block accuracies across the nonspeech auditory task, visual task, and Mandarin task (with logit transform).

Conclusions

- Similarities within individuals in learning for auditory and visual categories suggests that there are some domain-general components supporting learning.
- However, some important differences remain hinting at additional modality-specific processes.
- Further research is needed to understand the role of modality in perceptual category learning.

References and Acknowledgements

References

Francis, A. L., & Nusbaum, H. C. (2002). Selective attention and the acquisition of new phonetic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(2), 349–366. doi: 10.1037//0096-1523.28.2.349

Goudbeek, M., Swingley, D., & Smits, R. (2009). Supervised and unsupervised learning of multidimensional acoustic categories. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(6), 1913–1933. doi: 10.1037/a0015781

Maddox, W. T., Chandrasekaran, B., Smayda, K., & Yi, H.-G. (2013). Dual systems of speech category learning across the lifespan. *Psychology and Aging*, *28*(4), 1042–1056. doi: 10.1037/a0034969

Yi, H.-G., Maddox, W. T., Mumford, J. A., & Chandrasekaran, B. (2014). The Role of Corticostriatal Systems in Speech Category Learning. *Cerebral Cortex*, 1–12. doi: 10.1093/cercor/bhu236

Acknowledgements

This work was supported by a grant from the National Institute for Deafness and Communication Disorders to Bharath Chandrasekaran (R01DC013315A1).