Multiple sources of information contribute to novel category formation

Emily Cibelli

Department of Linguistics, University of California, Berkeley

Introduction

The acquisition challenge

Adult learners of a second language often struggle with the acquisition of phonemes in the target language, especially when they overlap with or crosscut phonemic categories in the native language [1, 2, 3].

Native attractors

One way to frame this difficulty is in terms of attractors: native categories act as stable, well-trained attractors that pull incoming acoustic information towards these familiar representations [2, 4]. Without intervention, non-native tokens may reinforce and strengthen these native categories [5], rather than be recognized as separate categories.

Pathways to learning

To break away from native representations, a number of interventions have been tested in laboratory training paradigms, including performance feedback [5, 6], the manipulation of attention [7, 8], cue enhancement [5, 9, 10] and high stimulus variability [8, 11, 12].

The current study

This study consists of a multi-day training paradigm targeted at teaching new learners several non-native contrasts in Hindi. It introduces multiple sources of cues to help learners form new categories, and tests the integration of perceptual and articulatory information in the learning process.

Study questions

1. Can native phoneme category biases be overcome with a combination of perceptual interventions in a short-term training paradigm?
   - Feedback: explicit information about performance
   - Adaptive fading: Transition from clear to less-distinct tokens
   - Repeated exposure
2. Do different non-native contrasts differ in learnability?
3. What does articulatory training [13] contribute to perceptual category formation in a short-term, integrated training design?

References

[1] Best, McRoberts, & Goodell. (2001). *JASA*, 109(2), 775 – 794.

[2] Kuhl, Conboy, & Coffey-Corina. (2008). *Philosophical Transactions of the Royal Society B*, 363, 979 – 1000.

[3] Flege. (1995). Speech perception and linguistic experience: Issues in cross-language research, 233-277.

[4] Köver & Bao. (2010). *PLoS ONE,* 5(5), 1 – 7.

[5] McCandliss, Fiez, Protopapas, Conway, & McClelland. (2002). *Cog., Affective and Behavioral Neurosci.*, 2(2), 89 – 108.

[6] Goudbeek, Cutler, & Smits. (2008). *Speech Communication*, 50(2), 109-125.

[7] Pederson & Guion-Anderson. (2010). *Journal of the Acoustical Society of America*, 127(2), EL54 – EL59.

[8] Kondaurova & Francis. (2010). *Journal of Phonetics*, 38(4), 569 – 587.

[9] Terrace. (1963). *Journal of the Experimental Analysis of Behavior*, 6(I), 1 – 27.

[10] Pruitt. (1995). Doctoral dissertation, University of South Florida.

[11] Sadakata & McQueen. (2011). *Interspeech 2011*, 873-876.

[12] Lim & Holt. (2011). *Cognitive Science*, 35(7), 1390 - 1405.

[13] Catford & Pisoni (1970). *The Modern Language Journal,* 54(7), 477-481.

Methods

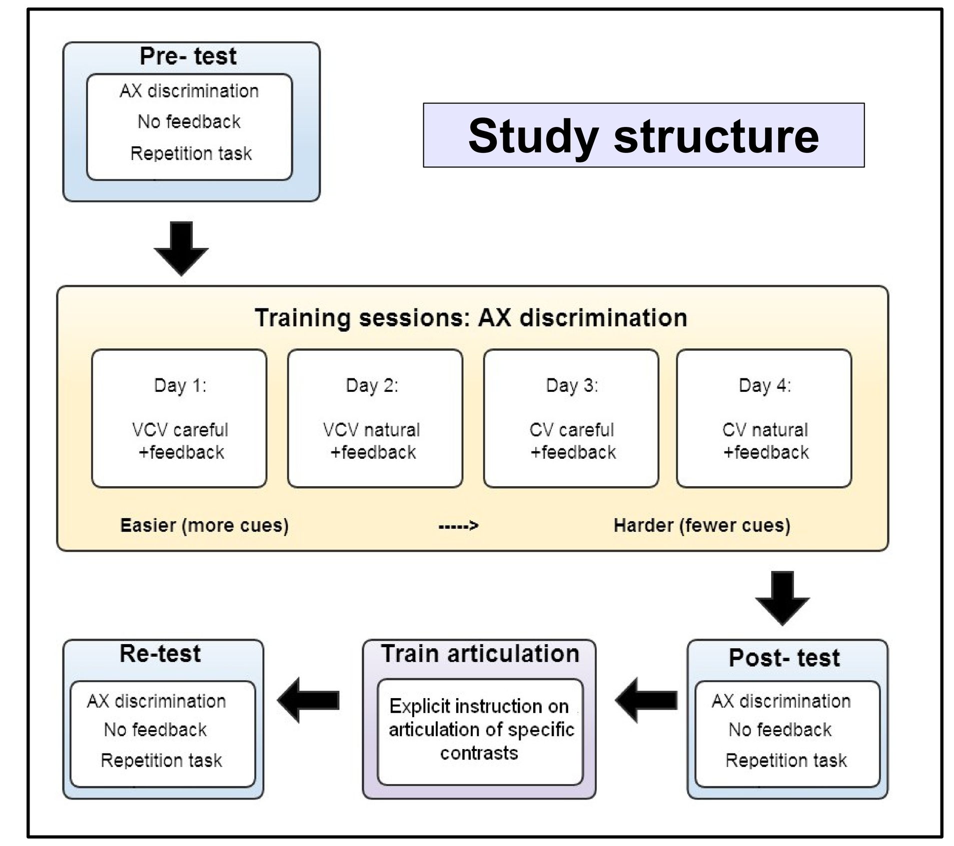
Experiment design

Tasks: AX discrimination, repetition

Perception training: adaptive fading, feedback on performance, repeated exposure to targets

Production training: explicit instruction about tongue/larynx control

Average days to complete: 12.7 (*s.d.* 9.4)



Subjects

11 native English speakers

No prior experience with Hindi or languages with phonemic contrasts similar to the target contrasts

Stimuli

VCV and CV syllables recorded by a native speaker of Hindi

Contrast types: voicing (e.g. /t̪/ - /d̪ʱ/), place (/t̪/ - /ʈ/), place-voicing (/t̪/ - /ɖʱ/), same (/t̪/ - /t̪/)

	Voiceless unaspirated	Voiceless aspirated	Voiced	Breathy	Vowels
dental	t̪	t̪ʰ	d̪	d̪ʱ	a i u
retroflex	ʈ	ʈʰ	ɖ	ɖʱ	

Results: Accuracy

Question: Does accuracy improve for different contrast types as a function of training?

Analysis: Mixed-effects model of d-prime values from "different" trials in test sessions (pre-test, post-test, re-test)

Model terms: contrast type, session, contrast type * session, random intercept for subject

Findings:
- Accuracy on voicing and placeVoicing trials higher than place trials (both *adj. p* < 0.001)
- Accuracy higher from pre-test to post-test (*adj. p* = 0.002) and pre-test to re-test (*adj. p* < 0.001), but not post-test to re-test
- Session differences driven by place trials (pre-post: *adj. p* = 0.017, pre-re: *adj. p* < 0.001)

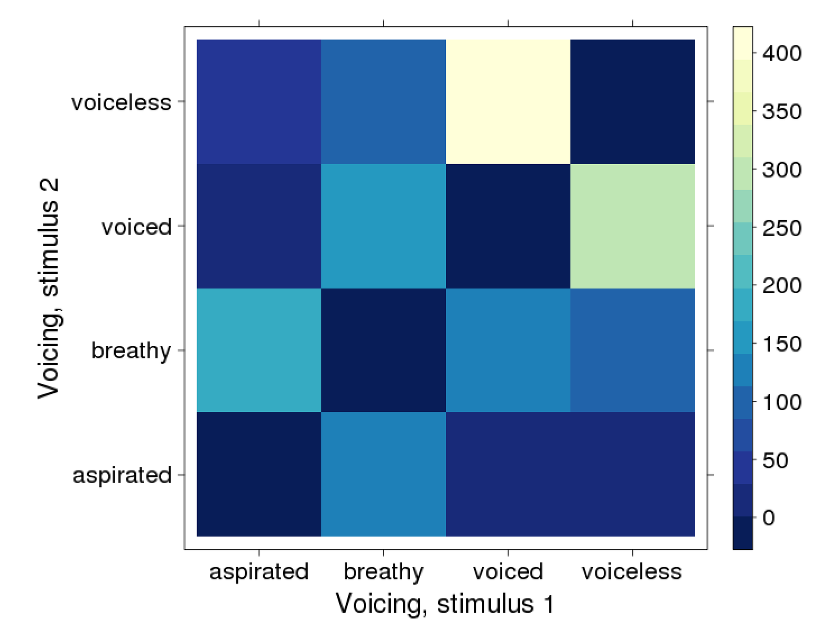


Figure 1: Confusion matrix of errors on voicing trials ("incorrect" responses only). Common mistakes reflect English phonology (e.g. conflating voiced and voiceless unaspirated tokens).

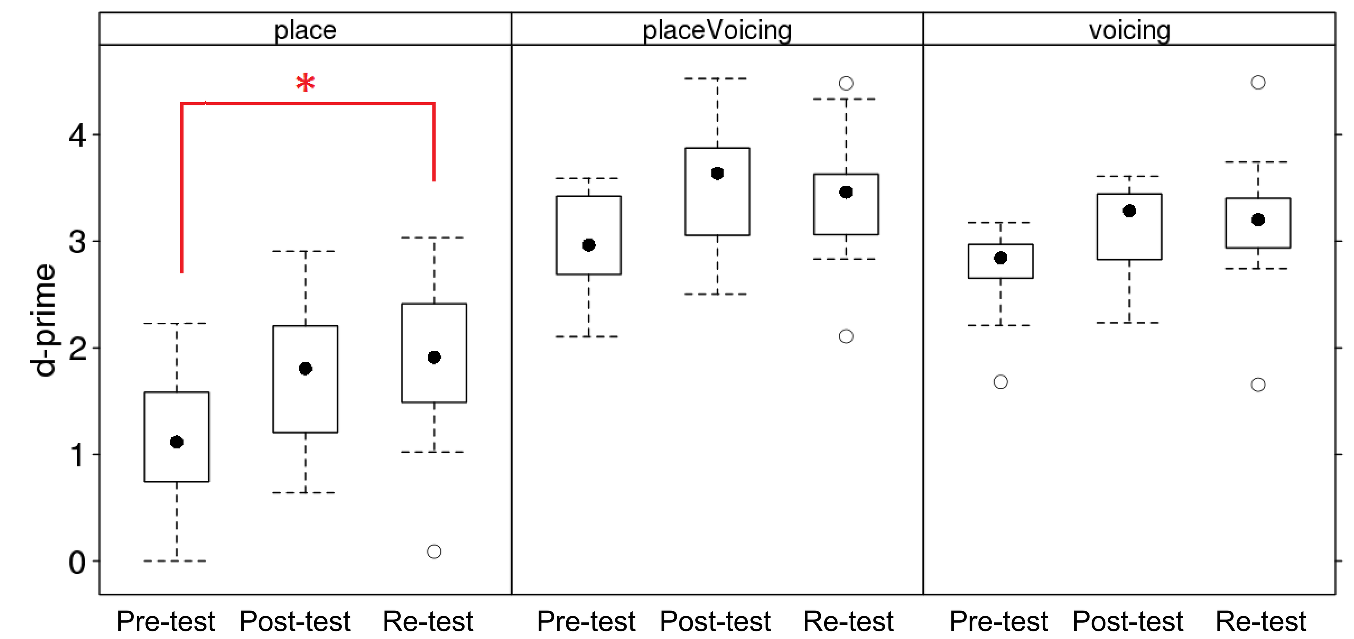


Figure 2: D-prime values by contrast type and test session.

Results: Individual differences

Performance accuracy varied across subjects, and initial performance did not necessarily predict end-of-study performance.

Range of performance on "different trials":
- pre-test: 42.8% - 79.8%
- post-test: 73.3% - 92.8%
- re-test: 66.2% - 95.1%

Correlation of performance across sessions:

pre-test ↔ post-test: $r = 0.544$, $p = 0.08$
post-test ↔ re-test: $r = 0.751$, $p = 0.008$
pre-test ↔ re-test: p = n.s.

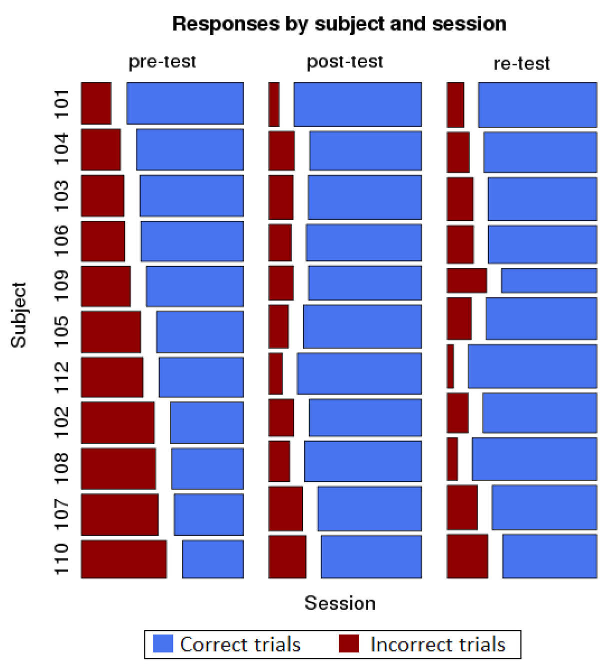


Figure 3: Accuracy by subject and session

Results: Reaction Time

Question: Can learning be indexed by speed of response?

Analysis: Mixed-effects model of reaction time in all "correct" trials

Model terms: contrast type, session, session * contrast type, vowel, trial count, # days to complete session, age, gender, other languages spoken, session-by-subject random slopes

Findings: Speed increases:
- over the course of a session
- in /a/ trials more than other vowels
- after training for all trial types (post-hoc test, Tukey correction)

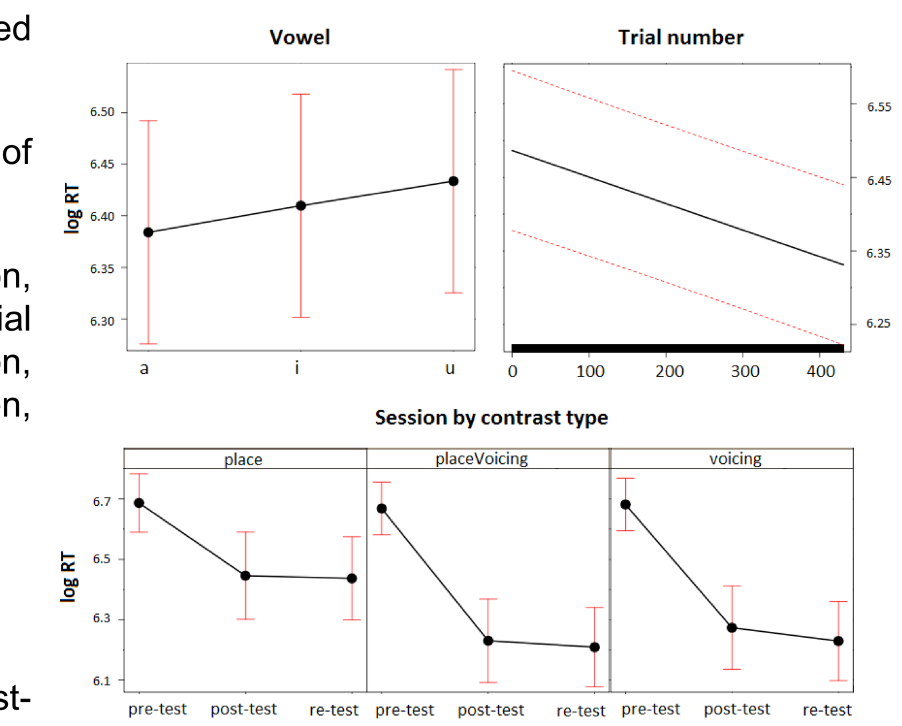


Figure 4: Selected partial effects of RT model

Discussion

- Improvement is possible in short-term training.
- Initial performance relates to, but does not determine, final outcomes.
- Speed may signal improvement once accuracy hits a near-ceiling.
- Production training does not add benefit at the end of perceptual training.

Future directions

- Production training: beneficial as stand-alone training?
- Benefit of perceptual training on quality of production data?
- EEG MMN study: is improvement detectable at a neural level before behavioral evidence is clear?

Acknowledgments

I am grateful to members of the UC Berkeley Phonology Lab and the UCSF Speech Neurophysiology Lab for suggestions on how to develop this work. Special thanks to Ritika Pandita and Usha Jain for assistance with stimulus development, and Dorothy Dao, Aaliyah Ichino, and Jocelyn Takahashi for data collection and analysis input. This project is supported by an NSF Graduate Research Fellowship.

ecibelli@berkeley.edu