Phonetic instruction as a tool to overcome L1 biases in L2 articulation

Emily Cibelli
Northwestern University, University of California, Berkeley

Introduction

How do adult learners acquire production targets in a new second language?

Acquisition of L2 production targets varies based on many factors, including L1, training type, and individual variation (e.g. Bradlow et al. 1997, Baese-Berk 2010)

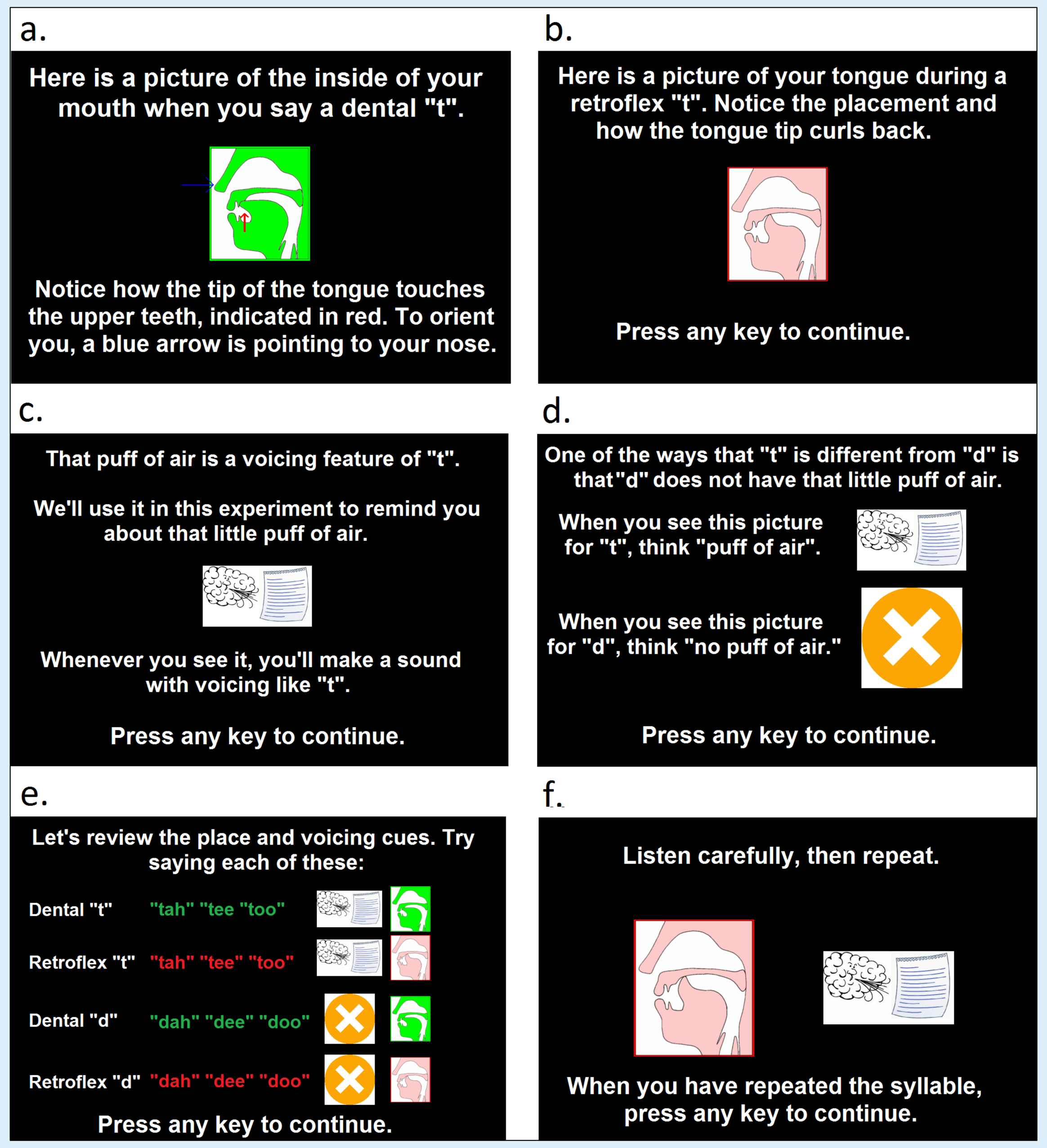
Classroom L2 learning rarely includes instruction on articulatory targets (Lord 2005), and it has not been investigated widely in the lab, although it may be beneficial for perceptual learning (Catford and Pisoni 1970)

Major study questions

1. Does production training focused on articulatory features aid in non-native phoneme production?
2. Can that learning be maintained when visual cues are no longer present?
3. Do phonetic features improve consistently across L2 targets that share them?

Production training

- Explicit training in target articulatory postures
- Repetition task with visual cues for reinforcement



Examples from production training. (a-b) Place of articulation training. (c-d) Voicing training. (e) Review of visual cues for two voicing categories. (f) Repetition trial: participants heard native speaker while seeing visual cues learned in training.

Study design and analysis

Study design: 8 sessions of training and testing with both perception training (AX discrimination + feedback, 4 sessions of adaptive fading, e.g. Terrace, 1963) and production training (1 session)

PRE-TEST ➡ PERCEPTION TRAINING ➡ POST-TEST ➡ PRODUCTION TRAINING ➡ RE-TEST

Test task: Repetition of CV syllables produced by native Hindi speaker during tests and production training (visual cues present in training only)

Stimuli: VCV and CV Hindi syllables with one of 8 coronal stops

	breathy	voiced	aspirated	voiceless
dental	d̪ʱ	d̪	t̪ʰ	t̪
retroflex	ɖʱ	ɖ	ʈʰ	ʈ

Participants: 21 native English speakers, no prior exposure to Hindi

Analysis: Manual segmentation of VOT, formant extraction from vowel, and spectral analysis of stop burst

Patterns of interest after training:
- Does positive VOT lengthen for breathy stops compared to other segments?
- Does negative VOT lengthen for breathy and voiced stops?
- Do dental and retroflex tokens dissociate across sessions?

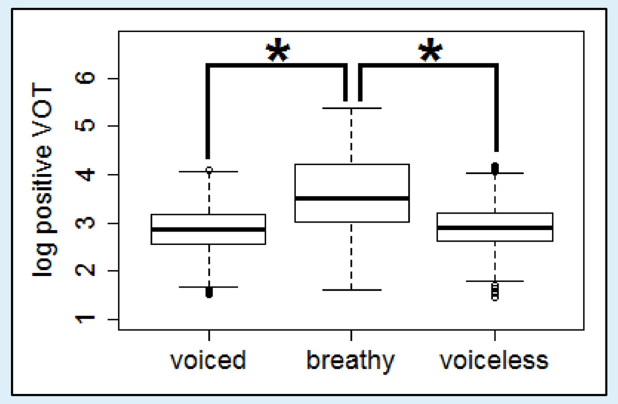
Results: Voicing contrast

Analysis of positive VOT (aspiration)

```
logPosVOT ~ session*voicing + (1 + session||subject) + (1|token)
```

Model selection procedure: maximal random effects structure containing components with non-zero variance (Bates et al. 2015)

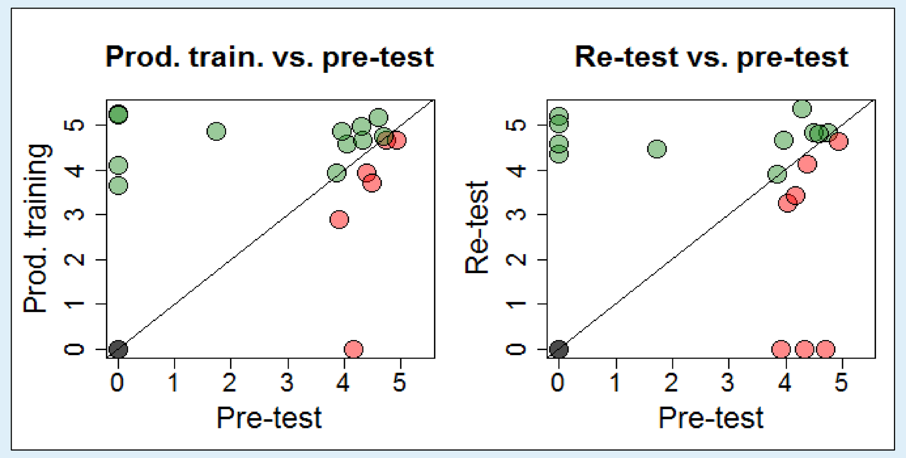
- **[A]** Breathy: more aspiration than voiced and voiceless across all sessions (p adj. < 0.0001, post-hoc tests with Tukey adjustment)
  - Breathy distinct from English /d/
- **[B]** Both voiced and voiceless marginally or significantly shorter (all p adj. < 0.10) after each session compared to baseline/pre-test
  - Voiced and voiceless (allophones of English /d/) dissociate from aspirated and breathy segments



[A]: Positive VOT durations in breathy tokens dissociated from both voiced and voiceless tokens across the experiment.

Analysis of negative VOT (prevoicing)

```
logNegVOT ~ session*voicing + (1|subject) + (1|token)
```



*[C]: Subjects showing **longer**, **equivalent**, or **shorter** negative VOT on breathy tokens in production training (**12**, **2**, **6**) and re-test (**11**, **2**, **7**) sessions, compared to pre-test. Each point represents one subject.*

[C] Breathy: marginally longer during production training than pre-test (p adj. = 0.069); pre-test difference maintained at re-test (p adj. = 0.020)

[D] Voiced: no change across any sessions (p adj. > 0.10); overall greater than breathy (p adj. = 0.005)

Results: Place of articulation contrast

- Production of place contrast tested with 2 classification analyses
  - Acoustic predictors (selected from preliminary GLM) entered into linear discriminant analysis (classify: dental/retroflex)
  - Classification above chance assessed with permutation tests

[E] Analysis 1: stop burst spectra

Predictors: burst centroid, burst standard deviation, burst kurtosis (following Forrest et al. 1988)
(Overall accuracy for stimuli with this analysis: 75.1%)

	Pre-test	Post-test	Production training	Re-test
Overall accuracy	**54.6% ***	**54.6% ***	**65.7% ***	**54.4% ***
Dental accuracy	61.0%	58.6%	69.7%	65.7%
Retroflex accuracy	48.2%	50.8%	61.6%	43.1%

[F] Analysis 2: vowel formants

Predictors: F2 and F3 at vowel onset (voiced/voiceless tokens only, following Sussman et al. 1993)
(Overall accuracy for stimuli with this analysis: 83.3%)

	Pre-test	Post-test	Production training	Re-test
Overall accuracy	52.3%	50.2%	**59.4% †**	49.8%
Dental accuracy	63.2%	55.0%	58.0%	54.1%
Retroflex accuracy	41.1%	45.5%	60.9%	49.8%

Overall accuracy significantly (, p < 0.05) or marginally (†, p < 0.10) above chance (50%), as assessed by permutation tests.*

Discussion

1. Does articulatory training aid non-native phoneme production?
- In some cases [C, E, F] improvement was detected across sessions
- In others, targets showed trait across all sessions [A, D]
- Training may also emphasize contrast with native categories [B]

2. Is learning maintained after visual cues are no longer present?
- In some VOT results [C] improvement maintained after training
- But for place of articulation [E, F], improvement not maintained to re-test
  - Past work (Cibelli 2015) shows perceptual learning maintained through re-test, so cannot be attributable to weak perceptual skills

3. Are phonetic features consistently shared across L2 targets?
- Not always [C vs. D]
- Different starting proficiencies may affect how much improvement is detectable during testing

References

Bates, D., Kliegl, R., Vasishth, S., & Baayen, H. (2015). Parsimonious mixed models. *arXiv* preprint arXiv:1506.04967.

Bradlow, A. R., Pisoni, D. B., Akahane-Yamada, R., & Tohkura, Y. I. (1997). Training Japanese listeners to identify English /r /and /l/: IV. Some effects of perceptual learning on speech production. *The Journal of the Acoustical Society of America*, 101(4), 2299-2310.

Baese-Berk, M. M. (2010). An examination of the relationship between speech perception and production (Doctoral dissertation, Northwestern University).

Catford, J. C., & Pisoni, D. B. (1970). Auditory vs. Articulatory Training in Exotic Sounds. *The Modern Language Journal*, 54(7), 477-481.

Cibelli, E. S. (2015). Aspects of articulatory and perceptual learning in novel phoneme acquisition (Doctoral dissertation, University of California Berkeley).

Forrest, K., Weismer, G., Milenkovic, P., & Dougall, R. N. (1988). Statistical analysis of word-initial voiceless obstruents: preliminary data. *The Journal of the Acoustical Society of America*, 84(1), 115-123.

Lord, G. (2005). (How) can we teach foreign language pronunciation? On the effects of a Spanish phonetics course. *Hispania*, 557-567.

Sussman, H. M., Hoemeke, K. A., & Ahmed, F. S. (1993). A cross-linguistic investigation of locus equations as a phonetic descriptor for place of articulation. *The Journal of the Acoustical Society of America*, 94(3), 1256-1268.

Terrace, H. S. (1963). Discrimination learning with and without "errors". *Journal of the Experimental Analysis of Behavior*, 6(1), 1-27.

Acknowledgments

This project was aided by helpful comments from the Berkeley Phonology Lab and the Northwestern Sound Lab. Aaliyah Ichino, Dorothy Dao, Amanda Geib, Charlotte Hoerber, and Jocelyn Takahashi assisted with data collection. This research was supported by an NSF Graduate Research Fellowship (DGE-1106400), and the Phi Beta Kappa Northern California Association Fellowship.

emily.cibelli@northwestern.edu