Do changes in shadowing predict those in subsequent productions?

Harim Kwon | **hkwon20@gmu.edu**
Dept. of English, Linguistics Program, George Mason University

Background

Spontaneous speech imitation

- Speakers **shift their production in the direction of what they just heard** without being told to "imitate" (e.g., Goldinger 1998).
- Imitative changes are **mediated by language-specific cue primacy**:
  Seoul Korean speakers imitate an exaggerated non-primary cue (**stop VOT**) by enhancing the primary cue (**post-stop *f0***) for the relevant phonological contrast (**aspirated stops**).

English (e.g., Shockley et al. 2004, Nielsen 2011)

heard:		*produced (shadowing/post-exposure):*
voiceless stops with **longer VOT**	→	**voiceless stops** with **longer VOT**

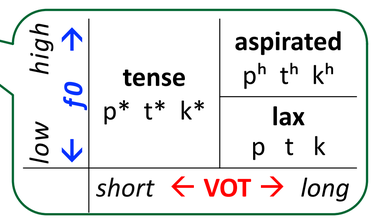
Seoul Korean (Kwon 2019)

heard:		*produced (post-shadowing):*
aspirated stops with **higher *f0***	→	**aspirated stops** with **higher *f0***
aspirated stops with **longer VOT**	→	**aspirated stops** with **higher *f0*** + **longer VOT**

Stop VOT and post-stop *f0* in Seoul Korean and English

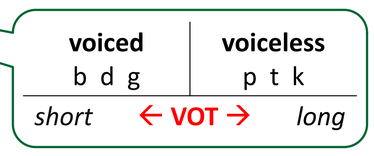
Seoul Korean: three-way laryngeal contrast

- Maintained by at least two separate acoustic cues, stop VOT and post-stop *f0*
- **Post-stop *f0*: primary cue** for aspirated stops (Kim et al. 2002, Kang & Guion 2008)



English voicing contrast

- **Stop VOT: primary cue** for English voicing contrast
- **Post-stop *f0*** plays a non-negligible role in native English listeners' perception of voicing contrast (e.g., Whalen et al. 1993).



Experiment 1: English stop VOT

Question 1.

Does a speaker who converges to, or diverges from, a model talker while shadowing retain the altered productions in her post-shadowing test productions?

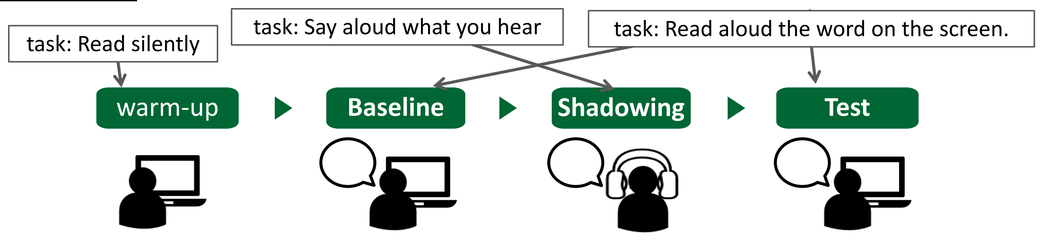
Methods

Participants: 14 American English speakers (7F/7M)

Stimuli: 25 English /t/-initial target words + filler words

- Target /t/ VOT extended (+60 ms): mean /t/ VOT after manipulation = 123.9 ms
- Disyllabic, initial stress, produced by a male American English speaker

Procedure:



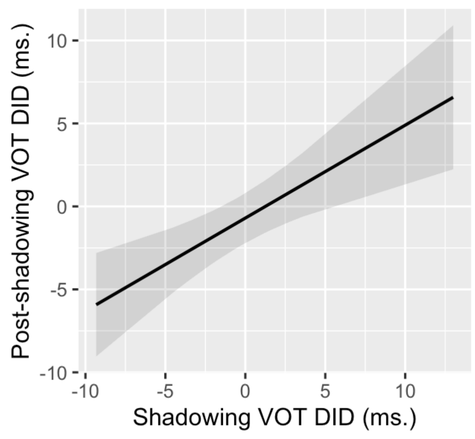
Measurements:

- /t/ VOT in targets measured and converted to difference-in-distance (DID) scores:
  DID = |**Baseline** – model| – |**Shadowing** *or* **Test** – model| (e.g., Pardo et al. 2017).

Results

- Shadowing DID predicts Test DID [β = 0.45, t = 6.58, p < .001]
- Perceptually induced changes during shadowing persisted to the subsequent productions without immediate targets to imitate.

Statistical analyses
- linear mixed-effects models
- DV: Test DID
- IV: shadowing DID (* manipulation: Exp.2)
- Random effects: (1|speaker) + (1|word)



Experiment 2: Seoul Korean stop VOT and post-stop f0

Question 2.

Does phonology mediate the relation between the convergence patterns in the shadowing productions and those in the post-shadowing productions?

Methods (same data reported in Kwon 2019)

Participants: 19 Seoul Korean speakers (12F/7M, living in MI, USA)

Stimuli: 25 Korean /tʰ/-initial target words + filler words

- Two manipulations
  - /tʰ/ VOT extended (+60 ms): mean /t/ VOT after manipulation = 119.8 ms
  - Post-/tʰ/ *f0* raised (+20%): mean post-/tʰ/ *f0* after manipulation = 176.2 Hz
- Disyllabic, produced by a male Seoul Korean speaker

Procedure:

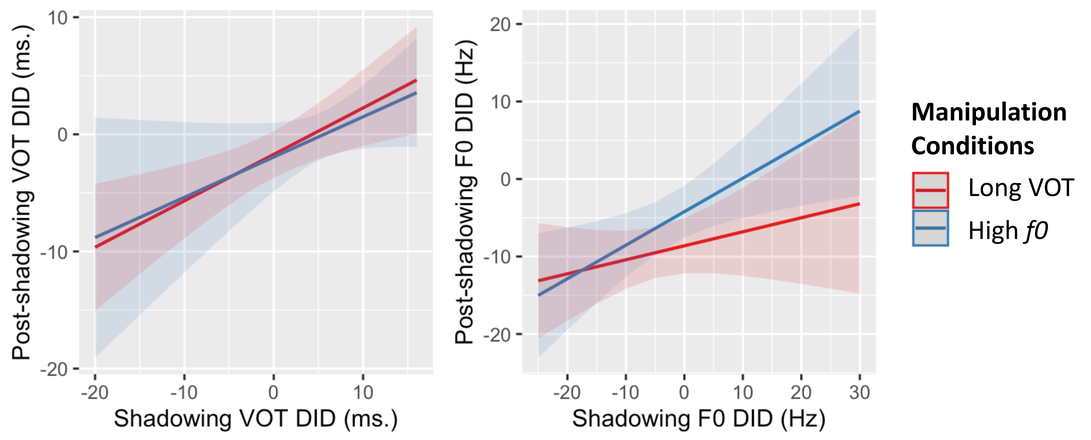
- Parallel to Experiment 1
- Each speaker tested twice for the two manipulations (at least 1 week apart, testing order counter-balanced)

Measurements:

- VOT DID and *f0* DID separately calculated
- Post-/tʰ/ *f0* measured at the temporal midpoint of the following vowel

Results

- Shadowing DID predicts Test DID **when the measured acoustic property matches the property enhanced in the stimuli**.
  - Stop VOT: Shadowing DID predicts Test DID in long VOT condition [β = 0.40, t = 3.06, p = .004], but not in high *f0* condition [β = 0.34, t = 1.70, p = .098]
  - Post-stop *f0*: Shadowing DID predicts Test DID in high *f0* condition [β = 0.43, t = 2.63, p = .013], but not in long VOT condition [β = 0.18, t = 1.09, p = .282]



Discussion

Summary of findings

Exp 1: English speakers

Heard:		*Shadowing:*		*Test:* (post-shadowing reading)
voiceless stops with **longer VOT**	→	**VOT DID**	**predicts**	**VOT DID**

Exp 2: Seoul Korean speakers

Heard:		*Shadowing:*		*Test:* (post-shadowing reading)
Aspirated stops with **higher *f0***	→	***f0* DID**	**predicts**	***f0* DID**
		VOT DID	~~predicts~~	**VOT DID**
Aspirated stops with **longer VOT**	→	***f0* DID**	~~predicts~~	***f0* DID**
		VOT DID	**predicts**	**VOT DID**

Changes during the shadowing are related to those in subsequent productions,

when the acoustic property matches the property enhanced in the stimuli.

- Imitative changes in the unmanipulated acoustic property in the post-shadowing productions, i.e., the increase in post-stop *f0* after shadowing aspirated stops with longer VOT in Seoul Korean (Kwon 2019), may not stem from the imitative changes during the shadowing.
- The changes targeting in the abstract phonological category arguably emerge in the post-shadowing productions distal from the immediate acoustic targets to imitate.

Selected References

Goldinger, S. D. 1998. Echoes of echoes? An episodic theory of lexical access. *Psychological Review* 105, 251-279.

Kang, K.-H., & Guion, S. G. 2008. Clear speech production of Korean stops: Changing phonetic targets and enhancement strategies. *Journal of the Acoustical Society of America* 124, 3909-3917.

Kim, M.-R., Beddor, P. S., & Horrocks, J. 2002. The contribution of consonantal and vocalic information to the perception of Korean initial stops. *Journal of Phonetics* 30, 77-100.

Kwon, H. 2019. The role of native phonology in spontaneous imitation: Evidence from Seoul Korean. *Laboratory Phonology* 10, 1-24.

Nielsen, K. 2011. Specificity and abstractness of VOT imitation. *Journal of Phonetics* 39, 132-142.

Pardo, J. S., Urmanche, A., Wilman, S., & Wiener, J. 2017. Phonetic convergence across multiple measures and model talkers. *Attention, Perception, & Psychophysics* 79, 637-659.

Shockley, K., Sabadini, L., & Fowler, C. A. 2004. Imitation in shadowing words. *Attention, Perception, & Psychophysics* 66, 422-429.

Whalen, D. H., Abramson, A. S., Lisker, L., & Mody, M. 1993. F0 gives voicing information even with unambiguous voice onset times. *Journal of the Acoustical Society of America* 93, 2152-2159.

Acknowledgments: I would like to thank Pam Beddor, Andries Coetzee, Kuniko Nielsen, and Yoonjeong Lee, for helpful discussions related to this work; Yuting Guo and Amy Lam for their assistance in English data collection and measurements; and the speakers who participated in this study.