



Evidence for raw acoustics as the target of phonetic imitation

Ivy Hauser, Emily Graham and Xinwen Zhang (University of Texas Arlington)



Phonetic imitation

- Imitation: process by which talkers alter production towards speech they hear (also accommodation or convergence)
- **Main question:** How does **speaker normalization** affect phonetic imitation?
 - Is convergence towards a normalized pattern or raw acoustic properties?

Case study: Imitation of English /s/ from a model talker with higher than average f0, formants, and spectral mean (SM)

Speaker normalization in imitation

- **Normalization:** perceptual process by which listeners identify phonologically identical sequences across different talkers (Johnson and Sjerps, 2021)
- Limited evidence on competing normalized and raw targets, but—
- English speakers decreased nasality after exposure to speaker with decreased nasality for that speaker but high raw nasality (Zellou et al., 2016)

→ What about **sibilant SM**, a primary cue?

Methods: Delayed shadowing

Two conditions: increased or decreased SM on model speech from talker with baseline high SM

- 80 /s/-initial target words balanced for frequency and following vowel rounding
- 40 sonorant-initial filler words
- Sibilant spectra shifted up/down 15%

Predictions

Direction of post-exposure shift depends on...

- **Normalized targets:** pattern in stimuli
- **Raw acoustic targets:** baseline relative to model

condition	target	predicted shift by baseline SM			
		low	avg	high	v. high
increased SM	both	{	/s/ →	}	model
decreased SM	acoustics	/s/ →	model	←	/s/
	pattern	←	/s/	model	←

Results: SM shift

24 L1 English participants randomly assigned to increased SM or decreased SM exposure

- **All participants exposed to increased stimuli** increased SM (left panel)
- **Direction of shift in decreased condition depended on participant baseline** (center and right panels)

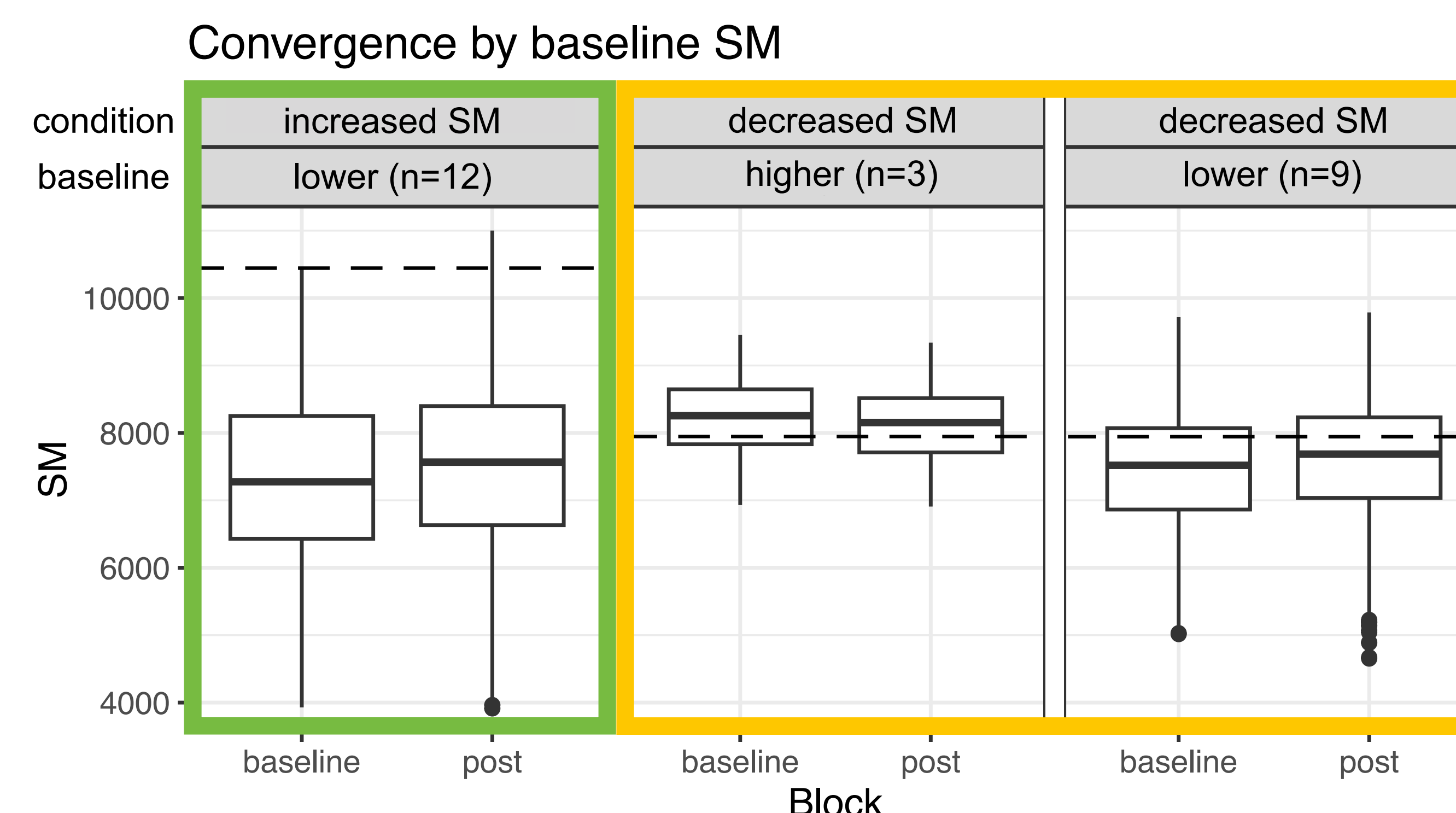


Figure 1. Shift by participant baseline. Dotted line is average model SM.

Results: Modeling

Metric of convergence and analysis of raw SM needed

Convergence analysis: Tests whether participants significantly converged, does not indicate direction of convergence

- Linear combination, alternative to DID (Priva & Sanker, 2019; MacLeod, 2021)
- Significance convergence in both conditions ($\beta = 0.27, p < 0.001^{***}$)

SM analysis: Indicates direction of shift, but not whether shift is convergent or divergent

- Best-fit mixed effects regression
- Significant increase in SM in both conditions ($\beta = 133.67, p < 0.001^{***}$)

Discussion and conclusion

Main finding: imitation patterns consistent with convergence towards raw (non-normalized) acoustics

- **Direction of shift depends on participant baseline**, not within-speaker pattern exhibited by model
- Increased SM is imitative and not related to increase in global hyperarticulation
- Decreased SM not necessarily blocked by phonological contrast with /ʃ/ (cf. Nielsen, 2011)

Implications: Imitation not necessarily mediated by speaker normalization

- Normalization may affect imitation for some phonetic dimensions but not others
- Potential effect of primary/secondary cue status