A Bayesian model of voice-onset time production

Mark VanDam[1] & Noah H. Silbert[2]

[1]Boys Town National Research Hospital

[2]Indiana University, Department of Psychological and Brain Sciences

mark.vandam@boystown.org www.vanDamMark.com

159th Acoustical Society of America 2pSC13

Baltimore, MD 20 April 2010

Main Point

VOT produced during a listen-and-repeat task was longer after training and after exposures to over-long tokens—but only in the voiceless series.

Background

Evidence suggests pervasive influence of linguistic experience in language storage, speech production, and speech perception.

Experiences affect linguistic memory representations or lexical storage (Pierrehumbert, 2001, Goldinger, 1996, 1998, Shockley, Sabadini, & Fowler, 2004, Bybee 2001, 2006)

Perceptual learning has shown recent exposure to specific words can affect representations (Norris, McQueen, & Cutler, 2003; Kraljic & Samuel, 2005).

Talkers' speech production is affected by ambient language (Sancier & Fowler, 1997) and can be affected by familiarity with specific words in conversational context (Pardo, 2006).

Method

Design: 56 target words with alveolar and velar stops in initial position, balanced by

a. lexical frequency: high, low

	low freq [g]	high freq [g]
low freq [k]	caulk – gawk coo – goo	cot – got curl – girl
high freq [k]	call – gall car – gar	could – good came – game

b. context: impoverished
could
say "could" very clearly
rich
Barry could shave in under 20 minutes
...before they could start to dine freely...

c. training condition: long, short
target tokens had synthetically lengthened or (slightly) shortened VoT
VoT was 180% or 80% of natural production for long or short, respectively

d. test phase: pretest, posttest

Task: Produce target words from orthographic cue.

Participants: 2 men, 4 women, 27.4 years old, monolingual American English talkers, normal hearing/vision, no speech or language deficits, white, right-handed.

Data: Each of 6 participants produced about 700 forms, yielding about 4200 total VOTs.

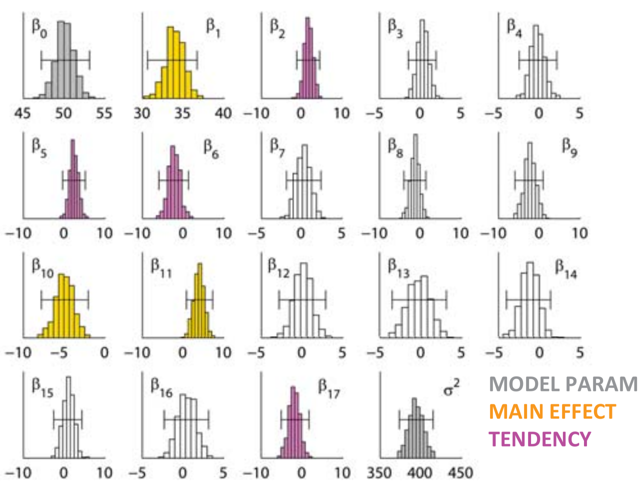
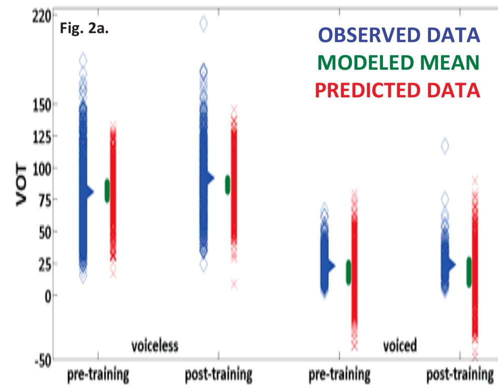


Fig. 1. Posterior distributions of model parameters. A central tendency different from zero indicates main effect and direction. Clear 'main effects' are shown in yellow and 'tendencies' are shown in pink. Model Parameters are described in Table 1.

Table 1: Model parameters, experimental factors, and effect coding scheme

Parameter	Purpose/Factor	Effect Code Values
σ^2	variance	
β_0	intercept	
β_1	voicing	−1 =voiced; 1 = voiceless
β_2	lexical frequency	−1 =low; 1 = high
β_3	context	−1 = impoverished; 1 = rich
β_4	test phase	−1 = pre; 1 = post
β_5	training condition	−1 = short; 1 = long
β_6	voicing × frequency	1 =voiceless & high
β_7	voicing × context	1 = voiceless & rich
β_8	frequency × context	1 = high & rich
β_9	voicing × frequency × context	1 =voiceless & high & rich
β_{10}	voicing × phase	1 = voiceless & post
β_{11}	voicing × condition	1 = voiceless & long
β_{12}	phase × condition	1 =post & long
β_{13}	voicing × phase × condition	1 = voiceless & post & long
β_{14}	frequency × phase	1 =high & post
β_{15}	voicing × frequency × phase	1 = voiceless & high & post
β_{16}	frequency × condition	1 =high & long
β_{17}	voicing × frequency × condition	1 = voiceless & high & long



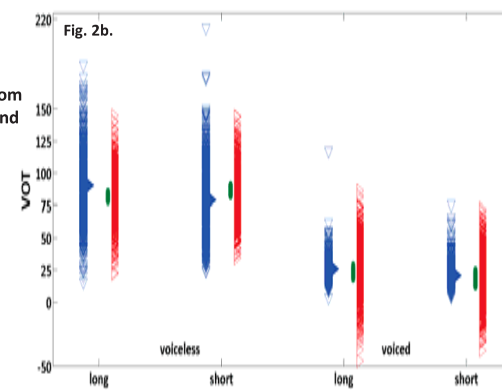


Fig. 2. Distributions and means of VoT production data for observed (blue) and modeled (green and red) data. The top panel shows voiceless VoTs were longer after training, but no change for voiced VoTs. The bottom panel shows shorter overall VoTs for those subjects trained on short exemplars.

Conclusions

1. Voiceless stops were longer after training, but voiced stops were not (Fig. 2a).
   ➔ The model did a good job here.

2. Subjects trained on short VoT exemplars had shorter VoT productions (Fig. 2b).
   ➔ The model did not do a good job here.

3. Tendencies lean toward possible
   a. longer VoT for high-frequency words (β_2)
   b. longer VoT when trained on long stops (β_5)
   c. shorter VoT for voiceless, high-freq words (β_6)
   d. shorter VoT for voiceless, high-freq, long-trained stops (β_{17})

4. Asymmetry and complexity is in the fine detail of stop production, especially after specific exposure. Supports previous work on fine-grained VOT production (Miller, Green, & Reeves, 1986, Volaitis & Miller, 1992) and perception (VanDam, 2007; VanDam & Port, 2009; VanDam, Silbert, & Port, 2009).

5. The Bayesian model suggests perhaps less complexity than other data modeling such as by traditional descriptive statistics (VanDam & Port 2009) or by resampling techniques (VanDam, Silbert, & Port 2009).

Funding: NIH/NIDCD T32-DC00012, T32 DC00013-30.