

Why /zname/ is easier to say than /pkamo/

Previous research has shown that not all non-native phonotactic patterns are equally difficult for speakers to produce. For example, Vietnamese learners of English had only moderate trouble producing coda /s/ and /f/, but had increasing difficulty with /v/ and /l/, followed by /ʃ/ (Hansen 2004). Davidson (2006a, 2006b) demonstrated that English speakers producing Slavic pseudowords with fricative-initial consonant clusters were more accurate on the /fC/-initial sequences, followed by /zC/ and then by /vC/. These studies have also shown that production of non-native phonotactics is not all-or-none, but rather gradient. In the current study, three variables are manipulated to determine what further factors contribute to gradient accuracy on non-native word-initial consonant clusters. The first two—consonant manner and place—are phonological factors, whereas the third factor—auditory-only or audio-visual presentation of the stimuli—is an extralinguistic factor that may play an important role in production patterns (e.g. Vendelin and Peperkamp 2006).

These influences were tested by having 23 English speakers produce 120 words containing Russian-possible C₁C₂ initial clusters. To examine the effect of manner, consonants were combined into stop-stop, stop-fricative, stop-nasal, fricative-stop, fricative-fricative, and fricative-nasal sequences. Place was examined with many different consonants in C₁ position. The role of the input medium was manipulated by presenting half of the stimuli, spoken by a native Russian speaker, in an auditory-only condition. The other half were presented both auditorily and visually on a computer screen in English-like orthography. Speakers repeated the target word aloud once.

The most common error for both presentation conditions was vowel insertion ($M = 22\%$), followed by C₁ deletion (8%), C₁ change (5%), and prothesis (5%). An analysis of manner and presentation showed a significant main effect of manner combination and presentation type, but no interaction. Speakers were significantly less accurate on audio-only stimuli. Post-hoc analyses indicated the following differences among manner combinations ('>' means significantly more accurate; '=' indicates no difference): FN = FS = SF > FF = SN = SS (see Figure 1). An analysis of C₁ showed a significant effect of the first consonant and presentation type. Again, accuracy was significantly worse for the audio-only condition. Results for C₁ (see Table 1) show that voiced consonants tend to be less accurate than voiceless ones, stops are generally less accurate than fricatives, and voiceless coronals are the most accurate of all.

These findings demonstrate that understanding the gradient production of consonant clusters depends on the level being examined. The examination of manner combination indicates that the six sequences under consideration fall into only two groups, whereas a breakdown by first consonant shows a finer-grained distinction. Like Davidson (2006a, 2006b), it can be argued that these findings have both articulatory and acoustic bases (e.g. the place of stops is better recovered by releasing into an approximant, Steriade 1997), but that they cannot be solely attributable to whether or not the same clusters are found in word-medial or final position in English. Not surprisingly, the presence of text improves accuracy by approximately 10%; however, the patterns of accuracy for manner combination and C₁ type are still largely the same regardless of whether or not the participants were presented with text. These findings suggest that while peripheral phonetic influences may cause a (relatively small) general decrement in performance, the same variable grammatical mechanism drives speech production independently from input type.

Figure 1. Proportion of accurate responses for CC stimuli broken down by manner combination.

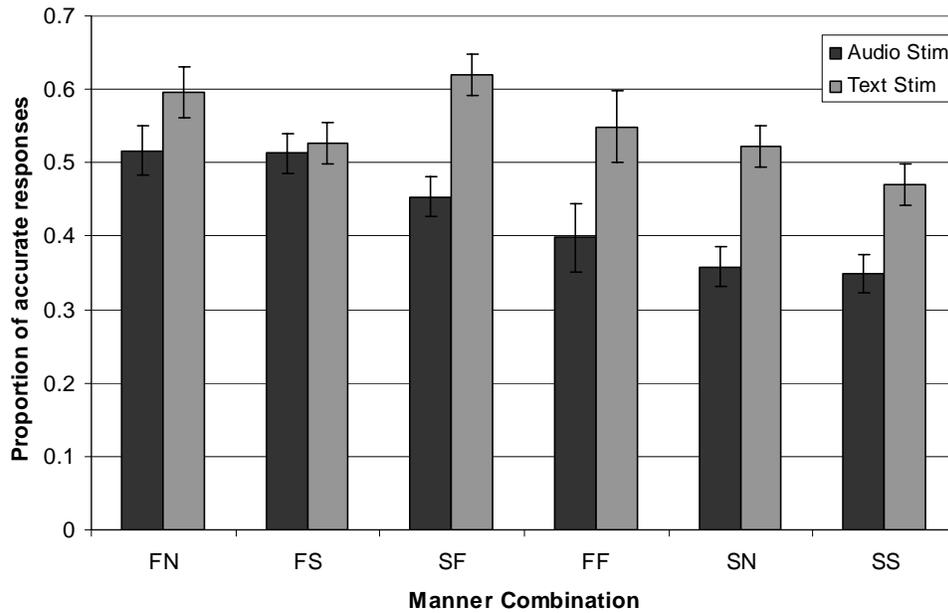


Table 1. Results of Student-Newman-Keuls post-hoc tests for C1 for both audio-visual and audio only stimuli. The values are proportion of correct utterances. Values in the same subset are not significantly different from one another.

C1	Subsets For Audio					Subsets For Text			
	1	2	3	4	5	1	2	3	4
d	0.24					0.45			
g	0.24					0.36			
v	0.27					0.35			
b	0.29					0.37			
p		0.42					0.59		
k		0.46	0.46					0.71	
z			0.54	0.54			0.56		
t				0.61				0.72	
f				0.64				0.74	
s					0.99				0.97

Davidson, L. 2006a. Phonology, phonetics, or frequency: Influences on the production of non-native sequences. *Journal of Phonetics* 34 (1), 104-137.

Davidson, L. 2006b. Phonotactics and articulatory coordination interact in phonology: evidence from non-native production. *Cognitive Science* 30 (5), 837-862.

Hansen, J. 2004. Developmental sequences in the acquisition of English L2 syllable codas. *Studies in Second Language Acquisition* 26, 85-124.

Steriade, D. 1997. Phonetics in Phonology: The Case of Laryngeal Neutralization. ms., UCLA.

Vendelin, I. and S. Peperkamp. 2006. The influence of orthography on loanword adaptations. *Lingua* 116 (7), 996-1007.