Supplementary Material for:
Chunking of phonological units in speech sequencing

Jennifer Segawa¹,², Matthew Masapollo², Mona Tong², Dante J. Smith³, & Frank H. Guenther²,³,⁴

¹ Departments of Neuroscience and Biology, Stonehill College
² Department of Speech, Language and Hearing Sciences, Boston University
³ Graduate Program for Neuroscience, Boston University
⁴ Department of Biomedical Engineering, Boston University

S.1 Effects of syllable position on speech motor learning

We conducted two additional sets of analyses to test for effects of syllable position (syllable-initial [onset] clusters vs. syllable-final [coda] clusters) as such effects are suggested by some theories of speech production (see, e.g., Browman & Goldstein, 1988, 1995; Byrd, 1995, 1996; Keating, Wright, & Zhang, 1999; Byrd, Lee, Riggs, & Adams, 2005; Bohland, Bullock & Guenther, 2010). S.1.1. Evidence of learning for practiced syllables.

Our first set of analyses was aimed at examining performance improvements over the course of the practice blocks on day one. We conducted an ANOVA (with syllable sequence type [native CC vs. non-native CC], syllable position [onset vs. coda], and time (first five trials vs. last five trials) as within-subjects factors) on the mean percentage of sequencing errors during the practice blocks (on day 1). Figure S1 shows the mean percentage of sequencing errors for each sequence type (native CC vs. non-native CC) as a function of syllable position (onset vs. coda) and time (first five trials vs. last five trials).

The ANOVA revealed a significant main effect of sequence type [$F(1,10) = 109.938, p < 0.001, \eta^2 = 0.917$], such that speakers produced fewer errors while executing the native CC syllables ($M = 3.0, SD = 5.1$) compared to the non-native CC syllables ($M = 44.0, SD = 19.9$). There was also a main effect of time, such that speakers produced more errors during the first five
trials (on day 1; $M = 29.6, SD = 28.1$) compared to the last five trials (on day 1; $M = 15.79, SD = 20.66$). The interaction effect between time and sequence type approached significance [$F(1,10) = 3.836, p = 0.079, \eta^2 = 0.277$]. To tease apart the interaction, difference scores were computed by subtracting the mean sequencing error rates averaged across the first five and last five trials for each sequence type. A post-hoc LSD $t$-test performed on these difference scores indicated that they were marginally significantly larger for the non-native CC syllables [$M = 18.1, SD = 21.4$] compared to the native CC syllables [$M = 6.1, SD = 5.9; t(10) = -1.930, p = 0.079$].

Critically, however, the main effect of syllable position [$F(1,10) = 2.518, p = 0.144, \eta^2 = 0.201$] did not reach statistical significance. In addition, none of the interaction effects involving syllable position were significant: syllable position $\times$ time [$F(1,10) = .457, p = 0.514, \eta^2 = 0.044$]; syllable position $\times$ sequence type [$F(1,10) = 2.758, p = 0.128, \eta^2 = 0.216$]; syllable position $\times$ time $\times$ sequence type [$F(1,10) = .444, p = 0.520, \eta^2 = 0.043$].

S.1.2. Generalization of learning to novel syllables.

Our second set of analyses were designed to examine the specificity of the motor sequence learning that occurred for syllables containing non-native onset and coda clusters during the practice blocks. We conducted separate ANOVAs (with syllable sequence type [practiced CCVCC vs. practiced CC vs. practiced CVC vs. novel CCVCC] and syllable position [onset vs. coda] as within-subjects factors) on the mean percentage of sequencing errors and the mean percentage of total errors (sequencing and non-sequencing) for the first five trials of the test session (to minimize practice effects during the test phase). Figure S2 shows boxplots of the mean percentage of sequencing errors (see main text for details) for the first five utterances of each syllable type as a function of syllable position (onset vs. coda). The ANOVA showed a significant main effect of sequence type [$F(3,30) = 3.306, p = 0.033, \eta^2 = 0.248$]. Post-hoc $t$-tests revealed that the mean sequencing error
rates for the *practiced CCVCC* syllables \([M = 30.4, SD = 42.7]\) were significantly lower than the *novel CCVCC* \([M = 52.2, SD = 29.2; t(10) = -3.425, p = 0.006]\) syllables and marginally lower than *practiced CC* \([M = 42.7, SD = 22.2; t(10) = -2.218, p = 0.051]\) and *practiced CVC* \([M = 44.5, SD = 25.4; t(10) = -1.985, p = 0.075]\) syllables. While sequencing error rates for the *practiced CC* syllables were not significantly lower than the *practiced CVC* \([t(10) = -0.281, p = 0.785]\), they were significantly lower than the *novel CCVCC* \([t(10) = -2.345, p = 0.041]\) syllables, suggesting some minor improvement from practicing the CC portion of the CCVCC syllables but less improvement than practicing the whole syllable. The *practiced CVC* and *novel CCVCC* error rates were not significantly different from each other \([t(10) = -1.376, p = 0.199]\). The effect of syllable position was marginally significant \([F(1,30) = 4.787, p = 0.054, \eta^2 = 0.324]\), such that more errors occurred during the execution of coda clusters \((M=30.3; SD=25.2)\) compared to onset clusters \((M=21.7; SD=17.3)\). The interaction did not reach statistical significance \([F(3,30) = 1.087, p = 0.369, \eta^2 = 0.098]\).

Figure S3 shows boxplots of the mean percentage of *total errors* (sequencing and non-sequencing) for the first five utterances of each syllable type as a function of syllable position (onset vs. coda). The ANOVA showed a highly significant main effect of sequence type \([F(3,30) = 4.636, p = 0.009, \eta^2 = 0.317]\). Post-hoc LSD paired *t*-tests showed that the mean error rates for the *practiced CCVCC* syllables \([M = 46.3, SD = 25.1]\) were significantly lower than the *practiced CC* \([M = 59.5, SD = 22.2; t(10) = -2.498, p = 0.032]\), *practiced CVC* \([M = 65.4, SD = 26.4; t(10) = -2.003, p = 0.043]\) and *novel CCVCC* \([M = 66.8, SD = 23.9; t(10) = -3.190, p = 0.010]\) syllables. The mean error rates for the *practiced CC* syllables were not significantly lower than either the *practiced CVC* \([t(10) = -0.643, p = 0.535]\) or *novel CCVCC* \([t(10) = -1.638, p = 0.132]\) syllables. The *practiced CVC* and *novel CCVCC* error rates were not significantly different from each other.
There was no significant effect of syllable position \([F(1,30) = 4.787, p = 0.054, \eta^2 = 0.324]\) or interaction \([F(3,30) = 0.631, p = 0.600, \eta^2 = 0.059]\).

Although previous kinematic studies have consistently found effects of syllable position on the spatial and temporal production of native-language consonant sequences (Browman & Goldstein, 1988; Byrd, 1995; Keating, Wright, & Zhang, 1999; Byrd, Lee, Riggs, & Adams, 2005), no clear effects emerged in the current experiment. It is important to note, however, that we measured motor learning using perceptual/transcription methods and simple acoustic measures of utterance duration. However, these measures provide a limited window on potential articulatory differences between onsets and codas since they cannot capture movement characteristics among cluster elements (such as inter-gestural timing), and how those characteristics may vary as a function of syllable position and motor practice. Furthermore, the model of speech sequencing presented in Figure 2 of the main article is agnostic to this distinction, largely because the current findings cannot fully illuminate this issue since a given cluster was always used as an onset or always used as a coda in our experiment. We plan to investigate this issue in future experiments.
Supplementary References


**Figure S1:** Boxplots of the mean percentage of *sequencing errors* for the first and last five utterances of each syllable type (on Day 1) as a function of sequence type (native CC vs. non-native CC) and syllable position (onset vs. coda).
Figure S2: Boxplots of the mean percentage of sequencing errors for the first five utterances of each syllable type as a function of syllable position (onset vs. coda).
Figure S3: Boxplots of the mean percentage of *total errors* (sequencing and non-sequencing) for the first five utterances of each syllable type as a function of syllable position (onset vs. coda).