Evidence for the consonant cluster as a basic unit of speech motor sequencing

Matthew Masapollo[1], Jennifer Segawa[1,2], Mona Tong[1], & Frank Guenther[1,3]
[1]*Department of Speech, Language & Hearing Sciences, Boston University;*
[2] *Departments of Neuroscience and Biology, Stonehill College;*
[3]*Department of Biomedical Engineering, Boston University*
mmasapol@bu.edu; jsegawa@stonehill.edu; monatong44@gmail.com; guenther@bu.edu

Fluent speech involves rapid sequencing and initiation of motor programs for the phonological units that make up an utterance. The exact nature of these motor sequencing units remains unclear (see Fig. 1). While many researchers have posited basic units that are gestural or phonemic ([1], [2]), others have pointed to larger units, which we will call sub-syllabic constituents (SSCs), that can contain multiple phonemes ([3]), such as syllable onsets, nuclei, and codas. Still others have posited that optimized motor programs exist for entire syllables ([4], [5]). In a prior study from our laboratory [6], English speakers were trained to produce novel phoneme sequences (monosyllabic CCVCC pseudowords) with consonant clusters that were phonotactically illegal in English but legal in other languages. Two days of practice led to measureable performance gains (e.g., reduced utterance durations) for these novel pseudowords, and it was postulated that these gains were primarily due to learning of motor "chunks" for the new consonant clusters, consistent with the view that the units of speech motor sequencing are SSCs. However, the design of that study could not distinguish this possibility from the possibility that the newly learned chunks were syllable-sized.

To address this issue, the current study investigated whether the learning gains found in [6] were specific to trained syllables or whether they generalized to novel syllables containing the newly learned consonant clusters. Generalization to untrained syllables would indicate that the learned motor chunks were smaller than the full syllable. We also compared the learning of novel pseudowords involving phonotactically legal English consonant clusters (e.g., 'flisk') to learning of novel pseudowords with novel (illegal) consonant clusters (e.g., 'gvasf'). If the consonant cluster is the motor chunk learned during training, pseudowords with illegal clusters should show performance gains with practice, whereas pseudowords involving legal clusters should show no gain as these clusters are already well-learned from prior linguistic experience.

We found that the illegal sequences were produced faster and with fewer errors over the two-day period, indicating that speech motor sequence learning occurred. In contrast, we found no significant behavioral gains for the legal sequences. Speakers started out at near-ceiling performance for the production of the legal sequences, presumably because they could produce these sequences by concatenating existing motor programs for native clusters in their production repertoire. Critically, speakers were also faster and more accurate at producing the novel illegal sequences with clusters that occurred in the learned illegal sequences. Moreover, this advantage for producing the previously learned illegal clusters fully generalized to vowel contexts not included in the training stimuli, indicating that learning gains were not specific to entire syllables; instead, once a novel cluster was learned, it could be efficiently produced in new sequences. Collectively, these findings indicate that, at some level, the speech production process entails learning and executing optimized sequences of vocal tract movements that correspond to phonological units smaller than an entire syllable but larger than an individual phoneme.

Keywords: speech production; phonotactics; sub-syllabic constituents; speech sound sequencing; speech motor learning
Theme: Perception-production dynamics

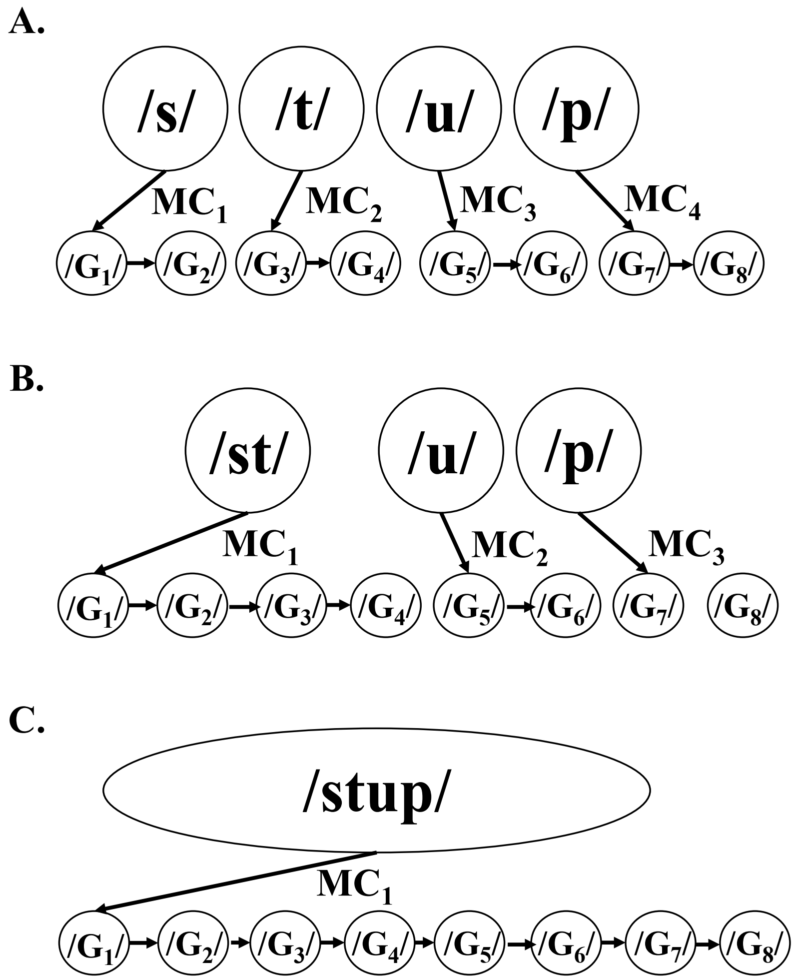

Figure 1. *Three possible accounts of the motor* "*chunks*" *underlying the production of the CCVC sequence* "*stoop.*" *(A.) Four separate motor chunks, one for each individual phoneme. (B.) Three separate chunks, one for each SSC (onset, nucleus, coda). (C.) A single motor chunk for producing the entire syllable. See text for further details. G = gesture, MC = motor chunk.*

[1] Browman, C.P. & Goldstein, L. (1992). Articulatory phonology: an overview. *Phonetica*, 49, 155-180.

[2] Guenther, F.H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, 102 (3), 594 – 621.

[3] Shattuck-Hufnagel, S. (1983). Sublexical units and suprasegmental structure in speech production planning. In: MacNeilage, P. (Ed.), The Production of Speech. Springer-Verlag, New York, pp. 109 – 136

[4] Levelt, W.J., Roelofs, A., & Meyer, A.S. (1999). A theory of lexical access in speech production. *Behavioral and Brain Sciences*, 22 (1), 1 – 38.

[5] Guenther, F.H. (2016). *Neural Control of Speech*. Cambridge, MA: MIT Press.

[6] Segawa, J., Tourville, J.A., Beal, D.S., & Guenther, F.H. (2015). The neural correlates of speech motor sequence learning. *Journal of Cognitive Neuroscience*, 27(4), 819-831.