

Temporal Coordination of Quantity and Stress in Estonian Prosody

Argyro Katsika¹, Eva Liina Asu², Matthew Gordon¹, Pärtel Lippus², Anton Malmi², Yining Qu¹,
Sherry Chien¹

¹*UC Santa Barbara, CA, USA;* ²*University of Tartu, Estonia*

The Estonian ternary quantity system has long challenged models of prosodic representation, as phonological length, lexical stress, and foot structure are tightly intertwined. Quantity contrasts are realized on stressed vowels and adjacent consonants, and their phonetic implementation has been described as distributed across a disyllabic foot, yielding inverse temporal relations between stressed and post-tonic syllables [1–4]. However, because these contrasts occur almost exclusively in word-initial stressed syllables, it remains unclear whether the observed timing patterns reflect foot structure, stress-driven prominence, phonemic quantity, or their interaction.

This study uses electromagnetic articulography (EMA) to disentangle the articulatory and temporal consequences of quantity and stress in Estonian. We examine kinematic properties of onset consonant gestures across words that vary systematically in vowel quantity (short, long, overlong), word length (two, three, and four syllables), and stress placement. Crucially, in addition to native vocabulary with initial stress, we analyze loanwords in which primary stress falls on the second syllable. This design allows quantity and stress to be dissociated in otherwise comparable prosodic environments. The analysis focuses on formation duration and displacement of consonant gestures in the onsets of the first three syllables (C1–C3), providing a window into how timing relations unfold across overlapping and sequential gestural intervals. Data were collected from eight native speakers under controlled prosodic conditions.

The results show that quantity and stress share similar kinematic correlates, i.e., longer and larger gestures, but differ systematically in their temporal alignment. Quantity effects emerge early in the gestural sequence, scaling with the phonemic length of the stressed vowel and, when coarticulatory conditions permit, extending into overlapping and immediately following consonant gestures. These effects are robustly reflected in post-tonic consonants (C2), even across syllable and word boundaries, indicating that quantity is implemented through temporally expansive articulatory domains. In contrast, stress-related effects arise later in the sequence, shaping post-tonic timing rather than uniformly expanding the stressed syllable itself. This delayed stress modulation gives rise to inverse temporal relations in more distal material (C3), without requiring the prosodic foot to serve as a fixed timing domain (Fig. 1).

Taken together, these findings suggest that quantity and stress are implemented through partially overlapping but temporally staged control mechanisms. Rather than reflecting a single foot-level timing template, inverse timing relations appear to emerge from the interaction of early quantity-driven planning demands with later stress-related adjustments. We argue that this pattern is compatible with a mu-gesture account of prosodic timing [5], in which prosodic structure reflects the coordination of planning-level temporal intervals with segmental articulation. More generally, the results highlight how fine-grained articulatory dynamics can inform cognitive models of prosodic representation and speech planning, showing that prosodic functions traditionally treated as structurally coextensive may instead be distinguished by their temporal organization in production.

From a cognitive perspective, these findings argue for a representational distinction between quantity and stress that is expressed in the temporal architecture of speech planning rather than fixed prosodic domains. Quantity functions as an early, segmentally anchored timing demand shaping overlapping gestures, whereas stress emerges later as a modulatory control on subsequent material. This temporal stratification explains inverse timing patterns without invoking foot-level timing templates.

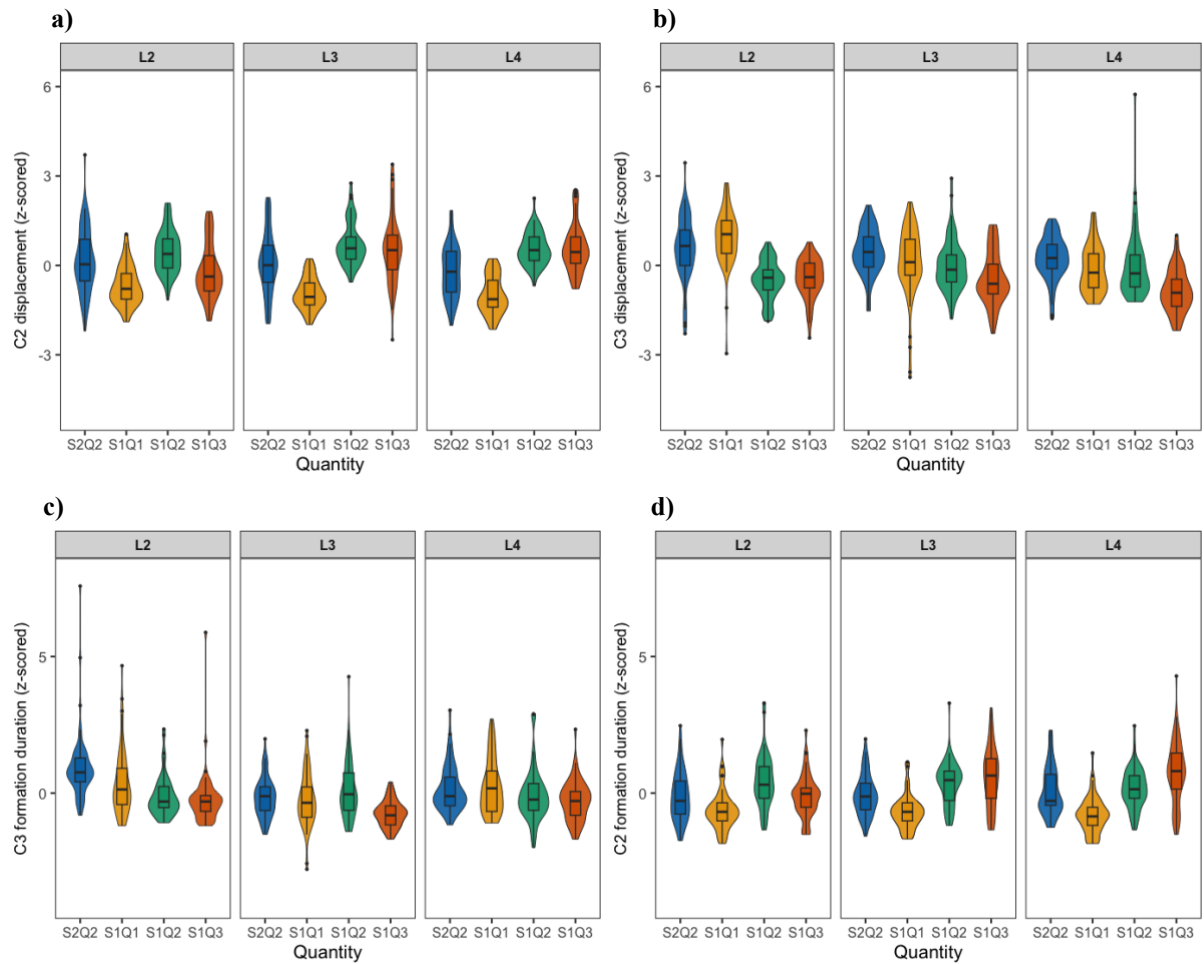


Fig. 1. Z-scored kinematic measures of onset consonant gestures as a function of vowel quantity in the first syllable, i.e., short (S1Q1), long (S1Q2), overlong (S1Q3), and long with stress on the second syllable (S2Q2), and word length (two, three, and four syllables: L2–L4). Panels show (a) C2 displacement, (b) C2 formation duration, (c) C3 displacement, and (d) C3 formation duration. Quantity-related modulation is expressed early and scales with phonemic length, while later inverse timing patterns emerge in more distal post-tonic gestures.

References

- [1] Lehiste, I. (1960). Segmental and syllabic quantity in Estonian. Indiana University Publications.
- [2] Lehiste, I. (1997). Search for phonetic correlates in Estonian prosody. Institute of Estonian Language.
- [3] Lippus, P., Asu, E. L., Teras, P., & Tuisk, T. (2013). Quantity-related variation of duration, pitch and vowel quality in spontaneous Estonian. *Journal of Phonetics*, 41, 17–28.
- [4] Katsika, A., Asu, E. L., Gordon, M., Lippus, P., & Malmi, A. (2024). The kinematic profile of the Estonian ternary quantity distinction. *Speech Prosody*.
- [5] Saltzman, E., Nam, H., Krivokapić, J., & Goldstein, L. (2008). A task-dynamic toolkit for modeling the effects of prosodic structure on articulation. *Speech Prosody*.