The link between sonority sequencing and gestural coordination

Overview: Previously, Crouch (2022); Crouch et al. (2023) suggested that sonority sequencing affects gestural overlap of CC clusters in Georgian. Specifically, they found that a sequence of two consonants in Georgian with a sonority rise exhibited less overlap than those with a sonority plateau, which in turn were less overlapped than those with a sonority fall. They suggested that this observation argues against a simple perceptual recoverability account of gestural coordination and challenges speech production theories which assume no variation of gestural coordination due to sonority (Browman & Goldstein, 1989, 1992, 2000; Nam & Saltzman, 2003; Goldstein et al., 2006; Xu et al., 2006; Hermes et al., 2013; Durvasula et al., 2021; Liu et al., 2022; Iskarous & Pouplier, 2022). While Crouch et al. (2023) speculated that the observed effect of sonority sequencing on consonant sequences is limited to Georgian, other researchers have also observed that the lag between a consonant and a vowel – CV lag – was significantly shorter in syllables beginning with a nasal stop than in syllables beginning with an oral stop (Shaw & Chen, 2019; Gao, 2008). We use these prior results to infer a more general relationship between sonority and segment sequencing beyond consonant sequences, to perhaps all relevant segment sequences within a syllable, including CV sequences. In line with this broader generalization, in the current study, we found that the CV lag in English correlates positively with the sonority difference between a C and V. The results suggest a sonority-based speech production model and potentially provide a novel unified basis to explain several markedness patterns.

Hypothesis & Variables: The current study tested the hypothesis that CV lag positively correlates with the sonority difference between C and V. To compute the CV lag, each timestamp of the consonant was subtracted from the corresponding timestamp of the vowel. Therefore, each syllable has four measurements: CV lag based on 1) gestural onset, 2) target onset, 3) target offset, and 4) gestural offset. To probe the effect of sonority difference, the sonority scale in Parker (2002, 2008, 2011) was adopted where each natural class has a value from 1-17 and a larger value indicates a more sonorous class. The sonority difference was quantified by subtracting the C sonority from the V sonority. For instance, [jɑ] has the sonority indexes 12 and 17 for glides and low vowels respectively, and its sonority difference would be 17-12=5. It is important to note that even though a numeric scale (which implies linearity) was used, we still consider sonority as abstract on a relative scale, and this consideration was reflected in the analysis discussed below.

Methods: The hypothesis was evaluated using measurements from 3346 utterances (57 different speakers, 34 stimuli) from the Wisconsin X-ray Microbeam Database, which has simultaneous acoustic and kinematic recordings (Westbury et al., 1990). Specifically, 17 nonce-words *uhCa* (e.g. *uhya* [jɑ], *uhma* [mɑ], with the target CV sequence underlined and transcribed here), 15 *sVd* words (e.g. *seed* [si], *sad* [sæ]) together with *back* [bɑ] and *been* [bi] were analyzed. The stimulus choice ensured that we could test both the same C with varying V and the same V with varying C. The kinematic data were annotated in Matlab using the default settings of the *lp_findgest* algorithm of the *mview* package (Tiede, 2005). To evaluate the hypothesis, descriptive plots were generated by the *tidyverse* package (Wickham et al., 2019), and mixed effects modeling was conducted by the *lme4* (Bates et al., 2014) and *lmerTest* packages (Kuznetsova et al., 2017) in R (R Core Team, 2017).

Results: Figure 1 shows that CV lag based on gestural onset statistically significantly increased with the sonority difference for all the stimuli in an appropriate mixed effects model ($\hat{\beta}_{\text{sonority diff.}}$=22 ms, t(1490)=16.4, p<0.00001). While an analysis with *all* the data and the use of difference scores as a predictor is easier to statistically model and has higher statistical power, it does assume that the sonority scale used is linear and not just relative (which is contrary to most phonologists' belief), and it collapses across different articulators or gestures. For these reasons, we also analyzed more nuanced sub-groups of the data. Specifically, we looked at sets of stimuli that control for the place of articulation or gesture of the consonant, such as 1) lip aperture: [wɑ, mɑ, bɑ, pɑ]; 2) tongue tip: [lɑ, nɑ, zɑ, dɑ, sɑ, tɑ]; 3) bV: [bi, bɑ]. Crucially, the analyses showed similar (statistically significant) positive relationships as the overall pattern. Finally, other measurements of CV lag based on target onset, target offset, and gestural offset all show similar patterns.

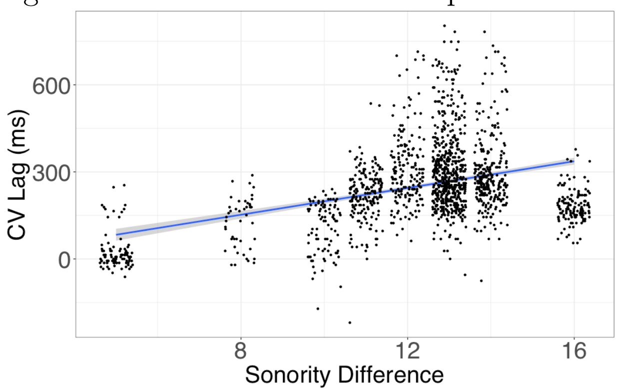

Figure 1: CV lag increases with sonority difference for all data based on gestural onset. The *geom_jitter* option (size=0.9) is used to spread out the overlapping dots for clarity.

Implications: The findings of the current study support the proposal of a specific type of speech production model wherein sonority determines gestural coordination patterns. Such a sonority-based speech production model entails that gestures sequentially map from abstract sequential segmental representations within a syllable according to the sonority differences with the adjacent segment. Note that this sonority difference is calculated in an ordered way within a syllable (latter segment minus the former), so the sonority difference can be negative. Though it needs further evaluation, this model can account for the results in the current study and previously observed correlations in gestural timing discussed above.

The study also provides the potential basis of a unified explanation for several markedness constraints in human language, if we make a further assumption that larger lags are preferred (this is, perhaps, related to the functional reason of perceptual recoverability). Since a sonority rise has a larger gestural lag than a sonority plateau and a sonority fall, the preference towards a larger gestural lag surfaces as the Sonority Sequencing Principle (SSP), which states that, cross-linguistically, in syllable onsets, a sonority rise is preferred over a plateau, which in turn is preferred over a sonority fall (Clements, 1990; Greenberg, 1965; Berent et al., 2007; Ren et al., 2010; Zhao & Berent, 2016). Similarly, the above view could also explain the Sonority Dispersion Principle, which states that, in a CV sequence within a syllable, the onset and nucleus should differ from each other in sonority as much as possible (Clements, 1990; Parker, 2011). Finally, the cross-linguistic tendency that CV syllables are much more common (more "unmarked") than VC syllables (Ohala, 1990; Tabain et al., 2004; Nam et al., 2009) could also be explained by the preference towards a larger lag since 1) V is more sonorous than C; 2) CV has a larger sonority difference than VC.