UNIVERSITY OF CALIFORNIA

Santa Barbara

Postcards from the syllable edge: sonority and articulatory timing in complex onsets in Georgian

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Linguistics

by

Caroline Quigley Crouch

Committee in charge:

Professor Argyro Katsika, Chair

Professor Eric W. Campbell

Professor Ioana Chitoran, Université Paris Cité

Professor Matthew Gordon

September 2022

The dissertation of Caroline Quigley Crouch is approved.

Eric W. Campbell

Ioana Chitoran

Matthew Gordon

Argyro Katsika, Committee Chair

August 2022

This dissertation is dedicated to my mom, who told me in 2010 that I might like linguistics and who was, as she always is, completely right. I love you.

ACKNOWLEDGEMENTS

This dissertation would not exist in its current form or have been completed at this time without the help and support I received from my friends, family, and fellow linguists. I hope that these acknowledgements can express even a fraction of my gratitude.

To my committee as a whole: Argyro Katsika, Ioana Chitoran, Eric Campbell, and Matt Gordon, for their flexibility and generosity with their time as I grappled with writing a laboratory dissertation during a pandemic, and to each of them individually as well.

To Eric for his guidance in my documentary fieldwork, which is integral to the way I do and think about linguistics and for telling me, way back in field methods class when I wasn't sure I 'had what it takes', that I did.

To Matt, for his insightful feedback that always directed me to research and analyses that greatly improved this dissertation and for his mentorship as my advisor during my MA.

To Ioana, who kindly agreed to serve as my external committee member at our first meeting and met weekly with me despite the seven-hour time difference, whose expertise on Georgian, syllable structure, sonority and laboratory phonology were invaluable to this dissertation, and whose previous research was critical for the design and analyses of my own work.

To Argyro, for being the best dissertation advisor I can imagine. I started this dissertation knowing almost nothing about laboratory phonology, Articulatory Phonology, or EMA, and my growth in these areas would not have been possible without her and the amount of time, care, and patience she devoted to me as her student and to my research. Working with her has made me a better and more thoughtful writer and a more rigorous scientific thinker and this dissertation would not be close to what it is now without her guidance. In 2019, someone told me I was lucky to be Argyro's student. He was right, and I told him so, and over the course of working on this dissertation I have re-realized so many times exactly how right he was and how lucky I am.

To Tamra Wysocki-Niimi, my Georgian instructor in undergrad, whose class was my favorite every year and the source of my love for Caucasian languages and phonetics and phonology.

To everyone at the Haskins A93 meetings for the feedback, especially to Doug Whalen for inviting me, and to Jason Shaw and Mark Tiede for their help with deciding on an anchor point. To Louis Goldstein, for writing Matlab code for me to analyze modulation functions, even though this analysis did not make it into the dissertation.

To Nestan for her help with designing stimuli and connecting me with other Georgian speakers in Southern California. To Elene, Nikoloz, and Tamta for participating in my experiments, and to the Georgian Cultural and Educational Center of Southern California for circulating my call for participants.

To my friends before grad school, who throughout my entire graduate career have always believed in me. In Savannah: Jess, Maria, James, Debbie, and most of all Lindsey, my oldest friend and, frankly, my non-biological sister, whose love and support mean the world and who inspires me constantly. In Chicago: Will and Michelle, who hosted and fed and entertained me so many times throughout graduate school. To Miche specifically, for over a decade of near-constant chatting, joking, chilling, hivemind-sharing, and friendship.

To Karen and Jiyoung, my labmates, who helped me run my dissertation experiments, and with whom I spent many, many wonderful hours in the lab, working and not working, and who are also great travel buddies.

To all of my other fellow UCSB ling grads, but especially Megan Lukaniec, Heather Simpson, Nate Sims, and Chris Brendel, for years of shooting the breeze about life, linguistics, and everything else. To Kevin Schaefer, my colleague, roommate, fieldwork companion, and ‘platonic life partner’, for every minute we spent together. *Naruwox romba’ txaine ka taxkox mbonembone hlamel ka taxndum ma’ taxyoxoxo noxoine a kıtax*.

And last, but most certainly not least, to my family, whose emotional and material support is the bedrock for everything. To my grandmother Ruth Quigley, to my little brother Stephen, and my mother Stephanie, who all endured and supported me during the final push to finish. To my late grandfather Stephen Quigley, whose belief in the value of education is the reason I am where I am. To my dad, the original Dr. Crouch, who I miss every day. There aren’t words enough.

VITA OF CAROLINE QUIGLEY CROUCH
August 2022

EDUCATION

Bachelor of Arts in Linguistics, University of Chicago, June 2013 (general honors)
Master of Arts in Linguistics, University of California, Santa Barbara, June 2016
Doctor of Philosophy in Linguistics, University of California, Santa Barbara, September 2022 (expected)

PROFESSIONAL EMPLOYMENT

2013-2021: Teaching Assistant/Teaching Associate, Department of Linguistics, University of California, Santa Barbara
Winter 2016: Teaching Assistant, Department of Linguistics, University of California, Los Angeles
Winter 2016-Spring 2017: Hewitt Project Temporary Student Assistant, Cataloging and Metadata Services, University of California, Santa Barbara

PUBLICATIONS

Crouch, Caroline. 2019. Acoustic properties of the lateral contrast in Ninde. In Sasha Calhoun, Paola Escudero, Marija Tabain and Paul Warren (eds.) *Proceedings of the 19th International Congress of Phonetic Sciences, Melbourne, Australia 2019* (pp. 1288-1291). Canberra, Australia: Australasian Speech Science and Technology Association Inc
.
Crouch, C., Katsika, A., Chitoran, I. 2020. The role of sonority profile and order of place of articulation on gestural overlap in Georgian. *Proc. 10th International Conference on Speech Prosody 2020,* (pp. 205-209). DOI: 10.21437/SpeechProsody.2020-42.

Crouch, C., Katsika, A., Chitoran, I. (2022) Georgian Syllables, Uncentered?. *Proc. 11th International Conference on Speech Prosody 2022,* (pp. 215-219). DOI: 10.21437/SpeechProsody.2022-44.

AWARDS

Department of Linguistics Graduate Fellowship, University of California, Santa Barbara 2013-2014

Endangered Languages Documentation Programme Small Grant, 2016-2017

Firebird Foundation Grant (with Kevin Schaefer), 2018

ABSTRACT

Postcards from the syllable edge: sonority and articulatory timing in complex onsets in Georgian

by

Caroline Quigley Crouch

This dissertation investigates how syllables are organized in space and time to form coherent units. To do so, I engage with two theoretical approaches to the syllable: one based on the Sonority Sequencing Principle, and the coupled oscillator model advanced within Articulatory Phonology. I focus on Georgian because Georgian has extremely permissive phonotactics: onsets can be up to seven consonants and there are minimal restrictions on the combinations of consonants that can occur. I address this question via a trio of Electromagnetic Articulography experiments that examine the effect of sonority shape on the temporal relationships between gestures in the onset.

The first experiment examines the relationship between sonority and the global timing of gestures in the onset. In the Articulatory Phonology framework, syllable onsets are coordinated as a unit with respect to the nucleus. I find that Georgian does not show evidence of global coordination regardless of the sonority shape of the onset and provide motivations for this unexpected behavior from other aspects of linguistic structure.

The second experiment examines the relationship between sonority and local timing of consonant gestures in the onset. Here I find a systematic relationship between sonority and timing: in sonority rises (e.g., /br/) consonant gestures are the least overlapped, plateaus (e.g., /bg/) are an intermediate case, and in sonority falls (e.g., /rb/) consonant gestures are the most overlapped. I argue that long lags between gestures are the default in Georgian, and that in sonority falls the lag is shortened in order to prevent intrusive vocoids, which are more threatening to the syllable parse in onsets with initial sonorants.

The third experiment examines the relationship between morphological structure and gestural timing and finds that the presence of a morphological boundary in the onset has no effect on timing. I then present a unified account of gestural timing in Georgian onsets in which I argue that 1) Georgian does not need to use temporal coordination to distinguish onsets and codas because of phonotactic patterns, and 2) the absence of the predicted global coordination facilitates the slotting-in of consonant-only morphemes, which are common in Georgian. I also provide typological predictions based on this account and suggest a mechanism by which local timing modulations are achieved.

TABLE OF CONTENTS

Chapter 1. Introduction

1. Introduction

This dissertation investigates the organization of complex syllable onsets in Georgian. In it, I engage with two theoretical approaches: one that provides a spatial definition of the syllable, and one that offers a temporal definition. The first is based on the Sonority Sequencing Principle (SSP), and the second is the definition advanced by the coupled oscillator model (Nam and Saltzman, 2003; Nam et al., 2009). By bringing these two approaches to bear together on Georgian, a language with syllable structure that poses challenges for both theories, I aim to expand both approaches in tandem, as well as provide a comprehensive account of syllable organization in Georgian.

Syllables are critical prosodic units in spoken language. They serve as tone-bearing units in languages like Mandarin (Xu, 1999), as the constituents of prosodic feet (e.g., Hayes, 1995), and as the docking site for phrasal pitch accents in languages including English (Beckman and Pierrehumbert, 1986), Korean (Jun, 2005), and Chickasaw (Gordon, 2003). Speakers are generally very proficient at syllable-counting tasks, though speakers of a given language may not agree either on the number of syllables in a word or on the syllabic affiliation of a given segment. Language games also provide evidence that speakers are aware of and can manipulate syllables (and sub-syllabic units) (e.g., Campbell, 2020; and see Botne and Davis, 2000, for an in-depth review of this topic). This body of research shows that syllables are coherent phonological constituents both from the position of the outside analyst and the speaker. The open question, which this dissertation aims to address, is how these coherent units are organized in speech production.

I approach the question of syllable organization from two directions. The first comes from sonority and the SSP. Although the SSP does not posit a mechanism for organizing syllables in production, it does make predictions about what the syllabic affiliations of different strings will be, and we can connect these predictions to specific acoustic and articulatory properties of speech units. The second approach is that of the coupled oscillator model (Nam et al., 2009), which proposes a temporal mechanism for organizing the syllable in speech production.

I probe the systematic relationship between sonority sequencing and articulatory timing, and demonstrate how, in the case of complex onsets in Georgian, the spatial and the temporal dimension are clearly connected. This is the major theoretical and methodological contribution of this dissertation. Sonority shape has never been included as an experimental factor in work dealing with intra-syllabic gestural timing. Accounts of the motivation behind the SSP (discussed in more detail in Section 1.1) implicate gestural timing as both an organizing mechanism of the syllable and also as having a synergistic relationship with SSP-based well-formedness. By engaging with both of these frameworks in the context of a language whose phonotactics permit all sonority shapes, I probe these implications and make explicit the space-time relationship in syllabification, with two specific goals.

Goal 1 (G1): Describe how syllable onsets are organized in Georgian.

Goal 2 (G2): Demonstrate what Georgian can tell us about how space and time interact in syllable organization in general.

Ultimately, I propose a unified account of how space and time interact to produce viable syllables in Georgian regardless of consonant composition in the onset and discuss how these findings can enrich and expand our understanding of sonority sequencing, intra-syllabic timing, and the syllable as a unit in speech production.

1.1 Sonority

Sonority is best understood as an abstract, scalar property of segments which can be invoked to explain phonotactic patterns and phonological processes in many of the world's languages. Two relatively reliable physical correlates of sonority have been identified. The first is intensity: sonority and intensity are positively correlated (Parker, 2002, 2008). For a more recent account based on periodic energy, see Albert and Nicenboim (2022). Sonority can also be crudely correlated with degree of openness of the vocal tract; the more sonorous a segment, the more open the vocal tract (e.g., Parker 2002).

I employ a coarse-grained sonority hierarchy, shown in (1).

(1) vowel > glide > liquid > nasal > fricative > stop (adapted from Parker, 2011)

The sonority hierarchy is the basis for the Sonority Sequencing Principle (Hooper, 1976; Kiparsky, 1979, 1980). The Sonority Sequencing Principle (SSP) defines well-formed syllables as having local sonority minima at the edges and a local sonority maximum at the nucleus. The resulting curve is show in Figure 1.

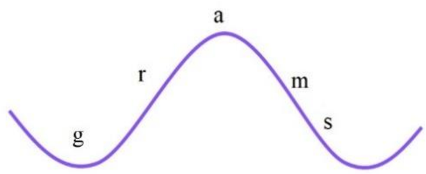


Figure 1. SSP-conforming syllable /grams/: gram-DAT, in Georgian

Because of the correlation between sonority and vocal tract openness, we can conceptualize the SSP as a spatial definition of the syllable. Well-formed syllables have a closed vocal tract at the edges and an open vocal tract at the nucleus. Although this bears some superficial similarity to the frame-content approach which equates syllables with jaw oscillations (MacNeilage, 1998), a more open vocal tract does not require a lower jaw position (e.g., Stone and Vatikiotis-Bateson, 1995), and in rapid speech the correlation between jaw oscillations and syllables weakens.

Sonority rises like the one shown in Figure 1 are the most common type of complex onsets cross-linguistically, and all languages that allow complex onsets allow sonority rises. However, exceptions to the SSP are well known. In many languages that otherwise permit only SSP-conforming syllables, /s/+ stop clusters are common. As a result, they have been the focus of much research (e.g., Engstrand and Ericsdotter, 1999; Morelli, 2003; Yavaş et al., 2008; Goad, 2011; Yavaş, 2013). These analyses are primarily focused on reconciling /s/+stop clusters with the otherwise SSP-conforming phonotactics of languages like English, by emphasizing the unique characteristics of sibilant fricatives which allow them to appear syllable-initially, or by showing that the /s/ is extrasyllabic. In other languages, exceptions to the SSP are much more widespread and cannot be dealt with via appeals to unique properties of specific consonants. Georgian is a prime example: essentially any syllable onset is permitted, both SSP-conforming and not. In order to understand how Georgian permits cross-linguistically uncommon and, from an SSP perspective, strongly dispreferred onsets such as /rg/ and /lt'/ it is important to consider proposed motivations for the SSP.

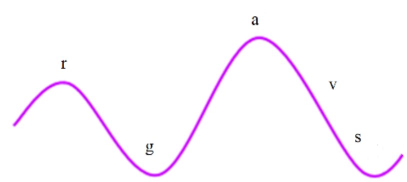


Figure 2. SSP-nonconforming syllable /rgavs/: plant(v.)-3SG.SUBJ, in Georgian

The preference for sonority rises has been attributed to both production and perception-based factors. Henke and colleagues (2012) argue that sonority sequencing is an epiphenomenon arising around cue preservation. In this account, sonority rises and /s/+stop clusters are best suited to preserving the identity of each consonant. Henke et al. also draw on studies of stop-stop clusters in Georgian (Chitoran et al., 2002) and Tsou (Wright, 1996, 1999). In these studies, longer lags between consonants are reported, and Henke and colleagues argue that this is evidence that gestural timing in non-SSP-conforming clusters is modulated to ensure consonant recoverability. Implicit in this argument is that there is a default timing pattern used in sonority rises and that it has more overlap than stop-stop clusters do, but it is not actually described.

Sonority rises also best support what Mattingly (1981) calls the "parallel transmission of information": the simultaneous conveyance of information about multiple segments. Parallel transmission, in Mattingly's account, is critical for both speech production and perception. The articulatory prerequisites for the parallel transmission of information are increasingly open constrictions followed by increasingly closed constrictions, which sounds remarkably like the well-formed syllable defined by the SSP. In this account, we can argue that sonority rises are preferable because they allow for a high degree of overlap between consonants with minimal loss of information. As with Henke et al. (2012) there is an implication that there is a default timing pattern for complex onsets and that rises are best suited to this timing. Chapter 2 systematically investigates the relationship between sonority shape and gestural overlap between consonants in order to directly addressed this implied relationship.

By focusing on Georgian, which allows onsets of all sonority shapes, I can systematically examine this implied relationship between space and time in syllable organization. Additionally, I approach my analysis through the framework of Articulatory Phonology, which includes temporal information in the phonological representation.

1.2 Articulatory Phonology

Sonority sequencing offers a spatial account of the syllable, though the temporal dimension is implicated in discussions of the motivations behind sonority sequencing. Articulatory Phonology and the coupled oscillator model of syllable structure provide an explicitly time-based definition of the syllable, though space is implicated through gestural overlap.

In Articulatory Phonology (AP), the phonological and phonetic levels are considered, respectively, the macro- and micro-level descriptions of a single self-organizing system (e.g., Browman and Goldstein, 1986, 1989, 1992; Goldstein and Fowler, 2003). The phonological primes are gestures, which are abstract, phonologically relevant events in the vocal tract. Gestures are defined using task-dynamic equations, which is the case for many motor tasks (e.g., Saltzman, 1986). Any given gesture controls a specified articulator (a full list is proposed in Browman and Goldstein, 1989) and has a constriction degree and constriction location specified. Dynamical systems definitionally change over time, so gestures are inherently spatiotemporal. The span of time over which a gesture occurs is its activation interval, and gestures are timed with respect to one another on the phonological level. We can represent these activation intervals and their temporal relationships in gestural scores, as in Figure 3.

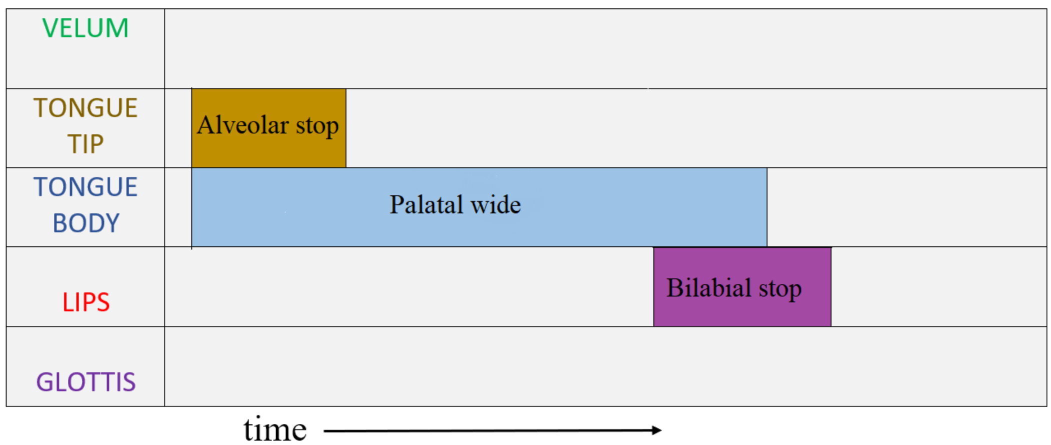


Figure 3. Gestural score for the English word 'dab'. No glottal gesture is shown because the assumed default for the glottis is to be closed for voicing.

Figure 3 shows the gestural score for the English monosyllable /dæb/. Each gesture is represented with a filled bar extending to the right of the articulator it controls. Gestures are labeled with their specified constriction location and degree in that order. The x-axis represents time. The longer the filled bar is on the horizontal dimension, the longer the activation interval of the gesture. The glottis is also active during the entire word, but no gesture is represented on the score because voicing is assumed to be the default state. There is also no velic gesture represented as a raised velum is assumed to be the default state.

The temporal relationship between gestures is modeled via dynamic coupling (Saltzman and Byrd, 2000; Nam and Saltzman, 2003). The initiation of gestures is controlled by oscillators, which are coupled together in specific ways. There are two phasing relationships that are stable enough to be used in speech: in-phase (0°) and anti-phase (180°). In-phase coupling is characterized by the simultaneous onset of gestures, and anti-phase coupling by the sequential onset of gestures. Figure 4 shows the same gestural score as in Figure 3, but with the coupling relationships marked. The solid arrow marks the in-phase relationship, and the dotted arrow marks the anti-phase relationship.

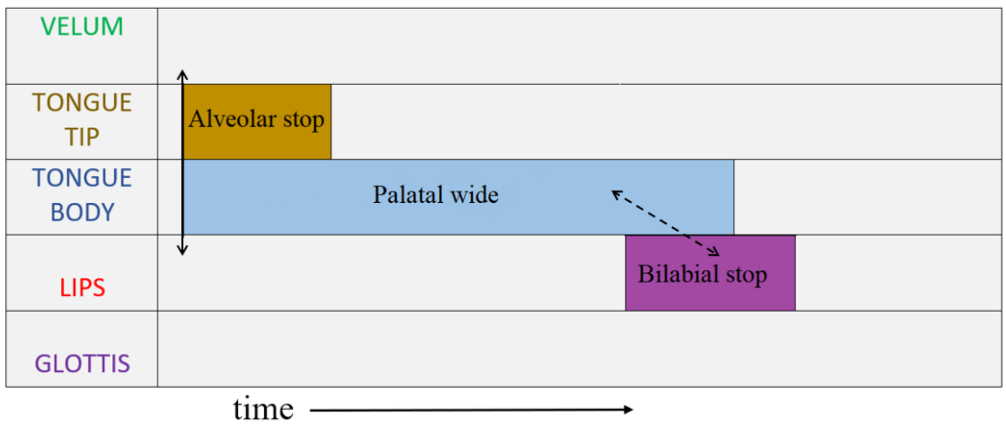


Figure 4. Gestural score for the English word 'dab' with phasing relationships marked. The solid arrow marks in-phase, and the dotted line marks anti-phase.

In the coupled oscillator model of syllable structure (Nam et al., 2009), these phasing relationships give rise to syllable structure in speech production.

1.2.1 Syllable structure in Articulatory Phonology

Early work in Articulatory Phonology identified different organizational patterns that characterized syllable onsets versus codas. Browman and Goldstein (1992) found that coda C gestures and onset C gestures are related differently to the nucleus V gestures. Codas have *local*-only timing relationships; the first coda consonant follows the V gesture, and all subsequent coda consonants follow the preceding consonant with no relationship to the vowel.

In terms of phasing relationships, coda consonant gestures are anti-phase coordinated with the preceding gesture.

Onsets are characterized by *global* timing. Browman and Goldstein (1992) is the originating paper for the c-center effect, which, in the AP framework, has been considered a major diagnostic for complex onset coordination. The c-center (short for consonant center) is a theoretical point in the onset which is equidistant from all C gestures. This c-center is in-phase with the vowel gesture, meaning that it coincides with the onset of the V gesture. In the coupled oscillator model (Nam et al., 2009) the c-center effect is the result of *competitive coupling* of gestures. Each consonant gesture is anti-phase with the preceding consonant gesture (as in codas) but is also in-phase with the vowel gesture. Figure 5 shows a coupling graph representing a complex onset, and Figure 6 shows a gestural score.

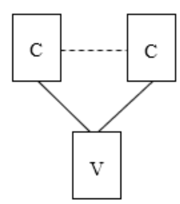


Figure 5. Coupling graph for a CCV syllable. Solid lines mark in-phase coupling and the dotted line marks anti-phase

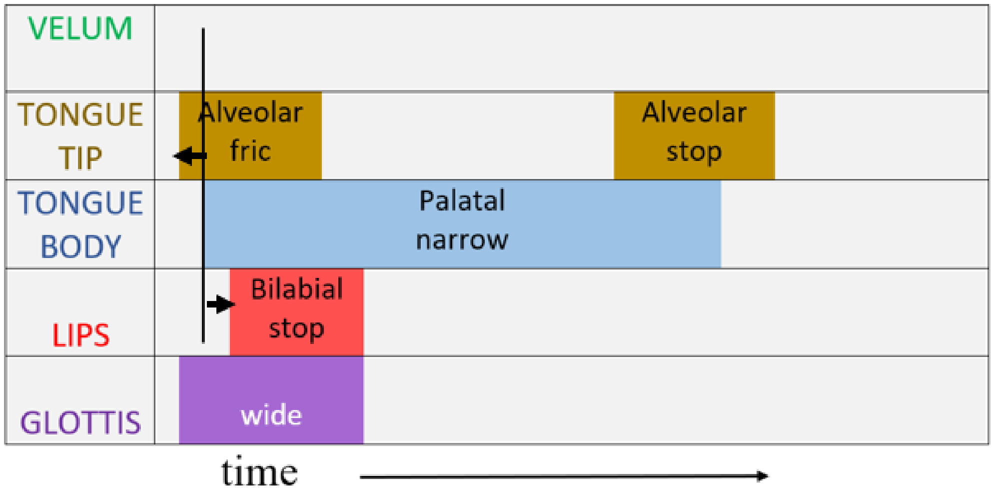


Figure 6. Gestural score for the English word 'speed' with the c-center marked by a solid vertical line. Horizontal arrows indicate that the C gesture onsets are equidistant from the c-center.

The c-center effect has been documented in a range of languages, including French (Kühnert et al., 2006), American English (Marin and Pouplier, 2010), and Italian (Hermes et al., 2013). Larger onsets are theorized to have the same underlying competitive coupling structure, but a study looking at onsets of up to seven consonants in Georgian (Hermes et al., 2020) found evidence for the c-center effect only in two- and three-consonant onsets. Languages where CCV sequences are syllabified as C.CV do *not* show the c-center effect (Goldstein et al., 2007 on Tashlhyit Berber; Shaw et al., 2011, on Moroccan Arabic).

In these studies, the presence of the c-center effect is assessed by a measure of rightward shift. Rightward shift refers to the later onset of the prenuclear consonant gesture in increasingly larger onsets. For example, the bilabial stop in Figure 6 occurs later relative to the V gesture than the alveolar stop in Figure 3 does. Marin and Pouplier (2010) also measure the stability of the c-center relative to an anchor point in the coda consonant. The c-center should remain equidistant from an anchor point, and simultaneous with the vowel gesture onset, regardless of the number of onset consonants. Chapter 3 focuses on using these metrics to assess the presence of global organization in complex onsets in Georgian, and to determine if and how the sonority shape of an onset affects this organization.

This global pattern can withstand a great deal of perturbation at the local level. Kühnert and colleagues (2006) report that different CC clusters in French show cluster-specific amounts of overlap between the two consonant gestures and display the c-center effect across the board. Recent work by Sotiropoulou and colleagues (2020) demonstrates that even in the absence of observed c-center stability or rightward shift there is still articulatory evidence for global organization in the onset. They find shortening of a lateral C2 in a CCV syllable as compared to a CV, early initiation of the V gesture in CCV as opposed to CV, and a compensatory relationship between the duration of the later C2 and the C1-C2 transition duration. Sotiropoulou and colleagues demonstrate that global organization is indexed simultaneously over many different parameters which can be measured in a variety of ways. This means that we may predict different C-C overlap patterns based on sonority without needing to assume different syllable parses, and that even if such differences cause so much perturbation that the c-center effect cannot be detected via traditional measures (e.g., rightward shift) there can still be abundant evidence of global coordination.

1.3 Georgian

Georgian is a South Caucasian language spoken by approximately four million people in the Republic of Georgia and in diaspora in Europe and North America, including in southern California. Georgian has a modest phoneme inventory by areal standards: the 5 vowels /a e i o u/ and 28 consonants, shown in Table 1. Georgian has no phonological vowel reduction, and

VV sequences are invariably disyllabic, not diphthongs or long vowels in cases of identical adjacent vowels (Shosted and Chikovani, 2006).

Table 1. Georgian consonant inventory (following Shosted and Chikovani, 2006)

	bilabial	labio-dental	dental	alveolar	post-alveolar	velar	uvular	glottal
plosive	p^h p’ b		t^h t’ d			k^h k’ g	q’	
affricate				ts ts’ dz	ʧ ʧ’ ʤ			
nasal	m			n				
tap/trill				ɾ/r				
fricative		v		s z	ʃ ʒ	x ɣ		h
lateral approximant				l				

The uvular ejective has a variably fricated release but is treated in phonological descriptions as an underlying stop. The labio-dental varies between /v/ and /w/, and there is some phonological evidence to suggest that it is underlyingly an approximant. When sequences of three approximants would arise as a result of syncope, that process is blocked, and /v~w/ can be one of those approximants. For example, from Chitoran (1999):

mts’rali -> mts’rlis

but: mtvrali -> mtvralis (where mtvrlis is expected)

For this reason, the labio-dental is excluded from all test clusters.

Lexical prominence in Georgian is weak and carries essentially no functional load (Borise, 2017). All accounts agree that the initial syllable bears stress in words of all length and some accounts posit secondary stress on the antepenult in words of four or more syllables, but evidence for secondary stress is hard to find reliably and may be inconsistent (Hewitt, 2005). All analyses in this dissertation are concerned only with word-initial syllables, so I do not expect any confounds related to lexical prominence in the results.

Many recent studies have addressed phrasal prosody (Borise, 2019; Borise and Zientarsky, 2018; Skopeteas and Féry, 2016; Vicenik and Jun, 2014). Although the results and analyses vary, there is a general agreement that prosody, syntax, and information structure are tightly linked in Georgian. Borise (2019) presents acoustic evidence that focus is not reliably marked by prominence on the stressed syllable of the focused word, but instead by prosodic grouping. I also do not expect confounds related to phrasal prominence.

1.3.1 Syllable structure in Georgian

Georgian permits syllable onsets that are typologically unusual both in their size and composition. The minimal Georgian syllable is a single V, such as either of the syllables in /i.a/ 'violet', but onsets can contain up to seven consonants, as in /gvptsxvnis/ 'they (sg) peeled us'. Virtually any sequence of consonants is permitted in the onset, including sonority plateaus such as those in /t'q'deba/ 'it breaks' and /mnatoba/ 'luminary', and sonority falls such as those in /sp'oba/ 'destroying (n)', /mdinare/ 'river' and /rʧeva/ 'advice'.

Complex codas are far less common and are almost universally the result of morphological affixation (see Butskhrikidze, 2010 for a quantitative breakdown). These all occur word finally. Intervocalic consonant sequences are, prescriptively, syllabified as simplex coda-complex onset sequences (see Harris, 2001 for a survey of descriptions of this syllabification). However, experimental data (e.g., Chitoran et al., 2002) show that speakers differ in their syllabification of VCCV sequences. Word-edge clusters, on the other hand, are unambiguous. There is no evidence that any consonants in large word-initial onsets are extrasyllabic or appendices in any way. Georgian also does not permit consonants to serve as syllable nuclei. For this reason, the experiments detailed in Chapters 2 and 3 focus only on word-initial onsets.

It is clear even from the few examples cited above that the SSP is not relevant for syllabification in Georgian. Sonority itself however, as a property of a given consonant or vowel, is relevant. Both the syncope blocking process described in section 1.4.1 and the restriction on syllable nuclei refer to sonority. Georgian does not ignore the sonority value of a given segment or 'assign' the same sonority value to every segment; it simply does not adhere to the SSP. This raises the question, naturally, of why. There are clear arguments for the advantages of sonority rises over other kinds of syllable onsets, as detailed in section 1.2. In Chapter 4, I will return to this question and argue that the timing patterns in Georgian syllables are such that sonority rises are not uniquely well-suited to being onsets, as they are in other languages.

The challenges syllabification in Georgian poses to a sonority-based account of syllable structure are clear. The diversity of sizes and sonority shapes found in onsets also raise questions for the Articulatory Phonology approach. First, the coupled oscillator model has only been applied to onsets of up to three consonants. It is not a given that larger onsets will display the same timing pattern, if only for practical reasons; with so many consonants, how do you ensure that there is still a clear acoustic vowel? Secondly, only sonority rises and the widespread SSP exception of /s/+stop clusters have been documented displaying the c-center effect because they are the only complex onsets permitted in languages that have been analyzed in this framework.

Georgian morphology also differs from other languages in which the temporal structure of syllables has been studied. Affixation occurs in both nominal and verbal morphology and can be both inflectional and derivational. There are no cases where affixation does not occur; any syllable in the right position could be expanded. Here I use inflectional verbal morphology as an example, both because it is extremely common in everyday speech and because it provides excellent examples of the kind of consonant-only morphemes. Georgian verbs obligatorily mark all core arguments using both prefixes and suffixes, many of which are of the form C or CC. As a result, the same lexical item in Georgian can have syllable onsets of different sizes depending on the core arguments, as well as on tense-aspect-mood variables. Tables 2 and 3 show the two sets of pronominal affixes and inflected examples for each set of affixes. These inflected forms are in the present tense.

Table 2. 'Subject' markers (Set A) and inflected example 'X writes'

	Singular	**Plural**			
1	v-	v-…-t		v-ts'er	v-ts'er-t
2	Ø(/x)-	Ø(/x)-…-t		ts'er	ts'er-t
3	-s/a/o	-en/an/es/nen		ts'er-s	ts'er-en

Table 3. 'Object' markers (Set B) and inflected example 'he kills X'

	Singular	**Plural**			
1	m-	gv-		m-k'lav-s	gv-k'lav-s
2	g-	g-…-t		g-k'lav-s	g-k'lav-t
3	h/s/Ø-	h/s/Ø-…-(t)		k'lav-s	k'lav-s

One experiment, discussed in Chapter 3, tests whether or not the presence of a morpheme boundary in the onset affects timing. This is an important point to establish, as I aim to provide an account of syllabification in Georgian that can apply to all syllables.

1.4 Organization of the dissertation and research questions

Chapters 2 and 3 each report the results of Electromagnetic Articulography (EMA) experiments that investigate the relationship between the sonority shape of an onset and gestural timing. EMA was selected because EMA data has the fine-grained temporal resolution necessary to answer questions about speech timing. Chapter 2 focuses on local timing relationships between consonant gestures in the onset, and addresses the following questions:

Research Question 1 (RQ1): How are consonant gestures in a complex onset in Georgian timed with respect to one another?

Research Question 2 (RQ2): Since combinatorial possibilities in Georgian onsets are not constrained by the sonority hierarchy, does the timing of the consonantal gestures systematically reflect sonority sequencing?

I find a significant effect of sonority shape on timing, where sonority falls (e.g., /rg/) are the most overlapped sonority shape. This is entirely unexpected given what we know about sonority rises and the parallel transmission of information, and Chapter 2 addresses possible motivations for this timing pattern.

Chapter 3 focuses on global timing in the onset and has the following research questions:

Research question 3 (RQ3): What coordinative pattern is found in onsets in Georgian?

Research question 4 (RQ4): What is the relationship between the sonority shape of a complex onset in Georgian and the global organization of the constituent gestures?

Research question 5 (RQ5): What is the relationship between the morphological composition of an onset in Georgian and the global organization of the constituent gestures?

As with Chapter 2, the answers to these questions are largely unexpected. Neither morphological structure nor sonority shape affect global timing in the onset, most likely because I find no evidence of any global coordination at all.

In Chapter 4 I synthesize the results of all three reported experiments and show how the local and global timing patterns in Georgian make sense when considered together and in light of the phonotactic and morphological structure of Georgian. The proposed relationship between syllable organization and morphological structure makes typological predictions, which I detail as a line of further research. I then discuss how this unified account of syllable organization in Georgian can point towards a more comprehensive definition of the syllable as a unit in speech production, with reference to future work that will expand on my proposal.

References

Albert, A., & Nicenboim, B. (2022). Modeling sonority in terms of pitch intelligibility with the Nucleus Attraction Principle. *Cognitive Science*, *46*(7), DOI: 10.1111/cogs.13161.

Beckman, M. E., & Pierrehumbert, J. B. (1986). Intonational structure in Japanese and English. Phonology, 3, 255-309.

Borise, L. (2017). Prosody of focus in a language with a fixed focus position: Evidence from Georgian. In *Proceedings of WCCFL* (Vol. 34, pp. 89-96).

Borise, L. (2019). *Phrasing is key: the syntax and prosody of focus in Georgian*. PhD dissertation, Harvard University.

Borise, L., & Zientarski, X. (2018). Word stress and phrase accent in Georgian. *Proceedings of TAL*, *6*, 2018-42.

Botne, R., & Davis, S. (2000). Language games, segment imposition, and the syllable. *Studies in Language. International Journal sponsored by the Foundation "Foundations of Language"*, *24*(2), 319-344.

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology*, *3*, 219-252.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201-251.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180.

Butskhrikidze, M. (2010). The nature of consonant sequences in Modern Georgian. In J. Wohlgemuth & M. Cysouw (Ed.), Rara & Rarissima: Documenting the Fringes of

Linguistic Diversity (pp. 23-46). Berlin, New York: De Gruyter Mouton. https://doi.org/10.1515/9783110228557.23

Campbell, E. W. (2020). Probing phonological structure in play language: Speaking backwards in Zenzontepec Chatino. *Phonological Data and Analysis*, *2*(1), 1-21.

Chitoran, I. (1999). Accounting for sonority violations: the case of Georgian consonant sequencing. In *Proceedings of the 14th International Congress of Phonetic Sciences,* 101-104.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. *Papers in laboratory phonology*, *7*(4-1), 419-447.

Engstrand, O., & Ericsdotter, C. (1999). Explaining a violation of the sonority hierarchy: Stop place perception in adjacent [s]. In *Proceedings from the XIIth Swedish Phonetics Conference*, 49-52.

Goad, H. (2011). The representation of sC clusters. In *The Blackwell companion to phonology*, 1-26.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007, August). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In *Proceedings of the XVIth international congress of phonetic sciences*, 241-244.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In Niels Schiller, Niels Olaf Schiller, and Antje Meyer (Eds.). *Phonetics and phonology in language comprehension and production: Differences and similarities*, 159-207.

Gordon, M. (2003). The phonology of pitch accents in Chickasaw. *Phonology*, *20*(2), 173-218.

Hayes, B. (1995). Metrical stress theory: Principles and case studies. University of Chicago Press.

Henke, E., Kaisse, E. M., & Wright, R. (2012). Is the sonority sequencing principle an epiphenomenon. *The sonority controversy*, *18*, 65-100

Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology*, *30*(1), 1-25.

Hewitt, G. (2005). *Georgian: A learner's grammar*. Routledge.

Hooper, Joan B. (1976). An introduction to natural generative phonology. New York: Academic Press.

Jun, S. A. (2005). Korean intonational phonology and prosodic transcription. *Prosodic typology: The phonology of intonation and phrasing*, *1*, 201.

Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic inquiry*, *10*(3), 421-441.

Kiparsky, P. (1980). Remarks on the metrical structure of the syllable. *Phonologica*, *3*.

Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and C-center in selected French consonant clusters. In *7th International Seminar on Speech Production (ISSP)*.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behavioral and brain sciences*, *21*(4), 499-511.

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*(3), 380-407.

Mattingly, I. G. (1981). Phonetic representation and speech synthesis by rule. In *Advances in Psychology* (Vol. 7, pp. 415-420). North-Holland.

Morelli, F. (2003). The Relative Harmony of/s+ Stop/Onsets Obstruent Cluslers and the Sonority Sequencing. In Caroline Féry and Reuben van de Vijver (Eds.). *The syllable in optimality theory*, 356-371.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. In Proceedings of the 15th international congress of phonetic sciences, 2253-2256.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. Approaches to phonological complexity, 16, 299-328.

Parker, S. (2002). *Quantifying the sonority hierarchy*. PhD dissertation. University of Massachusetts Amherst.

Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of phonetics*, *36*(1), 55-90.

Parker, S. (2011). Sonority. In *The Blackwell companion to phonology*, 1-25.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators: A preliminary model. *Experimental brain research series*, *15*, 129-144.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499-526.

Shaw, J. A., Gafos, A. I., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure: a case study of Moroccan Arabic consonant clusters. *Phonology*, *28*(3), 455-490.

Shosted, R. K., & Chikovani, V. (2006). Standard Georgian. *Journal of the International Phonetic Association*, *36*(2), 255-264.

Skopeteas, S., & Féry, C. (2016). Focus and intonation in Georgian: Constituent structure and prosodic realization. *Ms.*

Sotiropoulou, S., Gibson, M., & Gafos, A. (2020). Global organization in Spanish onsets. *Journal of Phonetics*, *82*, 100995.

Stone, M., & Vatikiotis-Bateson, E. (1995). Trade-offs in tongue, jaw, and palate contributions to speech production. *Journal of Phonetics*, *23*(1-2), 81-100.

Vicenik, C., & Jun, S. A. (2014). An autosegmental-metrical analysis of Georgian intonation. *Prosodic typology II: The phonology of intonation and phrasing*, 154-186.

Wright, R. A. (1996). *Consonant clusters and cue preservation in Tsou*. PhD dissertation. University of California, Los Angeles.

Wright, R. (1999). Tsou consonant clusters and auditory cue preservation. In *Selected papers from the eighth international conference on Austronesian linguistics* (Vol. 1, pp. 227-312).

Xu, Y. (1999). Effects of tone and focus on the formation and alignment of f0 contours. Journal of phonetics, 27(1), 55-105.

Yavaş, M., Ben-David, A., Gerrits, E., Kristoffersen, K. E., & Simonsen, H. G. (2008). Sonority and cross-linguistic acquisition of initial s-clusters. *Clinical Linguistics & Phonetics*, *22*(6), 421-441.

Yavaş, M. (2013). What explains the reductions in/s/-clusters: Sonority or [continuant]?. *Clinical linguistics & phonetics*, *27*(6-7), 394-403.

Chapter 2. Sonority sequencing and its relationship to articulatory timing in Georgian

1. Introduction

The syllable is an important unit in utterance organization and speech planning. Syllables are psychologically salient units into which sequences of individual segments are organized (e.g., Selkirk, 1982). They are also the base units for higher levels of prosodic organization such as feet and phonological phrases, and are generally the docking sites for lexical stress, tone, and pitch-accent (Ladd, 2008; Yip, 2002). The work presented here aims to further our understanding of how segments are organized into syllables in production. In doing so, we engage with two theoretical approaches to the syllable; the first is sonority-based (e.g., Clements, 1990) and the second relates to Articulatory Phonology (Browman and Goldstein, 1986). Although the definitions of the syllable advanced by these two theories can capture the syllable structure of many of the world's languages, as shown below, there are data that challenge both. We propose that these two approaches to the syllable should be taken together: we expect that sonority sequencing and gestural timing interact systematically with one another as part of the mechanism of grammar that determines syllable organization. We investigate this proposal through production experiments using electromagnetic articulography (EMA) in Georgian. Our focus is on complex syllable onsets, which have more combinatorial possibilities and can be longer than codas, both within Georgian and across languages generally. We selected Georgian as our test language because this language poses a challenge to both definitions of the syllable. Georgian permits essentially any combination of consonants in the onset, including sonority falls with left-edge approximants, thus amply defying sonority-based principles of syllable organization. From an Articulatory Phonology perspective, syllabification in Georgian is also difficult to capture. Georgian onsets can have up to seven consonants, and previous articulatory research has been unable to offer a unified account based on timing (Goldstein et al., 2007; Hermes et al., 2020): the limited data on this issue provide evidence that the timing pattern predicted of complex onsets holds at most for onsets of up to three consonants. Here, we systematically examine the relationship between sonority sequencing and gestural timing in two-consonant onsets in Georgian, and we show that the two relate. On the basis of these findings, we offer an account of how syllabification in Georgian works, that extends to all onset types and lengths. Ultimately, we provide a novel proposal of how the two approaches to the syllable – i.e., the sonority sequencing approach and the gestural timing approach—can be expanded in tandem with one another to capture syllable organization despite cases of apparent lack of conformity to one or the other.

The organization of the paper is as follows: in the rest of Section 1, we provide the relevant background information on sonority-based approaches to the syllable, Articulatory Phonology (AP) and its Task-Dynamic modeling approach, and on the Georgian language; in Section 2, we present details on the experimental design and the methodologies used for data collection and analysis; in Section 3 we present the stimuli and results of our experiment; and in Section 4, we discuss how the results can inform our understanding of the syllable, and the relationship between sonority and articulation. Section 5 concludes the paper and restates our key findings. Below is a concise statement of our research questions and hypotheses.

Research questions:

1) How are consonants in a complex onset in Georgian timed with respect to one another?

2) Since combinatorial possibilities in Georgian onsets are not constrained by the sonority hierarchy, does perhaps the timing of the consonantal gestures systematically reflect sonority sequencing?

Our broad proposal is one initially set forth in Chitoran (2016): that the sonority hierarchy and deviations from it can best be understood through articulatory timing. Our primary hypothesis (H1) predicts that we will see a hierarchical relationship between sonority shape and gestural timing, measured here on two dimensions. The first dimension is the actual amount of temporal overlap between the constriction gestures making up the syllable's onset, and the second dimension is how variable the observed temporal overlap is. Sonority rises are expected to be the most overlapped because they are well-suited for parallel transmission of information (e.g., Mattingly, 1981), a notion which we expand on in Section 1.1. For the same reason, we expect rises to also be the least variable shape. The main point of this argument is that less variability in such measures would indicate a more stable timing relationship between the constriction gestures in question. We predict the same direction of the hierarchical effect for both overlap and stability for the reasons noted above, but it is entirely possible for these two dimensions to interact differently with sonority because they are not dependent on one another. As an alternate hypothesis (H2), we consider a solely perceptual basis for overlap patterns. As we will argue in Section 1.1, these patterns will result in some significant differences between sonority categories, but their prediction is not motivated by sonority *per se* in the same way that the first hypothesis predicts. Instead, we base these predictions on previous research on Georgian (Chitoran et al., 2002), where patterns of overlap in different stop-stop clusters show clear perceptual motivations.

> **Hypothesis 1 (referred to as H1):** Both the temporal overlap of consonantal gestures in complex onsets and the variability of that overlap depend on a hierarchical effect of increasingly open constrictions (cf. Mattingly, 1981). Sonority rises (e.g., /br/ or /gl/) will be more overlapped than plateaus (e.g., /mn/ or /p^ht^h/), which will be, in turn, more overlapped than falls (e.g., /rb/ or /md/). Rises will also be the least variable sonority shape, while falls will be the most variable, and plateaus will again be an intermediate case.
>
> **Hypothesis 2 (referred to as H2):** Timing of consonantal gestures in complex onsets depends on perceptual recoverability, not necessarily related to sonority shape (cf. Chitoran et al., 2002). Plateaus (e.g., /mn/ or /p^ht^h/) will be the least overlapped sonority shape because the plateau sonority category is the only one in which we find onsets where the first consonant needs an audible release, and yet is not followed by a more sonorous segment, one with a more open constriction. Sonority rises (e.g., /br/ or /gl/) and sonority falls (e.g., /rb/ or /md/) will be more overlapped than sonority plateaus, but will not be significantly different from one another.

The overarching goal of this project is to better understand the articulatory patterns of complex onsets in Georgian and how such a wide range of sonority shapes are found throughout the lexicon. This also provides an important contribution to our understanding both of why sonority sequencing principles are able to explain so many phonotactic generalizations cross-linguistically, and, at the same time, why and how some languages are able to essentially disregard sonority as a factor for permissible complex onsets.

1.1 Sonority

Sonority is an abstract property of speech sounds that can be invoked to explain a wide variety of phonotactic patterns and phonological processes across languages. Speech sounds can be organized in a hierarchy based on their sonority values. The most widely accepted sonority hierarchy is shown below (adapted from Parker, 2011).

vowel > glide > liquid > nasal > fricative > stop

Sonority is often used to explain patterns of syllable structures across the world's languages, through the Sonority Sequencing Principle (Hooper, 1976; Kiparsky, 1979, 1980). The Sonority Sequencing Principle (SSP) states that each syllable has a sonority peak and is preceded and followed by sequences increasing and decreasing in sonority, respectively, with local sonority minima at syllable edges. Following the SSP, two other major sonority-based principles of syllable organization have been proposed: Minimum Sonority Distance (Steriade, 1982) that states that languages may impose a minimum difference in sonority values, assumed to be language specific, between adjacent segments; and the Dispersion Principle (Clements, 1990) which states that, all else being equal, languages will maximize sonority differences in onsets and minimize them in codas.

Any language that permits complex syllable onsets will allow the kind of complex onset illustrated in Figure 1, where each successive consonant in the onset is more sonorous than the one preceding it. But many languages that permit tautosyllabic CC(...) sequences also inevitably allow some non-SSP-conforming syllable shapes. Figure 2 illustrates a non-SSP-conforming syllable from Georgian.

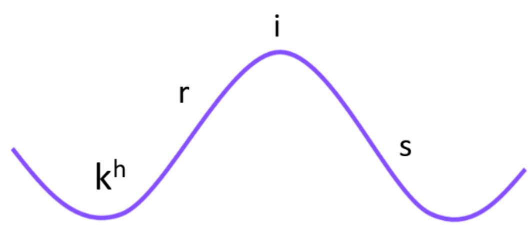


Figure 1. The Georgian word /kʰris/ 'it blows' (blow-3SG.SUBJ) follows the SSP

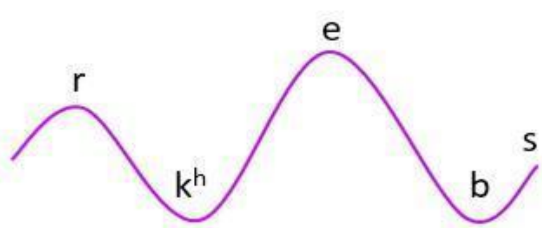


Figure 2. The Georgian word /rkʰebs/ 'antlers' (antler-PL-DAT) does not follow the SSP

The most common non-SSP-conforming cluster across the world's languages is the /s/-stop cluster, which has received significant attention in the literature (e.g., Engstrand and Ericsdotter, 1999; Morelli, 2003; Yavaş et al., 2008; Goad, 2011; Yavaş, 2013) and in many languages is the only tautosyllabic sonority 'reversal' permitted. Phonological analyses of /s/+stop clusters have therefore been focused on reconciling their presence within the otherwise SSP-conforming phonotactics of languages like English. This body of work explains the exceptionality of /s/-stop clusters and focuses primarily either on the unique properties of sibilants which allow them to appear at the syllable left-edge or on analyses that treat left-edge sibilants as extrasyllabic. However, the focus of the current study is not on individual exceptional clusters, but on an entire system in which essentially all clusters are permitted. To this end, i.e., to understand how Georgian and other languages permit syllable onsets with such cross-linguistically dispreferred shapes (e.g., /rg/ or /lb/), we consider accounts of what drives the general cross-linguistic preference for sonority rises (e.g., /gr/ or /bl/).

Henke et al. (2012) argue that the SSP and related principles are epiphenomena that emerge from purely perceptual concerns. Specifically, /s/-stop clusters and those sequences licensed by the SSP are those where the cues to each consonant's identity are most recoverable. Their account also considers Georgian and other languages with pervasive SSP violations, building on prior proposals that in these languages, gestural timing is modulated in ways that ensures consonant recoverability in non-SSP-conforming clusters (Chitoran et al., 2002 on Georgian; Wright, 1996, 1999 on Tsou). We expand upon this argument in our alternate hypothesis (H2), which assumes that gestural timing is motivated purely by concerns about consonant recoverability. This hypothesis, which as a reminder is not our primary hypothesis, proposes that the decreased overlap between stop-stop clusters observed in Chitoran et al. (2002) in Georgian will be found in sonority plateaus overall. Fricative-fricative clusters will be less overlapped to prevent coarticulation that would lessen the distinction between fricatives, and nasal-nasal clusters will also require some degree of audible release for recoverability of C1, especially because word-initial /m/ in Georgian is often devoiced (Harris, 2002). Sonority falls will behave like sonority rises, since in CCV sonority falls, C1 is not an oral stop, and therefore has at least some internal cues to segment identity. C2 is also easily recoverable, since it is released into the vowel, allowing for formant transitions to cue the consonant's identity.

Our primary hypothesis (H1) is, however, based on the notion of parallel transmission, another possible motivation for the prevalence of SSP-conforming onsets cross-linguistically. Parallel transmission—the simultaneous conveyance of information about multiple segments—is introduced by Mattingly (1981) as a crucial organizing element of speech production and perception. Mattingly (1981, p. 481) describes the articulatory prerequisites for parallel information transmission as a sequence of increasingly open constrictions followed by one of increasingly closed constrictions. This, as Mattingly notes, is strikingly similar to sonority sequencing. Sonority rises are those shapes which best allow for parallel transmission of information because a high degree of gestural overlap does not obscure cues to segment identity. This differs crucially from the perception-only account of Henke et al. (2012) in that the preference for rises is not solely because they best preserve consonant identity, but because they allow for that preservation *while* also allowing for a high degree of overlap or coarticulation between adjacent consonants. In short, the parallel transmission of information account places equal weight on production and perception, as opposed to claiming purely perceptual motivations for sonority sequencing phenomena.

In both of these accounts gestural timing is invoked to explain the cross-linguistic preference for sonority rises in complex syllable onsets. However, the relationship between sonority and gestural timing has not been directly examined previously. Here, we make a first step towards probing this relationship. More specifically, we designed an EMA experiment to measure the amount of gestural overlap as a function of sonority shape, following the assumptions of Articulatory Phonology. In the next section we discuss Articulatory Phonology, a phonological framework in which gestures are the phonological primes and temporal information is part of the phonological representation.

1.2 The role of articulatory timing in syllable organization

1.2.1 The syllable in Articulatory Phonology

In Articulatory Phonology (AP), the syllable is defined on the basis of temporal organization (Browman and Goldstein, 1988). Articulatory Phonology considers the cognitive-linguistic (or phonological) dimension and the biomechanical (the phonetic) one as the macroscopic and the microscopic description respectively of the same complex self-organizing system (e.g., Browman and Goldstein, 1986, 1989, 1992; Goldstein and Fowler, 2003). In AP, gestures, i.e., discrete movements of the speech articulators forming and releasing phonologically relevant constrictions in the vocal tract, are the phonological primitives. These gestures are abstract units that can be defined in a tract-variable space using a set of task-dynamic equations in the same way as many other skilled motor tasks can be defined (Saltzman, 1986; Saltzman and Kelso, 1983). The goals in these tract-variable spaces are achieved through synergies of different articulators. Different speakers may tune these synergies differently to achieve the same goal, or may employ different compensatory mechanisms to adapt these synergies when external perturbations are introduced (e.g., Kelso and Tuller, 1984; Shaiman, 1989; Ito et al., 2000; Golfinopolous et al., 2011). Each articulator has a set of dimensions associated with it, such as constriction location and constriction degree for the tongue tip, tongue body and tongue dorsum articulators, among others (see Browman and Goldstein, 1989, p. 210 for the full proposed set of articulators and their dimensions).

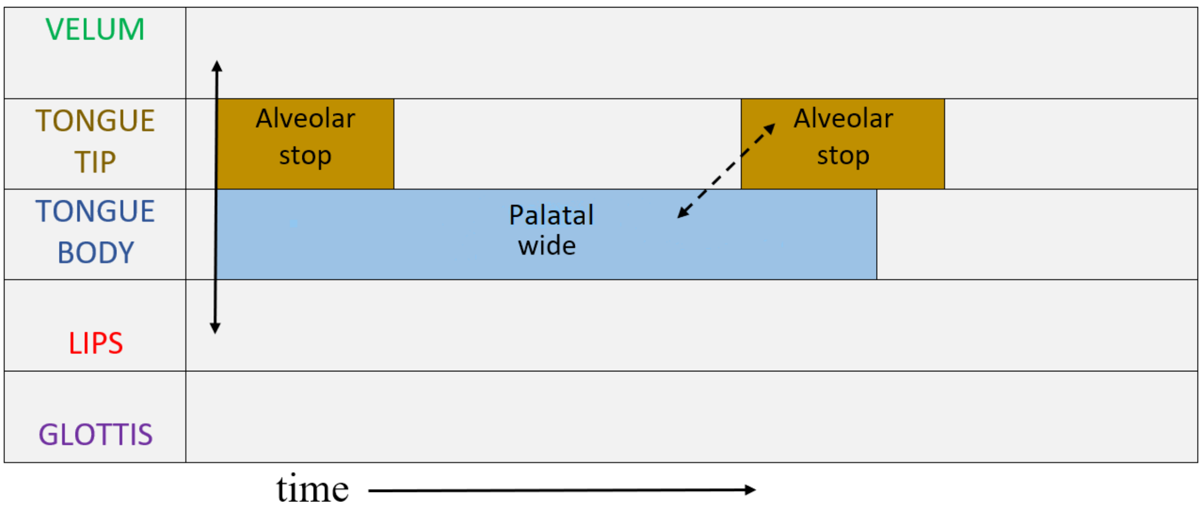


Figure 3. Gestural score for the word 'dad' with phasing relationships marked. Solid line marks in-phase coordination and the dotted line marks anti-phase coordination.

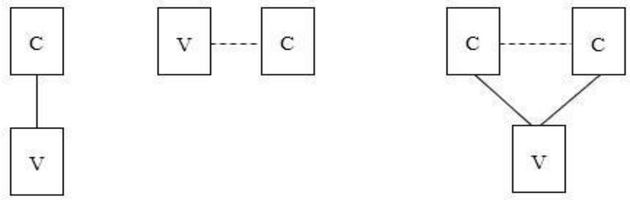
Gestures have not only spatial but also temporal properties, as dynamical systems by definition change over time. Specifically, they are active for a given temporal interval, called activation interval, and are timed with respect to each other. Figure 3 shows a gestural score for the English word ‘dad.’ The activation interval for each gesture is marked with a filled bar on the row for the relevant articulator and is labeled with the constriction location (upper line) and degree (lower line). For example, the tongue tip is making an alveolar (location) stop (degree) early in the utterance for the onset consonant of the word dad, and another such constriction later in the utterance for the coda /d/. The x-axis represents time; the longer the bar for the activation interval, the longer the duration of the gesture.

Later work in AP advanced a dynamic coupling model (e.g., Saltzman and Byrd, 2000; Nam and Saltzman, 2003) to explain inter-gestural timing patterns. In this model, timing of gestures is controlled by oscillators, which are coupled to one another in specific ways. There are two types of "coupling targets" (Nam et al., 2009, p. 4) found in these kinds of models: the first is an in-phase coupling, and the second is anti-phase. Activation intervals for gestures that are in-phase are initiated simultaneously. Oscillators in anti-phase coordination are linked in time, but in this case the activation interval for the second gesture is initiated after the first gesture has reached its target. In Figure 3, the solid line marks in-phase coordination, and the dotted line marks anti-phase coordination. Section 1.3 expands on how these coupling relationships account for syllabic structure.

1.2.2 Syllable affiliation and c-centers

Browman and Goldstein (1992) investigated the mechanism by which C(C)(C) sequences with different possible syllable affiliations are differentiated. They found different patterns of organization for the onset and the coda C gestures with respect to the V gesture in the nucleus. Coda C gestures are organized using a local metric: the first coda C gesture is anti-phase coordinated with the V gesture, and each following C gesture is anti-phase coordinated with the preceding C gesture. Onset consonants, on the other hand, are organized globally, and this observed pattern was termed the c-center organization (for "consonant center"). In onsets, all C gestures are in phase with the V gesture but antiphase with each other. The conflicting phasing results in all onset C gestures being temporally equidistant from the mean value of the midpoints of all of the gestural plateaus of the onset consonants (Browman and Goldstein, 1999, p. 144). This midpoint is called the c-center. The shift in timing seen between single C gestures (with respect to the V gesture) and the respective C gestures in onsets consisting of multiple consonants (with respect to the V gesture) is called the c-center effect or c-centering (compare the Tongue Tip gesture for onset /d/ in Figure 3 to the Tongue Tip gesture for onset /s/ in Figure 5).

The coupling model introduced by Nam and Saltzman (2003) accounts for the c-center using competitive coupling graphs. In these competitive coupling graphs, each individual oscillator representing an onset consonant has an in-phase coupling relationship with the vowel oscillator as well as an antiphase coupling relationship to the other consonant oscillators. The competition between these two coupling modes results in the rightward movement of the c-center effect as described in Browman and Goldstein (1992). Figures 4a-c show the oscillator coupling graphs for a simplex onset (4a), a simplex coda (4b), and a complex onset (4c); the gestural score for the word ‘speed’ (Figure 5) is the result of the coupling pattern shown in 4c; the vertical line indicates the c-center. Although the originating paper defines the c-center as the mean of gestural constriction midpoints, the coupled oscillator model makes predictions about gestural onsets specifically, so the c-center can also be calculated as the mean of all gestural onsets. This is how we calculate the c-center, as also illustrated in Figure 5.



Figures 4a, 4b, and 4c. Oscillator coupling graphs

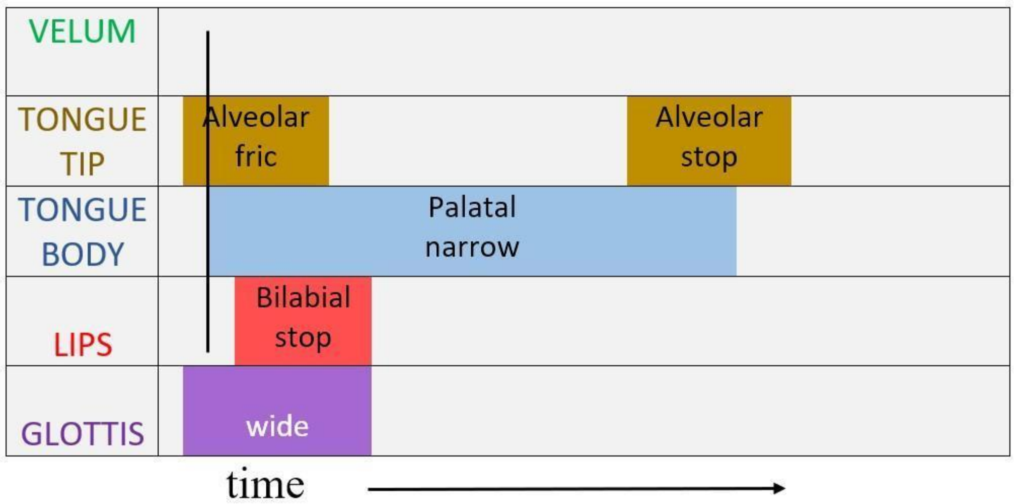


Figure 5. Gestural score for the word 'speed'. Vertical line marks the c-center. The c-center is equidistant from the onset (shown as the beginning of activation interval) of the Tongue Tip gesture (constriction location: alveolar; constriction degree: fricative) for /s/ and the onset of the Lips gesture (constriction location: bilabial; constriction degree: stop) for /p/.

The c-center effect has been investigated in several languages with different syllable structures and the results have confirmed the presence of the effect and the coupling model hypothesis (e.g., Hermes et al., 2013 on certain clusters in Italian; Kühnert et al., 2006 on French). Research looking at timing within syllables has also confirmed previous analyses of syllable structure in certain languages, such as Imdlawn Tashlhiyt Berber. To illustrate, Tashlhiyt Berber does not permit complex onsets, and indeed, no c-center is found in CCV or CCCV structures (Goldstein et al., 2007; Hermes et al., 2017). Instead, the data support the C.CV or

CC.CV parses proposed in non-AP based analyses of Tashlhiyt syllabification (Dell and Elmedlaoui, 1985, 2012). Shaw et al. (2011) show the same for Moroccan Arabic, which has been claimed to disallow complex onsets (Dell and Elmedlaoui, 2002).

Since there is no evidence to support a multisyllabic parse of CC(…)V sequences in Georgian, the assumption under this model is that Georgian will display the c-center effect in complex onsets. There is some evidence in support of this (Goldstein et al., 2007; Hermes et al., 2020) although neither study is conclusive. In the design and analysis of the experiment we discuss here, we are operating under the assumption that these sequences are tautosyllabic onsets and that therefore, the consonant gestures we analyze are temporally coordinated with one another and with the following vowel gesture to give rise to a syllable as defined in AP. We predict, however, in both hypotheses, that consonant gestures in different onsets will have different patterns of overlap depending on the general motivations for intrasyllabic gestural timing. There is a substantial body of research showing that gestural overlap and other phonetic indices can vary depending on a variety of factors without disrupting the phonological syllabic structure underpinning the coordination of the gestures in question. In the next section, we review the literature on this topic.

1.2.3 Gestural overlap

Syllable onsets of different composition can show significant differences in gestural overlap while still sharing the same phonological structure. Kühnert et al. (2006) show that different consonant sequences in French differ with regard to gestural overlap but show a stable c-center effect. This indicates that global timing pattens (i.e., the c-center effect) can be consistent within a language despite major differences in local timing patterns (e.g., consonant gesture overlap). Sotiropoulou et al. (2020 on Central Peninsular Spanish, 2022 on German) show that global reorganization occurs in CV versus CCV even when local changes result in so much perturbation that static indices of phonological structure such as a stable interval between onset and vowel or rightward shift cannot be found.

Overlap between consonants is both language-specific (Pouplier, 2012; Bombien and Hoole, 2013) and determined by a variety of factors (Bombien and Hoole, 2013; Bombien et al., 2013; Pouplier et al., 2022). Order of place of articulation affects degree of overlap; research on Georgian (Chitoran et al., 2002), French (Kühnert et al., 2006) and Korean (Son et al., 2007) has shown that back-to-front clusters are less overlapped than front-to-back clusters. Kochetov et al. (2007) compared degree of consonant overlap in VCCV structures in Korean and Russian to test the hypothesis that languages with more or stronger assimilation patterns will show more overlap. Using plateau lag, the time between the release of the first C gesture (C1) and the achievement of the target for the second C gesture (C2), as a measure of overlap, they showed that in Russian C1 is released prior to the achievement of the C2 target for all sequences tested, but that Korean showed both positive and negative lag depending on the composition of the consonant sequence. In all cases except /kt/ in Russian, more lag (and therefore less overlap) was found for back-to-front sequences.

In general, the identity of each consonant in the cluster could affect overlap, in some part due to different degrees of coarticulatory resistance for different consonants. Pastätter and Pouplier (2014) find that sibilant consonants overlap less with adjacent vowels than other consonants do, which has implications for the c-center effect in these sequences, and propose that coarticulation resistance plays a role in timing. Because of the sibilants' higher resistance to coarticulation (Recasens et al., 1997; Recasens, 2012), they resist the rightward movement associated with the c-center effect that normally results in a higher degree of overlap with the nucleus. By resisting this movement, the timing patterns normally associated with the c-center effect are disrupted, although they still play a role.

Hoole et al. (2013) find that obstruent-rhotic sequences show less overlap than obstruent-lateral sequences, and that nucleic consonants show different overlap patterns, measured from the offsets and onsets of gestural targets, with the rightmost onset consonant than vocalic nuclei do. Work on German has found that stop-nasal clusters and stop-liquid clusters display different patterns of overlap and are differently affected by changes in voicing status of the obstruent member of the cluster (Bombien and Hoole, 2013). Further work on German reveals a hierarchy of overlap in different clusters that seems to rely on a variety of factors including manner and order of place of articulation; /kl/ has the most overlap, followed by /pl/ > /ps/ > /ks/ > /kn/ (Bombien et al., 2013).

The present study aims to determine if sonority sequencing and gestural overlap are systematically related. Since in Georgian onsets of all sonority shapes share the same underlying syllabic structure, we predict that differences in local timing, and specifically in gestural overlap, are the mechanism by which Georgian allows onsets of any sonority shape. We also include order of place of articulation as an independent variable, since previous research has shown that this factor does affect overlap in Georgian (Chitoran et al, 2002). In Section 3, we report on two measures of overlap. W use two measures in order to make this work comparable to a wider range of research on other languages, and to understand how different factors may influence different types of overlap between gestures.

1.3 Georgian

The following section presents a few relevant facts about Georgian phonology. We selected Georgian for this study because the language permits in principle any combination of up to seven consonants as a syllable onset. Syllable codas, even simplex ones, are uncommon, and the vast majority of complex codas are multimorphemic (Butskhrikidze, 2002). Many complex onsets are also multimorphemic, as Georgian verbal morphology includes a variety of consonant only prefixes.

Georgian has a typical five vowel inventory: /a e i o u/. It is important to note that there is no phonemic schwa in Georgian, nor are there any phonological processes that would reduce underlying phonemic vowels to schwa (e.g., Aronson 1997). The consonant inventory is presented in Table 1.

Table 1. Georgian consonant inventory (following Shosted and Chikovani, 2006)

	bilabial	labio-dental	dental	alveolar	post-alveolar	velar	uvular	glottal
plosive	pʰ p’ b		tʰ t’ d			kʰ k’ g	q’	
affricate				ts ts’ dz	ʧ ʧ’ ʤ			
nasal	m			n				
tap/trill				ɾ/r				
fricative		v		s z	ʃ ʒ	x ɣ		h
lateral approximant				l				

Georgian ejectives have a short lag between the release of the oral closure and the release of the glottal closure (Vicenik, 2010), and the uvular ejective is often fricated to some degree (Shosted and Chikovani, 2006). The labio-dental fricative /v/ has a wide range of realizations which can be more or less approximant-like. For this reason, it has been excluded from test sequences to avoid possible variation in sonority shape both across and within speakers.

Word-level stress is found on the initial syllable. Vicenik et al. (2014) report a weak increase in F0 on the initial syllable; Borise and Zentarski (2018) report lengthening on the initial syllable but no evidence of a pitch target. Since stress is both fixed and weakly realized, we do not expect any confounds related to lexical prominence.

2. Methods

2.1 Participants and procedure

In order to answer our questions regarding gestural overlap and sonority, we collected kinematic data from three native speakers of Georgian, two female (F1 and F2) and one male (M1).[1] All three were in their twenties and living in southern California at the time of their participation. They were recruited via word of mouth and an announcement circulated by the Georgian Cultural and Educational Center of Southern California. Data were collected using an Electromagnetic Articulograph AG501 (Carstens Medizinelektronik GmbH). Electromagnetic articulograph (EMA) data allows us to track the movement of the tongue and

[1] Data collection was abruptly interrupted due to the Covid-19 pandemic.

lips at a millisecond level of temporal resolution. This tracking is done by affixing sensors to key points in the vocal tract. The sensors' movement is recorded as they move through a weak electromagnetic field generated by the EMA. Participants are seated in the center of the field, so all movements of interest occur within the electromagnetic field and are recorded.

For this study, we attached sensors at three points along the midsagittal line of the tongue: one on the tongue tip (TT), one on the tongue dorsum (TD), and one on the midpoint between TT and TD, referred to as TB. This means that TB was on the anterior and TD on the posterior part of the tongue dorsum. Vertical displacement of the TT sensor was used for coronal (dental and alveolar) segments; the vertical displacement of the TB sensor for palato-alveolar segments, and the vertical displacement of the TD sensors was used for velar segments. Two sensors were attached on the lip, one on the upper and one on the lower lip. The Euclidean distance between these sensors at each timepoint was calculated during data analysis, and the resulting variable—Lip Aperture (LA)—was used for labial segments. Sensors were also attached on the upper and lower incisor for reference and jaw movement respectively, and on the bridge of the nose and behind each ear for head correction. Audio was recorded with a Shure SCM262 microphone mixer at a 16 kHz sampling rate, with a Sennheiser shotgun microphone positioned a foot away from the participant's mouth. Kinematic data was automatically synchronized with the external audio data.

Participants read the stimuli off of a computer screen placed approximately three to four feet from the EMA. Stimuli were presented in the Georgian orthography, but we present here the IPA transcription. Georgian orthographic versions are available in the Appendix.

2.2 Experimental design and stimuli

Each test word appeared first in isolation and then in a frame sentence, without the isolated word disappearing from the screen. The phrase-medial tokens were analyzed for this study. The following frames were used:

(1) _____. Kal-ma _____ mo-m-ts'er-a. [Speakers M1 and F2]

_____. woman.ERG ____ PRVB-1SG.OBJ-write-3SG.SUBJ.AOR

_____. The woman wrote _____ to me.

(2) _____. K'idev _____ v-tkv-i. [Speaker F1]

_____. again ____ 1SG.SUBJ-say-AOR.

_____. I said ______ again.

The frame sentence was changed following the pilot participant (F1) in order to simplify the phonetic material immediately following the test word (i.e., removing the somewhat unwieldy four consonant onset in 'I said') and to bracket the test word with syllables of the form mV to avoid as much gestural blending as possible.

Tables 2 and 3 show the test words used in the experiment. Speakers F1 and F2 produced the test words eight times each, and M1 seven times. Of the resulting 442 phrase-medial tokens, nineteen were removed from the data set due to issues with sensor tracking or participant speech error, and two outliers that were more than three standard deviations from the mean were removed as well. The final data set analyzed had 413 total observations. In both sets of test words, words in opposite corners of the table mirror one another. For example, front-to-back sonority rises contain the same consonants as back-to-front sonority falls. Plateau clusters across order conditions are also mirrors of one another where possible.

Table 2. Test words for speaker F1

	Front to back	Back to front
Sonority rise	**br**elo 'chaff' **p'l**edi 'rug'	**tm**aze 'hair.in' **dm**anisi 'Dmanisi (town in Georgia)'
Sonority plateau	**mn**axe 'see me' **bk'**ich'i 'raisin' **bg**era 'sound'	**t'b**aze 'lake.in' **k'b**ili 'tooth' **k'b**ena 'sting'
Sonority fall	**mt**aze 'mountain.in' **md**are 'worthless'	**rb**ena 'running (n)' **lp'**eba 'decaying (n)'

Table 3. Test words for speakers F2 and M1

	Front to back	Back to front
Sonority rise	**br**egi 'mound' **bn**eli 'darkness' **bn**eda 'epilepsy' **bl**epi 'bluff' **ml**ode 'waiting'	**tm**aze 'hair.in' **dm**anisi 'Dmanisi (town in Georgia)' **gr**evi 'gift' **gl**exi 'peasant'
Sonority plateau	**ʃx**ama 'poison' **bg**era 'sound' **pt**ila 'hair lock' **mn**axe 'see me'	**xʃ**avs 'shut off' **gd**eba 'lie about' **tb**eba 'warm up/warming'
Sonority fall	**mt**aze 'mountain.in' **md**are 'worthless' **rg**eba 'benefitting'	**rb**ena 'running (n)' **lb**eba 'softening (n)' **lm**oba 'feeling sadness'

2.3 Analyses

Collected data were post-processed; they were head-corrected and rotated relative to the occlusal plane, which was recorded by having participants bite down on a hard plastic plane designed precisely for this purpose. Then, all data were semi-automatically labeled based on velocity criteria using the custom software Mview (Mark Tiede, Haskins Laboratory). Constriction maxima were identified using velocity minima, and other timepoints were identified using thresholds of velocity ranges between alternating velocity extrema (i.e., between a maximum and a minimum or vice versa). For all timepoints and tokens the velocity threshold was set to 20%.

Figure 6 shows the relevant timepoints used for the analysis. From left to right, they are: the onset of the gesture; the target achievement; the constriction release; and the offset of the gesture. The figure also shows the point of maximum constriction during the closure, as well as the points of peak velocity for the formation and release of the gesture. The stretch of time between the target and the release is referred to hereafter as the constriction duration.

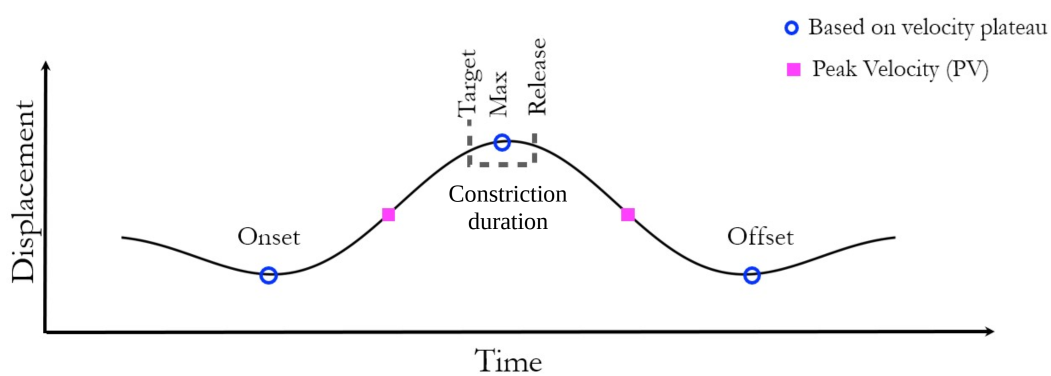


Figure 6. Timepoints labelled for each gesture

We calculate two measures of overlap for every test cluster, and our hypotheses apply to both measures. The first quantifies what we are terming *relative overlap*, following Chitoran and colleagues (2002) and Gafos and colleagues (2010): when C2 gesture starts with respect to C1 constriction, measured as the temporal interval between the onset of C2 and the target achievement of C1 relative to the constriction duration of C1. Equation 1 and Figure 7 show

how relative overlap was calculated, and how that measure can be visualized. The second measure quantifies *constriction duration overlap*, per Hoole and colleagues (2013): the overlap of the constriction duration of the two consonant gestures, as defined in Equation 2 and Figure 8.

$$(1)\ \frac{(C_2 Onset)-(C_1 Target)}{C_1 Constriction\ Duration}$$

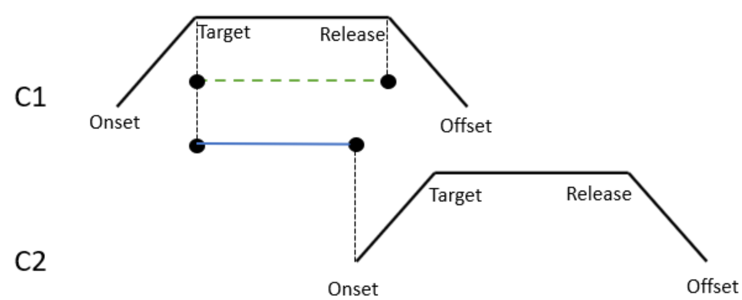


Figure 7. Relative overlap schematization. Solid horizontal line (blue) indicates the interval measured, dotted horizontal line (green) marks the normalizing interval.

For the measure of relative overlap, negative values indicate more overlap; C2 begins before C1 reaches its target. A value of zero means that C2 begins when C1 reaches its target, and positive values indicate that C2 begins after C1 reaches its target.

$$(2)\ \frac{(C_1 Release - C_2 Target)}{(C_2 Release - C_1 Target)}$$

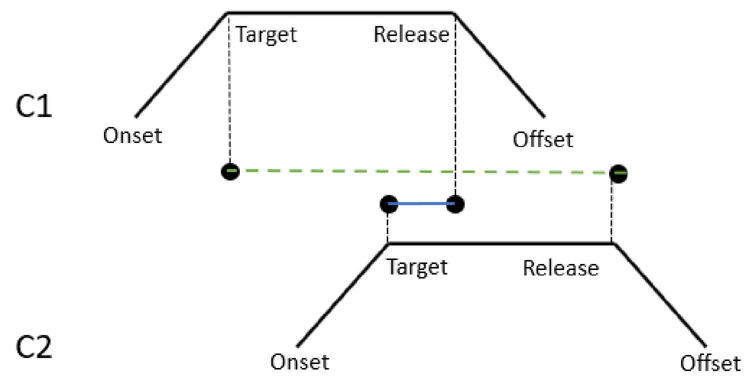


Figure 8. Constriction duration overlap schematization. Solid horizontal line (blue) indicates the interval measured, dotted horizontal line (green) marks the normalizing interval.

For constriction duration overlap, positive values indicate overlap between the two constriction durations. A value of zero means that C2's target and C1's release coincide. Negative values indicate that C2 achieves its target after C1 is released.

Statistical analysis was done in R (R Core Team, 2020) using linear mixed effects models from the lmerTest package (Kuznetsova et al., 2017), followed by post-hoc pairwise comparisons with Holm corrections, using the emmeans package (Length et al., 2020). Plots were generated using the ggplot2 package (Wickham, 2016).

Model selection for both measures was done by starting with the maximal model both in terms of random and fixed effects. The maximal fixed effects structure was Sonority (with three levels: Rise, Plateau, and Fall), Order (with two levels: Front-to-back and Back-to-front), and their interaction. The maximal random effects structure had random slopes by Word, Speaker, and Frame Sentence. The minimal adequate random effects structure was determined using rePCA from the lme4 package (Bates et al., 2015), and the minimal adequate fixed effects structure was determined using drop1 from the basic R stats package (R Core Team, 2020).

3. Results

3.1 Relative overlap results

The final model for relative overlap has fixed effects of Sonority and Order, and random intercepts by Speaker, Word, and Frame Sentence. Sonority ($F(2)=17.56$, $p<.0001$, with Falls as the baseline) and Order ($F(1)=4.45$, $\beta=-.53$, $SE=.25$, $p<.05$, with back-to-front as the baseline) are significant in the overall model. Table 4 summarizes the fixed effects results. Each pair of sonority shapes is significantly different from one another ($\beta=-.89$, $SE=.31$ $p<.01$ for Plateau vs. Rise; $\beta=-.99$, $SE=.35$, $p<.01$ for Fall vs. Plateau; $\beta=-1.89$, $SE=.35$ $p<.001$ for Fall vs. Rise). Falls are the most overlapped and rises are the least. This means that in falls, C2 starts earlier relative to C1 than in other sonority shapes. This is not the pattern predicted by either hypothesis. In fact, it is an exact reversal of the prediction of H1.

Table 4. Summary of the linear mixed effects model for relative overlap, with Falls as the baseline for Sonority, and Back to front as the baseline for Order.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.08	-0.59 – 0.74	0.823
SONORITY [Plateau]	.97	0.30 – 1.65	**0.005**
SONORITY [Rise]	1.95	1.27 – 2.62	**<0.001**
ORDER [Front to Back]	-0.59	-1.11 – -0.04	**0.028**

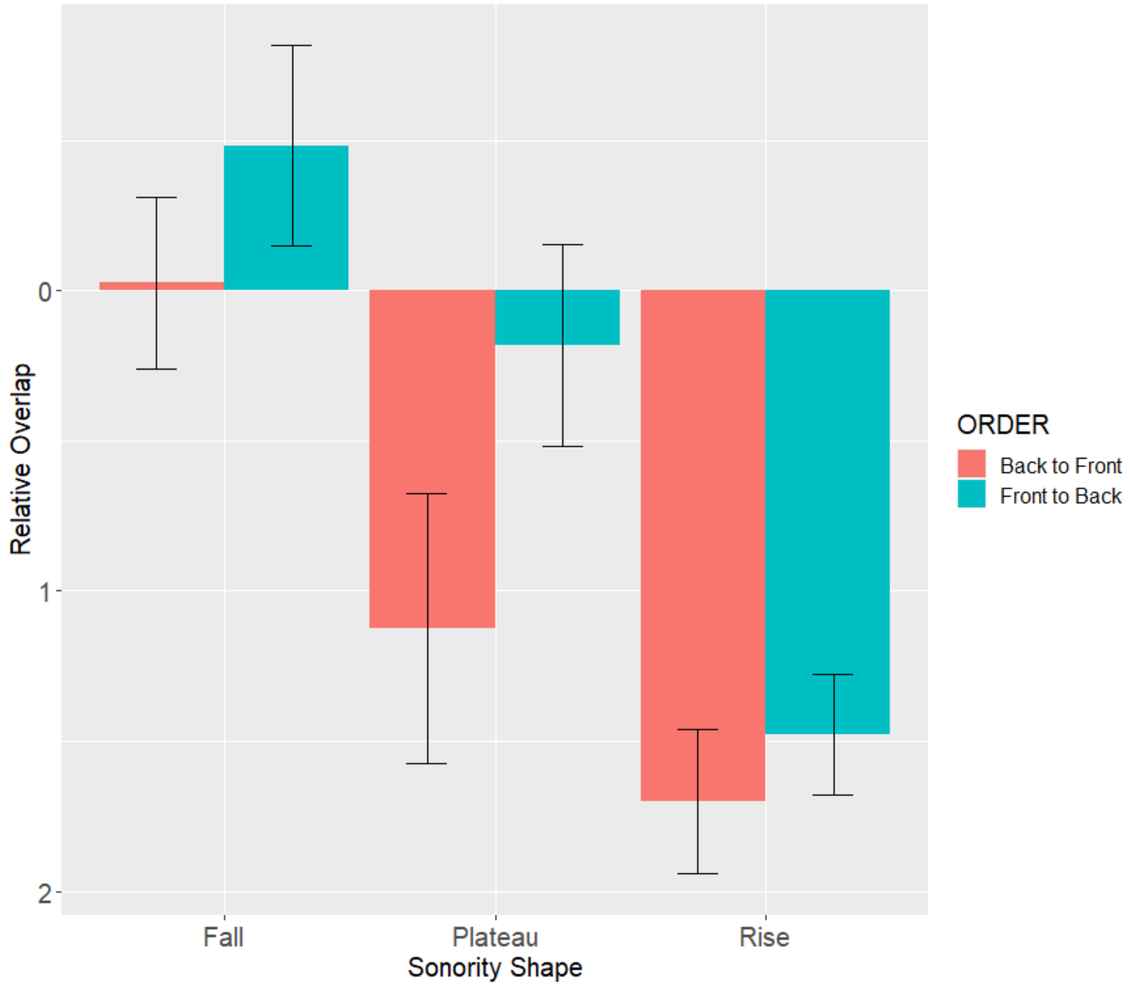


Figure 9. Relative overlap as a function of Sonority Shape. The y-axis is inverted, with positive values indicating less overlap

Back-to-front clusters (such as [**tʰm**] or [**gd**]) are less overlapped than front-to-back clusters (such as [**mtʰ**] or [**bg**]). This replicates findings for stop-stop sequences in previous research on Georgian (Chitoran et al., 2002) and extends them to a wider range of cluster types, even those where neither consonant requires an audible release, like fricative-fricative clusters. This can be taken as evidence that the order effect is a general timing pattern rather than one supporting perceptual recoverability in specific clusters. As seen in Figure 9, this pattern occurs in all three sonority shapes.

3.2 Constriction duration overlap results

The final model for constriction duration overlap had a fixed effect of Sonority, and random intercepts by Speaker, Word, and Frame Sentence. Table 5 summarizes this model. The only significantly different pairwise comparison is between sonority rises and sonority falls (β=.21, SE=.08 p<.05). However, as figure 10 illustrates, we do see a trace of the same hierarchy, in the same direction, that was found in the relative overlap measure. In other words, constriction overlap presents the tendency to increase from Falls to Plateaus and then to Rises. It is possible that with more data each pairwise comparison would reach significance. As mentioned in section 3.1, this is not what is predicted by either hypothesis but is an exact reversal of the pattern predicted by H1.

What is of further special interest here is that the values of constriction duration overlap are negative across all sonority shapes; these negative values indicate a lag between the constriction durations of C1 and C2. The uniformity of these negative values suggests that this may be a language-wide timing pattern, consistent with findings presented in Pouplier et al. (2022). Rises have more lag (less overlap) than falls; this makes sense given the results from the relative overlap measure discussed in Section 3.1, which show that C2 starts much later relative to the onset of the C1 constriction in sonority rises.

Table 5. Summary of the linear mixed effects model for relative overlap, with Falls as the baseline for Sonority.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	-.011	-0.33 – 0.11	0.327
SONORITY [Plateau]	-.024	-0.48 – 0.00	0.052
SONORITY [Rise]	-.043	-.067 – -0.20	**<0.001**

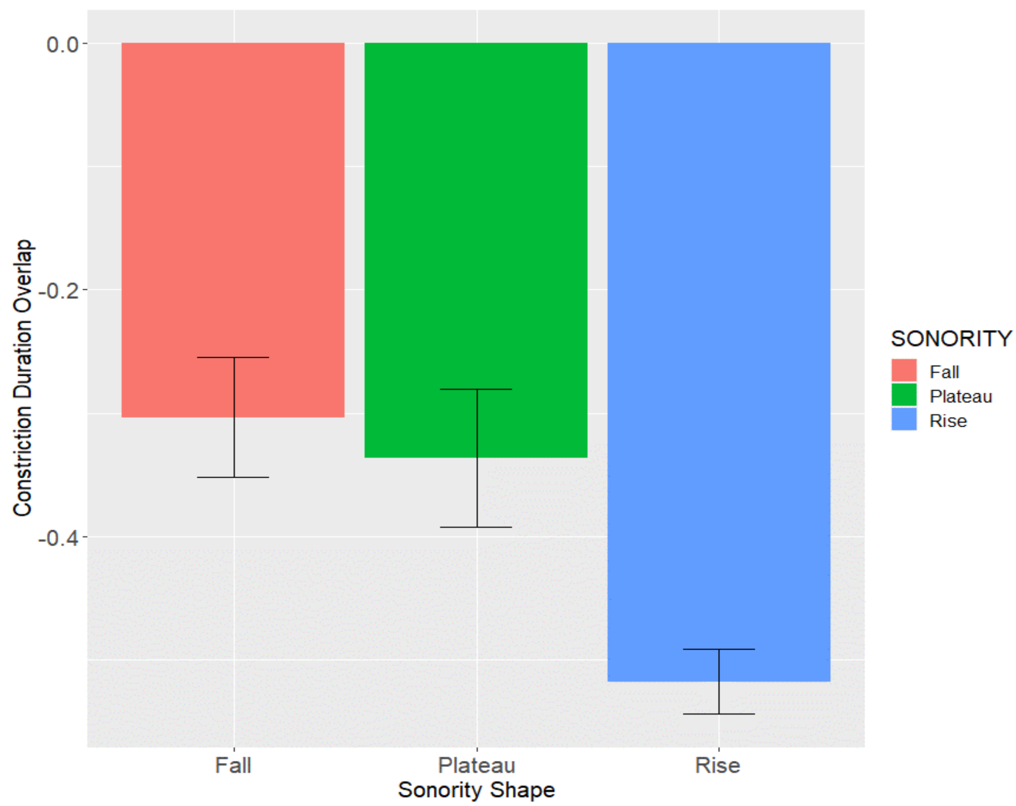


Figure 10. Constriction overlap by sonority shape. The y-axis shows normalized overlap values. They are all negative, indicating constriction duration lag in all sonority shapes.

Figure 11 shows a schematization similar to that in Figure 8, but with lag rather than overlap between the constriction durations.

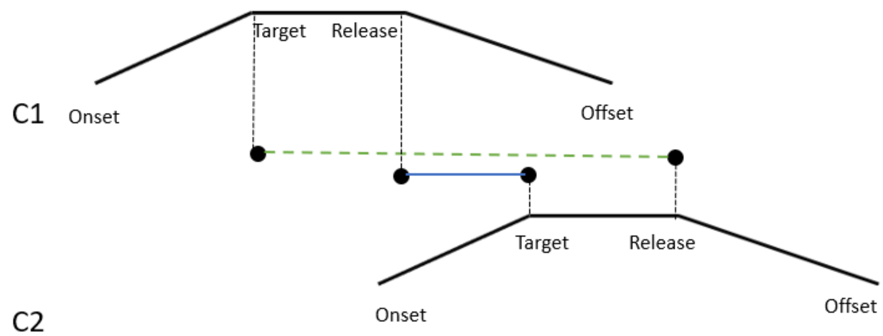


Figure 11. Constriction duration overlap schematization with lag between the constriction durations. Solid horizontal line (blue) indicates the interval measured, dotted horizontal line (green) marks the normalizing interval.

3.3 Variability

Our primary hypothesis (H1) proposes that the cross-linguistic preference for sonority rises will also be reflected in Georgian in the amount of variability in the timing of each sonority shape. We specifically predict that sonority rises will be the least variable shape, sonority plateaus an intermediate case, and sonority falls the most variable. In order to test this, we calculated the standard deviation for each measure of overlap (discussed in 3.1 and 3.2 above) for each word for each speaker and ran linear mixed effects models on the resulting data set. We opted to use the standard deviation rather than a normed measure of variability (e.g., residual standard deviation) because the measures of overlap we used are already normalized.

For relative overlap (3.1), there was no difference in variability between sonority shapes. For constriction duration overlap (3.2), sonority rises were significantly less variable than plateaus and falls, but plateaus and falls were not significantly different from one another. Table 6 shows the means and standard deviations for constriction duration overlap by speaker and by sonority shape. Table 7 shows the model summary. The model had sonority falls as the baseline. A post-hoc pairwise comparison with a Holm correction showed that rises and plateaus were significantly different (p=.01). Also of note is the similarity across speakers, despite the use of different stimuli.

Table 6. Means and standard deviations of constriction duration overlap presented by speaker and by sonority shape.

	Speaker			Sonority shape		
	F1	F2	M1	Rise	Plateau	Fall
mean	-0.44	-0.39	-0.36	-0.52	-0.34	-0.30
standard deviation	0.28	0.26	0.29	0.17	0.34	.27

Table 7. Summary of the linear mixed effects model for standard deviation of constriction duration overlap, with Falls as the baseline for Sonority.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	0.20	0.11 – 0.28	**<0.001**
SONORITY [Plateau]	0.01	-0.07 – 0.09	0.768
SONORITY [Rise]	-0.10	-0.18 – -0.02	**0.013**

This suggests that in complex onsets, the timing pattern for moving from more constricted to less constricted articulatory targets (i.e., as in rises) is more stable than other timing patterns (i.e., as in falls and plateaus.)

4. Discussion

Our results show a hierarchical effect of sonority shape on both measures of gestural overlap. We find lag across the board for constriction duration overlap (Section 3.2) and in sonority rises and plateaus for relative overlap (Section 3.1). For this reason, in the following discussion, we will refer to lag rather than overlap. Sonority rises show the greatest lag for both measures, sonority plateaus are an intermediate case, and sonority falls show the least lag. Sonority rises are also significantly less variable than either plateaus or falls for the measure of constriction duration overlap. The observed hierarchical effect allows us to reject our alternate hypothesis (H2) which predicted that sonority plateaus would show more lag than rises or falls and that rises and falls would not differ significantly from one another. H2 proposed consonant recoverability as the sole motivation for timing patterns, and although the long lag in Georgian allows for consonant recoverability, this is an across-the-board pattern, and not a case of timing modulation occurring in specific clusters to preserve consonant identity.

Our primary hypothesis (H1) did predict a hierarchical effect of sonority shape on both the temporal overlap of gestures and the variability of that overlap. This allows us to posit a systematic relationship between sonority and gestural timing. However, the observed direction of the effect for gestural overlap—falls have the least lag, rises the most—is the opposite of what we predicted. Variability (Section 3.3) follows our prediction to an extent; we predicted rises would be the least variable shape, and they are, but only for one measure of overlap. Additionally, falls and plateaus are not distinguished in the variability measure, while we predicted that they would be. In the rest of the discussion (Sections 4.1 and 4.2) we propose an interpretation of this unexpected hierarchy with a focus on the underlying motivation for timing patterns in Georgian, and how syllables are organized.

4.1: Hierarchical effect of sonority shape on gestural overlap

Our results suggest a principled relationship between sonority shape and gestural timing, but not one motivated by maximizing parallel transmission of information, according to which sonority rises should have been the most overlapped because they are best suited to parallel transmission (Mattingly, 1981). Further evidence that maximizing parallel transmission does not drive gestural timing in Georgian is the fact that both our measures of overlap actually show lag between adjacent consonant gestures across the board. A major consequence of these long lags is the appearance of intrusive vocoids. Vocoids have been reported before in research on Georgian (Goldstein et al., 2007), and we find vocoids both in the data presented here and in other data collected from these same participants for a different study (Crouch et al., 2022). By vocoids we mean intrusive, schwa-like elements that appear between consonants where no phonemic vowel is present and are not the result of a phonological process.

In the data reported in this paper, 56% of sonority rises produced have intrusive vocoids, as do 25% of sonority plateaus. Only 9% of sonority falls have intrusive vocoids, and over half of those occur in productions of the word /mdare/. Vocoids occur in words with all possible laryngeal settings for C1. However, our test words all have voiced consonants in C2 positions, so it is possible that a voiced C2 is a prerequisite for intrusive vocoids in Georgian. As exemplified by Figure 12, these vocoids are much shorter than full vowels: the mean duration of all intrusive vocoids in the data reported in this paper is 29 ms, with a standard deviation of 12.5 ms. We do not investigate any further acoustic properties (e.g., formant structure) of the intrusive vocoids, because our experiments were not designed for this purpose. Crucially, knowledge of these properties, such as the exact quality of the vocoid, does not affect our following discussion. This is, however, a fruitful avenue for future research.

Figure 12. Spectrogram of /bge/ from /bgera/ with vocoid labeled.

The vast difference in vocoid occurrence by sonority shape indicates that not only is constriction duration lag necessary for vocoid intrusion, but also a later onset of C2. When C2 begins after C1 reaches its target, as is the case for all sonority rises and plateaus, the vocal tract is more open for C2's formation. If phonation for C2 has already begun during the movement to target, a vocoid appears. The relatively earlier C2 onsets in sonority falls prevent these conditions from arising even when the presence of a voiced C2 means that they could technically be met.

Vocoid intrusion can improve consonant perceptibility by allowing for an audible release with formant transition information, which is most crucial for stops. When a stop is the first consonant in a CC sequence, that sequence is either a sonority rise or a plateau. Stops in CC sonority falls are in second position, released into the vowel nucleus. Their perceptibility is not at risk, and thus, sonority falls do not gain any specific advantage for consonant recoverability from the presence of a vocoid. The overall prevalence of open transitions (Catford, 1988) and tolerance of ensuing vocoids in Georgian is likely due to both the timing patterns discussed above as well as to the lack of phonemic schwa and of any widespread pattern of phonological vowel reduction. In a language without any of these features, we might expect less opportunity for open transitions because the resulting intrusive vocoid is more likely to be interpreted as a reduced phonemic vowel. We argue that a similar concern is what motivates the hierarchical effect of sonority on gestural overlap that we observe in these data.

First, we want to note that the long constriction duration lag is likely a language-wide setting which may have first arisen to ensure consonant recoverability in stop-stop sequences and was then phonologized and generalized to all sequences. Although this original motivation is the same one we predict in our alternate hypothesis (H2), the fact that this high-recoverability pattern occurs even in clusters where it is not necessary indicates that consonant recoverability is not the concern that drives observed variation in timing patterns.

We see in sonority falls a significantly earlier C2 onset than in other sonority shapes. This works to shorten and counterbalance the constriction duration lag and drastically reduce the likelihood of an intrusive vocoid appearing. Because of their greater stability, sonority rises most likely represent the default interconsonantal timing pattern in Georgian onset. The timing patterns seen in plateaus and falls are adjustments to that pattern, similar to those described in other languages in Section 1.4. We propose that sonority falls are specifically timed differently—more closely together—than other sonority shapes precisely in order to prevent intrusive vocoids, because vocoids can also affect the syllable parse. The potential sonorant-vocoid sequence that would emerge in sonority falls could be more likely to be perceived as a phonological CV sequence, and therefore result in incorrect syllabification and possible lexical misidentification. The vulnerability of these particular sequences to mis-syllabification could be because the sonorant-vocoid sequence would result in a sonority peak equivalent to that of a phonemic vowel, or because a vowel in that context would have a formant structure more similar to those of Georgian's phonemic vowels.

4.2 Typological versus language-specific perspectives

There are two perspectives that we can adopt with respect to the relationship between sonority and overlap in any given language. The first is typologically oriented; generalizations about linguistic phenomena derived from cross-linguistic data (like the SSP) will be applicable to specific languages as well. A hard form of this hypothesis would be one advanced in generative grammar, that there are universal markedness constraints active for all speakers of all languages (e.g., Berent et al., 2011). A more nuanced view would be that these patterns may be further affected by language-specific factors and in languages like Georgian, which have sonority plateaus and falls, the efficient production-perception balance afforded by sonority rises can be outweighed or interfered with by other factors, such as the need to ensure perception of all consonants in multimorphemic CC(…) sequences.

A second perspective is a strictly language-specific one. If we consider only Georgian, markedness is no longer relevant to the discussion about sonority. Georgian has no sonority restriction beyond blocking sequences of three sonorants (example from Chitoran 1999 p. 102).

e.g., mts'rali -> mts'rlis

BUT mtvrali -> mtvralis (where *mtvrlis* is expected)

From the viewpoint of a Georgian speaker, there is nothing ill-formed about a syllable onset like [rb] or [lp'] or longer sequences with sonority reversals, such as [k'rb]. Taking this viewpoint generates a completely different hypothesis. Instead of a hierarchical effect based on sonority shape, we should see no effect of sonority shape on the articulatory dimension of the syllable.

Although the overlap results reported here do not confirm this hypothesis, we do find this noneffect of sonority when we consider global timing measures (i.e., the c-center effect) in Georgian (Crouch et al., 2022). As discussed at the end of Section 1.4, it is not unusual to see variation in local timing across syllable onsets that have the same global organization. This is also the case in Georgian, where the extremely long lags that we see between consonants are consistent with anti-phase (sequential) coordination that, according to the coupled oscillator model, exists between consonants in the onset. However, these longer lags are more difficult to reconcile with the in-phase coordination that the model posits between each onset consonant gesture and the vowel gesture. In a concurrent study (Crouch et al., 2022), preliminary results show no evidence of the c-center effect in Georgian from the same speakers, and instead, we found evidence of anti-phase coordination between all gestures in the syllable, even between the C and V gestures in CV syllables, where simple in-phase coordination is predicted. Furthermore, evidence also suggests anti-phase coordination even in CV sequences. The conclusion of this is that the entire syllable string in Georgian is organized in the same way: sequentially. This is consistent with the observation that Georgian has no restrictions on possible onsets. There is no need to coordinate different parts of the syllable differently because there are no phonotactic violations to be concerned with.

However, neither of the approaches detailed above accurately predict our results on overlap: 1) the typologically based approach predicts the opposite direction of the effect that we actually observe , and 2) sonority shape is not irrelevant for gestural overlap. Rather than being sensitive to the markedness of sonority shapes or completely uniform in cluster timing, Georgian speakers are modulating timing according to a different concern. These data on overlap suggest that speakers are sensitive to the possibility of incorrect syllabification due to vocoid intrusion. Speakers are not simply inserting phonetic material; these are not epenthetic vowels, but intrusive vocoids that appear as a result of timing patterns. Speakers manipulate gestural timing, which is an organizing principle of the syllable, in order to avoid vocoids in syllable onsets that are particularly vulnerable to a disyllabic mis-parsing. This is where the local effect of sonority is seen. Falls of the CCV form are for some reason considered more likely to be interpreted as CVCV disyllables than other sonority shapes, possibly because of their left-edge sonorants. If a vocoid occurs after a stop, it is more likely to be interpreted as a stop release. But when it follows a sonorant, this option is not available, and if it is long enough it may be interpreted as a full vowel, leading to a CVCV interpretation and loss of lexical recoverability. We argue that this is the reason for the higher overlap (i.e., shorter lag) found in sonority falls in our results.

This account fits nicely with the global timing patterns discussed above. There are no phasing differences between parts of the syllable in Georgian. Intervocalic sequences are syllabified as a simplex coda followed by a complex onset regardless of the size of the sequence, although there is often disagreement among speakers (Harris, 2002). Word-edge phonotactics lend themselves to this syllabification, though, as even simplex word-final codas are rare in the lexicon, and virtually all complex codas are multimorphemic. It is possible that syllables are identified solely by the presence of a vowel. If a vocoid is too similar to a phonemic vowel, then there is no other timing pattern that can signal the intended syllabification. For this reason, recoverability of the syllable as a unit is the primary factor behind consonantal overlap in Georgian. Ensuring a felicitous syllabification also works to maintain lexical recoverability, since there are minimal or near-minimal pairs such as /p^ht^hila/ and /p^hit^hila/

The absence of the c-center effect in other data from the same speakers (Crouch et al., 2022) and the failure of H1 to correctly predict the nature of the relationship between sonority and timing are likely entwined. In languages with c-centering in complex onsets, syllables can be and likely are distinguished by having distinct right- and left-edge timing patterns (shown in section 1.4). The competitive coupling in complex onsets specifically results in considerable overlap between the onset consonants (Figure 5). In these languages we would also expect to see the hierarchical effect of sonority on overlap as predicted in H1; sonority rises are better able to preserve consonant recoverability while allowing for overlap. Georgian, however, does not clearly distinguish onsets and codas in this way, and instead the relationship between sonority shape and timing is modulated by a need to preserve the syllable parse in the absence of other coordinative clues.

This perspective also provides a set of testable hypotheses about perception, to be examined in future research. For example, if our syllable perception-driven hypothesis is correct, we should expect different rates of CVCV (mis)transcriptions for CCV sequences based on sonority shape and relative overlap value. To illustrate further, sonority rises with low relative overlap will be less likely to be perceived as CVCV disyllables than sonority falls with low relative overlap, as also suggested by the results of Fleischhacker (2001) and Berent et al. (2011), which show that sonority rises with an intervening schwa are judged to be more similar to the uninterrupted sequence than other sequences with intrusive schwa.

5. Conclusion

By using two measures of gestural overlap, we demonstrate a) that multiple timing relationships need to be examined in order to fully understand how syllables are organized, and b) that the typologically unusual phonotactic system of Georgian is maintained through a careful balance of timing relationships. There is a language-specific setting of constriction duration lag which applies to all clusters and works to preserve the identity of consonants in sequences. The measure of relative overlap, which tells us when the constriction gesture of the second consonant starts relative to the target achievement of the first consonant, is significantly different across sonority shapes. We see earlier C2 onset in sonority falls than in plateaus, and earlier C2 onset in plateaus than in rises. This prevents intrusive vocoids in sonority falls, which we hypothesize are much more vulnerable to perceived resyllabification than other kinds of CC sequences.

In light of these results, we propose that sonority sequencing can best be understood by considering articulatory overlap. Sonority sequencing violations in Georgian are possible due to a combination of large constriction duration lags with different lags between C2's onset and C1's target achievement across sonority shapes. These timing patterns ensure recoverability not only of each consonant in an onset but also of the syllable as an entire unit. The success of these measures in describing the Georgian system is certainly due in part to the fact that they are inherently relational measures. Although sonority itself as a property is ascribed to individual segments, sonority sequencing is about the relation between segments. In order to better understand sonority sequencing, we need to ground our investigation in relational properties.

Acknowledgements

This work was supported by the National Science Foundation [#1551428] and National Institutes of Health [NIDCD-DC-002717].

Appendix: Stimuli in Georgian orthography

(1) _____. ქალ-მა _____ მო-მ-წერ-ა. [Speakers M1 and F2]

_____. woman.ERG ____ PRVB-1SG.OBJ-write-3SG.SUBJ.AOR

(2) _____. კიდევ _____ ვ-თქვ-ი. [Speaker F1]

_____. again ____ 1SG.SUBJ-say-AOR.

Table A1. Test words for speaker F1

	Front to back	Back to front
Sonority rise	**ბრ**ელო 'chaff' **პლ**ედი 'rug'	**თმ**აზე 'hair.in' **დმ**ანისი 'Dmanisi (town in Georgia)'
Sonority plateau	**მნ**ახე'see me' **ბკ**იჭი 'raisin' **ბგ**ერა 'sound'	**ტბ**აზე 'lake.in' **კბ**ილი 'tooth' **კბ**ენა 'sting'
Sonority fall	**მთ**აზე 'mountain.in' **მდ**არე 'worthless'	**რბ**ენა 'running (n)' **ლპ**ება 'decaying (n)'

Table A2. Test words for speakers F2 and M1

	Front to back	Back to front
Sonority rise	**ბრ**ეგი 'mound' **ბნ**ელი 'darkness' **ბნ**ედა 'epilepsy' **ბლ**ეფი 'bluff' **მლ**ოდე 'waiting'	**თმ**აზე 'hair.in' **დმ**ანისი 'Dmanisi (town in Georgia)' **გრ**ევი 'gift' **გლ**ეხი 'peasant'
Sonority plateau	**შხ**ამა 'poison' **ბგ**ერა 'sound' **ფთ**ილა 'hair lock' **მნ**ახე 'see me'	**ხშ**ავს 'shut off' **გდ**ება 'lie about' **თბ**ება 'warm up/warming'
Sonority fall	**მთ**აზე 'mountain.in' **მდ**არე 'worthless' **რგ**ება 'benefitting'	**რბ**ენა 'running (n)' **ლბ**ება 'softening (n)' **ლმ**ობა 'feeling sadness'

References

Aronson, H. I. (1997). Georgian phonology. *Phonologies of Asia and Africa, 2*, 929-939.

Ball, M. J., & Müller, N. (Eds.). (2016). *Challenging sonority: Cross-linguistic evidence*. Equinox Publishing Limited.

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, 67(1), 1–48. doi: 10.18637/jss.v067.i01.

Berent, I., Harder, K., & Lennertz, T. (2011). Phonological universals in early childhood: Evidence from sonority restrictions. *Language Acquisition*, *18*(4), 281-293.

Bombien, L., & Hoole, P. (2013). Articulatory overlap as a function of voicing in French and German consonant clusters. *The Journal of the Acoustical Society of America*, *134*(1), 539-550.

Bombien, L., Mooshammer, C., & Hoole, P. (2013). Articulatory coordination in word-initial clusters of German. *Journal of Phonetics*, *41*(6), 546-561.

Borise, L., & Zientarski, X. (2018). Word stress and phrase accent in Georgian. *Proceedings of TAL*, *6*, 2018-42.

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology, 3*, 219-252.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2-4), 140-155.

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201-251.

Browman, C. P., & Goldstein, L. (1992). Articulatory phonology: An overview. *Phonetica*, *49*(3-4), 155-180.

Butskhrikidze, M. (2002). The consonant phonotactics of Georgian. Netherlands Graduate School of Linguistics.

Catford, J. S. (1988). *A practical guide to phonetics*. Oxford: Clarendon Press.

Chitoran, I. (1998). Georgian harmonic clusters: phonetic cues to phonological representation. *Phonology*, *15*(2), 121-141.

Chitoran, I. (1999). Accounting for sonority violations: the case of Georgian consonant sequencing. In *Proceedings of the 14th International Congress of Phonetic Sciences*, 101-104.

Chitoran, I. (2016). Relating the sonority hierarchy to articulatory timing patterns : A cross-linguistic perspective. In Martin J. Ball and Nicole Müller (Eds.). *Challenging sonority: Cross-linguistic evidence* (pp. 45-62). Equinox Publishing.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. *Papers in laboratory phonology*, *7*(4-1), 419-447.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in laboratory phonology*, *1*, 283-333.

Crouch, C., Katsika, A., & Chitoran, I. (2022). Georgian syllables, uncentered. In *Proceedings of Speech Prosody 2022.*

Dell, F. & Elmedlaoui, M. (1985). Syllabic Consonants and Syllabification in Imdlawn Tashlhiyt Berber. *Journal of African Languages and Linguistics*, *7*(2), 105-130. https://doi.org/10.1515/jall.1985.7.2.105

Dell, F., & Elmedlaoui, M. (2002). *Syllables in Tashlhiyt Berber and in Moroccan Arabic*. Dordrecht, Netherlands, and Boston, MA: Kluwer Academic Publishers.

Engstrand, O., & Ericsdotter, C. (1999). Explaining a violation of the sonority hierarchy: Stop place perception in adjacent [s]. In *Proceedings from the XIIth Swedish Phonetics Conference*, 49-52.

Gafos, A. I., Hoole, P., Roon, K., Zeroual, C., Fougeron, C., Kühnert, B., D'Imperio, M., & Vallée, N. (2010). Variation in overlap and phonological grammar in Moroccan Arabic clusters. *Laboratory phonology*, *10*, 657-698.

Goad, H. (2011). The representation of sC clusters. In *The Blackwell companion to phonology*, 1-26.

Goldstein, L., & Fowler, C. A. (2003). Articulatory phonology: A phonology for public language use. In Niels Schiller, Niels Olaf Schiller, and Antje Meyer (Eds.). *Phonetics and phonology in language comprehension and production: Differences and similarities*, 159-207.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007, August). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In *Proceedings of the XVIth international congress of phonetic sciences*, 241-244.

Golfinopoulos, E., Tourville, J. A., Bohland, J. W., Ghosh, S. S., Nieto-Castanon, A., & Guenther, F. H. (2011). fMRI investigation of unexpected somatosensory feedback perturbation during speech. *Neuroimage*, *55*(3), 1324-1338.

Harris, A. (2002). The word in Georgian. In Robert Dixon and Alexandra Aikenvald (Eds.). *Word: A cross-linguistic typology*, 127-142.

Hermes, A., Greisbach, R., & Chitoran, I. (2020). Probing the limits of coordination: Word onset clusters in Georgian. *12th International Seminar on Speech Production*.

Hermes, A., Mücke, D., & Auris, B. (2017). The variability of syllable patterns in Tashlhiyt Berber and Polish. *Journal of Phonetics*, *64*, 127-144.

Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology*, *30*(1), 1-25.

Hoole, P., Pouplier, M., Benus, S., & Bombien, L. (2013). Articulatory coordination in obstruent-sonorant clusters and syllabic consonants: Data and modelling. In Lorenzo Spreafico and Alessandro Vietti (Eds.). *Rhotics. New data and perspectives*, 81-98.

Hooper, Joan B. (1976). *An introduction to natural generative phonology*. New York: Academic Press.

Ito, T., Gomi, H., & Honda, M. (2000). Task dependent jaw-lip coordination examined by jaw perturbation during bilabial-consonant utterances. *5th Seminar on Speech Production.*

Jany, C., Gordon, M., Nash, C. M., & Takara, N. (2007). How universal is the sonority hierarchy?: a cross-linguistic acoustic study. In *Proceedings of the ICPhS*, 1401-1404.

Jun, S. A., Vicenik, C., & Lofstedt, I. (2007). WPP, No. 106: Intonational Phonology of Georgian.

Kelso, J. A., & Tuller, B. (1984). Converging evidence in support of common dynamical principles for speech and movement coordination. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, *246*(6), R928-R935.

Kiparsky, Paul (1973). Phonological representations. In Osamu Fujimura (ed.) *Three Dimensions of Linguistic Theory*, Tokyo: TEC. 1–136.

Kiparsky, P. (1979). Metrical structure assignment is cyclic. *Linguistic Inquiry*, *10*(3), 421-441.

Kiparsky, P. (1980). Remarks on the metrical structure of the syllable. *Phonologica*, *3*.

Kochetov, A., Pouplier, M., & Son, M. (2007). Cross-language differences in overlap and assimilation patterns in Korean and Russian. *Proceedings of the XVI International Congress of Phonetic Sciences*, 1361-1364.

Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and C-center in selected French consonant clusters. In *7th International Seminar on Speech Production (ISSP)*.

Kuznetsova, A., Brockhoff, P.B., Christensen, R.H.B. (2017). "lmerTest Package: Tests in Linear Mixed Effects Models." *Journal of Statistical Software*, 82(13), 1-26.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.

Length, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.2. https://CRAN.R-project.org/package=emmeans

Mattingly, I. G. (1981). Phonetic representation and speech synthesis by rule. In *Advances in Psychology* (Vol. 7, pp. 415-420). North-Holland.

Morelli, F. (2003). The Relative Harmony of/s+ Stop/Onsets Obstruent Cluslers and the Sonority Sequencing. In Caroline Féry and Reuben van de Vijver (Eds.). *The syllable in optimality theory*, 356-371.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. In *Proceedings of the 15th international congress of phonetic sciences*, 2253-2256.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. *Approaches to phonological complexity*, *16*, 299-328.

Parker, S. (2008). Sound level protrusions as physical correlates of sonority. *Journal of phonetics*, *36*(1), 55-90.

Parker, S. (2011). Sonority. In *The Blackwell companion to phonology*, 1-25.

Parker, S. (Ed.). (2012). *The sonority controversy* (Vol. 18). Walter de Gruyter.

Pastätter, M., & Pouplier, M. (2014). The temporal coordination of Polish onset and coda clusters containing sibilants. In *Proceedings of the 10th International Seminar on Speech Production, Cologne,* Germany, 312-315.

Pastätter, M., & Pouplier, M. (2015). Onset-vowel timing as a function of coarticulation resistance: Evidence from articulatory data . In *Proceedings of the 18th International Congress of Phonetic Sciences*.

Pouplier, M. (2012). The gestural approach to syllable structure: Universal, language- and cluster-specific aspects. In Fuchs, Weihrich, Pape, Perrier (ed.s) *Speech Planning and Dynamics*, 63-96.

Pouplier, M., Lentz, T., Chitoran, I., & Hoole, P. (2020). The imitation of coarticulatory timing patterns in consonant clusters for phonotactically familiar and unfamiliar sequences. *Journal of Laboratory Phonology, 11(1)*.

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Recasens, D. (2012). A study of jaw coarticulatory resistance and aggressiveness for Catalan consonants and vowels. *The Journal of the Acoustical Society of America*, *132*(1), 412-420.

Recasens, D., Pallarès, M. D., & Fontdevila, J. (1997). A model of lingual coarticulation based on articulatory constraints. *The Journal of the Acoustical Society of America*, *102*(1), 544-561.

Saltzman, E. (1986). Task dynamic coordination of the speech articulators: A preliminary model. *Experimental brain research series*, *15*, 129-144.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499-526.

Saltzman, E. L., & Kelso, J. S. (1983). Toward a dynamical account of motor memory and control. In *Advances in Psychology* (Vol. 12, pp. 17-38). North-Holland.

Selkirk, Elizabeth O. (1982). Syllables. In Harry van der Hulst and Norval Smith (eds.) *The structure of phonological representations*, *Part II*. Dordrecht, Foris, 337–383.

Shaiman, S. (1989). Kinematic and electromyographic responses to perturbation of the jaw. *The Journal of the Acoustical Society of America*, *86*(1), 78-88.

Shaw, J. A., Gafos, A. I., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure: a case study of Moroccan Arabic consonant clusters. *Phonology*, *28*(3), 455-490.

Shosted, R. K., & Chikovani, V. (2006). Standard Georgian. *Journal of the International Phonetic Association*, *36*(2), 255-264.

Son, M., Kochetov, A., & Pouplier, M. (2007). The role of gestural overlap in perceptual place assimilation in Korean. *Papers in laboratory phonology*, *9*, 504-534.

Sotiropoulou, S. & Gafos, A., (2022) Phonetic indices of syllabic organization in German stop-lateral clusters. *Laboratory Phonology 13(1)*. doi: https://doi.org/10.16995/labphon.6440

Sotiropoulou, S., Gibson, M., & Gafos, A. (2020). Global organization in Spanish onsets. *Journal of Phonetics*, *82*, 100995.

Steriade, D. (1982). *Greek prosodies and the nature of syllabification*. PhD diss., Massachusetts Institute of Technology.

Vicenik, C. (2010). An acoustic study of Georgian stop consonants. *Journal of the International Phonetic Association*, *40*(1), 59-92.

Vicenik, C., & Jun, S. A. (2014). An autosegmental-metrical analysis of Georgian intonation. In *Prosodic typology II: The phonology of intonation and phrasing*, 154-186. .

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis.

Wright, R. (2004). A review of perceptual cues and cue robustness. In Bruce Hayes, Robert Kirchner & Donca Steriade (eds.) *Phonetically based phonology*, 34–57.

Yavaş, M., Ben-David, A., Gerrits, E., Kristoffersen, K. E., & Simonsen, H. G. (2008). Sonority and cross-linguistic acquisition of initial s-clusters. *Clinical Linguistics & Phonetics*, *22*(6), 421-441.

Yavaş, M. (2013). What explains the reductions in/s/-clusters: Sonority or [continuant]?. *Clinical linguistics & phonetics*, *27*(6-7), 394-403.

Yip, M. (2002). *Tone*. Cambridge University Press.

Chapter 3. “A mountain of tongues”: syllables without c-centers in Georgian

1. Introduction

Syllables are important units in the prosodic hierarchy (Hayes, 1985; Ladd, 2008) and have a psychological reality (Selkirk, 1982; Treiman, 1986) that makes them relevant to speech processing, and we can define the syllable as a unit in these two aspects. However, we lack a similarly well-formed understanding of the syllable as a unit of speech production, and this gap prevents us from defining the syllable in its complex entirety. The work presented here reports on one of the lines of investigation we have pursued to identify the organization of syllables in speech production in Georgian. We bring together two dominant approaches to the syllable, one based on sonority sequencing (Kiparsky, 1980; Clements, 1990), and the other on the coupled oscillator model (Nam et al., 2009) in Articulatory Phonology (AP).

We bring together these two approaches in order to engage with both the spatial and temporal dimensions of the syllable. The coupled oscillator model (section 1.2) provides a temporal definition of syllable structure, and sonority sequencing gives us a primarily spatial account of syllable structure (section 1.3) By combining both approaches, we are able not only to provide a spatiotemporal account of syllable structure in Georgian, but to further enhance both approaches by exploring their relationship to one another.

We use Georgian as our test case because it permits syllables that challenge both sonority- and coupling-based approaches. Georgian has minimal phonotactic restrictions on both segmental composition and overall size of onsets. It also allows for any sonority shape in complex onsets. Georgian is also an extreme case for the coupled oscillator model; onsets of up to seven consonants are permitted. An additional major difference between Georgian and other languages analyzed under the assumptions of the coupled oscillator model is that syllable onsets in Georgian can be multimorphemic.

We know from a concurrent study of our larger project (Crouch et al., 2020) that the sonority shape of an onset systematically affects the local timing of the constriction gestures forming the consonants in the onset—by local timing, we mean temporal relationships between single adjacent constriction gestures. The current paper focuses specifically on global timing in Georgian, i.e., on the temporal relationship of the onset as a whole with the syllable's nucleus. We test the prediction that the onset relates to the nucleus as a single unit, regardless of the number of consonants and also ask, whether, and if yes, how the sonority shape of an onset affects the temporal relationship between the onset and the nucleus. Specifically, we address three questions about syllable timing in Georgian: 1) how does Georgian coordinate complex onsets, 2) how does the sonority shape of an onset interact with its coordination, and 3) how does the presence of a morphological boundary in the onset affect syllable coordination?

The paper proceeds as follows: the rest of Section 1 presents relevant information and previous research on the phonology and morphology of Georgian, sonority sequencing, and the coupled oscillator model of syllable structure. We then present more detailed version of our research questions formulated above, and our hypotheses. Section 2 details the methodologies used for data collection and analysis. Section 3 presents the design and results of our first experiment, which examines the effects of sonority and onset size on syllabic timing. Section 4 is an interim discussion of those results. Section 5 presents the design and results of our second experiment, which examines the effect of a morphological boundary on syllabic timing. Section 6 discusses C-V coordination and presents gestural scores based on those data for which we were able to measure the tongue dorsum gesture for the vowel. Section 7 brings together all of our results to present an overall account of syllable organization in Georgian, and section 8 concludes, and offers typological predictions based on this work.

1.1 Georgian phonology and morphology

This study focuses on Georgian because syllable structure in Georgian poses a challenge to both the sonority-based and the coupled oscillator definitions of syllabification. Georgian permits syllable onsets of up to seven consonants, has no restrictions on the segmental makeup of an onset, and only permits vowels to serve as syllable nuclei. Codas are more restricted. Simplex word-final codas are uncommon, and virtually all word-final complex codas are multi-morphemic (Aronson, 1997). Word-medial consonant sequences are syllabified as a simplex coda followed by a complex onset, so there are no word-medial complex codas at all (Harris, 2002). The only notable phonotactic restriction is an avoidance of sequences of three sonorants, as illustrated in the example below from Chitoran (1999) where syncope processes are blocked when they would result in such an illicit sequence.

(1) e.g., mts'rali -> mts'rlis

but: mtvrali -> mtvralis (where mtvrlis is expected)

Sonority sequencing (discussed in more detail in Section 1.3) obviously plays little to no role in syllabification in Georgian. A definition of the syllable as a sonority curve—with a maximum in the nucleus and minima at the syllable edges—excludes many well-formed syllables in Georgian (e.g., /rʧe.u.li/ ‘chosen’ and /mze/ ‘sun’). Sonority itself as a scalar property of speech sounds is relevant to some extent in Georgian, as evidence by both the restriction on possible syllable nuclei and the syncope-blocking mentioned in (1). Sonority also plays a role in local articulatory timing in Georgian (Crouch et al., 2020). The amount of overlap between consonant gestures in a CC onset differs significantly according to the sonority shape of the onset. All C-C sequences in Georgian show lag between constrictions, but in sonority falls C2 begins earlier than in either plateaus or rises.

The sheer size of permissible onsets in Georgian challenges the model of syllable structure advanced within Articulatory Phonology (discussed in more detail in Section 1.2) as well. Onsets larger than three consonants are not found in languages analyzed in this framework, and previous studies on syllable organization in Georgian have led to mixed findings (Goldstein et al., 2007; Hermes et al., 2020). Georgian onsets also differ from those of other languages studied in this paradigm in their morphological composition. Georgian verbal morphology includes obligatory person-marking prefixes, which have the form C or CC. As a result, the same lexical item can have a syllable with different onset sizes depending on the participants in the verbal action.

Georgian has five vowels: /a e i o u/ and no phonological vowel reduction. The consonant inventory of Georgian is presented in IPA in Table 1.

Table 1. Georgian consonant inventory (following Shosted and Chikovani, 2006)

	bilabial	labio-dental	dental	alveolar	post-alveolar	velar	uvular	glottal

plosive	p^h p’ b		t^h t’ d			k^h k’ g	q’	
affricate				ts ts’ dz	ʧ ʧ’ ʤ			
nasal	m			n				
tap/trill				ɾ/r				
fricative		v		s z	ʃ ʒ	x ɣ		h
lateral approximant				l				

Word-level stress in Georgian is weakly implemented and carries a minimal functional load. The generally accepted analysis states that in words of up the three syllables, stress is on the initial syllable, and in words of four or more syllables, primary stress falls on the antepenultimate and secondary stress falls on the initial syllable (Hewitt, 2005). Therefore, we do not predict any stress differences among the test words, which are all of three syllables or less. The marking of phrasal prominence has been the subject of many recent studies (e.g., Borise and Zientarsky, 2018; Skopeteas and Féry, 2016; Vicenik and Jun, 2014), which suggest that prosodic prominence, syntax, and information structure are closely linked in Georgian. Skopeteas and Féry (2016) propose that focus specifically is marked through prosodic phrasing rather than pitch-accent presence or contour.

1.2 Articulatory Phonology and the coupled oscillator model

Articulatory Phonology (AP) is a phonological theory in which the phonological primes are gestures (Browman and Goldstein, 1986). Gestures are abstract, phonological relevant events in the vocal tract which are inherently spatiotemporal in nature. As a result of their spatiotemporal nature, gestures are considered to be active during a specified interval, straightforwardly called the activation interval. Gestures can be treated as oscillating systems, which are coupled in different ways to achieve the different timing relationships observed in speech (Byrd and Saltzman, 2000). In the coupled oscillator model (Nam and Saltzman, 2003; Nam et al., 2009), phasing relationships between gestures are the primary mechanism of syllable organization. There are two types of phasing relationships: in-phase and anti-phase. In-phase coupling results in synchronous gestures, and anti-phase coupling results in sequential gestures. Figure 1a contains coupling graphs representing these two phasing relationships in the left panel and the gestural timing patterns associated with them in the right panel. The in-phase gestures C-V (upper) are initiated simultaneously. The anti-phase V-C gestures are not: the C gesture is initiated after the V gesture has reached its target (not shown in gestural scores).

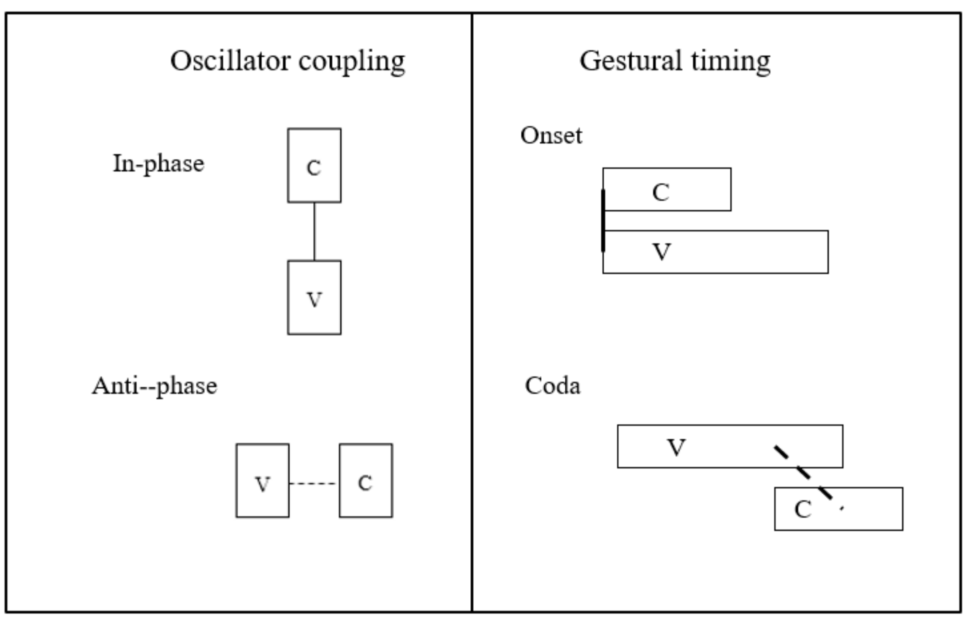


Figure 1a. Oscillator couplings (left) and their associated gestural timing patterns (right). C indicates consonant gestures and V vowel gestures. The solid line marks in-phase coordination, and the dotted line marks anti-phase coordination.

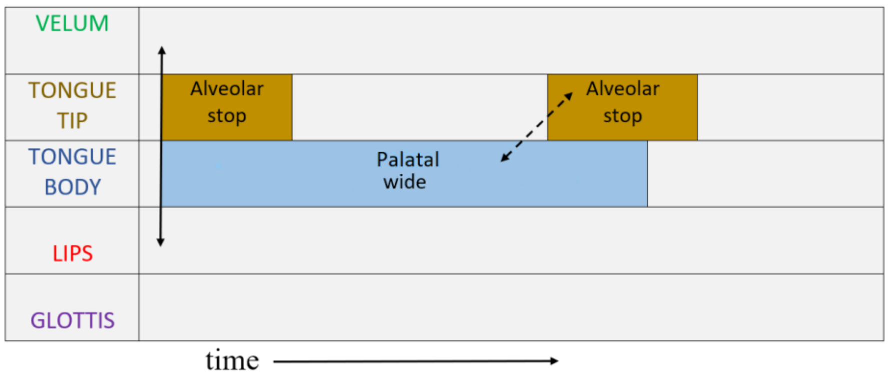


Figure 1b. Gestural score for the word 'dad' with phasing relationships marked. The solid line marks in-phase coordination, and the dotted line marks anti-phase coordination.

Specific parts of the syllable have different phasing relationships: onsets are in-phase with the nucleus, and codas are anti-phase. In Figure 1a, the graph on the left illustrates the in-phase CV relationship and the right illustrates the anti-phase VC relationship. Figure 1b is a gestural score for the English word /dæd/. The gestural score is read left-to-right and shows the duration of the activation intervals for gestures as well as the phasing relationships between gestures. The solid line marks in-phase coordination, and the dotted line marks anti-phase coordination. When the oscillators are in-phase, gestures are initiated simultaneously, as shown by the onset consonant (alveolar stop) and the vowel (palatal wide). When the oscillators are anti-phase, the gestures are initiated sequentially, as shown by the vowel and the coda consonant (alveolar stop).

Complex codas are straightforward to account for in this model. They are a series of anti-phase relationships, so that each coda consonant follows one after the other. Complex onsets, on the other hand, earn the descriptor 'complex'. Onsets are characterized by simultaneity with the nucleus. In a simple CV sequence, the consonant (C) and vowel (V) gestures are activated simultaneously, but the V gesture is slower than the C gesture and the acoustic result is a CV sequence, as show in Figure 1b. When additional C gestures are in the onset, they cannot all simply be in-phase with the nucleus; it wouldn't be feasible to initiate all onset C gestures simultaneously and maintain their perceptibility. Instead, complex onsets display competitive coupling, which is illustrated in Figures 2a (coupling graph) and 2b (gestural score).

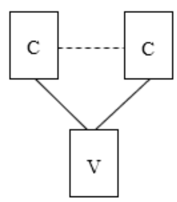


Figure 2a. Competitive coupling in a complex onset. The solid lines mark in-phase coordination and the dotted line marks anti-phase coordination.

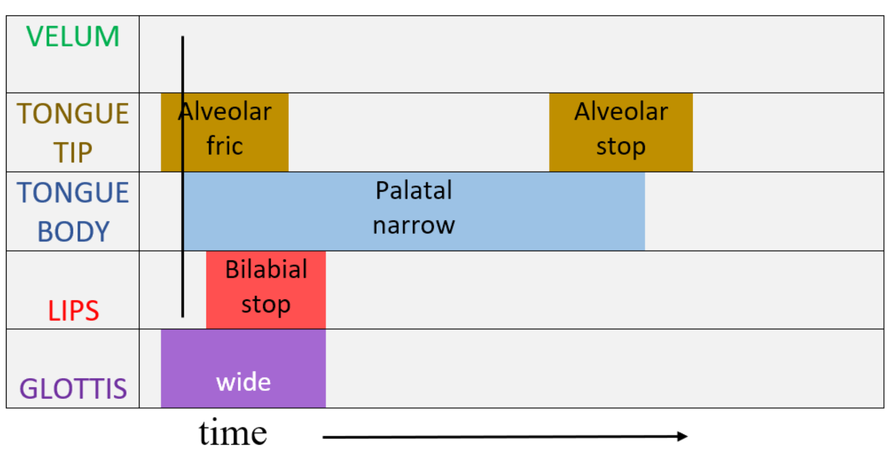


Figure 2b. Gestural score for the word 'speed' with c-center marked

Each consonant gesture is in-phase with the nucleus gesture but is also anti-phase with each of the other consonant gestures. The graph in Figure 2a shows the coupling graph for a CCV sequence, but the same principle is predicted to hold for larger onsets as well.

In this model, competitive coupling is the mechanism behind the c-center effect. The c-center effect was first described in Browman and Goldstein (1988) before the coupled oscillator model was proposed, and subsequently lead to the development of the coupled oscillator model. The c-center refers to a theoretical center of the onset, which is invariantly in-phase with the nucleus regardless of onset size. Each onset consonant is equidistant in time from the c-center. Browman and Goldstein (1988) define the c-center as the timepoint that is equidistant in time from the midpoint of each consonant's constriction. However, the coupled oscillator model makes predictions specifically about the onsets of gestures. For this reason, we employ a measure of the c-center that uses gesture's onsets rather than their constriction midpoints. Figure 2b shows the c-center, defined using gestural onsets, marked for the complex onset in the English word 'speed'. Because in this model the onset is coordinated as unit, rather than a sequence of single gestures like a coda, we refer to *global timing* of the onset.

Complex onsets are characterized in this model by global stability, illustrated by the c-center. In other words, regardless of the onset's size, the c-center is expected to be in-phase with the nucleus gesture. Local timing measures, however, will show differences across onsets of different sizes. The most prominent of these is rightward shift. Although the c-center remains in-phase with the nucleus gesture, the prenuclear consonant occurs later in larger onsets. Since time is traditionally represented on the x-axis, the prenuclear consonant appears further to the right in gestural scores, such as the one in Figure 2b. In many studies this rightward shift is used either as a proxy measure for the presence of the c-center effect or alongside measures that use a defined c-center point (Brunner et al., 2014; Hermes et al., 2008, 2013; Pouplier and Beňuš, 2011; Shaw et al., 2011).

Rightward shift has been used to assess the presence of the c-center effect in Georgian (Goldstein et al., 2007; Hermes et al., 2020). Goldstein et al. (2007) compares coordination in Georgian and Tashlhyit Berber CV, CCV, and CCCV sequences. In Georgian these are expected to be complex syllable onsets, while in Berber CCV and CCCV sequences are expected to be multisyllabic due to the prohibition on complex onsets in the language. Data from two speakers of Georgian were presented: one speaker consistently displayed rightward shift, while the second speaker did not. This difference between the speakers was related to the presence of intrusive vocoids in the CCV and CCCV words. Vocoids were less frequent in the speech of the first speaker compared to the second speaker. Both patterns differed distinctly from those displayed by the Tashlhyit speaker, who clearly did not display the c-center effect, and for whom CCV and CCCV sequences were multisyllabic. By contrast, Georgian was implied to have complex onsets and display the c-center effect. Hermes and colleagues (2020) present data from four Georgian speakers producing onsets of up to seven consonants. They find that for all speakers, rightward shift is present in CV, CCV, and CCCV words, but ceases to occur once onsets are four consonants long.

1.3 Sonority

No study in the AP framework has addressed sonority or sonority sequencing in syllable timing. A major contribution of our current study is to systematically investigate the relationship between the sonority shape of a syllable onset and the gestural timing in that syllable.

Sonority is an abstract property of speech segments and is correlated with both intensity and, more crudely, with degree of vocal tract openness. For a more recent proposal, based on periodic energy, see Albert and Nicenboim (2022). It is a scalar property, and we use the sonority scale show in (1).

(1) vowel > glide > liquid > nasal > fricative > stop (adapted from Parker, 2011)

This sonority scale is the basis for the Sonority Sequencing Principle (SSP), which provides a definition of a well-formed syllable based on sonority criteria (Kiparsky, 1980; Clements, 1990). Under the SSP a well-formed syllable has local sonority minima at its edges, and a sonority maximum in the nucleus. In all languages that permit complex onsets, SSP-conforming onsets are found, and in many languages are the only onsets permitted. Figure 3 shows an example of such a syllable in Georgian.

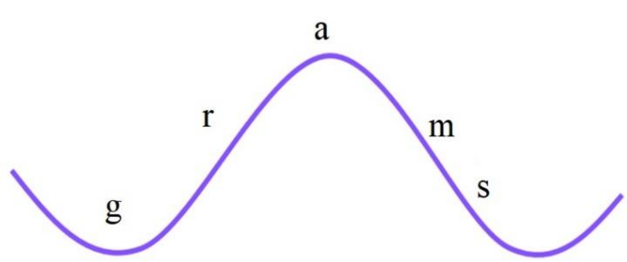


*Figure 7. Sonority curve for the monosyllable /grams/ gram-*DAT *in Georgian*

We refer to these onsets as sonority rises, because the sonority value of the consonants in the onset increases as they are closer to the nucleus. In sonority plateaus, all consonants have the same sonority value (e.g., /bg/), and in sonority falls, the initial consonant is the most sonorous, and the prenuclear consonant the least (e.g., /ld/).

We know very little about the relationship between sonority and timing. Crouch and colleagues (2020) find a systematic relationship in Georgian between the sonority shape of an onset and the timing between consonant gestures in the onset. Theories of the motivations behind sonority imply that sonority rises are better suited to some timing pattern that best support cue preservation (Henke et al., 2012) or the parallel transmission of information (Mattingly, 1981). However, this has never been explicitly studied.

1.4 Morphology and articulation

As mentioned in section 1.1, complex onsets in Georgian are often multimorphemic, so the interaction between morphology and articulatory timing needs to be considered when investigating complex onsets in Georgian. For this reason, our second experiment focuses on the relationship between morphological structure and syllabic timing in Georgian syllables. The exact nature of the morphology-articulation relationship is as yet unclear, although phonetic research has found some cases of an interaction between the two. Shaw et al. (2010) illustrate coordinative differences in C-C sequences in Moroccan Arabic when there are different types of intervening morphological boundaries, i.e., when concatenative versus nonconcatenative morphology is in play. Cho (2001) shows an effect of morpheme boundaries on intergestural timing in the context of derived environment effects in Korean. Song et al.

(2013) show tongue height differences in mono- versus bi-morphemic /ks/ codas in English for both children and adults. They suggest that in the bimorphemic cluster the plural /s/ morpheme could be the primary target of the cluster, while /k/ is the primary target of the monomorphemic cluster. Bell et al. (2021) show that lower paradigmatic diversity is associated with longer durations for geminate consonants at compound-internal boundaries in English.

These studies do suggest that morphology and articulation are related, but they do not allow us to further refine our predictions, both because they are examining different morphological structures than those we are focusing on, and because they are not explicitly concerned with syllable organization. The study we discuss in section 5 is the first to examine the effect of morphological boundaries on global syllable timing.

1.5 Research questions and hypotheses

This section presents our three research questions and their hypotheses.

Research question 1 (RQ1): What coordinative pattern is found in onsets in Georgian?

Hypothesis 1 (H1): Georgian will display the coordinative pattern predicted by the coupled oscillator model. Onset consonants will be in-phase with the nucleus and anti-phase with one another. This will result in the c-center effect in complex onsets. In simple onsets, an in-phase relationship between the onset and nucleus will be observed.

Research question 2 (RQ2): What is the relationship between the sonority shape of a complex onset in Georgian and the global organization of the constituent gestures?

Hypothesis 2 (H2): Onsets of all sonority shape will display the c-center effect. The effect of sonority will be seen in the variability—measured by standard deviation—for each measure. Sonority rises will be the least variable shape and falls the most variable. Sonority plateaus will be an intermediate case. We predict this direction for the sonority effect because sonority rises are the shape best suited to the parallel transmission of information (Mattingly, 1981) and are the only sonority shape permitted in other languages where the c-center effect has been observed.

Research question 3 (RQ3): What is the relationship between the morphological composition of an onset in Georgian and the global organization of the constituent gestures?

Hypothesis 3 (H3): The presence of a morphological boundary in a complex onset will have no effect on global timing. First, there is no evidence from previous research on Georgian that multimorphemic consonants sequences are in any way different from their monomorphemic counterparts. Additionally, Georgian does not have derived boundary effects or both concatenative and non-concatenative morphology, and these are the only cases in which there is strong evidence for a relationship between morphological structure and articulatory timing (Cho, 2001; Shaw et al., 2010).

2. Methodology

2.1 Procedure and participants

To test our hypotheses about gestural coordination, we targeted collection of kinematic data from native speakers of Georgian. Three speakers, two female (F1 and F2) and one male (M3), participated in our study before the Covid-19 pandemic interrupted data collection. Participants were recruited via word of mouth and an announcement circulated by the Georgian Cultural and Educational Center of Southern California. All three participants were in their twenties and had acquired Georgian as a first language, though all three had different linguistic backgrounds beyond that fact.

Data were collected using an AG501 Electromagnetic Articulograph (Carstens Medizinelektronik). Electromagnetic articulograph (EMA) data allows us to track the movement of the tongue and lips at a millisecond level of temporal resolution. This tracking is done by affixing sensors to key points in the vocal tract. The sensors' movement is recorded as they move through a weak electromagnetic field generated by the EMA.

For this study, we attached sensors at three points along the midsagittal line of the tongue: one on the tongue tip (TT), one on the tongue dorsum (TD), and one in the middle between these two sensors, i.e., at the center of the tongue body (TB). The vertical displacement trajectory of the TT sensor was used to label coronal (alveolar and postalveolar) consonants, and the vertical displacement of the TD sensor was used to label velar and uvular consonants. TD displacement using tangential velocity criteria was also used, where possible, to label vowel gestures. Two sensors were attached on the lips, one on the upper and one on the lower lip. The Euclidean distance between these sensors at each timepoint was calculated during data analysis, and the resulting variable—Lip Aperture (LA)—was used for labial segments. Sensors were also attached on the upper and lower incisor for reference and jaw movement respectively, and on the bridge of the nose and behind each ear for head correction.

Simultaneous audio data were collected using a Shure SCM262 microphone mixer at a 16kHz sampling rate, with a Sennheiser shotgun microphone positioned a foot away from the participant's mouth.

Participants read stimuli off a computer screen positioned approximately four feet away from the EMA. All stimuli were presented in the Georgian orthography. In Sections 3 and 4, we present the stimuli in IPA, but Georgian orthographic forms are available in the Appendix. Consent forms and instructions were given in English.

2.2 Analyses

Data were rotated relative to the occlusal plane and automatically head-corrected prior to labeling. Data were semi-automatically labeled using custom software in MATLAB (Mark Tiede, Haskins Laboratory). As mentioned above, labial consonants were labeled on the Lip Aperture (LA) trajectory, which was automatically calculated across the recordings using the Euclidean distance between the upper and lower lip sensors.

The following timepoints were identified using velocity criteria for all labeled gestures: gestural onset; constriction achievement; constriction release; and gestural offset. The timepoints of peak velocity for the formation phase (i.e., the interval between onset and release) and the release phase (i.e., the interval between release and offset) were also labeled for each gesture. Constriction maxima were identified using velocity minima, and other timepoints were identified using thresholds of velocity ranges between alternating velocity extrema (i.e., between a maximum and a minimum or vice versa). For all timepoints and tokens the velocity threshold was set to 20%. Figure 4 illustrates these labels.

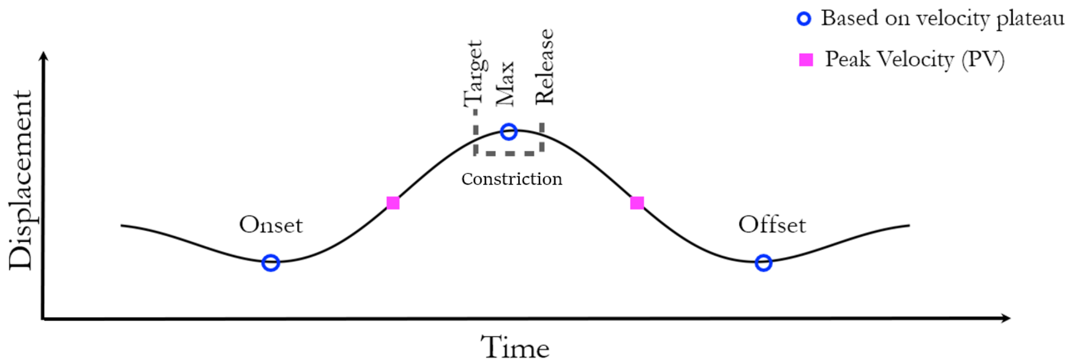


Figure 4. Timepoints labeled for each gesture.

These timepoints were used in the calculation of two measures. The first assesses the presence of rightward shift, and the second the stability of the c-center.

In theory these measures should be assessed by comparing the onset of Cn and the c-center point to the onset of the vowel gesture, but it was not possible in our data to consistently label the vowel gesture for all tokens of a test word. For this reason, we use an anchor point outside of the onset. The use of an anchor outside of the onset is widespread in studies investigation the c-center effect because of the difficulty of labeling a discrete gesture for the vowel. Many studies use anchors in the post-vocalic consonant (e.g., Browman and Goldstein, 1988; Kühnert et al., 2006; Pouplier and Beňuš, 2011; Shaw et al., 2011), but this was not an option for the present study because our test words include open monosyllables, such as /sma/.

Instead, we selected an anchor that occurs during the vowel of the syllable under analysis. This has the additional benefit of making our analysis more directly comparable, in terms of raw durational values, to other studies on Georgian syllable structure. Goldstein and colleagues (2007) use the articulatory target of the vowel gesture as their anchor point, and

Hermes and colleagues (2020) use the peak velocity of the lip aperture movement during the acoustic vowel. For our study, we have chosen to use the time of peak loudness of the vowel as the anchor. Peak loudness was defined as the time of highest root mean square value minus the zero-crossing rate (RMSZCR), which removed noise from frication. RMSZCR was calculated across each trial using Mview. This gives us a timepoint that can be considered the most sonorous part of the syllable. Figure 5 shows an example RMSZCR label. Post-hoc analyses also showed that RMSZCR also correlated well with the peak velocity of lip aperture during the vowel, as used by Hermes and colleagues.

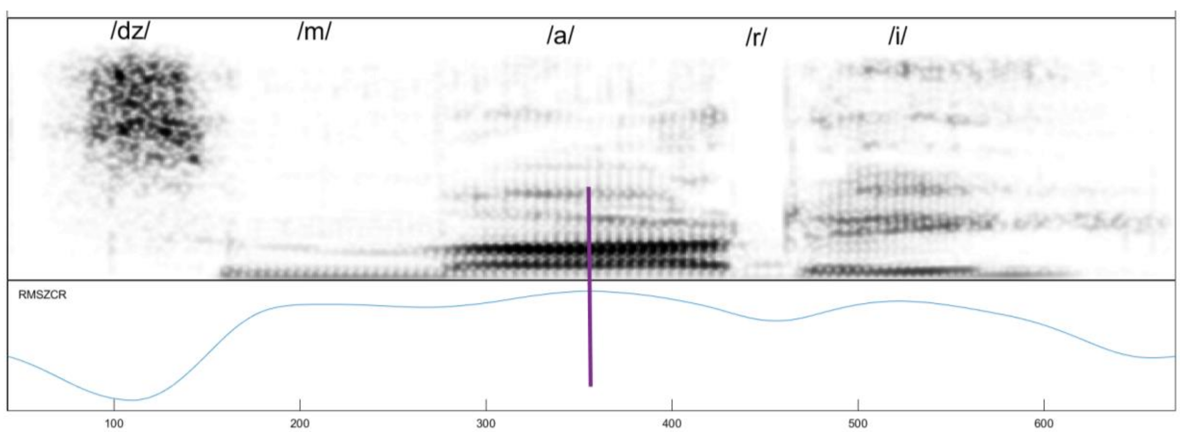


Figure 5. Anchor label for /dzmari/. The solid vertical line shows the timepoint of the RMSZCR maximum, used as the anchor point.

2.2.1 Rightward shift (Cn to anchor)

As discussed in section 1.1, rightward shift is often used as a proxy measure for the c-center effect. The coupled oscillator model predicts that as onset size increases, the Cn gesture will be initiated later and later relative to the C in a CV syllable. To assess the presence of this shift, we measured the time from the onset of Cn to the anchor point (see blue solid line in Figure 6). As onset size increases, the Cn gesture's onset is predicted to be closer to the anchor.

2.2.2 C-center to anchor

We also employ a more direct measure of the c-center effect. By direct, we mean that we calculate a timepoint that corresponds to the location of the c-center itself, and then measure the time from that point to the anchor (see green dotted line in Figure 6). We calculate the c-center as the average of all the consonant gesture onsets. The resulting value is the timepoint which we define as the c-center. We use this manner of calculating the c-center rather than one using constriction midpoints (e.g., Browman and Goldstein, 1989) because, as discussed in section 1.1, the coupled oscillator model makes predictions specifically about gestural onsets, and our goal here is to determine whether the coupled oscillator model's predictions hold true for Georgian.

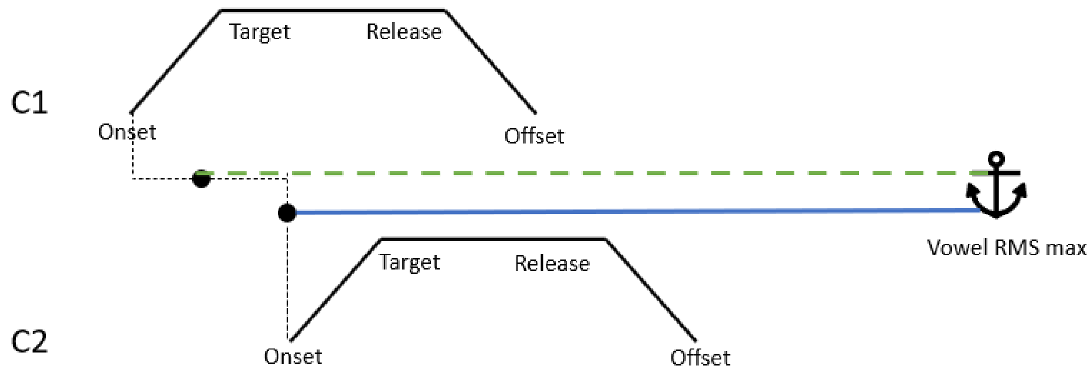


Figure 6. Schematic representation of measures used to assess the presence of the c-center effect. The representation illustrates two C gestures (C1 and C2) along with their labelled timepoints and the labeled anchor. The blue solid line indicates the interval measured for rightward shift, and the green dotted line indicates the interval measured for c-center stability.

2.2.3 Statistical analysis

All measures were evaluated using linear mixed effects models. In each case we began with the maximal random and fixed effects structures. Fixed effects in the maximal models were Sonority Shape (Rise, Plateau, Fall), Phrasal Position (Isolation, Quotative) and Onset Size (C, CC and CCC) for Experiment 1 (see details in Section 3), and Morphological Shape (Simplex, Complex), Sonority Shape (Rise, Plateau, Fall), and Phrasal Position (Isolation, Quotative) for Experiment 2 (see details in Section 4). The maximal random effects structures contained random slopes and intercepts by the following factors: Speaker, Word, Vowel Quality, Postvocalic Segment, and Prevocalic Segment.

Statistical analysis was done in R (R Core Team, 2018) using the lmertest package (Bates et al., 2015). Pairwise comparisons with a Holm correction were done using the emmeans package (Lenth, 2019), and all graphs were made using ggplot2 (Wickham, 2016). The optimal random effects structure was determined using the rePCA function, and the optimal fixed effects structure by the drop1 function.

Variability for each measure was determined by calculating the standard deviation for all words for all speakers, and then running linear mixed effects models on those data. The maximal models had Sonority Shape, Phrasal Position, and Onset Size as fixed effects and random slopes and intercepts by Word and Speaker. Procedure for model selection was the same as it was for the durational measures.

3. Experiment 1: Sonority, onset size, and global timing

Before diving into the particulars of each measure discussed above, we provide a few key generalizations from our results. First, sonority plays no role in either measure reported here. Crouch and colleagues (2020) demonstrate an effect of sonority on the timing between C gestures in the onset, but this appears to be the extent of sonority sequencing's influence on intergestural timing in Georgian. Although not reported here in detail, those results on gestural overlap are replicated for these data as well. Second, and more surprisingly, none of the measures presented here provide evidence for the c-center effect in Georgian.

3.1 Design and stimuli

Stimuli are presented below. The first female participant (F1) was the pilot participant and the test words used were expanded following her participation. The frame sentence used was changed after the pilot as well (see (1) below for Speaker F1, and (2) for Speakers F2 and M3) . The pilot data are included here as data collection was interrupted by the Covid-19 pandemic, making it impossible to bring further participants to the lab. Both sentences and sets of test words are presented below. Speakers produced each stimulus sentence eight times (seven for speaker M3). The frame gives two productions of the test words: the first in an isolation condition, and the second in a quotative condition. The context words were changed following the pilot participant (F1) in order to bracket the test word with /mV/ sequences to minimize blending of gestures and facilitate measurements of all test word gestures.

(1) _____. K'idev _____ vtkvi. [Speaker F1]

_____. again ____ 1sg.subj -say-aor

_____. I said ______ again.

(2) _____. Kalma _____ momts'era. [Speakers F2 and M3]

_____. Woman-erg ____ prvb-1sg.obj-write-3sg.subj.aor

_____. The woman wrote _____ to me.

Tables 2a and 2b show the test words used for the pilot participant and the following two participants, respectively. Consonants are kept the same across size conditions for each sonority shape. CCV words have both the C1C2 and C2C3 pairs from the corresponding CCCV word. Wherever possible, the same vowel is used across size conditions for each sonority shape as well.

Table 2a. Experiment 1 test words for speaker F1.

	C	CC	CCC
sonority rise	**dz**ala ‘strength’ **m**ada ‘hunger’ **r**ezi ‘Rezi (name)’	**dz**mari ‘vinegar’ **mr**etsi ‘downward slope’	**dzmr**iani ‘vinegary’
sonority plateau	**t’**apa ‘frying pan **b**ade ‘net’	**t’b**aze ‘lake.in’ **t’q’**eba ‘lamenting (n.)’	**t’q’d**eba ‘breaking (n.)’
sonority fall	**m**ogebi ‘prize’ **s**opeli ‘village’ **p’**ovna ‘finding (n.)’	**ms**oplio ‘world’ **sp’**oba ‘destroying (n.)’	**msp’**obeli ‘destroying (n).’

Table 2b. Experiment 1 test words for speakers F2 and M3.

	C	CC	CCC
sonority rise	**dz**ala ‘strength’ **m**ada ‘hunger’ **r**ezi ‘Rezi (name)’ **p**andi ‘wrestling move’ **s**ami ‘three’ **ʃ**avi ‘black’	**dz**mari ‘vinegar’ **mr**etsi ‘downward slope’ **sm**a ‘drinking (n.) **ps**alt’a ‘psalm’ **pʃ**avi ‘Pshavia (region)’ **ʃm**agi ‘maniac’	**dzmr**iani ‘vinegary’ **psm**a ‘peeing (n.)’ **pʃm**uis ‘cow snorting’
sonority plateau	**t’**eni ‘damp’ **q’**eli ‘throat’ **d**eba ‘placing (n.)’	**t’q’**eba ‘lamenting (n.)’ **q’d**ebi ‘bindings’	**t’q’d**eba ‘breaking (n.)’

sonority fall	**m**ogebi ‘prize’ **s**opeli ‘village’ **p’**ovna ‘finding (n.)’	**ms**oplio ‘world’ **sp’**oba ‘destroying (n.)’	**msp’**obeli ‘destroying (n).’

3.2 Rightward shift

The rightward shift measure compares the distance between the prenuclear consonant (Cn) and the anchor point across onsets of different sizes and sonority shapes. The hypothesis, based on the predication made by the coupled oscillator model (Nam et al. 2009), is that as onset size increases, the distance between Cn and the anchor point will decrease as a result of the reorganization associated with the c-center effect. In the framework of the spatial metaphor often adopted to talk about speech occurring in linear time, this looks like a rightward shift of Cn along the x-axis.

The final model for rightward shift has random intercepts by Speaker and Word, and fixed effects of Size ($F(2) = 4.186$, $p < .05$), Phrasal Position ($F(1) = 80.769$, $p < .001$), and their interaction ($F(2) = 19.003$, $p < .001$). The difference between phrasal positions is the same across measures for all speakers: phrase-medial productions are closer together in time than isolation productions. No within-phrasal position Size pairs are significantly different; this means that although onsets of every size were affected by their phrasal position, there is no evidence for rightward shift. Figure 6 shows C_n onset time to anchor across Size and Phrasal Position conditions. Table 3 summarizes the fixed effects results.

Table 3. Summary of the linear mixed effects model for rightward shift, with C as the baseline for Size, and Isolation the baseline for Phrasal Position.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	269.92	244.40 – 295.45	**<0.001**
SIZE [CC]	-41.75	-80.57 – -2.93	**0.035**
SIZE [CCC]	-83.88	-130.74 – -37.01	**<0.001**
PHRASE [Quotative]	-52.66	-60.99 – -44.33	**<0.001**
SIZE [CC] * PHRASE [Quotative]	38.99	25.80 – 52.19	**<0.001**
SIZE [CCC] * PHRASE [Quotative]	31.73	15.66 – 47.80	**<0.001**

In the isolation condition we see a tendency for rightward shift as onset size increases, but this tendency disappears in the quotative condition. The pattern in the isolation condition is similar to what is reported in Hermes et al. (2020). That study includes only one repetition of the test word per trial, as a quotative. Because of the similarity between our isolation condition and Hermes and colleagues quotative (the two are similar because in both cases the test word appears as new information), we propose that the observed pattern is not "rightward shift" yielded by c-center, but lengthening effects induced by prosodic prominence likely associated with new information. Presumably, in Georgian that marks prominence through phrasing, gestures are longer and less overlapped initially in the focused word (e.g., Byrd and Saltzman, 1998; Byrd and Choi, 2010; Cho, 2006), and as onset size increases, initial C gestures undergo more prominence-induced effects than the following ones, resulting to patterns similar to rightward shift (see also Sotiropoulou et al., 2015).

The acoustic vowel duration is constant across onset sizes (section 3.5.1), which further supports the analysis that the trend observed in the isolation condition is not the same as the rightward shift associated with the c-center effect. The gestural scores (Figures 19-21a-

c) show how differences in constriction plateau lag and C gesture duration can yield results that, when measured as they were above, mimic rightward shift caused by the c-center effect. This means that we do not find compelling evidence for rightward shift, and this absence suggests that the c-center effect does not occur in our data.

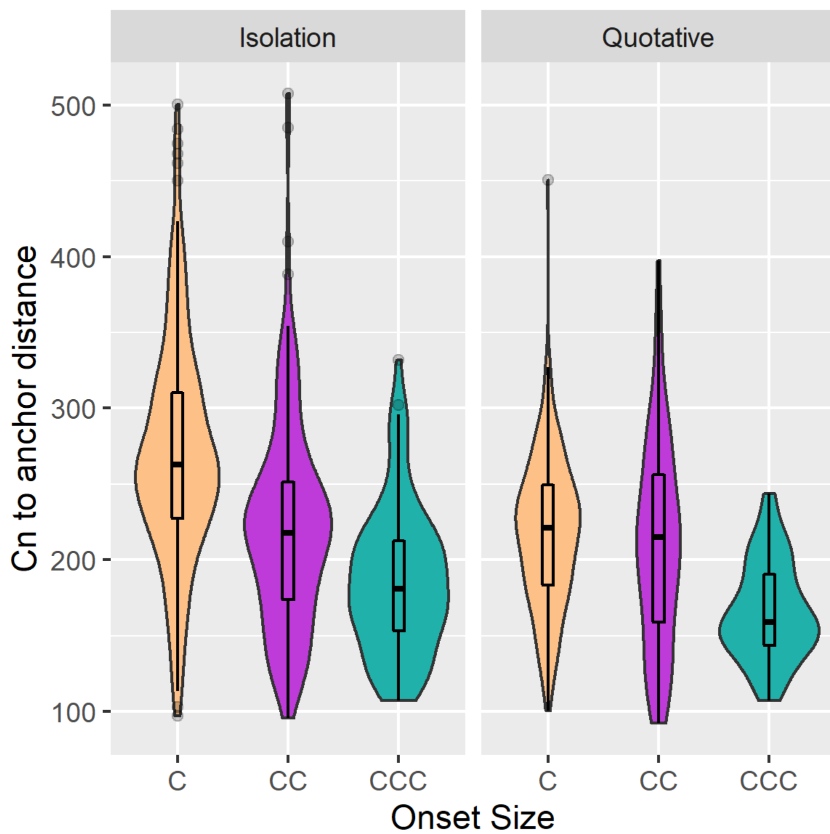


Figure 7. Rightward shift results as a function of onset size and phrasal position.

3.3 C-center to anchor

As a reminder, we use an onset-based measure of c-center because the coupled oscillator models makes predictions specifically about the onset of gestures relative to one another.

Specifically, we define the c-center as the average of the onset of each consonant gesture, as show in Figure 5.

The final model has random intercepts by Speaker and Word, and fixed effects of Size ($F(2) = 9.293$, p =. 001) and Phrasal Position ($F(1) = 209.462$, $p < .001$). Table 4 summarizes the fixed effects results. As with the rightward shift measure, the difference between phrasal conditions is that gestures in the isolation condition are further apart from one another. In terms of onset size, C-CC and C-CCC onsets are significantly different from one another (β = -42.1, SE = 16.2, $p < .05$ for C-CC, and β = -80.9, SE = 19.6, $p = .001$ for C-CCC), but after correction for multiple comparisons the CC-CCC difference barely fails to reach significance (β = -38.8, SE = 20.5, $p = .07$). The c-center is not a stable distance from the anchor point when the size of the onset changes. This is strong evidence against c-centering in Georgian.

Table 4. Summary of the linear mixed effects model for c-center to anchor distance, with C as the baseline for Size, and Isolation the baseline for Phrasal Position.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	267.81	246.60 – 289.02	**<0.001**
SIZE [CC]	42.13	10.31 – 73.94	**0.010**
SIZE [CCC]	80.93	42.54 – 119.33	**<0.001**
PHRASE [Quotative]	-48.39	-54.95 – -41.83	**<0.001**

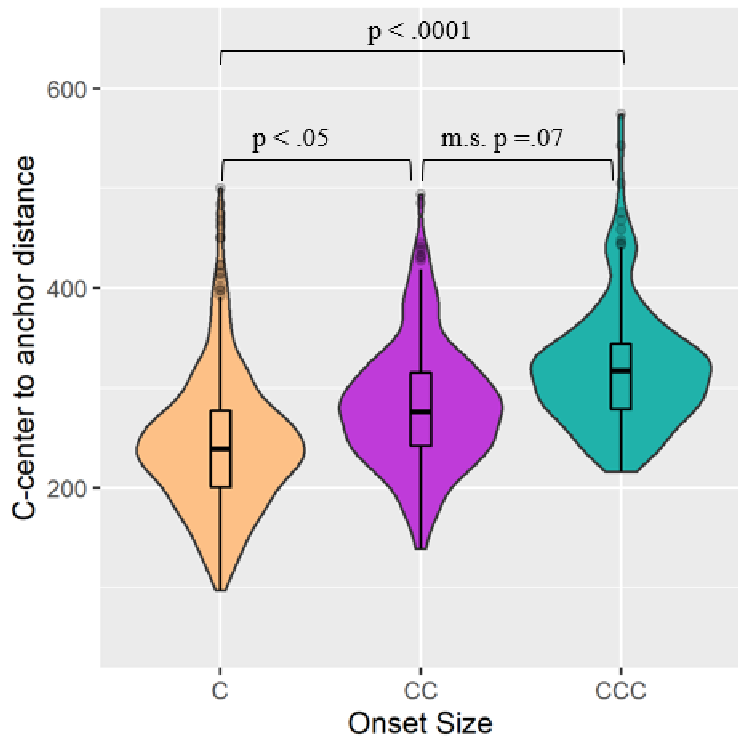


Figure 8. C-center to anchor results as a function of onset size

3.4 Variability

To assess the variability of each Size and Sonority category with respect to each measure, we calculated the standard deviation per word per speaker for both measures reported in section 3.2 and ran linear mixed effects models on those values.

The final model for rightward shift variability has fixed effects of Phrasal Position ($F(1) = 8.308$, $p < .01$) and Size ($F(2) = 8.460$, $p < .01$), and their interaction ($F(2) = 4.689$, $p = .01$), as well as a random intercept by Word. Post-hoc pairwise comparisons found that in the isolation condition one and two consonant onsets and one and three consonant onsets differ significantly, but two and three consonant onsets do not ($\beta = 15.31$, $SE = 4.96$, $p < .05$ for C-

CC, β = 27.13, *SE* = 5.84, $p < .001$ for C-CCC). The larger the onset, the less variable the measure of rightward shift. In the quotative there is no difference in variability between size conditions. In the quotative position rightward shift is less variable than in the isolation condition, which could be attributed to the prosodic factors—prominence marked exclusively by phrasing, and the new vs. given information distinction—discussed in 3.2 Variability in the c-center to anchor measure was only significantly different across phrasal positions. The final model included only Phrasal Position ($F(1) = 18.196$, $p < .001$) and a random intercept by Speaker. As with rightward shift, the isolation condition is more variable.

Sonority was not relevant for either measure, which disproves our second hypothesis, that the well-suitedness of sonority rises to the parallel transmission of information will result in their being the least variable. This suggests that sonority shape is not relevant at all for global onset organization in Georgian. Phrasal position was the most significant factor for both measures, and for both the quotative position was less variable. We can attribute this to the fact that the quotative was always the second repetition of the test word, and not focused, unlike the isolation condition. Gestures of greater magnitude, which occur in the focused condition, are more variable (e.g., Lammert et al., 2016). Repetition also results in less articulatory variability (Tomaschek et al., 2021).

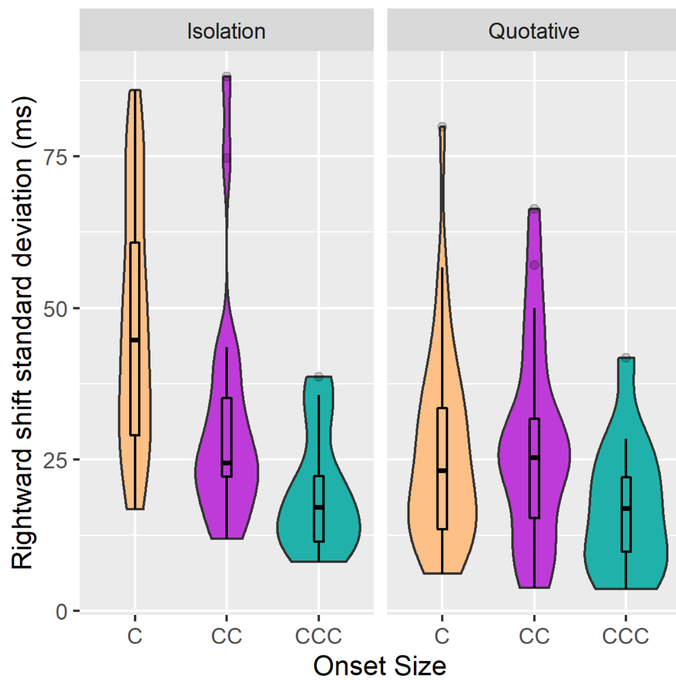


Figure 9. Standard deviation for rightward shift by onset size and phrasal position

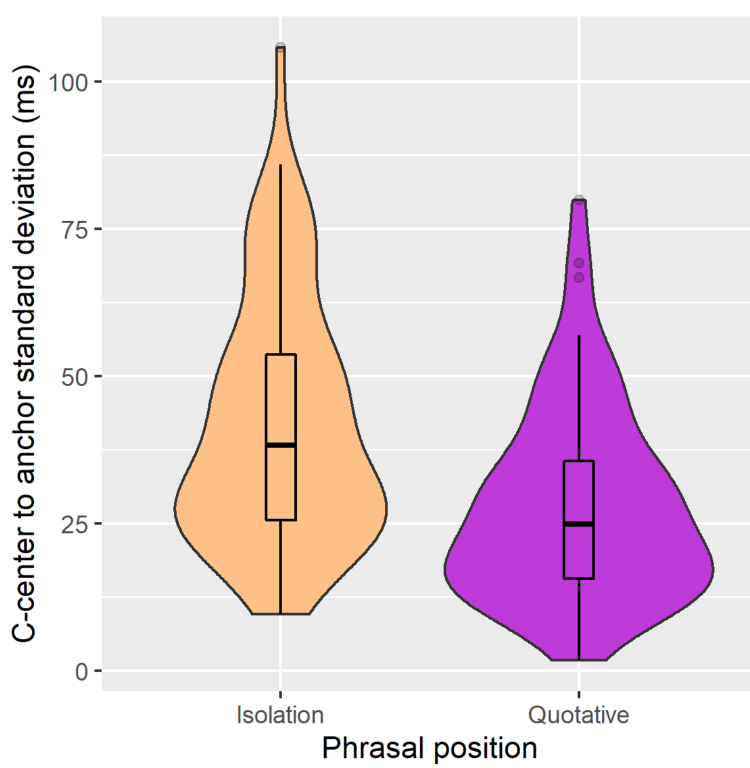


Figure 10. Standard deviation for c-center to anchor by phrasal position

3.5 Secondary measures

Neither the measure of rightward shift nor the measure of c-center stability detected the c-center effect. Given the unexpected nature of these results, we confirm them with two additional secondary measures of the c-center effect.

3.5.1 Acoustic vowel duration

An additional indicator of the presence of c-centering is the shortening of the acoustic vowel in syllables with longer onsets. Marin and Pouplier (2010) show that this is indeed the case in English. In larger onsets, as the prenuclear consonant shifts later in time, it overlaps more with the vowel gesture, resulting in a shorter acoustic vowel. In Georgian, we find no effect of onset size on the duration of the acoustic vowel. Sonority shape of the onset and phrasal position also have no effect.

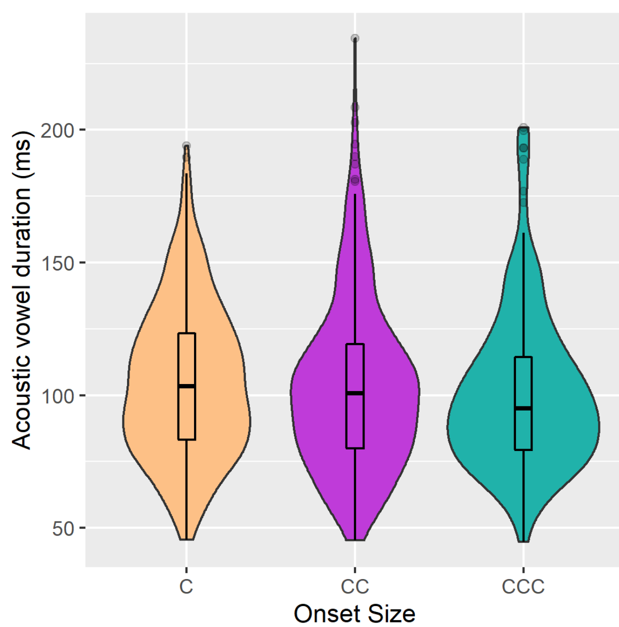


Figure 11. Acoustic vowel duration by onset size

3.5.2 C-gesture duration

Sotiropoulou and colleagues (2020) demonstrate, using data from Peninsular Spanish, that the global stability of the onset-nucleus relationship may not always be capturable through global heuristics like the c-center to anchor distance. They show instead that local variance—that is, durational differences in consonant gestures depending on their position in the onset—can be indirect indices of global organization in the onset.

The duration of each C gesture was measured in milliseconds, and the data were analyzed using lmertest in R, with fixed effects of consonant identity (levels: /d/, /dz/, /m/, /r/, /p/, /p'/ /q'/ /r/ /s/ /ʃ/ /t'/) and position in the onset (levels: sole, initial, medial, final). The position in the onset necessarily specifies the number of consonants in the onset as well. Random intercepts by Speaker and Word were also included. Consonant identity and its interaction with position in the onset were significant in the model ($F(15) = 10.175$, $p < .001$), but post-hoc pairwise comparisons showed that no single C gesture differed in its duration depending on its position in the onset. Figure 12 shows the duration of /m/, /s/, and /ʃ/ gestures in all positions. These three C gestures were selected because they are the most frequent, and appear in all four possible positions unlike, for example, /dz/, which is only the sole or the initial C gesture in the onset.

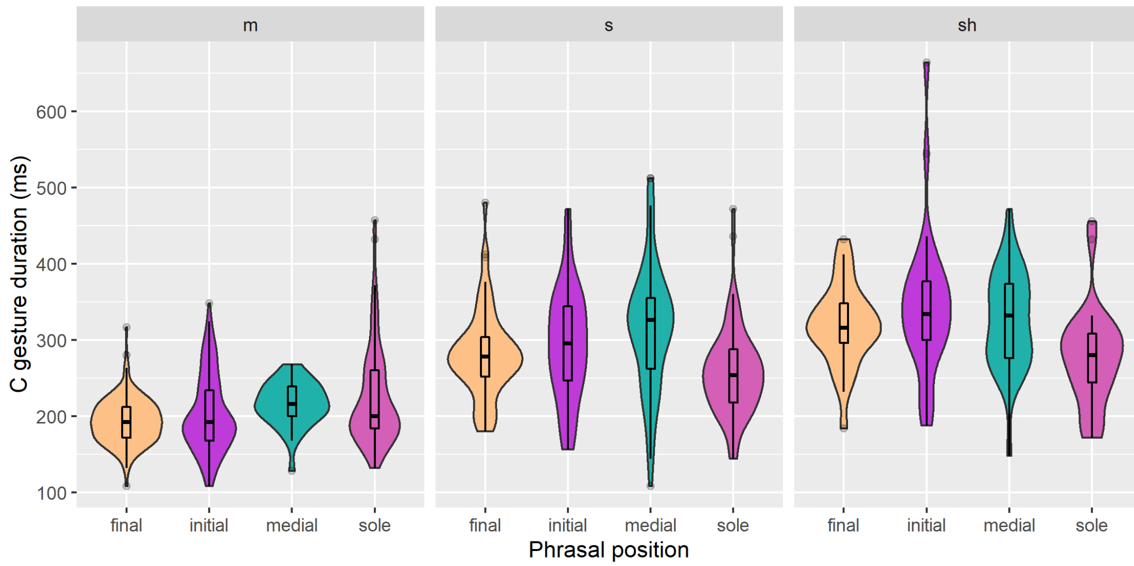


Figure 12. C gesture durations by position in the onset

3.6 Vocoids

To confirm that the failure to detect the c-center effect was not due to the presence of vocoids in several of our complex onsets, we *post hoc* subsetted our data based on the presence of an intrusive vocoid. We then ran the same analyses described in the previous subsections on the subset of our data with no vocoids. If the presence of vocoids were disrupting the c-center effect, then in the vocoid-free data set we should observe the timing patterns we originally predicted. CV words are present to serve as the baseline for Size condition comparisons.

26% of all rises (CCV and CCV) had intrusive vocoids, as did 60% of plateaus. For sonority rises, 21% of CC onsets and 37% of CCC onsets had intrusive vocoids. For sonority plateaus 66% of CC onsets and 57% of CCC onsets had intrusive vocoids. The plateau percentages are much higher, but this is possibly due to a) not controlling for order of place of articulation, which affects gestural overlap in Georgian (Chitoran et al. 2002, Crouch et al., to appear) and b) the dropping of /t'q'eba/ due to its having the only onset with exclusively ejective consonants. Table 5 summarizes the distribution of vocoids. No vocoids were found in sonority falls.

Table 5. Distribution of vocoids by sonority shape and onset size

	CC	CCC
Sonority rise	41/196 (21%)	34/91 (37%)
Sonority plateau	19/29 (66%)	25/44 (57%)

3.6.1 Rightward shift

The rightward shift results for the non-vocoid subset are strikingly similar to the results for the entire data set (Section 3.2). The final model has the same random and fixed effects structure as the model described in Section 3.2: fixed effects of Size ($F(2) = 4.356$, $p < .05$), Phrasal Position ($F(1) = 56.657$, $p < .001$), and their interaction ($F(2) = 14.742$, $p < .001$), and random intercepts by Word and Speaker. Table 6 provides a summary of the fixed effects. Rightward shift occurs even less in the non-vocoid set than in the entire data set. Figure 13 illustrates this. Although the isolation condition shows a tendency towards rightward shift, the pattern does not reach significance and the same tendency is not apparent in the quotative condition. As detailed in Section 3.2, the tendency towards rightward shift in presumably due to prosodic prominence and not syllabic structure. We again find no evidence for the presence of the c-center effect, and the strong similarity between the non-vocoid subset and the overall data set suggest that the absence of the c-center effect in our data is not due to the presence of intrusive vocoids.

Table 6. Summary of the linear mixed effects model for rightward shift for the no-vocoid subset, with C as the baseline for Size, and Isolation the baseline for Phrasal Position.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	270.80	243.47 – 298.13	**<0.001**
SIZE [CC]	-28.10	-51.16 – -5.05	**0.017**
SIZE [CCC]	-58.14	-87.90 – -28.38	**<0.001**
PHRASE [2]	-52.67	-61.28 – -44.05	**<0.001**
SIZE [CC] * PHRASE [2]	37.49	22.85 – 52.12	**<0.001**
SIZE [CCC] * PHRASE [2]	32.78	13.20 – 52.36	**0.001**

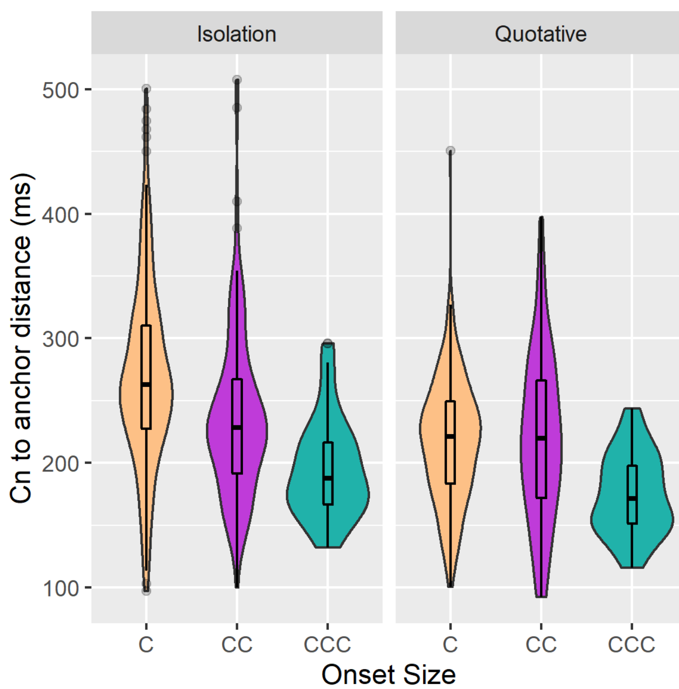


Figure 13. Rightward shift by onset size and phrasal position, for the non-vocoid subset

3.6.2 C-center to anchor

As with the entire data set (Section 3.3), the non-vocoid subset of the data does not display the expected stability of the c-center. The final model is identical to the final model for the entire data set: Size ($F(2) = 18.294$, $p < .001$) and Phrasal Position ($F(1) = 193.037$, $p < .001$) are significant. Table 7 summarizes the model results. All pairwise Size comparisons are significant ($\beta = -50.6$, $SE = 12.5$, $p < .01$ for C-CC, $\beta = -88.3$, $SE = 16.0$, $p < .001$ for C-CCC, $\beta = -37.7$, $SE = 16.0$, $p < .05$ for CC-CCC). As with the overall data set, the non-vocoid subset does not show c-center stability. This is further evidence that the c-center effect is not occurring in our data, and that this absence is not due to intrusive vocoids.

Table 7. Summary of the linear mixed effects model for c-center to anchor distance for the no-vocoid subset, with C as the baseline for Size, and Isolation the baseline for Phrasal Position.

Predictors	*Estimates*	*CI*	*p*
(Intercept)	268.66	244.29 – 293.03	**<0.001**
SIZE [CC]	50.63	26.60 – 74.67	**<0.001**
SIZE [CCC]	88.35	57.68 – 119.02	**<0.001**
PHRASE [2]	-48.76	-55.65 – -41.87	**<0.001**

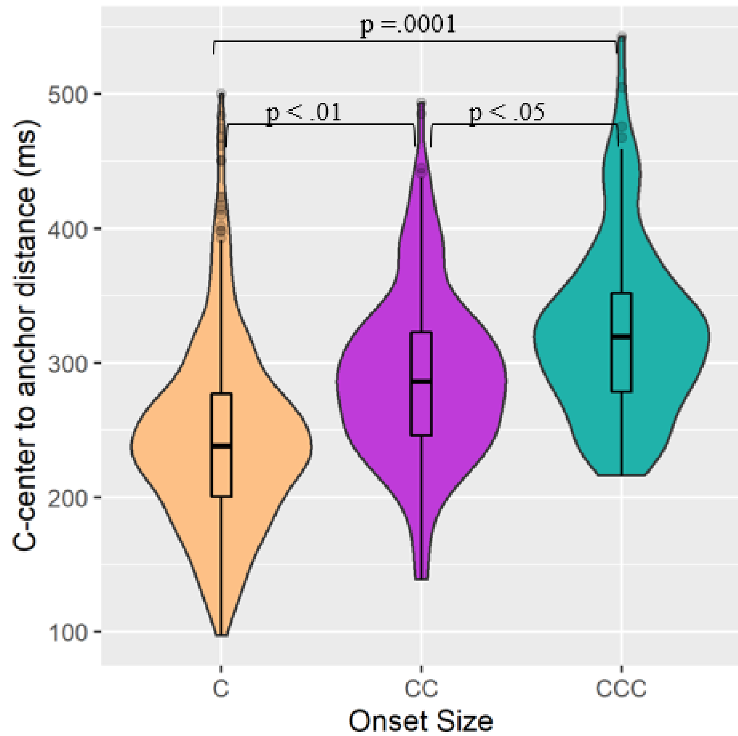


Figure 14. C-center to anchor distance by onset size, for the non-vocoid subset

3.6.3 Variability

Variability in the non-vocoid subset behaved exactly as variability in the overall data set. For rightward shift, the final model has fixed effects of Phrasal Position ($F(1) = 5.096$, $p < .05$) and Size ($F(2) = 5.354$, $p < .05$), and their interaction ($F(2) = 4.244$, $p < .05$), as well as a random intercept by Word. One and three consonant onsets differ significantly in the isolation condition, and the one and two consonant onset comparison nearly reaches significance ($\beta = 15.09$, $SE = 5.30$, $p = .06$ for C-CC, $\beta = 26.04$, $SE = 7.19$, $p < .001$ for C-CCC). In the quotative condition no pairwise Size comparisons are significant. In summary, larger onsets are less variable, and the quotative position is less variable than the isolation production. As with the overall data set, we can likely attribute this to the prosodic factors discussed in Section 3.2 In the c-center to anchor model, only phrasal position had any significant effect (β = -14.491, *SE* = 3.525, $p < .001$, with Isolation as the baseline). As with the rightward shift, the c-center to anchor measure is more variable in the isolation condition.

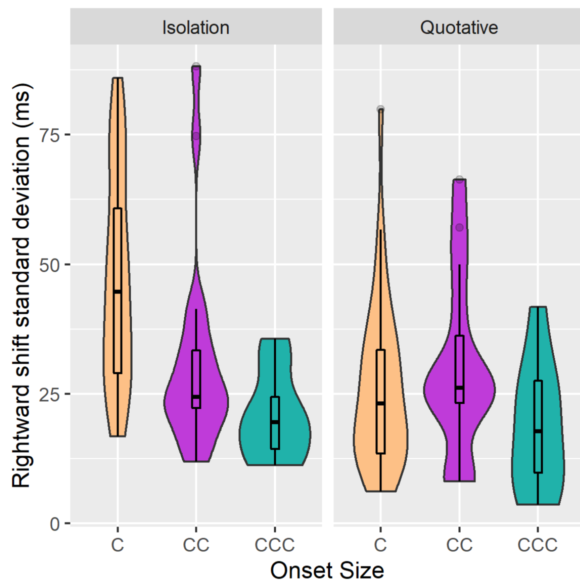


Figure 15. Standard deviation for rightward shift by onset size and phrasal position, for the non-vocoid subset

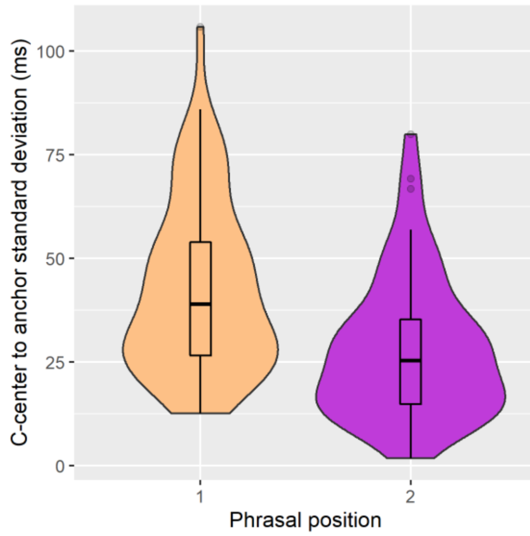


Figure 16. Standard deviation for c-center to anchor by phrasal position for the non-vocoid subset

4. Interim discussion

The analyses presented in Section 3 suggest that the c-center effect does not occur in Georgian. There are three possible scenarios that we can posit based on these results.

1) Georgian does not permit complex onsets. Instead, syllabification in Georgian is more similar to Tashlhyit Berber or certain Arabic languages (e.g., Moroccan Arabic), where CCV and CCCV sequences are di- or tri-syllabic.

2) The syllable is not a relevant prosodic unit in Georgian.

3) Syllable onsets in Georgian are coordinated in a novel way.

Articulatory evidence can be used to argue against (1). Goldstein et al. (2007) directly compared CV, CCV, and CCCV forms in Georgian and Tashlhyit Berber, and concluded that the two languages coordinate (and therefore syllabify) CC(…) sequences differently. The argument in our companion study (Crouch et al., 2020) and the articulatory patterns on which it is based also present evidence for exclusively monosyllabic parses of CC(…)V sequences. We find systematic differences in C-C overlap depending on the sonority shape of a CC onset. All C-C sequences in Georgian show lag between constrictions, but in sonority falls C2 begins earlier than it does in either plateaus or rises. This tighter overlap significantly reduces the number of intrusive vocoids that appear in falls versus rises and plateaus, where the open constriction allows for vocoids to emerge between consonant constrictions. We argue that vocoid avoidance is driven by the need to preserve a CCV syllable parse.

Multiple additional strands of evidence support an analysis of Georgian wherein the syllable is a relevant prosodic unit and that CC(…)V sequences are monosyllables with no appendices or extrasyllabic consonants. First, prosodic studies of Georgian (Borise and Zientarski, 2018; Bush 1999; Skopeteas and Féry, 2016; Vicenik and Jun, 2014) treat the syllable as it has traditionally been defined in Georgian and see no difference in the location or duration of tonal events based on the number of consonants in a syllable onset. Georgian-language haiku also show that syllables are identifiable to speakers and examination of haiku also provide evidence against the analysis in (1). CCCCV syllables, for example, appear in these poems and a non-monosyllabic parse of them would result in incorrectly metered haiku (c.f. e.g., http://arilimag.ge/მალხაზ-მაჭავარიანი-ჰაიკ).

In the following sections we will examine the relationship between morphological structure and global timing in light of the unexpected results in Section 3. We will then discuss C-V coordination and finally present a novel account of how syllables are coordinated in Georgian.

5. Experiment 2: morphology and timing

In this section we present the design and results of a second EMA experiment probing the relationship between morphological structure and global timing in the onset. Test words for this experiment are listed in Table 8. The independent variables for this experiment were Sonority Shape (3 levels: Rise, Plateau, Fall), and Morphological Shape (2 levels: Simplex, Complex).

Table 8. Experiment 2 test words for all speakers.

	morphologically simplex	morphologically complex
sonority rise	**gm**anavs ‘s/he fills it’	**g-m**alavs ‘s/he hides you’
sonority plateau	**gb**oba ‘boiling (n)’ **gd**eba ‘lie/lay about’ **mn**eoba ‘management’	**g-b**eravs ‘s/he inflates you’ **g-d**evs ‘s/he places you’ **m-n**ebavs ‘s/he allows me’
sonority fall	**md**are ‘worthless’ **mt**eli ‘whole, entire’	**m-d**arebs ‘s/he compares me’ **m-t**elavs ‘s/he tramples me’

5.1 Rightward shift and c-center to anchor

Morphological makeup of the onset has no effect on either of the timing measures used to assess the presence of the c-center effect. Sonority also has no effect, as was the case for the experiment reported in Section 3.

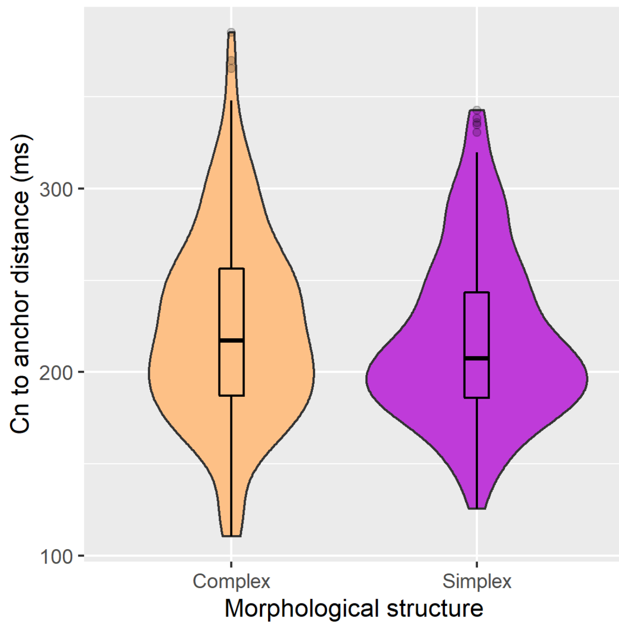


Figure 17. Rightward shift by morphological shape

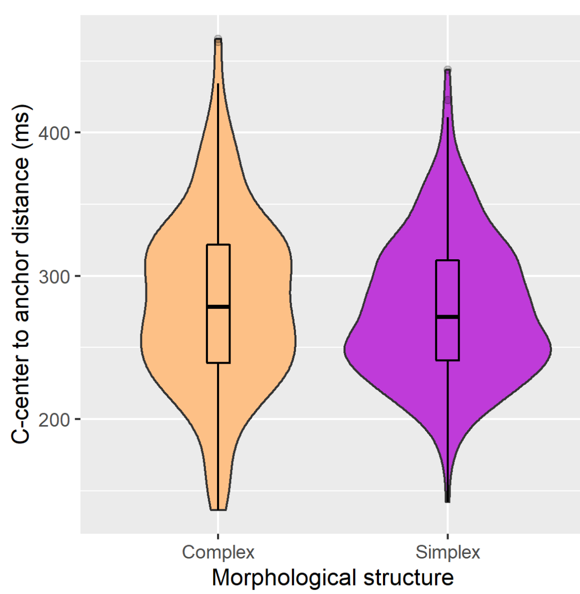


Figure 18. C-center to anchor distance by morphological shape

These values are comparable to those of the CC onsets in the previous experiment (Sections 3.2 and 3.3). The key finding of this experiment is that onsets with consonantal prefixes have the same timing pattern as monomorphemic onsets.

5.2 Interim discussion

The results of our second experiment confirm our hypotheses about timing and morphological structure (H3). The morphological makeup of a complex onset has no effect on the timing of the onset. As mentioned previously, inflectional morphology can increase onset size frequently in Georgian, so these results have important implications for experimental designs going forward. Since it has been established that mono- and bi-morphemic onsets are timed in the same way, future research can intermingle morphologically simplex and complex onsets in order to create data sets that contrast relevant factors while avoiding nonce words.

6. C-V coordination

The coupled oscillator model proposes an in-phase coordinative relationship between the C and V in simple CV syllables. This means that the two gestures (C and V) are predicted to start simultaneously. In languages where this interval has been reported (e.g., Katsika et al., 2014 on Greek), we see lags of around 20 milliseconds between the two gestural onsets. As we will illustrate in this section, we see very different lags in Georgian.

Because we could not label clear vowel gestures for all words and all tokens, we do not present a statistical analysis here. Instead, we construct gestural scores for an illustrative CV-CCV-CCCV trio. We present separate gestural scores for each speaker and each phrasal position. Tables 9, 10, and 11 show the measurements used to generate these gestural scores and provide the necessary data to construct gestural scores for any of the other words where vowel TD gestures could be labeled. The tables present the average duration for each measured interval for each phrasal position. The isolation condition value is presented first, and the quotative condition is presented second. They are divided by a slash. The number of tokens for each condition is included in parentheses below the word in the rightmost column of each table.

The measures used are the duration of each gesture, and the distance between each sequential pair of gestures. All distances are measured between gestural onsets, and the c-center was calculated as the mean of all consonant gesture onsets. Cn refers to the rightmost consonant. Greyed out cells indicate values that are not relevant for a given word. For CV words, the c-center-to-vowel cell is greyed out because that value is identical to the Cn-V distance.

Table 9. Values used to construct gestural scores for speaker F1

F1	Cn dur	Cn-1 dur	Cn-2 dur	V dur	c-center-V	Cn-V	Cn1-Cn	Cn2-Cn1
dzala (6/5)	291.5 / 287.2			143.8 / 236.8		186 / 192.8		
mada (4/6)	206 / 184.7			190 / 186.7		86 / 76		
rezi (4/3)	174 / 149.3			190 / 174.7		49 / 56		
dzmari (7/7)	213.1 / 190.3	306.8 / 302.2		215.4 / 225.7	180.6 / 167.1	105.7 / 84	149.7 / 166.3	
dzmriani (5/6)	156.8 / 170.8	227.2 / 215.3	328 / 274.7	223.2 / 212	105.1 / 97.1	53.6 / 40.7	178.4 / 168	202.4 / 166.7
t'eni (6/7)	249.3 / 201.7			286 / 241.1		189.3 / 136.6		

Table 10. Values used to construct gestural scores for speaker F2

F2	Cn dur	Cn-1 dur	Cn-2 dur	V dur	c-center-V	Cn-V	Cn1-Cn	Cn2-Cn1
dzala (7/8)	262 / 219.4			210.3 / 222.5		184.6 / 151.9		
mada (7/6)	266.3 / 212			297.7 / 226.7		39.5 / 76		
rezi (8/7)	149 / 157.7			175.5 / 180.6		81 / 77.1		
dzmari (8/8)	225 / 203.2	355 / 245.6		278 / 258.5	121.75 / 112.5	46 / 51.5	162.5 / 116	
dzmriani (8/8)	145 / 147.5	228.7 / 215.5	298 / 271.5	271 / 260.5	230.3 / 190.8	46.5 / 33.5	174.5 / 145	202.5 / 182
pandi (8/7)	259.5 / 203.7			282 / 243.4		130.5 / 102.5		
sami (4/6)	345 / 260			289 / 235.3		266 / 186.7		
sma (6/5)	224.7 / 194.5	397.3 / 333.6		322.5 / 209.1	193.25 / 186	114.7 / 141.6	244.7 / 152	
psalt'a (3/3)	329.3 / 298.7	308 / 194.7		188 /176	337.3 / 242.7	278.7 / 229.3	117.3 / 26.7	
psma (8/8)	207.5/164	325 /332	216 / 176.5	318.5 / 230.5	226.5 / 185.2	74.5 / 68	161.5 / 151	133 / 49.5
t'eni (4/7)	241 / 206.9			247 / 216		174 / 124		
deba (8/8)	272.5 / 213			264.5 / 235.5		192.5 / 139.5		

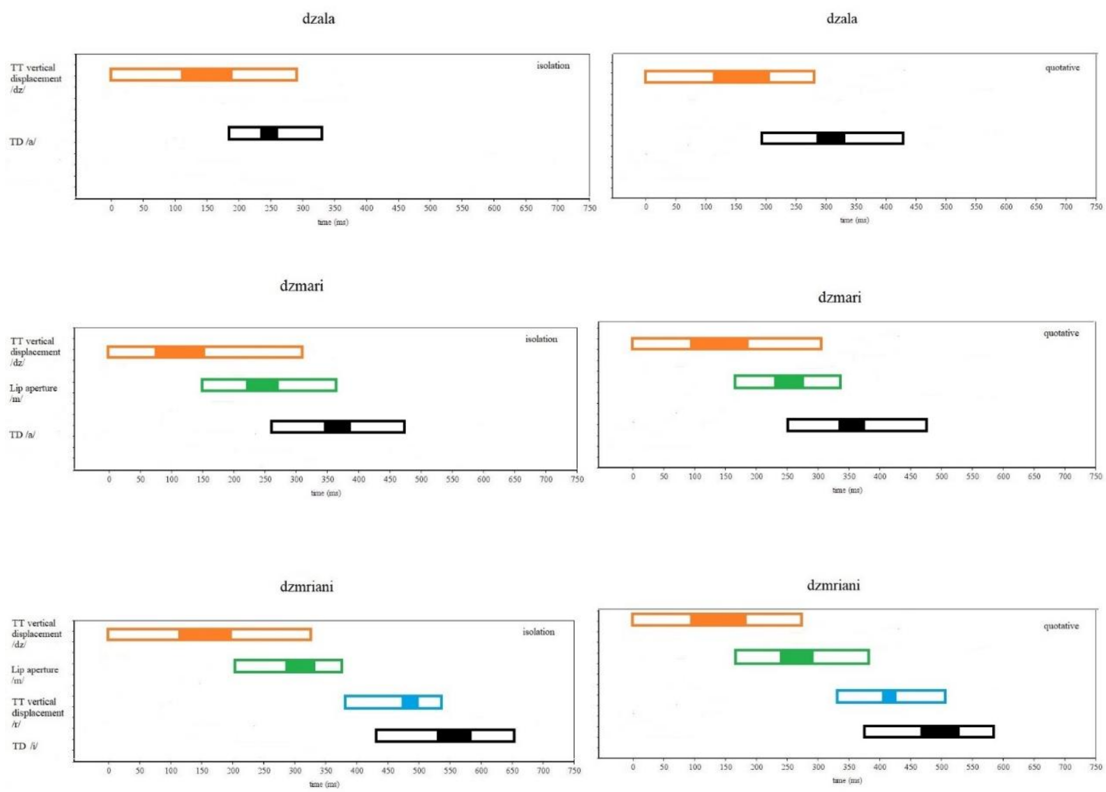
Table 11. Values used to construct gestural scores for speaker M1

M1	Cn dur	Cn-1 dur	Cn-2 dur	V dur	c-center-V	Cn-V	Cn1-Cn	Cn2-Cn1
dzala (1/4)	264 / 198.9			160 / 128		164 / 131		
mada (5/2)	264 / 182			236.8 / 178		109.6 / 76		
rezi (3/0)	162.7			168		76		
dzmari (1/3)	176.6 / 180	176 / 224		388 / 230.7	8 / 106	-44 / 54.7	104 / 102.7	
dzmriani (2/2)	166 / 161.9	184 / 220	278 / 364	200 / 168	70.7 / 82	66 / 54	132 / 166	250 / 248
sma (4/2)	199 / 176	254.7 / 180		244 / 234	121.5 / 110	53 / 64	137 / 128	
psalt'a (1/5)	320 / 246.4	200 / 182.4		172 / 157.6	312 / 198	248 / 186.4	128 / 23.2	
psma (5/2)	203 / 166	252 / 284	172 / 170	313.4 / 212	123.5 / 188	76.8 / 90	131.2 / 164	122.4 / 34
t'eni (1/5)	224 / 200			228 / 150.4		172 / 156.8		
deba (1/2)	264 / 170			240 / 220		160 / 104		

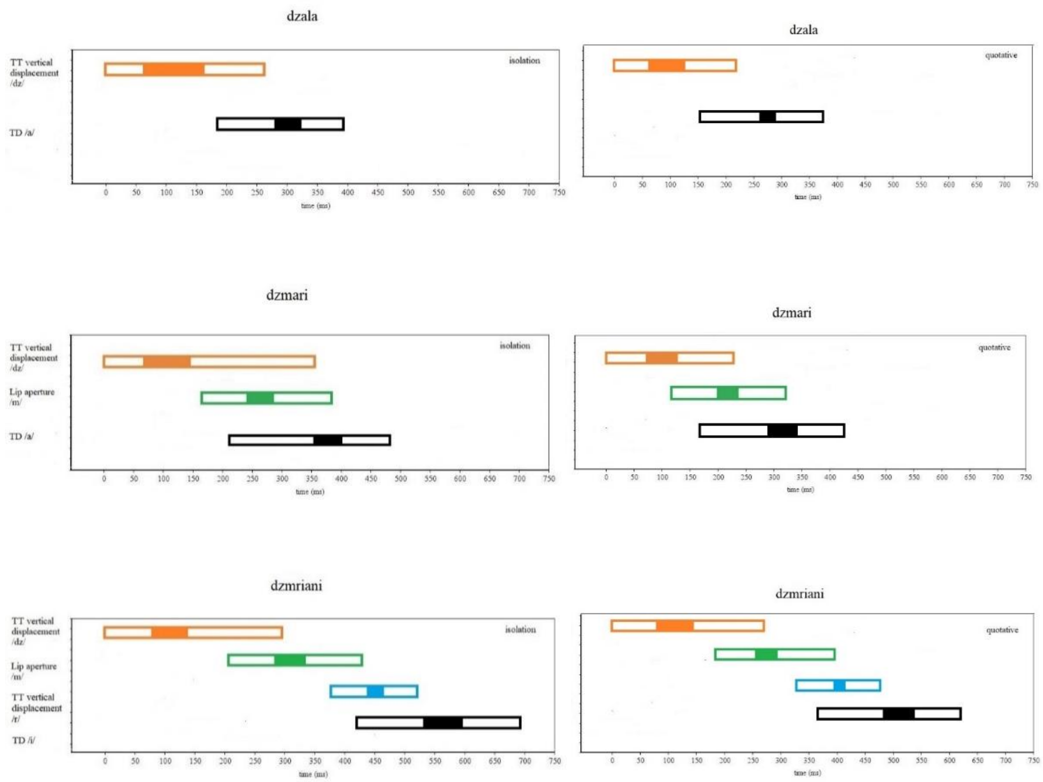
Figures 19-21a (first row of each set) are gestural scores for the first syllable of /dzala/ as produced by speakers F1, F2, and M1, respectively. The orange bar represents the TT /dz/ gesture, and the black bar represents the TD /a/ gesture. The gestural score on the left is for the isolation condition, and the score on the right is for the quotative. In both positions the vowel gesture is initiated after the consonant reaches its target, for all speakers. This is true for all CV words represented in the table above.

Figures 19-21b (second row) and 19-21c (third row) show gestural scores for the first syllables of /dzmari/ and /dzmriani/ in the same format described for Figures 19-21a. The orange bar represents the TT gesture for /dz/, the green bar represents the LA gesture for /m/, the blue bar represents the TT gesture for /r/, and the black bar represents the TD gesture for the vowel. In these figures it is clear that all gestures in the onset and nucleus occur sequentially. The notable exception is M1’s isolation production of /dzmari/. This gestural score is based on a single production, which does not reflect how M1 times other complex onsets, as can be seen in the quotative /dzmari/ gestural score. The two /dzmari/s provide an excellent contrast between the pattern we expected to find in all complex onsets, and the pattern we did find, instead.

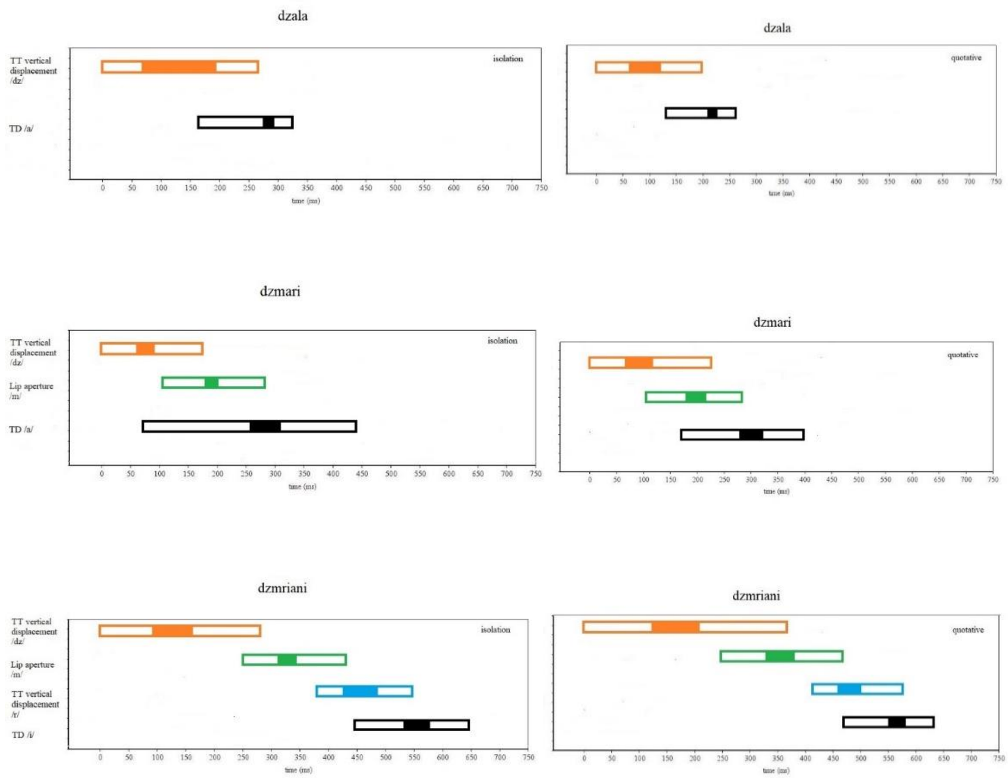
These gestural scores clearly do not show the c-center effect, as the score in Figure 2b does. The V gesture consistently begins after the onset of the prenuclear consonant, even in CV syllables. These data suggest that even simplex onsets, Georgian does not have in-phase coordination, which sets it apart from all other languages studied in the paradigm to date.



Figures 19a-c. Gestural scores for /dzala/, /dzmari/ and /dzmriani/ respectively from speaker F1. In each row the left gestural score is the isolation condition, and the right is the quotative.



Figures 20a-c. Gestural scores for /dzala/, /dzmari/ and /dzmriani/ respectively from speaker F2. In each row the left gestural score is the isolation condition, and the right is the quotative.



Figures 21a-c. Gestural scores for /dzala/, /dzmari/ and /dzmriani/ respectively from speaker M1. In each row the left gestural score is the isolation condition, and the right is the quotative

7. Overall discussion

As mentioned in the interim discussion (Section 4), we do not see the coordinative patterns in Georgian that are predicted by the coupled oscillator model (Nam et al., 2009). Georgian is not the first language where CC(…)V sequences are reported to not show the c-center effect, but it is the first language in which these sequences have traditionally been analyzed as complex onsets. Moroccan Arabic and Tashlhyit Berber show no evidence for the c-center effect in CCV or CCCV sequences, but in both languages phonological analysis has shown that these sequences are parsed as di- or tri-syllables. In Georgian, however, multiple strands of evidence suggest that CC(…)V sequences are monosyllables, as we discuss in section 4. This is not as troubling for previous analyses of Georgian as it may seem. Although CCV and CCCV sequences in Georgian, Moroccan Arabic and Tashlhyit Berber show some similarities, there is a critical difference: simple C-V coordination.

The coupled oscillator model predicts in-phase coordination for a simple CV syllable. In the abstract, this means that the consonant and vowel gestures will be activated simultaneously. In practice, there can be a lag between the two gestural onsets of around 20 milliseconds. As shown in Tables 9-11, average C-V onset lags in Georgian range from 40ms to 120ms depending on word and speaker, and standard deviations for these smaller values are around 20ms themselves. This strongly suggests that the C-V coordination in Georgian is not the same as C-V coordination in other languages. The time-to-anchor values that we report in Section 3 are similar to those reported in other research on Georgian (Goldstein et al. 2007; Hermes et al., 2020), but we offer a different interpretation here than those studies present. Figure 18 schematizes our proposal, and shows the differences between syllable coordination in Georgian, English, and Tashlhyit Berber. This is also consistent with the interpretation that Georgian is a long-lag language (Pouplier et al., 2022).

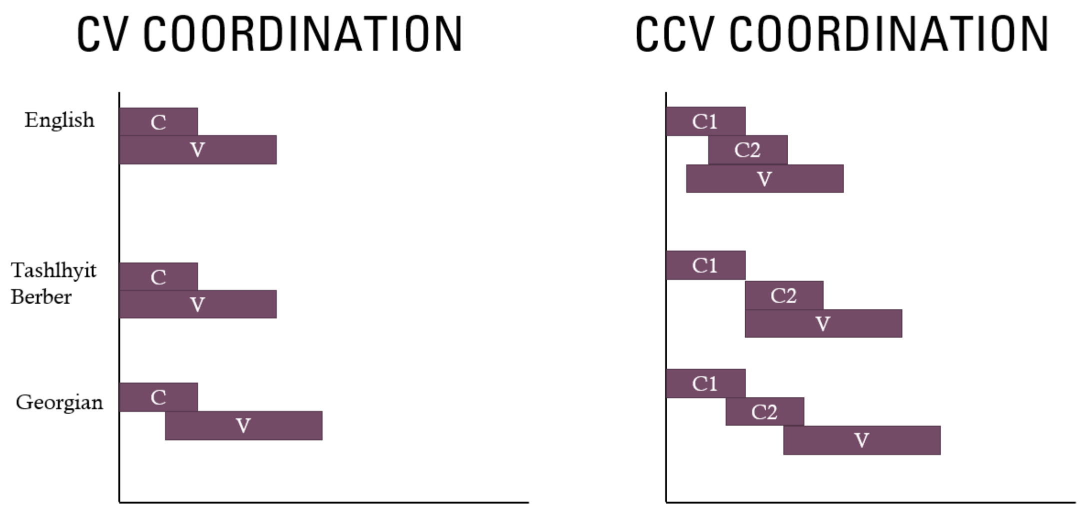


Figure 22. Proposed coordination patterns in English, Tashlhyit Berber, and (newly) Georgian

The pattern illustrated in figure 22 can be attributed to one of two timing patterns: 1) entirely anti-phase coordination that is also entirely local—that is, any given gesture is only coordinated with the immediately preceding gesture; or 2) the competitive coupling illustrated in Figure 2a, but with such weak in-phase coupling that it is essentially overruled by the anti-phase coordination among consonant gestures. While (2) could help explain the increasing overlap between Cn and V gestures in larger onsets, the lag in CV syllables is more difficult to account for. This will be tested by modeling in future research, but the results of both accounts are the same: sequential gestures. In both accounts, local C-C timing differences, such as those found between sonority shapes (Crouch et al., 2020) could be accounted for with a dynamical calculation of coupling strengths. This would function similarly to the dynamical model of syllabification proposed by Goldsmith and Larson (1990). Goldsmith and Larson provide an account of syllabification in Tashlhyit Berber in which syllable nuclei are selected based on their *derived* sonority. The derived sonority of a segment is calculated in relation to the inherent sonority values of adjacent segments, and competition between the two segments is mediated by language specific α and β values. These are local calculations done by a dynamical system with a specific phonological goal. In our model, the strength of the anti-phase coupling between adjacent gestures would be calculated, rather than the derived sonority of a segment.

The mode of Georgian syllabification we propose is clearly motivated by aspects of the grammar of Georgian. We argue that the morphological structure of Georgian is the primary motivator for the anti-phase only timing. The same lexical item may have different syllable shapes as a result of inflection. If coupling graphs are the stored representations of lexical items, then the same verb in Georgian would require multiple coupling graph representations to account for the various inflected forms. Forms where inflection can change onset size include those with first- and second-person core arguments in many tenses, and double-marking participants using overt pronouns is not the most common strategy in Georgian. For any given verb, then, speakers would need to have stored multiple different competitive coupling graphs associated with complex onsets in the coupled oscillator model, which are more and more intricate as onset size increases. In an anti-phase only pattern, however, these consonant-only morphemes can be essentially ‘slotted in’. If there is no difference in the phasing relationships between any two given gestures in Georgian, speakers would not need coupling graphs specific to various inflected forms. Instead, a general phasing pattern could be applied regardless of morphological structure. Anti-phase coordination also allows for more flexibility (see Kelso, 1995 on flexibility in natural systems).

Morphology also brings a perceptual motivation for the absence of c-centering. Both the SSP and the c-center effect support what Mattingly (1981) terms the parallel transmission of information without loss of perceptual recoverability. Although Mattingly (1981) posits that parallel transmission is crucial for both speech production and perception, the fact that Georgian and other languages permit clusters that are not optimal for parallel transmission tells us that parallel transmission is not *de rigueur* for either production or perception. In Georgian and languages with similar morphology perceptual recoverability of, e.g., a word-initial cluster, is not just about lexical retrieval but about transmitting grammatical information that may not be explicitly encoded elsewhere. This provides a powerful push away from what may appear to be more efficient modes of syllable organization.

Anti-phase only coordination is also connected to the phonotactic patterns of Georgian. Because there are no restrictions on onset shapes, there is no need to draw a distinction between codas and onsets in intervocalic sequences. No phonotactic violations can occur in this situation because Georgian essentially bars no sequences. Intervocalic sequences in Georgian are traditionally parsed as VC.C(…)V, although there is disagreement among speakers in syllabification tasks. This parse is likely informed by word-edge probabilities; Georgian has few monomorphemic word-final clusters. The disagreement among speakers for these tasks is expected if there is no phasing difference between codas and onsets; the line would be drawn simply based on word-edge frequencies.

8. Conclusion

We present evidence from a range of measurements showing that Georgian syllables have exclusively anti-phase coordination, and that the presence of a morphological boundary in the onset does not disrupt this timing. This is a novel type of syllable structure and goes directly against the predictions of the coupled oscillator model, but we argue that it is clearly motivated both by the morphological structure and the phonotactics of Georgian, and languages with similarities to Georgian in either of those domains are expected to show timing similar to what we have described for Georgian. Languages with consonant only morphemes would share the perceptual motivations that Georgian has for favoring cue preservation/recoverability more than parallel transmission of information, resulting in the same long lags as those observed in Georgian. Languages with minimal phonotactic restrictions might, like Georgian, not need to use timing to robustly distinguish onsets and codas, since there are no illicit clusters for word-medial consonant sequences to be errantly syllabified into. These results also highlight the contributions that can be made to laboratory phonetics and phonology by the inclusion of languages with variety in morphosyntactic structure, and of lesser-studied languages in general.

Appendix: Stimuli in Georgian orthography

(1) _____. კიდევ _____ ვ-თქვ-ი. [Speaker F1]

_____. Again ____ 1SG.SUBJ-say-AOR.

(2) _____. ქალ-მა _____ მო-მ-წერ-ა. [Speakers M1 and F2]

_____. woman.ERG ____ PRVB-1SG.OBJ-write-3SG.SUBJ.AOR

Table A4. Experiment 1 test words for speaker F1

	C	CC	CCC
sonority rise	**ძ**ალა ‘strength’ **მ**ადა ‘hunger’ **რ**ეზი ‘Rezi (name)’	**ძმ**არი ‘vinegar’ **მრ**ეცი ‘downward slope’	**ძმრ**იანი ‘vinegary’
sonority plateau	**ტ**აფა ‘frying pan **ბ**ადე ‘net’	**ტბ**აზე ‘lake.in’ **ტყ**ება ‘lamenting (n.)’	**ტყდ**ება ‘breaking (n.)’
sonority fall	**მ**ოგები ‘prize’ **ს**ოფელი ‘village’ **პ**ოვნა ‘finding (n.)’	**მს**ოფლიო ‘world’ **სპ**ობა ‘destroying (n.)’	**მსპ**ობელი ‘destroying (n).’

Table A5. Experiment 1 test words for speakers F2 and M1

	C	CC	CCC
sonority rise	**ძ**ალა ‘strength’ **მ**ადა ‘hunger’ **რ**ეზი ‘Rezi (name)’	**ძმ**არი ‘vinegar’ **მრ**ეცი ‘downward slope’ **სმ**ა ‘drinking (n.)’	**ძმრ**იანი ‘vinegary’ **ფსმ**ა ‘peeing (n.)’

	ფანდი 'wrestling move' **ს**ამი 'three' **შ**ავი 'black'	**ფს**ალტა 'psalm' **ფშ**ავი 'Pshavia (region)' **შმ**აგი 'maniac'	**ფშმ**უის 'cow snorting'
sonority plateau	**ტ**ენი 'damp' **ყ**ელი 'throat' **დ**ება 'placing (n.)'	**ტყ**ება 'lamenting (n.)' **ყდ**ები 'bindings'	**ტყდ**ება 'breaking (n.)'
sonority fall	**მ**ოგები 'prize' **ს**ოფელი 'village' **პ**ოვნა 'finding (n.)'	**მს**ოფლიო 'world' **სპ**ობა 'destroying (n.)'	**მსპ**ობელი 'destroying (n).'

Table A6. Experiment 2 test words for all speakers

	morphologically simplex	morphologically complex
sonority rise	**გმ**ანავს 's/he fills it'	**გ-მ**ალავს 's/he hides you'
sonority plateau	**გბ**ობა 'boiling (n)' **გდ**ება 'lie/lay about' **მნ**ეობა 'management'	**გ-ბ**ერავს 's/he inflates you' **გ-დ**ევს 's/he places you' **მ-ნ**ებავს 's/he allows me'
sonority fall	**მდ**არე 'worthless' **მთ**ელი 'whole, entire'	**მ-დ**არებს 's/he compares me' **მ-თ**ელავს 's/he tramples me'

References

Aronson, H. I. (1997). Georgian phonology. *Phonologies of Asia and Africa, 2*, 929-939.

Bates D, Mächler M, Bolker B, Walker S (2015). "Fitting Linear Mixed-Effects Models Using lme4." *Journal of Statistical Software*, 67(1), 1–48. doi: 10.18637/jss.v067.i01.

Bell, M. J., Ben Hedia, S., & Plag, I. (2021). How morphological structure affects phonetic realisation in English compound nouns. *Morphology*, *31*(2), 87-120.

Borise, L., & Zientarski, X. (2018). Word stress and phrase accent in Georgian. *Proceedings of TAL*, *6*, 2018-42.

Browman, C. P., & Goldstein, L. M. (1986). Towards an articulatory phonology. *Phonology*, *3*, 219-252.

Browman, C. P., & Goldstein, L. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica*, *45*(2-4), 140-155

Browman, C. P., & Goldstein, L. (1989). Articulatory gestures as phonological units. *Phonology*, *6*(2), 201-251.

Brunner, J., Geng, C., Sotiropoulou, S., & Gafos, A. (2014). Timing of German onset and word boundary clusters. *Laboratory Phonology*, *5*(4), 403-454.

Bush, R. 1999. Georgian yes-no question intonation. In *Phonology at Santa Cruz*, Vol. 6.
Santa Cruz, CA: US Santa Cruz. 1-11.

Byrd, D., & Choi, S. (2010). At the juncture of prosody, phonology, and phonetics—The interaction of phrasal and syllable structure in shaping the timing of consonant gestures. *Laboratory phonology*, *10*, 31-59.

Byrd, D., & Saltzman, E. (1998). Intragestural dynamics of multiple prosodic boundaries. *Journal of Phonetics*, *26*(2), 173-199.

Clements, G. N. (1990). The role of the sonority cycle in core syllabification. *Papers in laboratory phonology*, *1*, 283-333

Chitoran, I. (1999). Accounting for sonority violations: the case of Georgian consonant sequencing. In *Proceedings of the 14th International Congress of Phonetic Sciences,* 101-104.

Cho, T. (2001). Effects of morpheme boundaries on intergestural timing: Evidence from Korean. *Phonetica*, *58*(3), 129-162.

Cho, T. (2006). Manifestation of prosodic structure in articulatory variation: Evidence from lip kinematics in English. *Laboratory phonology*, *8*, 519-548.

Crouch, C., Katsika, A., Chitoran, I. (2020) The role of sonority profile and order of place of articulation on gestural overlap in Georgian. Proc. 10th International Conference on Speech Prosody 2020, 205-209, DOI: 10.21437/SpeechProsody.2020-42.

Gafos, A. I., Hoole, P., Roon, K., Zeroual, C., Fougeron, C., Kühnert, B., D'Imperio, M., & Vallée, N. (2010). Variation in overlap and phonological grammar in Moroccan Arabic clusters. *Laboratory phonology*, *10*, 657-698.

Goldsmith, J., & Larson, G. (1990). Local modeling and syllabification. In *Papers from the 26th Annual Regional Meeting of the Chicago Linguistics Society: Parasession on the Syllable in Phonetics and Phonology. Chicago Linguistics Society*, 129-152.

Goldstein, L., Chitoran, I., & Selkirk, E. (2007, August). Syllable structure as coupled oscillator modes: evidence from Georgian vs. Tashlhiyt Berber. In *Proceedings of the XVIth international congress of phonetic sciences*, 241-244.

Harris, A. (2002). The word in Georgian. In Robert Dixon and Alexandra Aikenvald (Eds.). *Word: A cross-linguistic typology*, 127-142.

Hayes, B. (1985, June). Iambic and trochaic rhythm in stress rules. In *Annual Meeting of the Berkeley Linguistics Society* (Vol. 11, pp. 429-446).

Henke, E., Kaisse, E. M., & Wright, R. (2012). Is the sonority sequencing principle an epiphenomenon. *The sonority controversy*, *18*, 65-100.

Hermes, A., Greisbach, R., & Chitoran, I. (2020). Probing the limits of coordination: Word onset clusters in Georgian. *12th International Seminar on Speech Production*.

Hermes, A., Grice, M., Mücke, D., Niemann, H., & Phonetik, I. (2008). Articulatory indicators of syllable affiliation in word initial consonant clusters in Italian. *rema*, *151*(177), 230.

Hermes, A., Mücke, D., & Grice, M. (2013). Gestural coordination of Italian word-initial clusters: the case of 'impure s'. *Phonology*, *30*(1), 1-25.

Hewitt, G. (2005). *Georgian: A learner's grammar*. Routledge.

Katsika, A., Krivokapić, J., Mooshammer, C., Tiede, M., & Goldstein, L. (2014). The coordination of boundary tones and its interaction with prominence. *Journal of Phonetics*, *44*, 62-82.

Kiparsky, P. (1980). Remarks on the metrical structure of the syllable. *Phonologica*, *3*.

Kühnert, B., Hoole, P., & Mooshammer, C. (2006). Gestural overlap and C-center in selected French consonant clusters. In *7th International Seminar on Speech Production (ISSP)*.

Ladd, D. R. (2008). *Intonational phonology*. Cambridge University Press.

Lammert, A. C., Shadle, C. H., Narayanan, S. S., & Quatieri, T. F. (2016). Investigation of Speed-Accuracy Tradeoffs in Speech Production Using Real-Time Magnetic Resonance Imaging. In *Interspeech* (pp. 460-464).

Lenth, R. (2019). emmeans: Estimated Marginal Means, aka Least-Squares Means. R package version 1.4.2. https://CRAN.R-project.org/package=emmeans

Marin, S., & Pouplier, M. (2010). Temporal organization of complex onsets and codas in American English: Testing the predictions of a gestural coupling model. *Motor Control*, *14*(3), 380-407.

Mattingly, I. G. (1981). Phonetic representation and speech synthesis by rule. In *Advances in Psychology* (Vol. 7, pp. 415-420). North-Holland.

Nam, H., & Saltzman, E. (2003). A competitive, coupled oscillator model of syllable structure. In *Proceedings of the 15th international congress of phonetic sciences*, 2253-2256.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. *Approaches to phonological complexity*, *16*, 299-328.

Pouplier, M., & Beňuš, Š. (2011). On the phonetic status of syllabic consonants: Evidence from Slovak. *Laboratory phonology*, *2*(2), 243-273

R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Saltzman, E., & Byrd, D. (2000). Task-dynamics of gestural timing: Phase windows and multifrequency rhythms. *Human Movement Science*, *19*(4), 499-526.

Selkirk, Elizabeth O. (1982). Syllables. In Harry van der Hulst and Norval Smith (eds.) *The structure of phonological representations, Part II*. Dordrecht, Foris, 337–383.

Shaw, J. A., Gafos, A. I., Hoole, P., & Zeroual, C. (2011). Dynamic invariance in the phonetic expression of syllable structure: a case study of Moroccan Arabic consonant clusters. *Phonology*, *28*(3), 455-490.

Shosted, R. K., & Chikovani, V. (2006). Standard Georgian. Journal of the International Phonetic Association, 36(2), 255-264.

Skopeteas, S., & Féry, C. (2016). Focus and intonation in Georgian: Constituent structure and prosodic realization. *Ms.*

Song, J. Y., Demuth, K., Shattuck-Hufnagel, S., & Ménard, L. (2013). The effects of coarticulation and morphological complexity on the production of English coda clusters: Acoustic and articulatory evidence from 2-year-olds and adults using ultrasound. *Journal of Phonetics*, *41*(3-4), 281-295.

Sotiropoulou, S., Gibson, M., & Gafos, A. (2020). Global organization in Spanish onsets. *Journal of Phonetics*, *82*, 100995.

Sotiropoulou, S., Pillmeier, E., Katsika, A., & Gafos, A. I. (2015). Dissecting the consonant duration ratio. In *ICPhS*.

Tomaschek, F., Tucker, B. V., Ramscar, M., & Harald Baayen, R. (2021). Paradigmatic enhancement of stem vowels in regular English inflected verb forms. *Morphology*, *31*(2), 171-199.

Treiman, R. (1986). The division between onsets and rimes in English syllables. *Journal of Memory and Language*, *25*(4), 476-491

Vicenik, C., & Jun, S. A. (2014). An autosegmental-metrical analysis of Georgian intonation. *Prosodic typology II: The phonology of intonation and phrasing*, 154-186.

Wickham, H. (2016). ggplot2: Elegant Graphics for Data Analysis.

Chapter 4. Discussion and future research

This dissertation investigates the relationship between the spatial and the temporal organization of syllable onsets in Georgian. Two broad questions guide the design and framing of this work:

Goal 1 (G1): Describe how syllable onsets are organized in Georgian.

Goal 2 (G2): Demonstrate what Georgian can tell us about how space and time interact in syllable organization in general.

The spatial dimension of syllable organization is considered through the lens of sonority and sonority sequencing, since sonority is correlated with degree of vocal tract openness. The temporal dimension is analyzed within the framework of Articulatory Phonology (AP), which includes timing at the level of phonological representation and defines syllable structure based on temporal criteria. The question of how sonority shape and gestural timing relate in Georgian was addressed by three Electromagnetic Articulography (EMA) experiments, detailed in Chapters 2 and 3. The research questions of those experiments are:

Research Question 1 (RQ1): How are consonant gestures in a complex onset in Georgian timed with respect to one another?

Research Question 2 (RQ2): Since combinatorial possibilities in Georgian onsets are not constrained by the sonority hierarchy, does the timing of the consonantal gestures systematically reflect sonority sequencing?

Research question 3 (RQ3): What coordinative pattern is found in onsets in Georgian?

Research question 4 (RQ4): What is the relationship between the sonority shape of a complex onset in Georgian and the global organization of the constituent gestures?

Research question 5 (RQ5): What is the relationship between the morphological composition of an onset in Georgian and the global organization of the constituent gestures?

In this chapter I will first discuss G1 by revisiting and synthesizing the results of RQ1-5 and present a unified account of gestural timing in Georgian syllables. Then, based on this account, I will address G2. The experimental results can be broken down into two categories: local timing patterns (RQ1 and RQ2), and global timing patterns (RQ3, RQ4, and RQ5). Following the general flow of the dissertation, I will first address local timing and then global timing.

1. Local timing

The general local timing pattern in Georgian onsets is one of long lag, indicating, as predicted by the coupled oscillator model, anti-phase coordination between C gestures (RQ1). We quantified the lag between C gestures in two ways: first, through what we term relative overlap (e.g., Chitoran et al., 2002) which measures the distance between the onset of the C2 gesture and the target of the C1 gesture; and second, through constriction duration overlap, which refers to the amount of overlap between the target constrictions of the two C gestures (see figures 8 and 11, respectively, in Chapter 2). In CC onsets (the only kind examined in Chapter 2), C2 begins, at its earliest, once C1 has reached its target, and the constriction plateaus of the two C gestures never overlap. Such a timing pattern entails open transitions (Catford, 1988) during which intrusive vocoids may occur. Intrusive vocoids are schwa-like acoustic elements which are not targeted phonological gestures; they are purely phonetic.

This local timing is significantly and systematically affected by the sonority shape of an onset. The existence of such a relationship is expected, but the observed direction of the effect is the opposite of our hypothesis. We predicted that consonant gestures in sonority rises would overlap more than gestures in either sonority plateaus or in sonority falls. Sonority rises are best suited to the parallel transmission of information (Mattingly, 1981), which has been argued to be a critical element of both speech production and perception. Instead, we found that sonority rises were the least overlapped sonority shape according to both measures we employed. Sonority falls were the most overlapped shape, and sonority plateaus were an intermediate case (RQ2). Optimizing parallel transmission of information is not the primary motivator for timing patterns in Georgian. This suggests that parallel transmission is only one of the important factors in speech perception and production, and that different languages prioritize it to different degrees.

Our proposal is that these timing modulations are motivated by a need to preserve the syllable parse. Sonority falls, which showed the earliest onset of the C2 gesture and the smallest lag between consonant constrictions, had the fewest number of intrusive vocoids as well. We argue that sonority falls are more vulnerable to being misparsed as CVCV when a vocoid emerges between C1 and C2. At the local level, then, intergestural timing in Georgian is modulated by perceptual concerns.

2. Global timing

The experiments detailed in Chapter 3 examine global timing in Georgian by testing the predictions of the coupled oscillator model (Nam et al., 2009). The coupled oscillator model defines syllabic structure based on coupling dynamics between oscillators. These oscillators trigger the activation intervals of gestures, which are the phonological primes in AP. As a reminder, the coupled oscillator model predicts the following coupling relationships and accompanying temporal organization of gestures:

(1) In-phase coordination in simplex CV. The C and V gesture are initiated simultaneously.

(2) Anti-phase in VC and between adjacent coda consonant gestures. The C gesture is initiated after the V gesture has reached its target.

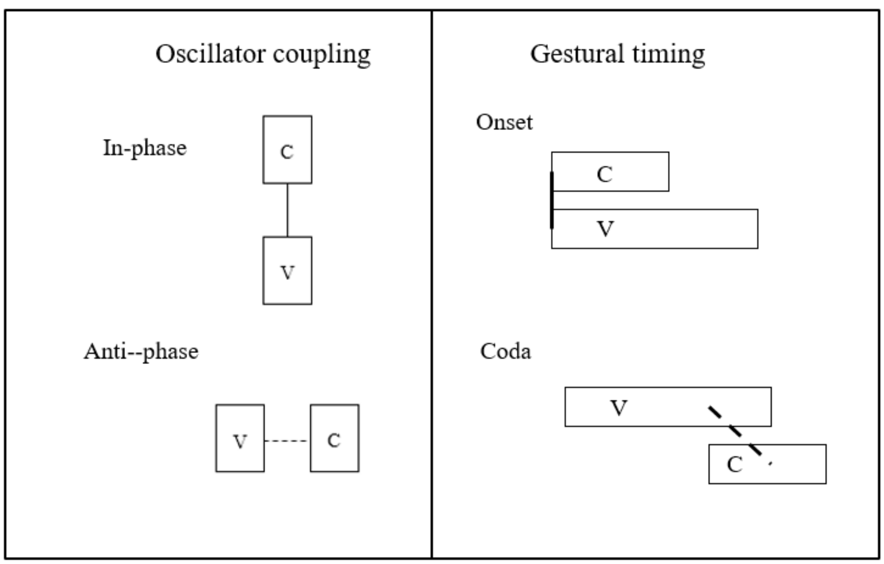


Figure 8. In-phase and anti-phase coupled oscillators and associated gestural timing.

(3) Competitive coupling in complex onsets: in-phase coupling between the V gesture and each C gesture, and anti-phase coupling between C gestures. Competitive coupling is the mechanism behind the c-center effect (Browman and Goldstein 1988). The c-center is the midpoint of the onset cluster—it is equidistant from the onsets of all C gestures, and the V gesture onset is synchronous with the c-center.

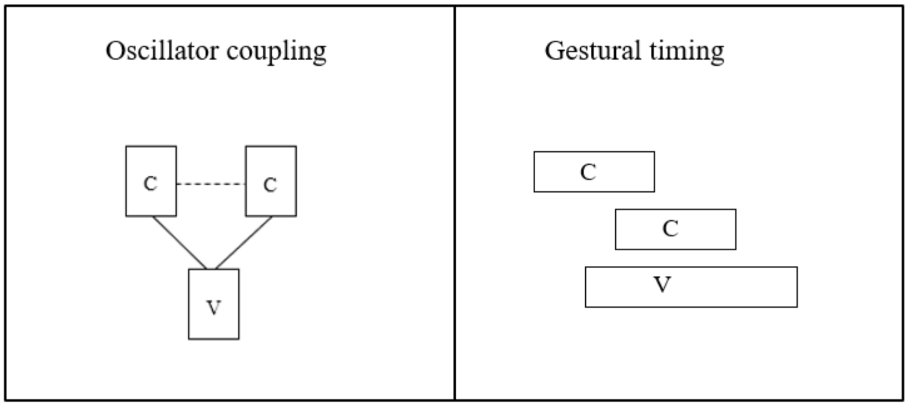


Figure 9. Competitively coupled oscillators and the associated gestural timing

Georgian does not display the predicted timing patterns in CV, CCV, or CCCV syllables.

The first major takeaway from this work is that Georgian does not organize its syllables in the way predicted by either the Sonority Sequencing Principle (SSP) or the coupled oscillator model of syllable structure. That Georgian does not adhere to the SSP is well-known and can be ascertained simply by looking at even a Swadesh list, for example: /mze/ 'sun', /mtsxoveli/ 'animal', /t'q'e/ 'forest', /p^{h}rt^{h}a/ 'feather', and /p^{h}rʧxili/ 'fingernail'. The results assessing the coupled oscillator model, on the other hand, are unexpected. When we consider the two together, however, it becomes clear that the unusual timing of gestures in the onset and the diversity of permissible consonant clusters are related. Work on sonority sequencing has suggested that sonority rises are best suited to some 'default' timing pattern (e.g., Henke et al., 2012), but in Georgian, where this timing pattern—presumably the one generated by competitive coupling—is not used, sonority rises are not any more well-suited for serving as syllable onsets than any other consonant clusters.

Georgian shows no evidence for the competitive coupling shown in Figure 2. Instead, all gestures in Georgian are in anti-phase relationship to one another, including the prenuclear C gesture and the V gesture (RQ5). The sonority shape of the onset has no effect on global timing (RQ6), nor does the morphological composition of the onset (RQ7). In this chapter we argue that the anti-phase only coordination is supported by both phonotactics and morphology. Georgian does not need to distinguish between word-medial onsets and codas, since there are virtually no illicit consonant sequences; this is likely the reason for the variable syllabification of intervocalic consonant sequences (see Harris, 2002 for an overview). Anti-phase only coordination also facilitates the 'slotting-in' of inflectional morphemes that can increase the onset size of what we would assume is the same lexical item. Anti-phase only also boosts the recoverability of each consonant in the onset, which is especially important in a language where a single consonant can be the only explicit source of certain grammatical information.

3. A unified account of Georgian syllable organization

Neither local nor global timing in Georgian behave in the ways predicted in Chapters 2 and 3, but these C(…)V sequences should still be treated as single syllables, as I argue in Chapter 3. We need, then an account of timing in Georgian syllables that accounts for both the local and lack of global timing patterns observed, and a definition of the syllable that moves beyond what is proposed either by the SSP or the coupled oscillator model.

All coordination in Georgian onsets is anti-phase, but the results from Chapter 2 show systematic differences across sonority shapes in the onset of C2 relative to C1's target. I propose that these differences are the result of different coupling strengths. These coupling strengths are calculated dynamically, as Goldsmith and Larson (1990) describe. As with

Goldsmith and Larson's account, this local dynamical calculation can be extended to all languages, which have their own specific parameters that then yield the coupling strengths and timing patterns found in those languages. Their model calculates derived sonority from inherent sonority, and their test case is Tashlhyit Berber. The derived sonority values determine which segments serve as syllable nuclei.

In the model I describe here, coupling strength is the output of the calculation instead of derived sonority, and the broad phonological goal would be not to assign syllabic affiliation, but to produce a single energy curve per syllable. The input to the model, however, could still be a sonority value. Sonority is correlated with intensity, and the articulatory results do suggest that sonority sequencing is not irrelevant for Georgian. In cases where there is a sonority drop across a C-C sequence the coupling strength would ensure that C2 begins when or before C1 reaches its target. Preventing intrusive vocoids in these cases—where C1 is often a sonorant—ensures that there won't be a sonorant-vocoid sequence followed by a stop, which could, if the vocoid were long and intense enough, be interpreted as a syllable on its own. In short, I propose that one possible treatment of syllables focuses on their correlation with pulses of resonant intensity.

Treating syllables as pulses in a continuous function is not without precedent. Goldstein (2019) presents initial work showing that syllables could be pulses in modulation functions of acoustic and kinetic energy. Goldstein (2019) investigates the correlation between change over time in the acoustic and kinematic signals in English and finds a correlation between the two, which both display repetitive pulse structures. Strikingly, the pulses in both signals appear to be related to syllable structure. There is likely also a correlation between the shape of the continuous intensity signal and the two modulation functions, and one may be a result of the other. Goldstein also suggests that kinetic energy could be considered an index of sonority change, which makes this an especially relevant theoretical avenue for future work.

4. Space and time in the syllable

This dissertation demonstrates that in Georgian, space and time interact systematically to form syllables in production. These findings can help expand and even integrate the definitions of the syllable advanced by the SSP and the coupled oscillator model.

The Georgian data here show a syllabic coordination pattern not previously described for onsets. The anti-phase only coordination, even in the CV case, goes entirely against the predictions of the coupled oscillator model. Additional phonological evidence, detailed in Chapter 3, supports both the existence of the syllable as a unit in Georgian and the continued analysis of C(..)V sequences as monosyllables. This means that we have to reconsider how the syllable is defined in time. Competitive coupling is not the defining characteristic of a complex onset. Instead, it is one way that languages can organize onsets and nuclei. Competitive coupling has obvious benefits: 1) when considered with the anti-phase only coordination in codas, it clearly differentiates onsets from codas and 2) the resulting gestural overlap supports the parallel transmission of information. It also likely supports the modulation function pulse structure described in Goldstein (2019) where, regardless of the number of consonants, onsets and nuclei are together associated with a single pulse.

Languages with the observed c-center effect or other indices of global timing are also languages that generally adhere to the SSP (e.g., English, French, Spanish, Italian, German). This is not a coincidence; sonority rises are the spatial contours that best support the parallel transmission of information, which competitive coupling also supports. The SSP and the coupled oscillator model are in that sense two sides of the same coin. The SSP defines the spatial contour best suited to the parallel transmission of information, and the coupled oscillator model defines the temporal coordination best suited to it.

The Georgian data show that, although the patterns defined above have clear benefits that explain their prevalence in the worlds languages, they do not exclusively define how syllables are organized in production. The results reveal some of the factors that can motivate more uncommon syllable structures, namely morphology and phonotactics. First, if a language does not need to robustly distinguish word-internal onsets from word-internal codas, the benefit of different temporal organization for onsets and codas is not relevant. Secondly, and, I argue, more importantly, the morphology of a language can affect how syllables are organized. Studies in the AP framework have focused primarily on languages with minimal morphology or with morphology that does not result in changes in syllable size. When we extend these lines of research to languages such as Georgian, however, new considerations are introduced that can outweigh the benefits conveyed by global organization of the onset.

Syllables as units are spatiotemporal in nature and the spatial and temporal dimensions of their organization interact in systematic ways. This dissertation has shown that the nature of this interaction can vary depending on language-specific differences, as illustrated by the case of Georgian. These findings allow us to expand both the sonority- and timing-based definitions of the syllable and to shed light on the motivations behind common patterns of syllabification observed in the world's languages.

5. Future directions

The account of Georgian syllable structure presented here makes two clear, testable predictions. The first is that the local timing differences observed are motivated by preservation of the syllable parse. Future work will test this via perception experiments run online. Participants will hear CV or CCV stimuli clipped from the production data analyzed here and will be asked to transcribe what they hear in Georgian orthography. If sonority falls are more vulnerable to being misparsed as CVCV, then they will have higher rates of mis-transcription than sonority rises, and mis-transcriptions will begin at lower lags in the range than in the analogous value for sonority rises. The second is that Georgian will show the same pulse structure in its acoustic and kinetic modulation functions as English does in Goldstein (2019). This can be tested by simply performing the same analysis on the Georgian data, and this analysis is already underway.

My future research on Georgian will also including modeling of gestural coordination in order to find out what coupling strength differences will yield the local timing patterns described in Chapter 2, and whether an anti-phase only pattern or one with very weak in-phase C-V coupling (as discussed in Chapter 3 section 7) better accounts for the global timing patterns observed.

This dissertation also makes typological predictions. Languages where consonant-only morphemes can change the size of a syllable onset, such as Abkhaz (Chirikba, 2003), Lak (Friedman, 1996), and Yateé Zapotec (Jaeger and Van Valin Jr., 1982) might also use an anti-phase only coordination. Languages with phonotactics as permissive as Georgian's, such as Bella Coola (Bagemihl, 1991), Pashto (Bell and Saka, 1983), and Russian (Redford, 1999) are predicted to modulate C-C timing relationships by sonority shape via changes in coupling strength. These predictions can be tested by extending this research paradigm to a range of languages with different morphologies and phonotactics. Even if the results do not support the analysis presented here, they will further expand our knowledge of the way in which languages organize syllables in speech production.

References

Bagemihl, B. (1991). Syllable structure in Bella Coola. *Linguistic Inquiry*, *22*(4), 589-646.

Bell, A., & Saka, M. M. (1983). Reversed sonority in Pashto initial clusters. *Journal of Phonetics*, *11*(3), 259-275.

Catford, J. S. (1988). A practical guide to phonetics (Clarendon Press, Oxford).

Chirikba, V. A. (2003). Abkhaz (Languages of the World/Materials 119). *München: Lincom Europa*.

Chitoran, I., Goldstein, L., & Byrd, D. (2002). Gestural overlap and recoverability: Articulatory evidence from Georgian. *Papers in laboratory phonology*, *7*(4-1), 419-447.

Friedman, V. (1996). Gender, class, and age in the Daghestanian highlands: Towards a unified account of the morphology of agreement in Lak. *Linguistic studies in the non-Slavic languages of the Commonwealth of Independent States and the Baltic Republics*, *8*, 187-199.

Goldsmith, J., & Larson, G. (1990). Local modeling and syllabification. In *Papers from the 26th Annual Regional Meeting of the Chicago Linguistics Society: Parasession on the Syllable in Phonetics and Phonology. Chicago Linguistics Society*, 129-152.

Goldstein, L. (2019). The role of temporal modulation in sensorimotor interaction. *Frontiers in Psychology*, *10*, 2608.

Harris, A. (2003). The word in Georgian. In R. Dixon & A. Aikhenvald (Eds.), *Word: A Cross-linguistic Typology* (pp. 227-242). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511486241.010

Henke, E., Kaisse, E. M., & Wright, R. (2012). Is the sonority sequencing principle an epiphenomenon. *The sonority controversy*, *18*, 65-100.

Jaeger, J. J., & Van Valin Jr, R. D. (1982). Initial consonant clusters in Yateé Zapotec. *International Journal of American Linguistics*, *48*(2), 125-138.

Mattingly, I. G. (1981). Phonetic representation and speech synthesis by rule. In *Advances in Psychology* (Vol. 7, pp. 415-420). North-Holland.

Nam, H., Goldstein, L., & Saltzman, E. (2009). Self-organization of syllable structure: A coupled oscillator model. Approaches to phonological complexity, 16, 299-328.

Redford, M. A. (1999). The mandibular cycle and reversed-sonority onset clusters in Russian. In *Proc. 14th ICPhS* (pp. 1893-1896).