Chloe Craft
Linguistic Analysis
Professor Cooper
April 24, 2024

An Analysis of the Sound System of Georgian

I. Introduction to Georgian

Georgian is an incredibly unique language, not only in its linguistic systems but also in its geographic, historical, and cultural context. Among the many linguistic qualities that make Georgian distinctive compared to other languages of the world, some particularly notable features are that it does not belong to any of the six major language families (What are the largest language families, para. 2), it employs the use of three unique writing systems which all remain in use today (Living culture of three writing systems of the Georgian alphabet, para. 1), and it contains incredibly heavy, complex consonant clusters that have captivated linguists (Butskhrikidze 2002:16).

1 Social and geopolitical context

Regarding the larger context surrounding Georgian, I thought it was particularly interesting to examine a language from a region that spans both Europe and Asia transcontinentally. Georgia, located in the Caucasus region, has historically faced the influence of an impressive amount and variety of empires. Most notably, these include the Greeks, Romans, Persians, Arabs, Ottomans, Armenians, and Russians, as well as most recently, the Soviet Union, as written by Asatiani and Janelidze (2009). I was interested to know how this complex history influenced the development and change of the Georgian language, especially given that despite a long history of conquest and oppression, the language has survived. The remarkably long time that the Georgian language has survived is particularly astonishing considering that the earliest record of its ancestor, Old Georgian, dates back to the year 430 A.D. in the form of the Bir el Qutt inscriptions (Khurtsilava 2017:133).

Today, Georgian is used in a variety of modern contexts in its home country of Georgia and beyond. According to Ethnologue (Georgian, n.d.), Standard Georgian has approximately 3,724,240 native speakers and 154,540 second-language speakers around the world as of 2020, with the vast majority of both residing in Georgia. Of the other nations where Georgian is spoken (all according to Ethnologue): Russia has the highest population of total users at 81,000, followed by Iran with 67,200, Azerbaijan, Greece, Germany, and Spain with 20-30,000, Ukraine, Kazakhstan, and Uzbekistan with 5-15,000, and Canada, Turkmenistan, Tajikistan, Sweden, and Latvia with under 4,000. These nations all use a variety of other languages native to them, and Georgian is considered a minority language.

The only country where Georgia is considered a national or official language is Georgia, where it is used in many contexts, most notably in government, education, and media, though it should be noted that some radio stations and television channels broadcast content in Russian as well as Georgian, making Russian the primary competitor for Georgian in these contexts (The World Factbook: Georgia, para. 8). Albeit this is not a strong competitor, as Russian speakers compose only 1.2% of Georgia's population, presiding over a mere 1% of other language speakers, while 3.9% of the Georgian population speak Armenian, 6.2% speak Azeri, and a majority of 87.6% speak Georgian, according to the CIA's World Factbook on Georgian.

2 Linguistic genealogy

As previously mentioned, one of the features that make Georgian particularly unique is its origin; as opposed to belonging to a larger and more well-known language family, it is a member of the Kartvelian language family, which has no known relation to any other language family, according to Dalby (2003). Georgian is the most widely spoken of the mere five languages in the Kartvelian language family, with its closest relatives being the Laz and Mingrelian languages, as seen in Figure 1.

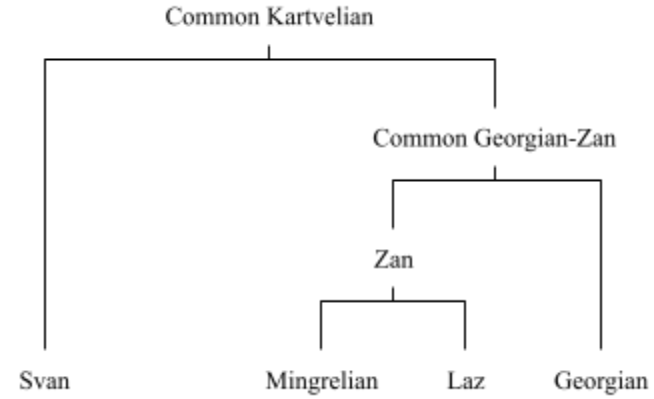


Figure 1: The Kartvelian Languages (Hillery 1994:1)

However, perhaps an even closer relative to Georgian would be its frequently contested daughter dialect, Judeo-Georgian, which has an estimated 80,000 speakers as of 2019 (Judeo-Georgian). The status and classification of Judeo-Georgian have been debated often by linguists (Boeder, 2005:14), but it is the most significant of the dialects of Georgian.

Overall, Georgian is an incredibly unique language, individually and comparatively. Not only does it feature a variety of uncommon and fascinating linguistic features, but it also exists within a rich and complicated historical and geopolitical context. Georgian is a resilient language, surviving through countless occupations and invasions, but it manages to balance its vibrant history with its progressive modernity. This paper aims to analyze the expansive, unique features of the sound system of the Georgian language, considered within the rich context of the language.

II. Georgian phonetics

1 Introduction

The modern Standard Georgian language is intricate in its phonetic system. This section will illustrate the different phones in Georgian, provide linguistic examples to explain the use of particular phones and describe the phonetic typology of Georgian.

2 Phonetic inventory

Georgian is quite complex in phonetic variety. It has 34 distinct phonetic segments, 28 of which are consonants and 6 vowels (Georgian sound inventory (UPSID) 2019). Many of these have various distinct allophones contributing to a high quantity of phones compared to other languages, which will be explained further later in this paper. This section will focus on the

inventory of phones in Georgian, as well as example data defending the prevalence of each phonetic unit.

2.1 IPA chart of Georgian phones

All possible phones used by speakers of Georgian can be represented in the IPA. This is shown below, with consonants in Figure 2 and vowels in Figure 3. In Figure 2, voiceless consonants are represented on the left side of a column, while voiced consonants are on the right side.

		Bilabial		Labiodental		Dental		Alveolar		Postalveolar		Velar		Uvular		Glottal	
Plosive	*Standard*		b				d						g				
	Weak voice		b̥				d̥						g̊				
	Aspirated	pʰ				tʰ						kʰ					
	Ejective	pʼ				tʼ						kʼ		qʼ			
Affricate	*Standard*								dz		dʒ						
	Aspirated							tsʰ		tʃʰ							
	Ejective							tsʼ		tʃʼ							
Nasal			m						n								
Fricative	*Standard*			f	v			s	z	ʃ	ʒ	x	ɣ	χ		h	
	Ejective													χʼ			
Tap									ɾ								
Lateral approximant	*Standard*								l								
	Velarized								ɫ								

Figure 2: Consonant phones of Georgian (Shosted and Chikovani 2006:255) (Butskhrikidze 2002:44-59) (Georgian sound inventory (UPSID) 2019)

Vowels	Front	Near-Front	Central	Near-Back	Back
Close	i				u
Near-Close					
Close-Mid	e				o
Mid			ə		
Open-Mid					
Near-Open					
Open	a				

Figure 3: Vowel phones of Georgian (Testelets 2021:498) (McCoy 1999:447-450)

2.2 Illustrative data of phonetic units

The following exemplifies each phone illustrated in Figures 2 and 3 as they appear in common Georgian words and phrases. They are arranged with the phone illustrated, followed by the

phonetic transcription of the Georgian example in IPA, the Georgian example in its original form, and the English gloss of the example. All words can be attributed to John Torikashvili's *Georgian-English, English-Georgian Dictionary* (Torikashvili 1992).

[p']	[p'uɾi]	პური	'bread'	[f]	[ʤalefs]	დალევს	'to drink'
[pʰ]	[pʰeɾo]	ფერო	'color'	[v]	[veɾ]	ვერ	'no'
[b]	[abi]	აბი	'pill'	[s]	[satʰ]	სად	'where'
[ʤ]	[ʤabua]	ბაბუა	'grandfather'	[z]	[zɾda]	ზრდა	'growth'
[t']	[t'averna]	ტავერნა	'tavern'	[ʃ]	[ʃaɾpʰi]	შარფი	'scarf'
[tʰ]	[ʤaamotʰ]	გაამოთ	'bon appetit'	[ʒ]	[ʒanɾi]	ჟანრი	'genre'
[d]	[ʤidi]	დიდი	'large'	[x]	[xili]	ხილი	'fruit'
[ʤ]	[ʤaisi]	დაისი	'sundown'	[ɣ]	[ɣori]	ღორი	'pig'
[k']	[k'ałami]	კალამი	'pen'	[χ]	[d͡ʒiχuɾi]	ჯიხური	'kiosk'
[kʰ]	[makʰ]	მაგ	'that'	[χ']	[ʤaχ'ali]	ბაყალი	'grocer'
[g]	[k'aɾgi]	კარგი	'okay'	[h]	[haeɾi]	ჰაერი	'air'
[ʤ]	[ʤadis]	გადის	'to leave'	[ɾ]	[ɾaime]	რაიმე	'anything'
[q']	[q'ava]	ყავა	'coffee'	[l]	[lek'ʷi]	ლეკვი	'puppy'
[dz]	[d͡zeli]	ძელი	'tree'	[ɬ]	[k'ała]	კალა	'tin'
[dʒ]	[d͡ʒado]	ჯადო	'magic'	[ʷ]	[vepʰχʷi]	ვეფხვი	'tiger'
[ts']	[t͡s'eɾs]	წერს	'to write'	[ɪ]	[iład͡ʒi]	ილაჯი	'strength'
[tsʰ]	[ʤat͡sʰi]	ბაცი	'bright'	[e]	[elʷa]	ელვა	'lightning'
[tʃ']	[t͡ʃ'oɾi]	ჭორი	'gossip'	[a]	[aguɾi]	აგური	'brick'
[tʃʰ]	[t͡ʃʰaidani]	ჩაიდანი	'kettle'	[o]	[otʰaxi]	ოთახი	'room'
[m]	[muzeumi]	მუზეუმი	'museum'	[u]	[uazɾo]	უაზრო	'meaningless'
[n]	[nana]	ნანა	'lullaby'				

3 Conclusion

Overall, Georgian is a phonetically unique language with a variety of phones that provide insight into what makes the language so distinctive. The diversity of the particular phones present in modern Georgian as well as the extent of their variation from genetically related languages are significant features of Georgian phonetics, albeit the credibility and relevance of the sources

providing this information are also noteworthy in this analysis. These phonetic features will certainly play into later examinations of the phonology of the Georgian language in this paper.

III. Georgian phonology

1 Introduction

The Georgian language is unique and complex in its phonological systems, from phonemic inventory to allophonic variation to syllable structure to phonotactics and more. This section will briefly illustrate each of these features with accompanying data to support a description of the phonology of the modern Georgian language.

These range from the diverse phonemic inventory of the language and the instances of contrast and allophonic variation within that inventory, to the uniquely complex syllable structure and phonotactics that govern the language, to an examination of the typology of Georgian by itself and compared cross-linguistically, to Georgian's adherence and violation of commonly accepted phonological principles.

2 Phonemes and allophones

Georgian has a large and complex phonemic inventory (Maddieson 2013) with several allophonic variations. This section will illustrate an inventory of all phonemic units, provide illustrative data to support contrast, and describe allophonic variation among said units.

2.1 Phonemic Inventory

There are 28 consonant phonemes in the Georgian language (Shosted and Chikovani 2006:255-257) as illustrated in Table 4, as well as five vowel phonemes (Testelets 2021:498) as shown in Table 5.

Consonants		Bilabial	Labiodental	Dental	Alveolar	Postalveolar	Velar	Uvular	Glottal
Stop	*Standard*	b		d			g		
	Aspirated	pʰ		tʰ			kʰ		
	Ejective	pʼ		tʼ			kʼ	qʼ	
Affricate	*Standard*				dz	dʒ			
	Aspirated				tsʰ	tʃʰ			
	Ejective				tsʼ	tʃʼ			
Nasal		m			n				
Fricative			v		s	z	ʃ	ʒ	x	ɣ			h	
Trill					r				
Lateral approximant					l				

Table 4: Consonant phonemes of Georgian (Shosted and Chikovani 2006:255-257)

Vowels	Front		Near-Front		Central		Near-Back		Back	
Close	i									u
Near-Close										
Close-Mid	e									o
Mid										
Open-Mid										
Near-Open										
Open	a									

Table 5: Vowel phonemes of Georgian (Testelets 2021:498)

These phonological inventories contain noticeably fewer units than in Georgian's phonetic inventory, which consists of 36 distinct consonant phones (Georgian Sound Inventory (UPSID) 2019) and six distinct vowel phones (McCoy 1999:447-450). This is because the phonemic units in Georgian are governed by numerous rules with various instances of allophony, which will be discussed further in their designated section shortly.

The primary contrastive features between the phonemic and phonological inventories are in regard to voicing, aspiration, ejection, and velarization of consonants (Easterday 2019:405). Additionally there is the insertion of the mid central vowel [ə], which is not represented in the phonological inventory, to break up dense consonant clusters (McCoy 1999:448), the nature of which will be described in relation to syllable structure, phonotactics, and the typology of such later in this paper.

2.2 Illustrative data of phonemic contrasts

There are numerous instances where phonemic contrast must be acknowledged within the phonemic inventory of Georgian. The following illustrates some minimal pairs and near-minimal pairs in the Georgian language, grouped into a set of five examples of contrastive consonant phonemes in a) through e), and two sets of two examples of contrastive vowel phonemes, in f) and g) as well as h) and i). All examples are sourced from Shosted and Chikovani (Shosted and Chikovani 2006:256-261).

	Phoneme	IPA	Georgian	Gloss
a)	/z/	/zari/	ზარი	'call'
b)	/ʃ/	/ʃari/	შარი	'quibbling'
c)	/x/	/xari/	ხარი	'bull'
d)	/k^h/	/k^hari/	ქარი	'wind'
e)	/k'/	/k'ari/	კარი	'door'
f)	/i/	/gira/	გირა	'why'
g)	/o/	/gora/	გორა	'hill'
h)	/a/	/gada/	გადა	'move on'

i) /u/ /guda/ გუდა ‘leather bag’

2.3 Allophonic variation

One allophonic feature to note is the variation in the bilabial, dental, and velar stops. Here, the voicing of these stops is subject to change — and often, the presence of certain diacritics in addition — based on the placement of the phone.

As written by Howard Aronson, the voiced stops /b/, /d/, and /g/ become weakly voiced and are represented with the diacritics indicating weak voicing as [b̥], [d̥], and [g̊] when they are placed word-initially (Aronson 1990:15). This is exemplified in the example data in instances such as ბაბუა [b̥abua] meaning ‘grandfather’, as the word-initial bilabial stop is weakly voiced as [b̥], while the word-medial of the same is standard and presented as [b].

However, according to Brian Hewitt, when these same stops occur word-finally, they become completely voiceless and aspirated, then represented as [pʰ], [tʰ], and [kʰ] (Hewitt 1995:21). This would be exemplified in the example data in the instance of მაგ [makʰ] meaning ‘that’. If the final phone in this word were word-medial, it may have been pronounced [mga], but since it is word-final, the velar stop is devoiced and aspirated at the end to become [kʰ]. The phonological rules corresponding to these allophones are as follows:

/b/ → [b̥] / #_
[pʰ] / _#
[b] / elsewhere

/d/ → [d̥] / #_
[tʰ] / _#
[d] / elsewhere

/g/ → [g̊] / #_
[kʰ] / _#
[g] / elsewhere

These aforementioned stops in their voiceless forms — as well as other phones including the alveolar and post-alveolar fricatives which also include voiced forms, voiceless ejective forms, and voiceless aspirated forms (Shosted and Chikovani 2006:255) — can also appear as ejective in certain situations. Another stop that appears in an ejective form and displays some unique phonetic features is the uvular ejective stop [q’]. In the realm of phonetic features, it is unique in that it has many allophones, the most frequent being the uvular ejective fricative [χ’] according to Shosted and Chikovani, shown in the example data as ბაყალი [b̥aχ’ali] meaning ‘grocer’. Shosted and Chikovani cite these two phones as being in free variation (Shosted and Chikovani 2006:256-257) which do not mandate a rule to be written.

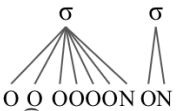
3 Syllable structure and phonotactics

Georgian has a remarkably complex syllable structure (Maddieson 2013) which contains numerous rules and restrictions. Known for its complex consonant clusters in both onsets and codas (Butskhrikidze 2002:197-205), there is a high amount of variation and many possible syllable structures in Georgian which cannot be fully explained by merely a few syllable trees, but this paper will provide six syllable trees in Figures 6-11 in which serve as a brief explanation

of syllable structures in Georgian. All six trees were created using words sourced from John Torikashvili's Georgian-English dictionary (Torikashvili 1992).

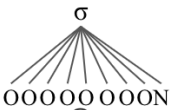
'north'

Figure 6: Syllable tree demonstrating CCCV, CV, and V syllable patterns.



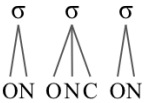
'trainer'

Figure 7: Syllable tree demonstrating CCCCCCV and CV syllable patterns.



'you peel us'

Figure 8: Syllable tree demonstrating CCCCCCCV syllable pattern.



'to break'

Figure 9: Syllable tree demonstrating CV and CVC syllable patterns.

'you (PL)'

Figure 10: Syllable tree demonstrating CCCVC syllable pattern.

'tuft of hair'

Figure 11: Syllable tree demonstrating CVCCCCCC syllable pattern.

The syllable trees in Figures 6-11 provide a brief insight into the many possible syllable structures of Georgian, most importantly featuring examples of the maximal onset (eight segments) in Figure 8 and the maximal coda (six segments) in Figure 11, both of which were indicated by Butskhrikidze (Butskhrikidze 2002:197-205) and Easterday (Easterday 2019:405). Though onsets tend to feature more prominently in Georgian (Butskhrikidze 2002:199-200), long, complex codas can appear in Georgian (VanDam 2004:35) though they are less common than onsets (Couch 2022:13). Given these maximal sizes and the data available according to previously discussed sources, a reasonable syllable template for Georgian would be as follows:

(C)(C)(C)(C)(C)(C)(C)(C)V(C)(C)(C)(C)(C)(C)

There is a large quantity of phonotactic restrictions on different segments of Georgian syllables which have been inventoried copiously by Easterday, and this paper will describe the most significant of such.

For onsets, all consonants appear as simple onsets while all complex and four or more segment onsets feature sonorants among the obstruents so that strings of otherwise consecutive obstruent consonants can be broken up and more easily pronounced, for instance, the liquid /r/ breaking up obstruent consonants in the onset of გვფრცქვნი /gvpʰrt͡skʰvni/ 'you peel us' in Figure 8 (Easterday 2019:405).

For codas, all consonants except /h/ occur in simple codas, and the same rule in onsets regarding the insertion of sonorants in more complex codas applies (Easterday 2019:405-406).

4 Conclusion

Overall, Georgian features a very unique phonological system. It is comprised of a diverse variety of phonemic units illustrating many instances of phonemic contrast, numerous examples of allophonic variation, a complex syllable structure and set of phonotactic restrictions, and notable typological features.

Some of the most notable phonological features of Georgian include its different allophonic rules, such as the variation of stop consonants to become weakly voiced, aspirated, or

ejective based on their position within a word. This is one of many instances of allophonic variation in Georgian but is a defining feature of the phonology of the language.

Additionally, Georgian boasts a uniquely complex syllable structure, containing a remarkable maximal onset of up to eight consonants and a maximal coda of up to six consonants. These dense consonant clusters can often create long, dense syllable segments, which add to the syllabic complexity of Georgian. This complexity has no cross-linguistic association other than its genetic relation to the other Kartvelian languages.

Another notable phonological feature of Georgian is its violation of the Sonority Sequencing Principle, wherein the most sonorous element of a syllable in Georgian does not serve as the nucleus, and less sonorous elements do not appear closer to the edges of the syllable. This is unique in that it rejects the ideas of a principle that was often believed to be a linguistic standard across languages.

IV. Typological profile of Georgian

1 Phonetic typology

In terms of typology, Georgian has 34 basic segments, according to the UCLA Phonological Segment Inventory Database, 28 of which are consonants and 6 of which are vowels (Georgian sound inventory (UPSID) 2019). Typologically, this is fairly close to the average overall number of segments among languages which is around 30, with 22 being consonant and 8 being vowels (UPSID Nr. of Segments), placing Georgian slightly higher than average for number of consonant segments and slightly lower than average for number of vowel segments. This is supported by the World Atlas of Language Structures Online, which says that Georgian has a "moderately large" consonant inventory compared to other languages, albeit an "average" sized vowel inventory (Language Georgian 2013).

However, it is in the type of phonetic segments that the Georgian language stands out. According to the UCLA Phonological Segment Inventory Database, the frequency index — which is presumably a quantified measurement of the frequency/commonality of the phonetic segments present in a given language — of Georgian is about 0.243. This is between 0.1, which UPSID labels as meaning "many very rare segments", and 0.39, meaning an "average" frequency of segments (Georgian sound inventory (UPSID) 2019). This means Georgian has a high frequency of phonetic segments that are considered rare. In examining UPSID's page on Georgian and database of phonetic segments, about 76.5% of segments in Georgian are represented in less than 25% of languages (Georgian sound inventory (UPSID) 2019).

In particular, Georgian contains some particularly uncommon sounds such as the voiceless ejective uvular plosive [q'] (Word List for Georgian), which UPSID claims is represented in only 3% of world languages (Georgian sound inventory (UPSID) 2019). An example in the provided data of this is in the word ყავა [q'ava] meaning 'coffee'. According to previous lectures, given the typological frequency of the uvular plosive, it is logical that a language like Georgian that has this phone would also have the more common velar, dental, and bilabial plosives.

A similar logic could apply to the less cross-linguistically common voiced alveolar affricate [dz], whose presence in Georgian (Word List for Georgian) makes more plausible the presence of the more common of the voiced affricates, the postalveolar [dʒ]. This particular phone occurs in common words such as ძელი [d͡zeli] meaning 'tree', while its postalveolar counterpart appears frequently, such as in ჯადო [d͡ʒado] meaning 'magic'. Similarly is the

presence of the voiced labiodental, alveolar, and postalveolar fricatives, which are less common and indicate the likely presence of their voiceless counterparts in Georgian.

While these phonetic traits of Georgian make it different from other languages in many ways, it still shares many features with what few languages to which it is genetically related. Its closest relatives, Svan and the two Zan languages, Mingrelian and Laz, all descend from Proto-Kartvelian (Hillery 1994). There are a few phonetic variations between Proto-Kartvelian and its daughter languages according to comparative research by German linguist and University of Tbilisi professor Heinz Fähnrich. According to Fähnrich, the phonetic inventory of the Kartvelian languages is largely consistent, with the only variation in vowel phones being the Zan languages, where the low-mid back rounded vowel [ɔ] takes the place of the low back unrounded vowel [ɑ], and the latter takes the place of the low-mid front unrounded vowel [ɛ] (Fähnrich 2002). For consonant phones, Fähnrich suggests just a few variations between the languages, almost all of which appear in the Svan and Zan languages, with Georgian rarely deviating from Proto-Kartvelian, save the replacement of the voiced labial-velar approximant [w] with the voiced labiodental fricative [v] in most situations.

Being a member of such an isolated language family means that Georgian has not been heavily influenced by non-Kartvelian languages. There are some instances of borrowing in its lexicon (Hewitt 1995), but its phonetics are attributed exclusively to genetically related languages according to sources referenced in this paper. This makes a cross-linguistic analysis of Georgian with any non-Kartvelian language largely illogical, as its phonetics are almost entirely unchanged from its parent language.

2 Phonological typology

Georgian is often atypical in its phonological system. Perhaps most notably abnormal is Georgian's phonotactics and syllable structure, which are classified as complex (Maddieson 2013). Georgian is particularly known for its complex consonant clusters, some of which being harmonic clusters that involve consonants of shared features of voicing, aspiration, or ejection which are pronounced with only a single release (Aronson 1990:33), such as the aspirated stops placed word-initially in თქვენ /tʰkʰven/ ('you' PL), as seen in Figure 10.

There are also many dense consonant clusters with upwards of eight consecutive consonants in a single syllable (Butskhrikidze 2002:197-205), such as გვფრცქვნი /gvpʰrt͡skʰvni/ 'you peel us' in Figure 8.

2.1 Cross-linguistic analysis

These abnormalities are rare, but are shared among the few languages genetically related to Georgian, Laz, and Mingrelian (which, along with the more distantly related Svan language, are the only languages in the Kartvelian language family, which is not related to any other language families) (Chikobava 1938).

Laz, for instance, is incredibly close to Georgian and shares all of its most common phonological processes as well as its complex syllable structure (Chikobava 1936). Similarly, Mingrelian shares an identical phonemic inventory with Georgian and the same phonological processes, and its phonotactics and syllable structure (Qipshidze 1914). Svan is the most distinct of the Kartvelian languages, as its most widely spoken dialects contain a larger phonemic inventory than its genetically related languages, with three more consonant phonemes than Georgian and a whopping 13 more vowel phonemes than Georgian (Siegel 2022).

However, the syllable structure and phonotactics of Svan once again are practically identical to that of Georgian (Tuite 1997). According to Winfried Boeder, a notable contributor to the body of linguistic research on the Kartvelian languages, these languages are incredibly similar in all respects, including phonologically (Boeder 2005).

All Kartvelian languages being so similar, they share a complex syllable structure (Maddieson 2013), illustrated by the given data. Across non-Kartvelian languages, this is out of the ordinary; according to the World Atlas of Language Structures Online (WALS), just under a third of surveyed languages have syllable structures classified as complex (Maddieson 2013). Geographically, WALS shows no significant concentration of syllable structure complexity in South Caucasian languages (that is, Georgia and surrounding regions) with a mix of complex and moderately complex syllable structures in the area. Considering this, with the knowledge that Georgian is not related to any languages outside of the Kartvelian family, one can assume that there is no cross-linguistic correlation regarding the complexity of syllable structures with Georgian.

2.2 Suprasegmental properties

In terms of suprasegmental properties, Georgian has no tone, and weak stress (Hewitt 1995:28-29). The rules regarding stress in Georgian are largely debated in the linguistic community; Aronson claims that a word with four or less syllables would require stress on the first or antepenultimate syllable, while a word with five or more syllables would require stress on the first and antepenultimate syllables (Aronson 1990:18). Robins and Waterson present a more complicated approach, wherein a two syllable word would place stress on the first syllable, a three syllable word on the first syllable or second syllable, a four syllable word on the second syllable or on the first and third syllables, a five syllable word on the first and third syllables or second and fourth syllables, and a six or more syllable word on the first and antepenultimate syllables (Robins and Waterson 1952:60). It has also been suggested that pitch accents on word-initial syllables could determine stress (Jun, S., et. al. 2007:54-56). Overall, Borise states that as stress is determined by the length of syllables as well as the position of a syllable in a word, it is thus predictable (Borise 2023:1-37).

When it comes to length, Kwon and Chitoran state that feature of length in Georgian is that vowels in syllables that are the first in a word have a longer length than vowels in medial or final syllables in a word (Kwon and Chitoran 2023:8). However, Easterday argues that there is no contrastive length in Georgian (Easterday 2019:405).

2.3 Adherence to phonological principles

Georgian does adhere to the Maximal Onset Principle, as the previous consonant clusters that appear word-initially can be syllabified into complex onsets — as per the syllable template — and since a consonant cannot be a syllable nucleus in Georgian, a consonant preceding a vowel will constitute its onset (Hewitt 1995:19-20).

However, it especially is unique in that it does not adhere to the Sonority Sequencing Principle. The Sonority Sequencing Principle (SSP) dictates that “the most sonorous element of a syllable serves as the nucleus; segments become progressively less sonorous towards the edge of the syllable” (Cooper 2024:3). Due to the complex syllable structure of Georgian allowing for practically any syllable onset to exist (Couch 2022:4-5) — especially the aforementioned consonant clusters as onsets — and the language’s tendency to omit codas, leaving the vowel nuclei at the end of the syllable (Butskhrikidze 2002:64), there are numerous instances in

Georgian where the SSP is violated. Take, for example, the single syllable word გვფრცქვნი /gvpʰr͡tsʰkʰvni/ ‘you peel us’ from Figure 8. If the SSP was complied with, as with many languages such as English, the most sonorous segments of this syllable would be located at the end. However, the sequence of segments of this syllable places an onset cluster comprised of low sonority segments, with the high sonority segments, particularly the vowel nucleus, placed at the very end. The sonority curve of this word is illustrated in Figure 7, where it is clear that it does not comply with the SSP due to the placement of the most sonorous segments. The word used in Figure 12 is sourced from Torikashvili’s *Georgian-English, English-Georgian Dictionary* and the analysis of sonority that guided this figure was sourced from Couch’s and Butskhrikidze’s discussions of sonority in Georgian phonology.

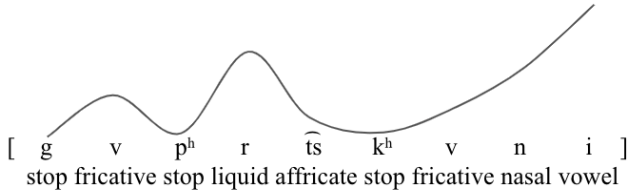


Figure 12: Sonority curve of non-SSP-compliant word გვფრცქვნი /gvpʰr͡tsʰkʰvni/ ‘you peel us’.

V. Project reflection

There were quite a few difficult decisions made in the research process, some of which related to choosing which example words to use for the multiple phonology-related data sets. These words, which were first featured as examples of syllable structure, were sourced originally from references in various sources as the maximal onset and coda sizes, and then found in Torikashvili’s dictionary with phonetic and phonemic transcriptions as well as syllable hyphenation so that I could create the trees (Torikashvili 1992). These were re-used for other examples, such as to explain the non-compliance with the SSP, because they were so unique cross-linguistically and yet featured many common trends within the Georgian language, which I rationalized in order to use them multiple times.

Additionally, preparing this paper consisted of much more of a deep dive into old and often foreign resources than I expected. This meant that one of the primary decisions I consciously made in preparing and presenting the linguistic content of my pape was to rely heavily on sources from Georgia, translated from Georgian, or otherwise produced by native Georgian speakers. This was vital to me, as I went into this knowing Georgia’s long history of conquest and influence by international powers. I wanted to ensure that the linguistic information I was analyzing was purely regarding the Georgian language, not Ottoman-Turkish-influenced Georgian, nor Russian-influenced Georgian, nor Arabic-influenced Georgian, or otherwise. Though the version of Georgian most widely spoken today is the standard Georgian, not any of the historically influenced versions of the language, I was concerned that referencing a source written in a particular historical context could lend itself to an implicit bias that would make it unreliable for the study of modern Georgian. Luckily, focusing on sources from Georgia — which I confirmed by researching the authors and publications from which I pulled information — showed no significant bias that would distance them from modern Georgian.

One of the most significant ways I employed this strategy was by obtaining my illustrative data of phonetic units from a Georgian-English dictionary produced by a Georgian linguist, John Torikashvili. I obtained this resource from the Open Language Archives

Community (OLAC) catalog of resources in and about the Georgian language, which is also where I found plenty of primary-source material; one of such materials was the Georgian translation of the Book of Genesis, which I originally intended to use for all of my example data, before deciding it was more efficient and informative to reference Torikashvili's dictionary with its included IPA transcriptions.

I found OLAC, as well as UPSID and many other resources that ended up not relevant enough to the sound system in particular to be used in this paper, from the resources page on Canvas. It was particularly helpful to have an academically credible jumping-off point in the form of these resources on Canvas, which I used to cross-reference with sources I found on my own, as well as gain a better understanding of linguistic information that may not be specific to Georgian.

Another way I drew from the material of our course on the phonological side was in defining the Maximal Onset Principle and, more prominently, the Sonority Sequencing Principle, with which I referenced the slides from class which were available on Canvas (Cooper 2024). This allowed me to provide a clear description of the SSP as I understood it, so that I could best describe how Georgian disobeyed it according to Couch (Couch 2022:4-5) and Butskhrikidze (Butskhrikidze 2002:64), both of which I found on OLAC.

The resources I found from OLAC among other sources were especially important, as a major resource-related constraint I found myself operating within the preparation of my paper was in the discrepancies across various sources in regard to certain phonetic data. One example of this is the phonemic inventory of vowel phones in Georgian. One of my more trusted sources, Shosted and Chikovani, asserted that there are five vowel phones in Georgian, those being [i], [ɛ], [ɑ], [ɔ], and [u] (Shosted and Chikovani 2006:262). However, another credible source, Aronson, claimed [i], [e], [ä], [o], and [u] comprised Georgian's inventory of vowel phones (Aronson 1990:18). Other sources, such as Luciano Canepari, claimed the vowel phones to be [i], [ɐ̯], [a], [ɒ̯], and [u], which was backed up by the credibility of being one of the more recently updated sources (Canepari 2007:385). The UCLA Phonetics Lab Archive even suggested an almost entirely different inventory of [ɪ], [ɛ], [a], [o], and [ʊ] (Word List for Georgian).

However, I ultimately decided to trust Yakov Testelets's inventory of vowel phones from 2021, which is what is displayed in Figure 3 (Testelets 2021). This was not only because Testelets provided the most recently updated source by far, but also because his paper accounted for the discrepancies across previously mentioned sources among others, and seemed to balance the most agreed-upon take on Georgian vowel phones. I also admired Testelets' inclusion of the schwa as it is used in consonant clusters, a primary focus of his research into Georgian, as this added some clarity to the pronunciation of Georgian and explained the complexities between research analyzing speaker data of Georgian with consonant clusters.

Similarly, there were many discrepancies across sources when it came to defining the syllable structure of Georgian. First of all, there were very few sources describing syllable structure that I could find, and those I could find were either potentially outdated, sourcing from the late 19th to early 20th century, or were in other languages that I could not find an English translation of to understand. However, upon finding Easterday's description of Georgian phonetics (Easterday 2019) and cross-examining with other sources, particularly Butskhrikidze (Butskhrikidze 2002), I was able to draw my own conclusions based on data provided.

VI. Project conclusion

Despite discrepancies among various sources regarding particular linguistic phenomena, there generally seems to be agreement on one thing: Georgian is, in countless ways, an incredibly unique language. The language is both defined by and unyielding to the complexities of its geopolitical, social, and historical contexts, and existing as one of very few members of a language family with no connection to any others sets it apart from the linguistic status quo.

When it comes to the sound system itself, Georgian displays many notable phenomena. Phonetically, Georgian has an inventory of phones that is just slightly larger than the surveyed average with a few unique units such as the voiceless ejective uvular plosive [q'] (Georgian sound inventory (UPSID) 2019). Phonologically, Georgian displays many interesting instances of phonemic contrast and allophonic variation. Perhaps most notable in terms of unique features is Georgian's infamously complex syllable structure, as well as the phonotactic rules that go along with it. Another notable feature is the lack of compliance with the SSP, rejecting the expected norm of most languages of the world.

All considered, Georgian is a fascinating language. Cross-linguistically, it stands out, in its larger context and in the specific instances where it defies commonly expected standards of a language such as the syllable structure and SSP non-compliance. Individually, it has many complex features and rules which define it, from phonetic units to phonetic contrast and allophonic variation. It is a vibrant and beautiful language, with an ancient history that runs through to its usage today, yet it is redefined by generation after generation as its community of speakers changes and modernizes through the ages. Linguistically and culturally, Georgian is not only interesting to study, but deeply worthwhile. There is so much to learn not only about the language, but also from it; it is defined by resilience, adaptability, history, and modernity, and undoubtedly stands out among the languages of the world.

VII. List of abbreviations

IPA	International Phonetic Alphabet
OLAC	Open Language Archives Community
SSP	Sonority Sequencing Principle
UPSID	UCLA Phonological Segment Inventory Database
WALS	World Atlas of Language Structures Online

VIII. References

Aronson, H. (1990). *Georgian: A Reading Grammar.* Columbus: Slavica.

Asatiani, N., and O. Janelidze. (2009). *History of Georgia: From Ancient Times to the Present Day.* Tbilisi: Publishing House Petite.

Boeder, W. (2005). The South Caucasian Languages. *Lingua, 115*(1-2).

Borise, L. (2023). Disentangling word stress and phrasal prosody: A view from Georgian. *Phonological Data and Analysis, 5*(1).

Butskhrikidze, M. (2002). *The Consonant Phonotactics of Georgian*. Utrecht: LOT.

Canepari, L. (2007). *Natural phonetics and tonetics: Articulatory, auditory, & functional*. München: Lincom Europa.

Chikobava, A. (1936). *Grammatical analysis of Laz with texts (in Georgian)*. Tbilisi.

Chikobava, A. (1938). *Chan-Megrel-Georgian Comparative Dictionary (in Georgian)*. Tbilisi.

Cooper, A. (2024). *Phonology Wrap-Up; Morphology* [PowerPoint Slides]. Canvas. https://northeastern.instructure.com/courses/167802/pages/class-14-f-feb-23-phonology-wrap-up-morphology?module_item_id=9722963.

Couch, C. (2022). Postcards from the syllable edge: sonority and articulatory timing in complex onsets in Georgian. *Journal of the International Phonetic Association, 53*(3).

Dalby, A. (2003). *Language in Danger: The Loss of Linguistic Diversity and the Threat to Our Future*. New York: Columbia University Press.

Easterday, S. (2019). Highly complex syllable structure: A typological and diachronic study. *Studies in Laboratory Phonology, 9.*

Fähnrich, H. (2002). *Kartwelische Wortschatzstudien*/*Georgian Language*. Jena: Friedrich-Schiller-Universität.

Georgian. (n.d.) Ethnologue, SIL International. https://www-ethnologue-com.ezproxy.neu.edu/language/kat/ (Accessed January 18, 2024).

Georgian sound inventory (UPSID). (2019). UCLA Phonological Segment Inventory Database. https://phoible.org/inventories/view/390. (accessed February 1, 2024)

Hewitt, B. (1995). *Georgian: A Structural Reference Grammar.* Amsterdam: John Benjamins Publishing Company.

Hillery, P. (1994). *Georgian: The Kartvelian Literary Language*. Pontypridd: Languages Information Centre.

Judeo-Georgian. (n.d.). Jewish Language Project, Sarah Bunin Benor. https://www.jewishlanguages.org/judeo-georgian (accessed January 20, 2024)

Jun, S., et. al. (2007). Intonational Phonology of Georgian. *UCLA Working Papers in Phonetics* (106).

Khurtsilava, B. (2017). The Inscriptions of the Georgian Monastery in Bi'r El-qutt and Their Chronology. *Christianity in the Middle East, 1*.

Kwon, H. and I. Chitoran. (2023). Perception of illusory clusters: the role of native timing. *Phonetica*.

Language Georgian. (2013). *The World Atlas of Language Structures Online, Max Planck Institute for Evolutionary Anthropology.* https://wals.info/languoid/lect/wals_code_geo (accessed January 17, 2024).

Living culture of three writing systems of the Georgian alphabet. (n.d.). UNESCO Intangible Cultural Heritage, UNESCO. https://ich.unesco.org/en/RL/living-culture-of-three-writing-systems-of-the-georgian-alphabet-01205 (accessed January 18, 2024).

Maddieson, I. (2013). Syllable Structure. *WALS Online* [Data set]. Zenodo.

McCoy, P. (1999). Harmony and Sonority in Georgian. *14th International Congress of Phonetic Sciences*.

OLAC resources in and about the Georgian language. (n.d.) Language Archives, Open Language Archives Community. http://www.language-archives.org/language/kat (accessed January 17, 2024).

Qipshidze, I. (1914). *The Grammar of Megrelian (Iver) Language with reader and dictionary.* St. Petersburg.

Robins, R. and N. Waterson. (1952). Notes on the Phonetics of the Georġian Word. *Bulletin of the School of Oriental and African Studies, 14*(1).

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

Siegel, W. (2022). Svan alphabet, language and prounciation. Omniglot. https://www.omniglot.com/writing/svan.htm. (accessed February 15, 2024)

Testelets, Y. G. (2021). Kartvelian (South Caucasian) Languages. *The Oxford Handbook of Languages of the Caucasus,* Oxford Handbooks.

The World Factbook: Georgia. (2024). The World Factbook, Central Intelligence Agency. https://www.cia.gov/the-world-factbook/countries/georgia/#communications (accessed January 19, 2024).

Torikashvili, J.J. (1992). *Georgian-English, English-Georgian Dictionary*. New York: Hippocrene Books.

Tuite, K. (1997). *Svan*. Languages of the World Materials. Munich: LINCOM-Europa.

UPSID Nr. of Segments. (n.d.). UCLA Phonological Segment Inventory Database.

http://www.phonetik.uni-frankfurt.de/upsid_nr_seg.html. (accessed February 1, 2024)

VanDam, M. (2004). Word Final Coda Typology. *Journal of Universal Language, 5*(1).

Word List for Georgian. (n.d.). UCLA Phonetics Lab Archive. http://archive.phonetics.ucla.edu/Language/KAT/kat_word-list_2000_01.html. (accessed February 1, 2024).

What are the largest language families. (n.d.). Ethnologue, SIL International. https://www.ethnologue.com/insights/largest-families/ (accessed January 18, 2024).