Abstract

This article surveys the major topics of Caucasian phonetics and phonology focusing on those aspects that bear broader implications for general phonetics and phonological theory. The article first presents an acoustic phonetic treatment of phonemic inventories in the three Caucasian families that involves both a review of recent instrumental data on the topic as well as a new analysis of experimental acoustic data. The focus of the phonetic analysis is obstruents with different laryngeal features, typologically unusual segments, small vocalic inventories, and pharyngealization. The second part reviews treatments of Caucasian phonotactics, primarily of South Caucasian consonant clusters that play a crucial role in the discussion on production vs. perception in phonology. The article concludes with a collection of phonological alternations that have the potential to bear broader theoretical implications.

Keywords: acoustic phonetics, naturalness in phonology, laryngeal features, ejective obstruents, double articulation, aspiration dissimilation, pharyngeals, epiglottals, consonant clusters, final voicing

1 Introduction

Among the most prominent characteristic features of Caucasian phonetics and phonology are large consonantal vs. small vocalic inventories in Northwest Caucasian, large consonant clusters in South Caucasian, and a large number of phonemes with place of articulation in the post-velar parts of the vocal tract in Northeast and Northwest Caucasian languages. Common to all three families is a three-way distinction in laryngeal features of obstruents that include ejective consonants. It is safe to say that beyond these focal topics, segmental phonetics and phonology of Caucasian languages are understudied fields, especially given their typological richness. This chapter aims to outline the main characteristics of Caucasian phonetics and phonology, focusing on processes that bear broader implications for general phonetics and phonological theory. Many questions remain open even for the well-studied aspects of Caucasian phonology. For some of these questions, the chapter aims to offer new insights with new experimental data or with reanalysis of existing data. The chapter also points to areas of Caucasian phonology that went largely unnoticed, but have the potential to provide insightful and broadly relevant results with future research.

The chapter begins with an overview of phonemic inventories. Most treatments of Caucasian phonology include detailed qualitative phonetic descriptions, but lack quantitative acoustic phonetic analyses. Section 2 of the chapter aims to fill this gap: for each language family, aspects of consonantal and vocalic inventories that are most relevant for phonetic typology are discussed; the
discussion is backed by a survey of existing acoustic analyses as well as by new analysis of the data. New experimental data is presented for South Caucasian: for the purpose of examining various aspects of South Caucasian phonetics, an experiment involving twelve speakers of Georgian and one speaker of Mingrelian who read 675 nonce words each was conducted (for the study in Beguš 2017). For Northwest Caucasian, new analyses of already existing data, primarily from Ubykh, are conducted and discussed. Section 3 of the chapter presents phonotactics of Caucasian languages and focuses on South Caucasian consonant clusters which played a major role in the development of articulatory approaches to phonology and are relevant for the discussion on the role of production vs. perception in phonology. The final section discusses active phonological alternations in Caucasian languages, both those that received due attention as well as those that went largely unnoticed, but are directly relevant to the focal topics of current phonological theories.

2 Phonemic Inventories

2.1 South Caucasian

Phonemic inventories are relatively similar in all four South Caucasian languages. The main characteristics of all four languages are symmetrical five-vowel systems with no length opposition, three-way oppositions in laryngeal features of stops and affricates (voiced, voiceless aspirated, ejective), and two-way opposition of laryngeal features in fricatives (voiced and voiceless). The common places of articulation are bilabial, labio-dental, dental or alveolar, post-alveolar, velar, uvular, and glottal. Table 1 illustrates the South Caucasian languages’ consonantal inventories while Table 2 presents their vocalic inventories.

<table>
<thead>
<tr>
<th></th>
<th>bilab.</th>
<th>l.-d.</th>
<th>dent.</th>
<th>alveo.</th>
<th>post-al.</th>
<th>pal.</th>
<th>vel.</th>
<th>uvul.</th>
<th>glot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td></td>
<td></td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap/trill</td>
<td></td>
<td></td>
<td></td>
<td>r/r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td></td>
<td></td>
<td></td>
<td>v</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approx.</td>
<td></td>
<td></td>
<td></td>
<td>s, z</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat. appr.</td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Combined consonantal inventories of South Caucasian languages (from Harris 1991, Tuite 1997, Hosted and Chikovani 2006, and Lacroix 2009). /x, u/ are usually analyzed as velar /x, y/ in Georgian and Mingrelian. /q/ is limited to Svan, /j/ is absent from Georgian, /ʔ/ is a phoneme only in Mingrelian. Svan has /w/ instead of /v/; in Laz and Mingrelian, [v] and [w] are allophonic.

<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i (y)</td>
<td></td>
<td>u</td>
</tr>
<tr>
<td>open-mid</td>
<td>ε (æ)</td>
<td>a</td>
<td>ø</td>
</tr>
<tr>
<td>low</td>
<td>(æ)</td>
<td>a</td>
<td></td>
</tr>
</tbody>
</table>

Table 2: South Caucasian vowel inventories (Holisky 1991, Hosted and Chikovani 2006). Georgian and Laz have five-vowel systems: /a, e, o, i, u/. /æ/ is limited to Svan and Mingrelian; in the latter it is a back vowel according to Harris 1991. Svan deviates from the system by introducing fronted /æ, ø, y/ (by umlaut; Tuite 1997). Upper Bal and Lashx dialects additionally feature length opposition (Tuite 1997).
Precisely laryngeal features are among the most well-studied aspects of South Caucasian phonetics: this is not surprising, as Georgian is one of the more accessible languages with ejective stops. This section focuses on laryngeal features in South Caucasian and discusses crucial information that these languages bring for phonetic typology of ejective obstruents. The section surveys previous acoustic data on laryngeal features in South Caucasian and provides new information on phonetics of ejective, aspirated, and voiced stops in Georgian.

2.1.1 Phonation

The exact phonetic realization of the three laryngeal features, especially in Georgian, is subject to extensive debate. The first point of dispute is the phonetic realization of the voiced series of stops in Georgian (Butskhrikidze 2002, Wysocki 2004, Shosted and Chikovani 2006, Vicenik 2010, Grawunder et al. 2010). While most studies analyze the voiced series as phonetically voiced (Shosted and Chikovani 2006, Vicenik 2010), recurring proposals have appeared in the literature that claim voiced stops are phonetically voiceless and characterize the Georgian opposition voiced vs. voiceless aspirated in terms of presence or absence of aspiration: aspirated vs. unaspirated (Robins and Waterson 1952, Wysocki 2004).

Wysocki (2004), for example, argues in favor of the latter approach and bases her claims on the fact that the phonetic realization of most absolute initial voiced stops features no phonation during closure. However, it is not clear that absolute word-initial position should be representative: phonation is dispreferred in absolute initial position for articulatory reasons (Solé 2011, Davidson 2016). Vicenik (2010) in his experiment measured amount of phonation into closure in Georgian post-vocalic stops, both word-initially and word-finally. On average 75% of closure is voiced in voiced stops, compared to 17% in voiceless aspirated and 27% in ejectives. The difference in voicing into closure between voiced vs. voiceless/ejective is significant, the difference among the latter two is not for most places of articulation. Moreover, Vicenik (2010) measures voicing into closure as a function of prosodic position: voicing into closure does not differ significantly in word-initial position after a vowel-final carrier phrase vs. word-medial intervocalic position: percent of voicing during closure does not significantly differ between the two conditions, except for velar stops. For velars, a significantly greater proportion of closure is voiced intervocically, but the absolute duration of voicing during closure remains the same (Vicenik 2010).

Our experimental data aligns well with these results. Twelve speakers of Georgian were instructed to read 675 nonce words of the structure CVCVCV. Voicing into closure was not quantified in the study, but acoustic inspection aligns well with the conclusion that the “voiced” series of stops in Georgian is in fact phonetically voiced in intervocalic position. Impressionistically, most speakers have at least half of the closure or more voiced (see Figure 1). Based on these observations and backed by the study in Vicenik (2010), we can maintain the analysis that the unaspirated series of Georgian stops is phonetically and phonologically voiced: amount of phonation into closure is substantial enough and significantly different for voiced stops compared to ejectives and voiceless aspirated that it likely does not result from automatic phonetic voicing in post-vocalic position (cf. Wysocki 2004). Our recordings also reveal an aspect of voiced stop realization that received less attention so far, but has the potential to shed further light on the system of laryngeal features in Georgian. There exists a good amount of variation in the realization of voiced stops among speakers, but relatively little within-speaker variation (although this observation is not quantified in our study). Speakers vary in the production of the intervocalic voiced stops to the degree that for some speakers closure is consistently fully voiced, whereas for some almost no phonation into the closure is present. Spectrograms and waveforms in Figure 1 illustrate this variation. Inter-speaker variation in the
production of voiced stops would warrant a further study. Variation in stop production is attested not only for the voiced series, but also for ejectives and aspirated stops. A study that would test potential correlations between inter-speaker variability in closure voicing with other phonetic parameters, such as VOT and burst intensity of the other two series of stops is lacking and would shed light on the otherwise unclear distributions.

A different issue, but related to the question of voiced stop realization, is what perceptual cues speakers employ in distinguishing stops with different laryngeal features (Vicenik 2010). It is possible that VOT duration (or presence/absence of aspiration noise) is a more prominent cue than voicing into closure. Such a hypothesis, however, should be tested with perceptual experiments. To my knowledge, no such studies exist for any of the four languages.

2.1.2 Phonetic properties of other laryngeal features

Three of more recent studies on Georgian phonetics present detailed acoustical analysis of stops with respect to their laryngeal features: Wysocki (2004), Vicenik (2010), and Grawunder et al. (2010). Of seven parameters measured (closure duration, VOT, voicing into closure, relative burst intensity, F0, phonation type of the following vowel, and spectral measures of the burst) in Vicenik (2010), only VOT duration and phonation type reliably distinguish all three stop types in Georgian. VOT is longest in the voiceless aspirated series, shorter in ejectives, and shortest in voiced stops. Vowels have significantly more creaky phonation after ejectives and more breathy phonation after voiceless aspirated stops. Voicing into closure distinguishes voiced stops from voiceless aspirated and ejectives, but not ejectives from voiceless aspirated. F0 in the following vowel falls for voiced and voiceless aspirated, but stays flat after ejectives; yet this difference does not reach significance for all places of articulation. Mean frequency, skewness, and kurtosis of the burst were significant predictors for only a subset of places of articulation, according to Vicenik (2010).

While several results in Wysocki (2004) and Vicenik (2010) are replicated in our experiment, some of the results are also different. Our analysis in Beguš (2017) confirms that stops with different laryngeal features have significantly different VOT durations. VOT of the third stop in nonce words of the CVCVCVCV structure was measured. Three places of articulation (bilabial, dental, velar) and three laryngeal features (voiced, voiceless aspirated, ejective) were tested in the study. VOT was significantly longer in voiceless stops compared to ejectives ($\beta = 36.1$ ms, $t = 9.8$, $df = 11$, $p < 0.0001$) and significantly shorter in voiced stops compared to ejectives ($\beta = -18.1$ ms, $t = -6.1$, $df = 11$, $p < 0.0001$) at the means of other predictors (Beguš 2017).
As already mentioned, some results from our study in Beguš (2017) also differ from those in Wysocki (2004), Vicenik (2010), and Grawunder et al. (2010). All these studies claim that closure duration does not significantly differ across stops with different laryngeal features. This is a surprising result as we know that closure duration of voiced stops is cross-linguistically shorter at least compared to closure of voiceless stops (Lisker 1957, Port 1981, Luce and Charles-Luce 1985). The results of our experiment, however, show a significant effect of laryngeal features on closure duration (Beguš 2017). Closure duration is significantly shorter in voiced stops than in ejectives ($\beta = -5.0$ ms, $t = -3.3$, df = 20, $p < 0.01$) and significantly longer in ejectives than in voiceless aspirated stops ($\beta = 3.7$ ms, $t = 3.0$, df = 17, $p < 0.01$) at means of other predictors (Beguš 2017).

This difference in closure durations between ejective and voiceless aspirated stops is, to my knowledge, reported for the first time in Beguš (2017), not just for Georgian, but for ejective stops in general. Warner (1996) measures closure duration in ejective vs. voiceless stops in Northeast Caucasian Ingush. Closure duration was indeed shorter in plain unaspirated voiceless stops, but the difference did not reach statistical significance. Her measurements are based on recordings of only one speaker. Gordon and Applebaum (2006) measure closure duration of stops in Northwest Caucasian Kabardian. They find a significant effect of laryngeal features on closure duration, but the difference is significant only for the difference between voiced and ejectives, not between ejectives and voiceless aspirated stops. Maddieson et al. (1996) measured closure duration and VOT of voiceless and ejective stops in NEC Tsez and found no significant differences between the two groups for neither of the measured parameters. McDonough and Ladefoged (1993) found no differences in closure duration between ejective and voiceless aspirated or unaspirated stops in Navajo (Athabaskan), but unaspirated stops have significantly longer closure durations than aspirated stops. Similar results are obtained for Athabaskan Witsuwit’en in Hargus (2007): closure is longest in unaspirated voiceless stops, but there is no significant difference in closure duration between voiceless aspirated and ejective stops.

The differences in closure durations across different stop types, despite being significant, are relatively small in Georgian. We can safely maintain that such small differences do not provide a prominent perceptual cue: in fact, they fall below the Just-Noticeable Difference ratio for vocalic speech stimuli (about 0.055 in Nooteboom and Doodeman 1980). The ratio of closure duration difference between ejective and aspirated stops compared to the full closure duration is 0.045 in our case (i.e. below 0.055).

Wysocki (2004) offers a detailed qualitative analysis of ejective stops in Georgian. Realization of ejective stops is highly variable across speakers. Ejective stops feature an audible oral release that is often higher in amplitude than the first period of the following vowel. Oral release is followed by a period of silence that is interrupted by one or more glottal releases that precede the vowel onset (Wysocki 2004). Figure 2 illustrates acoustic characteristics of an ejective stop in Georgian.
Shosted and Chikovani (2006) also provide a qualitative airflow analysis of different stop types in Georgian. Oral air flow of three initial dental stops are measured. Ejective stops feature a relatively small rise in airflow that rapidly drops back to the zero level before the onset of the vowel. This fall results from glottis being closed after the oral release: once the air flow between oral closure and glottal closure is released, glottal closure prevents further airflow until it is released and the vowel onsets. Volume of the ejective oral release is considerably lower than airflow volume of voiced stops or voiceless aspirated stops. The latter have the highest volume of air flow, although the distribution is not quantified in Shosted and Chikovani (2006).

2.1.3 Ejective stops and vowel duration

One phonetic parameter involving laryngeal features that has, to my knowledge, not been measured either for Georgian or for any other language and has the potential to shed light on a long standing question in phonetics, is duration of vowels before ejective stops. Vowel durations have long been known to differ before voiced and voiceless obstruents: vowels are longer before voiced than before voiceless obstruents and this generalization has been confirmed for over a dozen languages (cf. Chen 1970 among others). The causes of these durational differences are, however, poorly understood and several competing explanations have been proposed (Beguš 2017). The opposing explanations arise primarily because most studies measure vowel duration before voiced and voiceless obstruents exclusively and disregard vowel duration before obstruents with other laryngeal features. Only a subset of the studies measure vowel duration as a function of aspiration of the following stop. While most studies argue that aspiration lengthens the preceding vowel, some results are inconclusive (Maddieson and Gandour 1976, Ohala and Ohala 1992, Lampp and Reklis 2004, Durvasula and Luo 2014). Measurements of vowel duration before ejective stops are lacking altogether. Georgian is in this respect especially informative, as it features voiceless aspirated, ejective, and voiced stops in its inventory and thus allows measurements of vowel duration differences before stops with all three laryngeal features.

Beguš (2017) presents results from the experiment described above (in section 1) with nonce-words of the structure CVCVxCyV: vowel duration of Vx was measured along with the closure duration and VOT of Cy. Vx included three vowels ([a], [e], and [o]) and Cy involved all three laryngeal features and three places of articulation (3 × 3 levels).

A model with four predictors (Laryngeal Features, Vowel, Place, and Closure duration reveals that vowels are significantly longer before voiced stops compared to ejectives ($\beta =8.7$ ms, $t =7.1$, $df =10$, $p < 0.0001$) and significantly shorter before voiceless aspirated stops compared to ejectives ($\beta =-4.7$ ms, $t =-7.2$, $df =11$, $p < 0.0001$) (Beguš 2017). This generalization is called the “ejection effect” (in Beguš 2017) and is, to my knowledge, reported for the first time on Georgian material. Closure duration is significantly, but only slightly negatively correlated with preceding vowel duration (for ejective stops, $\beta =-1.8$, $t =-2.4$, $df =21$, $p < 0.05$; Beguš 2017). There is also a significant interaction between Laryngeal Features and Closure: closure duration is significantly more negatively correlated with preceding vowel duration in voiced stops compared to ejectives.

Our experimental design also allowed us to model duration of VOT and its effect on preceding vowel duration. To my knowledge, no studies before Beguš (2017) modeled effects of laryngeal features and VOT at the same time. As shown in 2.1.2 above, VOT significantly differs across stops with different laryngeal features. It would thus be conceivable to assume that VOT duration is the primary factor that determines preceding vowel duration. The results in Beguš (2017), however,
show that laryngeal features remain significant predictors even if we add VOT to the model. Vowels are longest before voiced, shorter before ejective, and shortest before voiceless aspirated stops. In addition, closure and VOT both inversely affect preceding vowel duration (Beguš 2017).

As argued in Beguš (2017), the “ejection effect” has broader implications: it shows that laryngeal features are significant predictors of preceding vowel duration, even when effects of closure and VOT are controlled for. Several competing proposals exist for the causes of vowel duration differences before different stop types (Chen 1970, Beguš 2017). The results show that voice feature or closure duration alone cannot be the cause of vowel duration differences. Moreover, because laryngeal features remain significant predictors even when VOT and closure duration are controlled for, and because the durational differences are small (smaller than the Just-noticeable difference threshold), perception is likely not the primary factor for durational differences. The “ejection effect” primarily supports two hypotheses: the Laryngeal Accommodation hypothesis (Halle and Stevens 1967) that claims laryngeal features require different more or less complex laryngeal gestures which in turn require different times to be achieved and (ii) the Timing hypothesis (Kozhevnikov and Chistovich 1967) claiming that timing across syllables tends to be constant: longer closure or VOT duration is compensated with shorter vowel duration. It is thus likely that several factors influence vowel duration and the new data from Georgian crucially contributes to the discussion on what exactly these causes of durational differences in vowels are (Beguš 2017).

2.1.4 Aspiration

Another aspect of Georgian phonetics that contributes to the general discussion on phonetics, and more precisely, to the discussion on the role of perception vs. production in phonetics, are voiceless aspirated stops. Aspirated stops in Georgian are characterized by long VOTs with high oral airflow (Shosted and Chikovani 2006). Least-square mean VOT duration of voiceless aspirated stops in our experiment (involving 2,630 tokens of aspirated stops in position CVVCV across the three places of articulation) is 75.4 ms (70.4 ms for labials, 72.8 ms for dentals, and 83.0 ms for velars). Georgian with its prominent aspiration provide insights into a process that has received increased attention, especially recently (Ohala 1993, Garrett 2015, Jatteau and Hejná 2016): aspiration dissimilation (AD) — dissimilation of two subsequent aspirated stops (T<sup>1</sup>...T<sup>b</sup> > T<sup>1</sup>...T<sup>b</sup> or T<sup>1</sup>...T). While AD is well-documented and relatively common process, mechanisms that underlie it are poorly understood. Two lines of thought emerge in the discussion. Ohala (1993) explains the dissimilation in terms of perceptual hypercorrection: speakers assume that aspiration of two subsequent stops is a result of assimilation and “undo” this assimilation. Garrett (2015), on the other hand, claims that motor planning errors are responsible for gradual shortening of aspiration, which over time results in aspiration dissimilation. The articulatory explanation predicts gradual and small phonetic differences in VOT duration, while the perception explanation predicts catastrophic and total dissimilation. Gradual and small-magnitude shortening of VOT supporting the articulatory explanation have already been reported for Aberystwyth English (Jatteau and Hejná 2016) and Halh dialect of Mongolian (Svantesson and Karlsson 2012). In both Halh and Aberystwyth English, however, the second stop is pre-aspirated and not post-aspirated, which means that the measured aspiration time in these languages surfaces on the same vowel. Differences in aspiration duration have not yet been established for a language with only post-aspirated stops (Jatteau and Hejná 2016), such as Georgian.

To test the effect of preceding aspiration on VOT duration of the following aspirated stop, we measured VOT duration of voiceless aspirated stops of three places of articulation after three vowels ([ə], [ɛ], and [o]) in Georgian. The VC<sub>x</sub> sequences were embedded into two frames: [rub-i] and [vot<sup>b</sup>-i], i.e. a frame in which the first part ends in a voiced stop and a frame with the first
part ending in a voiceless aspirated stop. VOT of $C_x$ was measured from the onset of oral release until first periodic vibration with clear formant structure of the following vowel. Altogether 213 tokens were analyzed. VOT was shorter in voiceless stops that were preceded by another voiceless aspirated stop. Figure 3 illustrates by-speaker differences in VOT duration as a function of preceding stop.

Figure 3: Boxplot of VOT duration across twelve speakers (left) and the effect of preceding consonant on VOT duration obtained from a linear mixed effects model with standard error bar (right)

The data were fit to a linear mixed effects model\(^1\) with preceding consonant Type, preceding Vowel, and Place of articulation with Vowel $\times$ Place interaction as fixed effects and a random intercept for Speaker. VOT is significantly shorter if another voiceless stop precedes ($\beta = -11.5$ ms, $t = -4.6$, $df = 192$, $p < 0.0001$).

This generalization is, to my knowledge, reported for the first time in Georgian and for the first time in a language with only post-aspirated stops in general. The results bear implications for understanding of mechanisms behind aspirate dissimilation process and for the discussion on perception vs. production in phonetics: the gradual effect of dissimilation of two post-aspirated stops in our experiment renders support for the articulatory explanation.

### 2.1.5 Mingrelian, Laz, and Svan

The other three South Caucasian languages have very similar phoneme inventories as Georgian (see Tables [1] and [2]), but have received considerably less attention, especially in the recent literature. A detailed phonetic analysis of Mingrelian and Laz phoneme inventory is offered among others in Imnadze (1981). More recently, Lacroix (2009) describes Arhavi dialect of Laz; Öztürk and Pöchtrager (2011) Pazar dialect of Laz.\(^2\)

While phonemic inventories are very similar in all four South Caucasian languages, exact phonetic realizations of phonologically identical segments can differ substantially, even in closely related languages or even between different dialects of the same language. Melikishvili et al. (2011) measure several acoustical parameters of ejective stops across Caucasian languages, including Northeast and Northwest Caucasian languages. The results suggest that exact phonetic realization of ejective stops differs across dialects and languages, although no detailed statistical analyses are presented in their study. VOT duration of ejective stops differs across dialects: the longest VOT was measured in Svan with 49 ms and shortest in Gurian dialect or standard Georgian measured at 26 and 25 ms (Melikishvili et al. 2011). The result of such contrastive studies also have the potential to shed light on the long-standing discussion of ejective stop typology (Kingston 2005, Grawunder et
al. 2010), but further studies with statistical analysis are needed to confirm differences in phonetic realization of ejectives across Caucasian languages and dialects.

Despite these dialectal phonetic differences, the “ejection effect” proposed in Beguš (2017) seems to hold true for Mingrelian as well. In addition to the twelve speakers of Georgian, Beguš (2017) also recorded a male speaker of Mingrelian reading the same 675 nonce words, but in a Mingrelian carrier phrase. The confound of the Mingrelian experiment is that it only includes one speaker who reported to speak Georgian as a first language and who spoke Mingrelian in his home village with grandparents and relatives. For the Mingrelian speaker too, vowels are shorter before voiceless stops than before ejectives ($\beta = -5.2$ ms, $t = -2.1$, $df = 623$, $p < 0.05$) and longer before voiced stops than before ejectives ($\beta = 28.4$ ms, $t = 8.4$, $df = 613$, $p < 0.0001$) at the means of other predictors. This suggests that the ejection effect may be universal, at least for languages with similar realization of ejective stops to those of Georgian and Mingrelian (Beguš 2017).

This section aimed to outline descriptive phonetic facts of South Caucasian phoneme inventories focusing on their most prominent aspect — laryngeal features. The discussion of descriptive facts is supplemented by experimental results that bear several implications for general phonetics. Acoustic studies of laryngeal features in SC provide crucial information for the effects of ejective stops on preceding vowel duration, their intrinsic closure duration as well as for VOT duration in general. Moreover, phonetic behavior of post-aspirated stops in Georgian in the environment with or without a preceding aspirated stop show support in favor of articulatory origins of the aspiration dissimilation process. Several aspects of South Caucasian laryngeal features merit further investigation, such as correlation in variation of inter-speaker stop realization or investigation of perceptual cues for laryngeal features disambiguation.

### 2.2 Northwest Caucasian

One of the most prominent features of Northwest Caucasian languages are their large consonantal and small vocalic inventories (for counts, see Catford 1977). Tables 3 and 4 illustrate the striking difference in the size between consonantal and vocalic inventories.

<table>
<thead>
<tr>
<th>stop</th>
<th>bilab.</th>
<th>l.-d.</th>
<th>alv.</th>
<th>alv.-pal.</th>
<th>post-al.</th>
<th>pal.</th>
<th>vel.</th>
<th>uvl.</th>
<th>phar.</th>
<th>glot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>labial</td>
<td>p, b, p' (p&lt;sup&gt;W&lt;/sup&gt;)</td>
<td>t, d, t'</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td>k, g, k'</td>
<td>k&lt;sup&gt;W&lt;/sup&gt;, g&lt;sup&gt;W&lt;/sup&gt;, k&lt;sup&gt;W&lt;/sup&gt;'</td>
<td>q, q'</td>
<td>q&lt;sup&gt;W&lt;/sup&gt;, q&lt;sup&gt;W&lt;/sup&gt;'</td>
<td>(?), (?&lt;W&gt;)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>palat.</td>
<td>p'&lt;sup&gt;W&lt;/sup&gt;, b', p' &lt;sup&gt;W&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pharyng.</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phar.&amp;lab.</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>t$\wedge$, d$\wedge$, t'&lt;sup&gt;W&lt;/sup&gt;</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td>j, ë, j'</td>
<td>t$\wedge$, d$\wedge$, t'&lt;sup&gt;W&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>m</td>
<td>m&lt;sup&gt;W&lt;/sup&gt;</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap/trill</td>
<td>t/ç</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td>l, (l'), [v] (l&lt;sup&gt;W&lt;/sup&gt;)</td>
<td>s, z, (s&lt;sup&gt;W&lt;/sup&gt;)</td>
<td>s&lt;sup&gt;W&lt;/sup&gt;, z&lt;sup&gt;W&lt;/sup&gt;, (s&lt;sup&gt;W&lt;/sup&gt;')</td>
<td>j', ë, (j') (l&lt;sup&gt;W&lt;/sup&gt;', ë&lt;sup&gt;W&lt;/sup&gt;')</td>
<td>š, ž, (š&lt;sup&gt;W&lt;/sup&gt;)</td>
<td>š&lt;sup&gt;W&lt;/sup&gt;, ž&lt;sup&gt;W&lt;/sup&gt;, (š&lt;sup&gt;W&lt;/sup&gt;')</td>
<td>(ç), j</td>
<td>x, ñ (x&lt;sup&gt;W&lt;/sup&gt;)</td>
<td>x&lt;sup&gt;W&lt;/sup&gt;, ñ&lt;sup&gt;W&lt;/sup&gt;</td>
<td>x&lt;sup&gt;W&lt;/sup&gt;'</td>
</tr>
<tr>
<td>labial</td>
<td>l</td>
<td>l&lt;sup&gt;W&lt;/sup&gt;</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>palat.</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pharyng.</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>phar.&amp;lab.</td>
<td>t&lt;sup&gt;W&lt;/sup&gt;, d&lt;sup&gt;W&lt;/sup&gt;, t&lt;sup&gt;W&lt;/sup&gt;'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat. fric.</td>
<td>ţ, (ţ), ţ'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approx.</td>
<td>w</td>
<td>w&lt;sup&gt;W&lt;/sup&gt;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>pharyng.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lab.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat. appr.</td>
<td>l</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

($\beta = \ldots$, $t = \ldots$, $df = \ldots$, $p = \ldots$)
Table 3: Consonantal inventory of Ubykh with some prominent phonemes limited to other Northwest Caucasian languages added in brackets (from Catford 1977, Smeets 1984, Colarusso 1988, Chirikba 2003, Hewitt 2004, Gordon and Applebaum 2006, 2013, Fenwick 2011). The “peculiar NWC” (Catford 1977) series of “laminal closed post-alveolar” sibilants (Ladefoged and Maddieson 1996:162) is transcribed with ‘ˆ’ (e.g. /ˆs/). For a discussion on phonetics of this series, see section 2.2.2. Some dialects of Adyghe (Hatkoy, Shapsugh) feature a four-way opposition in obstruents (plain voiceless, aspirated, voiced, ejective; Gordon and Applebaum 2013) which are not represented in the table. For a detailed phonetic study of the four-way opposition in Circassian, see Gordon and Applebaum (2013). Note that phonemic inventories can vary substantially across different analyses, especially for coronal fricatives and affricates. Phonetic values of these two series can vary substantially (see discussion in section 2.2.2 and in Gordon and Applebaum 2013).

Table 4: A typical Northwest Caucasian vocalic inventory consisting of only two phonemic vowels (Colarusso 1988). For the phonetic realization of the two phonemic vowels as well as for different proposals, see discussion in section 2.2.3.

This section focuses on acoustic phonetic aspects of large phonemic inventories: NWC languages feature several typologically highly unusual segments that were in many cases considered impossible or unattested until they were discovered precisely in these languages. Where available, I present a new acoustic analysis of typologically rare segments based on existing recordings of Ubykh from online databases LaCiTO and the UCLA Phonetics Lab Archive.10 I also provide statistical analyses of observations that have already been established, but have not yet received quantified treatments. The section further presents an acoustic analysis of small vocalic inventories that are highly variable and coarticulatorily influenced by the large consonantal inventories.

Among NWC languages, Ubykh stands out as the system with the highest number of consonantal phonemes: 80-85 (depending on the analysis, Fenwick 2011). Moreover, Ubykh has the highest number of consonantal phonemes of any language without clicks. It features three laryngeal features: voiced, voiceless, and ejective; four secondary articulations: palatalized, labialized, pharyngealized, and labialized and pharyngealized; six manners of articulation: stop, fricative, affricate, nasal, approximant, and trill; and ten places of articulation: bilabial, labiodental, alveolar, alveolo-palatal, two post-alveolar series (one of which is analyzed as retroflex or subapico-palatal), palatal, velar, uvular, and glottal (Colarusso 1988, Fenwick 2011). Different descriptions vary in their analysis of Ubykh places of articulation, but they all point to a disproportionately high number of segments in the post-alveolar and uvular regions. Other NEC languages additionally feature pharyngeal place of articulation.

2.2.1 Typologically rare segments

Unique to Ubykh is a complete series of plain, pharyngealized, labialized and labialized and pharyngealized uvulars that include voiceless aspirated and ejective stops as well as voiceless and voiced fricatives (Colarusso 1988). These oppositions result in segments as typologically unusual as the labialized pharyngealized uvular ejective /q̥w̥/. It has been hypothesized that “rounding and pharyngealization are never distinctive within a language” (Jakobson et al. 1951 in Colarusso 1988:221). As shown by [q̥w̥] and other labialized pharyngealized consonants, Ubykh (as well as in some dialect of Abkhaz and in Abaza; Chirikba 2003) can feature both secondary articulations on a single segment. To my knowledge, no language outside of the Northwest
Caucasian family features simultaneous labialization and pharyngealization. On the other hand, some of the most common stops, such as plain voiceless or voiced velars /k/ and /g/ are marginal in Ubykh (and in some other languages, such as Kabardian), appearing only in one loanword each (Colarusso 1988). Such mismatches between cross-linguistic frequency and markedness are relevant for general constraint architecture of Optimality Theoretic frameworks (Prince and Smolensky 1993/2004). Currently, the most widely-accepted version of Optimality Theory and Harmonic Grammar with restricted Con predicts less marked segments will be more frequent in a given environment (Beguš and Nazarov 2017, Hayes 2016). Languages such as Ubykh pose a problem for such predictions. Further experimental work is needed to confirm that speakers internalize such “unnatural” restrictions: if they do, the OT/HG family of theoretical approaches needs a revision in their constraint architecture.

Unique to the Abdzakh dialect of Adyge are glottal stops with two secondary articulations: the dialect features a contrast between plain /l/, labialized /lw/, and palatalized /lj/ glottal stops (Catford 1983). Secondary articulations on glottal stops are rare: palatalized /lj/ was even considered impossible (Merlingen 1977 in Catford 1983) until the segment was discovered in Abzakh. Catford (1983) presents spectrograms of plain vs. palatalized glottal stop: his analysis confirms an acoustic distinction between the two phonemes that is reflected in the formant structure of the following vowel, although the effects are phonetically weak.

A prominent feature of many NWC languages are labialized alveolar obstruents. Their phonetic realization bears implications for phonetic typology in general: labialization in alveolar stops is realized in NWC as labial closure and in fricatives as labial frication. For example, /tw/ is realized as a typologically rare doubly articulated bilabio-alveolar [tp].

Ubykh shares doubly articulated bilabio-alveolar stops with Abkhaz (Colarusso 1988). For the purpose of examining acoustic properties of labialized alveolars in NWC, we analyze recordings of Mr. Tevfic Esenç uttering words /dwa/ ‘awl’ and /da/ ‘now’ in isolation from the LaCiTO online database. Figure 4 shows spectrograms of the two words. The spectrograms show that the labialized alveolar is realized as a doubly articulated stop [dB]. While it is possible that the mild, but abrupt rise in energy at about 15 ms before the full release in the spectrogram for /dwa/ shows an alveolar release (before the labial release), further and more accurate recordings and articulatory data are needed to confirm this analysis. The formant structure provides cues for acoustic disambiguation of /d/ and /dw/: we observe lowering of all formants for the labialized stop, but especially of F3 and F4.
To test effects of doubly articulated stops on formants of the following vowel, we measured formant values in the same recording from the LaCiTO database that contains three repetitions of /d\textsuperscript{w}a/ and two repetitions of /da/. Formants at 5% of vowel duration were measured in Praat (Boersma and Weenink 2015) with a modified Vowel Analyzer Praat script (Riebold 2013). Labialization lowers F2 ($Z = 1.86, p = 0.1$) and F3 ($Z = 1.74, p = 0.2$), but with such small sample size, no differences are significant (according to Exact Two-Sample Fisher-Pitman Permutation Test, based on the one-way test function from the coin package, Hothorn et al. 2006). Further data with larger sample sizes in Abkhaz should yield more conclusive results.

Labialized alveolar stops (/d\textsuperscript{w}a/) that are realized as doubly articulated stops ([\textipa{db}]) also appear acoustically distinct from plain labials (/b/). A recording of Ubykh words /\textipa{abana}/ and /\textipa{ad\textsuperscript{w}ana}/ from the LaCiTO database was analyzed. The recording contains three tokens of each word. F1 is lower at 5% of vowel duration after /d\textsuperscript{w}/ compared to position after a plain labial ($Z = 2.06, p = 0.1$). F3 likewise lowers after /d\textsuperscript{w}/ ($Z = 2.20, p = 0.1$). F2, however, is higher after /d\textsuperscript{w}/ compared to /b/ ($Z = -2.05, p = 0.1$), but none of these differences reach statistical significance. Acoustically distinct is also the vowel preceding /d\textsuperscript{w}/, compared to the vowel preceding /b/, but again the differences are not statistically significant (according to the Permutation Test).

Based on the analysis above, we can conclude that Ubykh shows support for the existence of a rare type of double articulation: bilabio-alveolar doubly articulated stops. While doubly articulated stops involving velar and labial closure are not infrequent, those involving labial and alveolar closure are considered very rare. It has even been claimed that no language features a bilabio-alveolar doubly articulated stop as a contrastive phoneme (Maddieson 1983). In light of these claims, Ladefoged and Maddieson (1996) suggest that double articulation in Caucasian might be better analyzed as a secondary articulation, based on the articulatory description from Catford (1972) (via Ladefoged and Maddieson 1996) who claims that labial contact in the closure of doubly articulated stops is “light”, lips are protruded further forward and contact is made with the inner part of lips. It is not immediately clear, however, why these articulatory properties described in Catford (1972) would necessarily point to an analysis with secondary articulation or even what counts as a distinctive criterion for disambiguating secondary articulation with complete closure and “true” double articulation. Moreover, since Maddieson (1983), other cases of bilabio-alveolar doubly articulated stops were found: Ladefoged and Maddieson (1996) present an analysis of Yelentye (spoken in Papua New Guinea) as featuring truly bilabial-alveolar doubly articulated stops. Most analyses in the literature thus rightfully maintain the double articulation status of labialized stops in Northwest Caucasian.

While voiceless and voiced doubly articulated bilabio-alveolar stops are featured in other languages such as Yelentye, no language outside Northwest or Northeast Caucasian, to my knowledge, features doubly articulated bilabio-alveolar ejective stops, such as Ubykh [\textipa{\textsuperscript{\textipa{t}p}}].

Further articulatory studies of labialized alveolars, especially of the ejective series, in languages such as Abkhaz, where labialized alveolars are also realized as doubly articulated stops, should reveal further information about phonetic properties of these rare segments. Because the phenomenon is typologically rare and bears broader implications for phonetic typology of double articulations, further articulatory studies are a desideratum.

The realization of labialized alveolars in Northwest Caucasian also provides evidence for diachronic origins of doubly articulated stops, which has, to my knowledge, not been discussed so far. Ubykh and Abkhaz examples point to one potential source of doubly articulated stops: labialization as secondary articulation.
2.2.2 Fricatives

Another prominent feature of Northwest Caucasian phoneme inventories is presence of a high number of fricative phonemes. As summarized in Catford (1977) and Gordon and Applebaum (2013), a canonical NWC inventory includes four sibilant fricative series: “apico- or lamino-alveolar” /s/, alveolo-palatal /ɕ/, “apico-postalveolar (slightly velarized)” /ʃ/ (sometimes analyzed as retroflex, e.g., in Fenwick 2011), and the “peculiar NWC sibilant” /̂s/. This latter is described as “laminal closed post-alveolar” (in Ladefoged and Maddieson 1996:162) and as a “hissing-hushing sound” articulated with tongue tip “rest[ing] against the alveoles of the lower teeth” with “the main articulatory channel [...] at the back of the alveolar ridge” (Catford 1977:290) and “not produced with the sublingual cavity that often characterizes postalveolar fricatives cross-linguistically” (Gordon and Applebaum 2013). Adding laryngeal features to this four series of sibilants as well as labialization as secondary articulations, systems can feature up to 14 phonemic coronal sibilant fricatives, e.g. in Bzhedukh (Catford 1977) or Shapsugh (Gordon and Applebaum 2013) dialects of Adyghe. Figure 5 shows spectrograms of the four-way contrast with contrastive labialization in two series in Ubykh (the spectral analysis was performed in Praat, Boersma and Weenink 2015, based on the recordings from the UCLA Phonetics Lab Archive). The spectrograms show the generalization established for Kabardian (Gordon and Applebaum 2006) and for sibilants in general (Gordon et al. 2002): peak energy is higher for more anterior fricatives. For spectral studies of Kabardian fricatives and their influences of formant structure of the following vowel, see Gordon and Applebaum (2006); for spectrograms and X-ray tracings of Ubykh and other NWC languages, see Colarusso (1988).

Other NWC dialects can merge the four-way opposition into a three-, two-, or one-way opposition (Gordon and Applebaum 2013). The exact phonetic realization as well as analyses of the four series of fricatives can vary substantially. For example, some NWC sibilants are analyzed as “moderately retroflex”: /ʃ/ is analyzed as /ɕ/ in Ubykh in Fenwick (2010:18) (see also Gordon and Applebaum 2013). For a detailed phonetic treatment of NWC sibilants, see Colarusso (1988), Ladefoged and Maddieson (1996), Gordon and Applebaum (2006, 2013), Fenwick (2011:19), and literature therein.
Typologically even more informative are ejective fricatives. A detailed acoustic study of Kabardian fricatives is conducted in Gordon and Applebaum (2006). Kabardian features three ejective fricatives, /f’, ì’, and /S’/ which are typologically uncommon and for which very few phonetic descriptions are available. Labiodental ejective fricatives are particularly rare, attested only in Kabardian, Abaza, and some dialects of Abkhaz (Colarusso 1988); outside Caucasian, the PHOIBLE database (Moran et al. 2014) lists only North American isolate Yuchi to feature /f/. Gordon and Applebaum (2006) provide valuable acoustic and articulatory information about ejective fricatives. Ejective fricatives have shorter frication duration, smaller frication intensity, and greater degree of constriction compared to their plain counterparts (based on palatograms). In fact, some tokens show complete oral closure. A period with complete closure is not uncommon for ejective fricatives and is phonetically motivated (Kuipers 1960:46): closure increases intraoral pressure which facilitates an audible release of ejectives (Gordon and Applebaum 2006). Gordon and Applebaum (2006) measure intraoral pressure and oral flow in [f] and [f’], uttered by one Kabardian speaker. Intraoral pressure rises substantially in production of the ejective fricative and is comparatively lower for the plain fricative. Airflow data aligns well with airflow trajectories in ejective stops (Shosted and Chikovani 2006): for plain fricatives it rises with frication and decreases only slightly before vowel onset. For ejective fricatives, airflow raises and decreases back to the zero level before the onset of the following vowel (Gordon and Applebaum 2006).

Ubykh only features one ejective fricative /ì’/ and even this one is marginal in the system. Spectrograms of two words read in isolation, /p’ì’o/ ‘four’ (right) and /pì’o/ ‘red’ (left) from the LaCiTO database are given in Figure 6. The main characteristics of ejective vs. plain fricatives reported for Kabardian are observed for Ubykh as well: frication duration is substantially shorter for the ejective fricatives. The shorter frication is followed by a period of silence which results from constricted glottis blocking the airflow. Noise in the plain fricative, on the other hand, is steady throughout the frication duration.

While most scholars accept the analyses of large consonantal systems in NWC presented thus far, some attempts have been made to reduce the unusually large consonantal systems of Northwest Caucasian languages by analyzing secondary articulations such as labialization and palatalization as underlying sequences of consonants + glides /j/ and /w/, which would reduce
the number of phonemes across NWC languages substantially (discussion in Colarusso 1988:94). Colarusso (1988:94), however, provides crucial evidence against these proposals. The most convincing counterevidence lies in the fact that there exist sequences of stop + /j/ in NWC that do not change to the palatalized variant of the stop, but surface as two segments.

2.2.3 Vocalic inventories

The second most prominent feature of Northwest Caucasian languages are small, vertical vocalic inventories, mostly limited to two or three vocalic phonemes, commonly described as /a/, /a/ (and sometimes /a/). With only two-way phonemic contrast in the vowel system, /a/ and /a/, we can define vowels with a single feature value, [±high], hence the term vertical vowel system.

Ubykh is one such example: it features two vocalic phonemes, /a/ and /a/. Some analyses include a third vowel /aj/, but this low vowel can also be analyzed as underlying /ah/, /ha/, or /a/ (Colarusso 1988, Chirikba 2003). Some analyses, on the other hand, claim that the distribution of /a/ and /a/ is predictable (Kuipers 1960, Allen 1956, 1965) and posit only one vocalic element for Ubykh. In other words, the one-vowel analysis claims that /a/ is an automatic epenthetic vowel. It is true that /a/ is often optional, can be deleted under certain circumstances, and, as Colarusso (1988) notes, it “has a low functional load”. He, however, shows that there exists strong evidence in favor of the phonemic status of /a/ for all languages: Colarusso (1988:347-373) offers a detailed discussion of cases in which /a/ contrasts with ∅.

The rest of this section focuses on acoustic analyses of small vocalic inventories. That the two posited vowels have different phonetic values is confirmed by the analysis of Ubykh recordings. The LaCiTO database contains a minimal pair, /bla/ ‘eye’ and /bla/ ‘seven’, uttered in isolation. The spectrogram in Figure 7 clearly shows that the two vowels have different qualities. For the low vowel /a/ at midpoint of vowel duration in a relatively neutral phonetic environment, F1 = 745 Hz, F2 = 1,712 Hz; for /a/ F1 = 424 Hz, F2 = 1,958 Hz. These formant values suggest that /a/ is phonetically relatively front and high, /a/ is relatively high. Fenwick (2011) transcribes the two vocalic phonemes as /ə/ and /i/ to better reflect their phonetic values. Vowel /aj/, regardless of its phonemic status, is phonetically lower than /a/ and therefore transcribed as /a/ in Fenwick (2011). Very similar analyses of vocalic inventories and their phonetic values are proposed for other NWC languages, especially in Abaza and Abkhaz; for Circassian languages the overall phonetic realization of vowels is reported to be comparatively higher (Colarusso 1988). For spectral measurements of Kabardian vowels and their coarticulatory variation, see Choi (1991), Wood (1994), and literature therein.

Figure 7: Spectrogram of /bla/ ‘eye’ (left) and /bla/ ‘seven’ (right)

Small vocalic inventories are often prone to a high degree of coarticulatory influence of adjacent
consonants. Colarusso (1988: 295) even claims that “all tautosyllabic consonants tend to color vowel to a greater or lesser degree”, i.e. not only consonants with secondary articulation. For example, anterior consonants “produce front vowels,” high consonants high vowels, rounded anterior consonants rounded vowels, e.g. Ubykh /t@t/ → [tʰɛtʰ]; /sʰɑ/ → [sʰæ] (Colarusso 1988:296). In Circassian, alveolars are reported to affect tautosyllabic /tɑ/ to a high [i]. High variability of phonetic values of Ubykh vowels is illustrated in Figure 8 that shows formant transitions in an utterance /amGang@kanan/, available in the LaCiTO database and analyzed in Praat (Boersma and Weenink 2015). Initial /a/ has values of F1 = 753 Hz and F2 = 1,552 at midpoint, similar to the values for /a/ in isolation in Figure 7. All subsequent vowels are colored by palatalization: F1 lowers and F2 rises. The most radical coloring targets the schwa /ɑ/ between two palatalized velars /k@/ and /ɡ@/ to the degree that its F1 lowers to 415 Hz and F2 rises to 2,089 Hz with almost no transitions in formant structure. Despite this heavy coloring, formants still tend to transition back to their targets in the absence of coarticulatory influence of the following consonant. In pre-pausal position, formants transition to their underlying vocalic targets, even after consonants with secondary articulations. Figure 8 shows formants of /a/ in /ɡ@/ in pre-pausal position that clearly transition toward the target for /a/ in the second half of the vowel duration (F1 = 675 Hz, F2 = 1,626 Hz). In other words, consonants do not fully color vowels in all positions: underlying targets are still present and realized, despite some traditional descriptions implicitly suggesting that vowels are colored completely.

![Figure 8: Spectrogram of /amGang@kanan/ (left) and /ɡ@/ (right)](image)

In sum, Northwest Caucasian languages feature large consonantal inventories with many typologically highly unusual segments, from obstruents with labialization and pharyngealization as double secondary articulation, to doubly-articulated bilabio-alveolar stops and fricatives, ejective fricatives, and palatalized and labialized glottal stops. Large consonantal inventories contrast with small vocalic inventories. The typologically unusual aspects of NWC phoneme inventories warrant further articulatory and perceptual studies as many of the segments appear exclusively in this language family and are not attested outside of the Caucasus.

### 2.3 Northeast Caucasian

Many Northeast Caucasian phoneme inventories are also comparatively large, primarily due to a four-way distinction in stop type in some languages (that includes voiceless aspirated, voiced, ejective, and “fortis”), a high number of phoneme segments in the post-velar part of the vocal tract (including pharyngeal, epiglottal and laryngeal obstruents), and phonemic status of some secondary articulations, such as labialization (Kibrik and Kodzasov 1990). Compared to Northwest Caucasian, Northeast Caucasian languages feature larger vocalic inventories, often including
length distinction. The number of NEC languages and dialects and their variability in phonemic inventories is substantially greater compared to SC or NWC, which is why two illustrative languages were chosen: for a survey of phonemic inventories of most other languages, see Bokarov et al. (1967), Kibrik and Kodzasov (1990), Job (1994), Alekseev (1998), SMEETS (2004), Hewitt (2004), van den Berg (2005). Tables 5 and 7 (with data from Chechen for the Nakh and from Archi for the Dagestanian group) illustrate NEC consonantal inventories, Tables 6 and 8 their vocalic inventories.

Table 5: Chechen consonant inventory (from Nichols 1994, 1997, Sylak 2011). Ingush features a very similar consonant inventory: Ingush lacks /p/, /k/, and /k/, but features an additional series: palatalized velars /k, q, k/ (Nichols 2011).

<table>
<thead>
<tr>
<th></th>
<th>bilab.</th>
<th>l.-d.</th>
<th>dent.</th>
<th>post-al.</th>
<th>pal.</th>
<th>vel.</th>
<th>uvul.</th>
<th>phar/epigl.</th>
<th>glot.</th>
</tr>
</thead>
<tbody>
<tr>
<td>stop</td>
<td>p, b, p', p:</td>
<td>t, d, t', t:</td>
<td>t, d, t', t:</td>
<td>k, g, k', k:</td>
<td>q, q', q:</td>
<td>?</td>
<td>?</td>
<td></td>
<td></td>
</tr>
<tr>
<td>affricate</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>nasal</td>
<td>v</td>
<td>s, z, s:</td>
<td></td>
<td></td>
<td></td>
<td>j</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>tap/trill</td>
<td></td>
<td>r, r'</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>fricative</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>approx.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>lat. appr.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


<table>
<thead>
<tr>
<th></th>
<th>front</th>
<th>central</th>
<th>back</th>
</tr>
</thead>
<tbody>
<tr>
<td>high</td>
<td>i, i:</td>
<td>u, u:</td>
<td></td>
</tr>
<tr>
<td>open-mid</td>
<td>e, e:</td>
<td>ᴇ, ᴇ:</td>
<td>ᴥ, ᴥ:</td>
</tr>
<tr>
<td>low</td>
<td>a, a:</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 8: Archi vowels (Kibrik and Kodzasov 1990, Kibrik 1994).

Perhaps the most prominent and well-studied feature of Northeast Caucasian are pharyngeal segments and pharyngealization. The NEC family provides crucial information for the typology of pharyngealization: NEC languages feature both voiceless and voiced stops and fricatives of pharyngeal or epiglottal place of articulation as well as pharyngealization or epiglottization as secondary articulation.

This section aims to outline the main acoustic properties of NEC phoneme inventories with a focus on pharyngealization. Because pharyngealization is also present in NWC, data from NWC languages are included in this section to supplement the analysis. Several aspects of pharyngeal place of articulation and pharyngealization as secondary articulation are still unknown, both from acoustic as well as articulatory and perceptual perspectives. Further studies on these phenomena in Caucasian languages are needed.

2.3.1 Laryngeal features

As already mentioned, in addition to the “common” Caucasian three-way opposition in obstruent types with three laryngeal features (voiceless aspirated, voiced, and ejective), some Northeast Caucasian languages, such as Chechen, Ingush, Batsbi, Andi, Avar, Lak, Lezgian, Dargi, Tabasaran, Agul, Archi, and Tsakhur (Kibrik and Kodzasov 1990) feature the “fourth type” of stops. This “fourth” series receives various different analyses: they are labeled as intensive, fortis, unaspirated, or geminated consonants (Kibrik and Kodzasov 1990, Catford 1977, Hewitt 2004, Gaprindašvili 1966). The characteristic features of this series is lack of aspiration and phonetically longer closure. The exact phonetic realization differs across languages: in some languages “fortis” stops are realized as geminates intervocically (e.g. in Lak, Dargi, and Khinalug, Catford 1977, cf. Gaprindašvili 1966, Kibrik et al. 1972), in others as affricates (e.g. in Avar and Andi, Catford 1977). A four-way opposition in obstruent types is also present in the Hatkoy and Shapsugh dialects of Adyghe (for a detailed phonetic study of the contrast there, see Gordon and Applebaum 2006, 2013).

A recent quantified study of laryngeal features in Ingush is presented in Warner (1996). Most of the common phonetic generalizations of ejective stops are confirmed in this study. VOT in Ingush is, for example, significantly longer in voiceless stops (M = 45.1 ms) and shorter in ejectives (M = 26.2 ms). Ejective stops lack noise after the burst: there is a significant difference in average power of post-burst noise between ejective and voiceless series, but no significant difference in average peak burst power (also replicated in Grawunder et al. 2010). Warner also reports an audible effect of ejective stops on pitch of the following vowel. On average, the difference in pitch value at the onset of following vowel is 26.2 Hz: pitch is higher after ejectives compared to voiceless stops. The difference is significant both at the vowel onset as well as at fifth to seventh period of the vowel. On the other hand, Warner (1996) found no significant difference in peak burst power between the ejective and voiceless series. Likewise, no specific spectral characteristics of burst in ejectives were observed.
Hewitt (2004) makes an important observation with respect to laryngeal features in NEC. He reports a correlation between aspiration and ejection: ejectives appear “more glottalized” in those dialects that have less aspiration and vice versa. This observation is, however, not supported by any quantitative analysis. Further quantified treatments of correlations in phonetic realization of different stop types could yield important insights into a thus far unexplained variation (cf. also our observation of inter-speaker variation in stop production in Georgian in 2.1.1 above).

2.3.2 Lateral obstruents

Another notable feature of Northeast Caucasian languages is the presence of a large number of lateral obstruents. Lateral fricatives in NEC can be voiceless (l̩) or “fortis” (lː). Lateral affricates can be voiceless (l̩̊), ejective (l̩̊’), fortis (l̩ː), or fortis ejective (l̩ː’; Catford 1977, Kibrik and Kodzasov 1990). According to Catford (1977), the system with a highest number of lateral consonants is Akhwakh, featuring seven contrastive consonants with lateral articulation (including the approximant /l/).

Among NEC languages, Archi stands out as featuring typologically very rare velar lateral fricatives and affricates (Kodzasov 1977, Ladefoged and Maddieson 1996). Velar laterals are reported in very few languages: Ladefoged and Maddieson (1996) report only Zulu and Axuxlay beside Archi to feature contrastive velar laterals. Archi features voiceless, voiced, and geminate (or tense) pre-velar lateral fricatives (l̩i/, l̩i/, and l̩iː/) as well as voiceless and ejective pre-velar lateral affricate /l̩k̩i/, /l̩k̩i/. Moreover, the voiceless and ejective affricate and voiceless and geminate fricative can be phonemically labialized: /l̩k̩iːw/, /l̩k̩iːw’, /l̩iːw/ and /l̩iːːw/ (based on Kodzasov 1977). Ladefoged and Maddieson (1996) provide spectrograms of these segments and a qualitative description of acoustic properties. The pre-velar lateral fricative in Archi features strong frication and a close proximity between the second and third formants. Articulatorily, these segments are produced with tongue body constriction along the velum and the palate and with tongue tip “passively lowered to the lower teeth” (Kodzasov 1977 in Ladefoged and Maddieson 1996:206). For further acoustic measurements of these sounds, see Grawunder et al. (2010), who suggest that pre-velar realization of lateral fricatives and affricates is also found for some speakers of Bezhta and Avar.

2.3.3 Pharyngeal place of articulation

As already mentioned, one of the major topics in phonetics of Northeast Caucasian is pharyngealization. As Sylak (2011) points out, pharyngealization is rare in the world’s languages: only 5% of languages surveyed in Maddieson (1984) feature pharyngeal place of articulation or pharyngealization as secondary articulation. Most work on pharyngealization centers on Semitic and North American languages and comparatively fewer instrumental studies were done on Caucasian pharyngealization.

There are still many opposing views in the general phonetic literature on the topic of articulatory and acoustic aspects of consonants produced in the region between the uvula and the glottis: pharyngeal/epiglottal place of articulation (Laufer and Condax 1979a,b, Catford 1983, Esling 1996, 1999, for surveys of post-velar articulations and their phonetic and phonological properties, see Bessell 1992, Moisik 2013, Sylak-Glassman 2014). For example, no consensus has been reached on the question of whether languages can contrast epiglottal and pharyngeal place of articulation, exactly which articulators are active (and in what ways) during the production of these obstruents, or what the exact acoustic properties for each place of articulation are.

Northeast Caucasian languages provide crucial phonetic information for the typology of segments in the radical part of the vocal tract (Kibrik and Kodzasov 1990, Kodzasov 1987, Catford
1983, Ladefoged and Maddieson 1996, Esling 1999, Nichols 2000): different NEC languages have a large number of phonemes with primary articulation between the glottis and uvula, the most common of which are pharyngealized voiceless /h/ and voiced /ŋ/ fricatives. In some languages, such as Chechen, the NEC “pharyngeal” series is produced at the epiglottal place of articulation (Catford 1983). Despite the rich inventory of post-velar sounds in NEC, detailed and systematic instrumental acoustic and especially articulatory studies of pharyngeal/epiglottal obstruents are still lacking.

The most informative for the discussion on differences between pharyngeal and epiglottal place of articulation are the Burshag and Burkikhan dialects of Agul (Kodzasov 1987, Kibrik and Kodzasov 1990, cf. Šaumjan 1941, Magometov 1970). The dialects are reported to phonemically contrast pharyngeal and epiglottal places of articulation (Kodzasov 1987, Ladefoged and Maddieson 1996): Agul features voiceless /h/ and voiced /ŋ/ pharyngeal fricatives and voiceless /n/ and voiced /ŋ/ epiglottal fricatives. The Agul phonemic inventory also contains an epiglottal stop /ɾ/ (Catford 1983, Kibrik and Kodzasov 1990). Altogether, Agul features five phonemes between the uvular and glottal places of articulation. Ladefoged and Maddieson (1996) further show that the pharyngeal and epiglottal fricatives are acoustically quite different: the epiglottal series is noisier and its formant structure resembles neighboring vowels more than the pharyngeal series. The latter is characterized by a high first formant above 1,000 Hz and a small distance between the first and the second formant (Ladefoged and Maddieson 1996: 167-8). These observations are confirmed by spectral analysis: Figure 9 shows spectra of a /h/ and a /n/ in words /hatʃ/ ‘wolf’ and /hætʃ/ ‘apple’.

Recently, however, Esling (1999, 2010, 1997) and others (Heap 1997, Moisik 2013, Sylak-Glassman 2014) proposed that the distinction in Agul is not of place of articulation, but rather of manner of articulation: Esling argues that pharyngeal fricatives are fricatives, but that what Ladefoged and Maddieson (1996) analyze as epiglottal fricatives are in fact trills. In other words, all five obstruents in Agul are produced in the same epiglottal (aryepiglottal-epiglottal) region and their primary distinction is in manner of articulation (see also Moisik 2013). Sylak-Glassman (2014) and Moisik (2013) even question the phonemic status of the two series in Agul. The question of whether a language can phonemically distinguish pharyngeal and epiglottal place of articulation of the same manner of articulation thus remains open for further research.

Figure 9: Spectra of a /h/ (left) and a /n/ (right) analyzed in Praat (Boersma and Weenink 2015) with 25ms window length, at approximately midpoint of the fricative (where clear formant structure was visible) from recordings of Agul at the UCLA Phonetics Lab Archive.
2.3.4 Pharyngealization as secondary articulation

NEC languages also feature pharyngealization as a secondary articulation. Kibrik and Kodzasov (1990) distinguish two types of pharyngealization in NEC languages, which, according to Nichols (2000), corresponds to the distinction between pharyngeal and epiglottal places of articulation. The distinction between pharyngealization and epiglottalization as secondary articulation is, however, problematic (just like for the corresponding primary places of articulation) and several open questions remain to be answered (see Moisik 2013 and literature therein and discussion below).

There exist different analyses as to whether pharyngealization is a property of consonants or of vowels: in NEC, pharyngealization is most commonly analyzed as a property of the vowel or of the syllable, rather than a property of the consonant (Kodzasov 1987, Kibrik and Kodzasov 1990, Maddieson et al. 1996, Nichols 2011). Pharyngealization has been analyzed as an autosegmental/prosodic feature on the basis of pharyngealization spreading, where the pharyngealization feature spreads from one vowel/syllable in the word to neighboring vowels/syllables (Kibrik and Kodzasov 1990, Schulze 1997, Sylak-Glassman 2014), e.g. in Lak or Archi (Anderson 1997, Kibrik 1994; in Sylak-Glassman 2014 and Moisik 2013). Pharyngealization spreading can be sensitive to stress or blocked by other segments (Sylak-Glassman 2014, Moisik 2013). In Northwest Caucasian, on the other hand, pharyngealization is most commonly analyzed as a secondary articulation on consonants (Hewitt 2004, Catford 1983). In some languages such as Tsez or Rutul, pharyngealization is analyzed as a property of both vowels and consonants (see Maddieson et al. 1996, Sylak-Glassman 2014). Phonetic distinction between vocalic and consonantal pharyngealization is difficult to draw, because both involve similar gestures: “compression” or “contraction” of the pharynx or epiglottis that is simultaneous with vocalic or consonantal articulation (for complexity of articulations, see Moisik 2013). This means that analyses have to rely primarily on phonological data. Further articulatory studies of timing of pharyngeal gestures and distributional patterns should contribute to a better understanding of different realizations of pharyngealization across Northeast and Northwest Caucasian. For example, as will be shown below, pharyngealization in Ubykh, despite being analyzed as a consonantal feature, affects the following vowel throughout its duration by lowering the F1 and F3 values.

For both families, X-ray tracings and spectral analyses of pharyngealization exist (Gaprindašvili 1966, Dzejranišvili 1959 via Ladefoged and Maddieson 1996, Catford 1983, Colarusso 1988), but acoustic studies are ambiguous with sometimes opposing results. The analysis of acoustic effects of pharyngealization is complicated by several factors: pharyngealization is not homogeneous in terms of where and how it is realized (primarily on the vowel or primarily on the consonant) or in terms of place of articulation (pharyngealization vs. epiglottization). It is thus not surprising that different studies report different results.

Kingston and Nichols (1987) and Nichols (1997) examine general acoustic effects of pharyngealization as secondary articulation in Chechen and argue that pharyngealization primarily affects formants of the vowel and VOT duration of the consonant. The authors show that VOT duration of pharyngealized consonants is longer compared to plain voiceless or ejectives (reported in Sylak 2011 and in Nichols 2011). Nichols (1994) further reports “very noisy aspiration” or “murmured” phonation in the VOT of pharyngealized voiceless and voiced consonants. Moreover, F1 rises and F2 and F3 lower in the presence of pharyngealization. In other words, pharyngealization “produces compaction” of the spectrum: “lowering and backing of the low vowels and centralization of the others” (Nichols 2011).

Sylak (2011) measures formants of five male Chechen speakers and compares the measurements with predicted formant values. The predicted values are calculated based on source-filter theory and estimation of vocal tract parameters. His study confirms that pharyngealization raises
F1 values and lowers F2. No significant effects of pharyngealization were found on F3. For some places of articulation, his measurements suggest that the secondary articulation is in fact epiglottalization rather than pharyngealization: Nichols (2011) claims the same for Ingush. These claims can only be confirmed with further articulatory studies.

Ladefoged and Maddieson (1996) and Catford (1983) report a study of pharyngealization in Tsakhur and Udi (by Catford, ms., cf. Ibragimov 1968, 1990). The “most noticeable” acoustic effect of pharyngealization reported there is a substantial lowering of F3, for anything between 150–1200 Hz (depending on the vowel and language). Additionally, F1 is reported to raise, but not as considerably. F2 lowers in /ḙi/ and /i̭i/, but rises in /a̭i/, /o̭i/, and /ṷi/. X-ray tracing of pharyngeals in these two languages (from Gaprindašvili 1966) reveals the “curious tongue configuration. The tongue root at about the level of the tip of the epiglottis bulges backwards into the pharynx, while a depression is formed in the dorsal surface of the tongue approximately opposite the uvula, with a further upward bulge further forward in the tongue”, also called the “double bunching” (Catford 1983:349, see also Gaprindašvili 1966 and Bessell 1992, Moisik 2013).

In Tsez, pharyngealization is reported to affect formants of the vocalic counterpart of /a/ and /u/ as well as at their midpoint. The effects of pharyngealization differ across different vowel qualities (Maddieson et al. 1996). Pharyngealization is reported to raise F1 in all five vowels measured, although the effect of F1 raising is greater for /i/ and /e/ compared to other vowels. F2 is lower in pharyngealized /a̭i/, /ḙi/, and /i̭i/, but higher in pharyngealized /a̭i/, /o̭i/, and /ṷi/ compared to their non-pharyngealized counterparts. The measurements are reported to yield “complex” results for F3. Pharyngealization is reported to lower F3 significantly in /a/ and /u/. The magnitude of the effect of pharyngealization is greater at vowel onset compared to the midpoint position (Maddieson et al. 1996). Gaprindašvili (1966) further suggests that the main acoustic effect of vocalic pharyngealization in Dargi is the presence of the fourth formant in the region around 1280 Hz.

Moisik (2013) and Sylak-Glassman (2014) also report a number of studies that claim vowels adjacent to pharyngeal fricatives get fronted or pattern phonologically as palatalizing vowels. Fronting is reported for Avar in Sylak-Glassman (2014) (based on Charachidze 1981), Kryz (based on Authier 2009), and Agul (based on Magometov 1970). For example, /o/ and /u/ are realized as [o] and [y] in Avar if they are adjacent to a voiceless or voiced pharyngeal fricative: [goh] for /goh/ ‘mountain’ (Sylak-Glassman 2014). Moisik (2013) reports that pharyngealized vowels in Lak (based on Andersen 1997) and Bezhta (based on Kibrik and Testelets 2004) cause phonological palatalization. For example, in Lak /k/ and /l/ palatalize before pharyngealized fronted vowels.

Effects of pharyngealization are not only observed in formant structure and VOT, but also in other parameters. Grawunder et al. (2010), for example, report that burst spectra differ for pharyngealized stops compared to plain stops in Tsez; the same observation is made for Tsez in Maddieson et al. (1996). Maddieson et al. (1996) also measure formant transition of vowels before pharyngealized uvulars: the only consistent result they report is lowering of F3 of the preceding vowel (e.g. in /naq̄a/ vs. /naq/).

The most comprehensive study of effects of pharyngealization in NWC is provided in Colarusso (1988) with spectrograms, but with no statistical studies of the effects. Colarusso (1988) concludes that acoustic effects are “complex”, but identifies the most noticeable ones: noisy energy in 400-600 Hz range (and sometimes around 1,100 Hz and 2,000-2,400 Hz; Colarusso 1988:222).

To my knowledge, no quantified measurements with statistical tests of Ubykh pharyngealization have been presented in the literature and Colarusso (1988) does not measure formants of vowels following pharyngealized segments. For the purpose of analyzing pharyngealization in NWC, a recording of two Ubykh words, /q̄a:p’a/ ‘handful’ and /q’a:p’a/ ‘hand’, each uttered eight times in isolation in apparent random order by Tevfik Eșenç and recorded by Georges Dumézil.

22
(available at LaCiTO) was analyzed. The following parameters of the stop following /q/ and /q'/ were measured: F0, F1, F2, F3 (all at 20%, 50%, and 80% of vowel duration).

Pharyngealization significantly affects F1 and F3 values, but not F2 values. Contrary to the effects of pharyngealization observed for Chechen, Ubykh pharyngealization lowers F1 significantly at 20%, 50%, and 80% of vowel duration (at 20%, $t = 9.1$, df = 10.2, $p < 0.0001$). There is no significant effect of pharyngealization on F2 at any of the three points measured (at 20%, $t = 0.76$, df = 8.2, $p = 0.47$). Pharyngealization significantly lowers F3 at all three points (at 20%, $t = 7.9$, df = 9.1, $p < 0.0001$). Figure 10 illustrates the effects of pharyngealization on the first three formants with standard errors.

![Figure 10](image)

**Figure 10:** Differences in vowel formants with standard errors for pharyngealized vs. plain low vowel (obtained with `lm()` function and `effects` package, Fox 2003)

Furthermore, vowel durations are significantly shorter if the vowel is pharyngealized ($t = 4.7$, df = 13.9, $p < 0.001$; Welch Two Sample t-test). The data, however, shows no effect of pharyngealization on F0, either at 20%, 50%, or 80% of vowel duration (at 50%, $t = -1.67$, df = 13, $p = 0.12$ with Welch two-sample t-test).³

The effects of pharyngealization on the following vowel in Ubykh differ in some aspects from the effects summarized for Northeast Caucasian languages above, but note that the data in Ubykh is limited to one speaker and a single minimal pair. F3 does indeed lower substantially after pharyngealized stops, in accordance with results from other languages. F1, however, lowers significantly too, which differs from most other studies. This discrepancy reveals that effects of pharyngealization can be quite different across languages and that detailed descriptions of acoustic effects of pharyngealization as well as their articulatory and perceptual properties are still lacking.

### 3 Phonotactics

Consonant clusters are probably the most well-studied aspect of Caucasian phonotactics, especially in the South Caucasian family. Articulatory and perceptual research on Georgian clusters provides crucial information for the discussion on the role of production vs. perception in phonology. This section reviews the discussion on Georgian clusters and their relevance for phonological theory as well as points to aspects of South Caucasian phonotactics that remain largely unnoticed in the current phonological literature despite their potential broader relevance.
3.1 Georgian clusters

Traditional grammarians identify two kinds of clusters in South Caucasian: harmonic and non-harmonic (Akhvlediani 1949, i.a., via Butskhrikidze 2002, i.a.). Harmonic clusters are clusters of two obstruents, a dorsal and a non-dorsal that agree in laryngeal features. Non-harmonic are all other clusters that do not belong to the harmonic group, but are permitted phonotactically in Georgian. Table 10 shows harmonic clusters in Georgian (from Chitoran 1998, Butskhrikidze 2002).

<table>
<thead>
<tr>
<th>Voiced</th>
<th>Voiceless</th>
<th>Ejective</th>
<th>Voiced</th>
<th>Voiceless</th>
<th>Ejective</th>
</tr>
</thead>
<tbody>
<tr>
<td>bg</td>
<td>p\textsuperscript{h}k\textsubscript{h}</td>
<td>p\textsuperscript{'k}'</td>
<td>bg\textsuperscript{h}</td>
<td>p\textsuperscript{h}x</td>
<td>p\textsuperscript{'\chi}'</td>
</tr>
<tr>
<td>d\textsuperscript{g}</td>
<td>t\textsuperscript{h}k\textsubscript{h}</td>
<td>t\textsuperscript{'k}'</td>
<td>d\textsuperscript{g}</td>
<td>t\textsuperscript{h}x</td>
<td>t\textsuperscript{'\chi}'</td>
</tr>
<tr>
<td>d\textsuperscript{g}</td>
<td>ts\textsuperscript{h}k\textsubscript{h}</td>
<td>ts\textsuperscript{'k}'</td>
<td>d\textsuperscript{g}</td>
<td>ts\textsuperscript{h}x</td>
<td>ts\textsuperscript{'\chi}'</td>
</tr>
<tr>
<td>d\textsuperscript{g}</td>
<td>tf\textsuperscript{h}k\textsubscript{h}</td>
<td>tf\textsuperscript{'k}'</td>
<td>d\textsuperscript{g}</td>
<td>tf\textsuperscript{h}x</td>
<td>tf\textsuperscript{'\chi}'</td>
</tr>
<tr>
<td>sk\textsuperscript{h}</td>
<td></td>
<td></td>
<td>z\textsuperscript{g}</td>
<td>sx</td>
<td></td>
</tr>
<tr>
<td>j\textsuperscript{k}\textsubscript{h}</td>
<td></td>
<td></td>
<td>\textsuperscript{5}g</td>
<td>jx</td>
<td></td>
</tr>
</tbody>
</table>

Table 9: Georgian harmonic clusters (from Butskhrikidze 2002: 103)

Harmonic clusters are distinguished from non-harmonic ones both phonetically and phonologically. Butskhrikidze (2002) summarizes phonological arguments for distinction of harmonic and non-harmonic clusters: harmonic clusters can appear in stem final position, never feature optional r-insertion, are copied in reduplication (e.g. /tsxel-tsxeli/ for /R\textsuperscript{red}-tsxeli/ ‘hot’), and syllabify into the same syllable, according to speakers’ judgments (e.g. /si.t\textsuperscript{q}’va/ ‘word’) (Butskhrikidze 2002:103-105). This latter claim is, however, disputed. Syllabification data is based on speaker’s judgment exclusively and Chitoran (1998) reports inconsistent judgments in her experiment.

Phonetic status of harmonic clusters has been subject to even more debate. Harmonic clusters are either analyzed as complex, doubly-articulated segments or as sequences of simple segments. Some analyses divide harmonic clusters even further: corono-dorsal clusters are analyzed as complex segments, whereas labio-dorsal cluster can be either clusters or complex segments with lexicalized distribution between the two (discussion in Chitoran 1998). The basis for this distinction is the observation that corono-dorsal clusters can surface in three-stop clusters, e.g. /t\textsuperscript{'k}bili/ ‘sweet’, whereas labio-dorsals cannot.

Acoustic studies of Georgian harmonic clusters yield little support for the complex-segment analysis. Chitoran (1998) analyzed harmonic clusters and corresponding sequences of two obstruents across word boundary. Her study found no significant difference in proportion of released vs. unreleased stops between harmonic clusters and “harmonic” sequences of stops across word-boundary: the first and second element in clusters are equally frequently released in harmonic clusters compared to sequences of stops across word boundary. Harmonic clusters almost always have two releases, contrary to what has been described impressionistically in earlier literature. Similar results are obtained in McCoy (1999), who reports that all clusters in her experiment feature two releases and that in voiced harmonic clusters a presence of an automatic transitional vowel was detected (e.g. [d\textsuperscript{gas}] for /d\textsuperscript{gas}/). Moreover, measurements of duration do not support the complex-segment analysis either. Duration of harmonic clusters is not shorter compared to sequences of stops across word-boundary (Chitoran 1998, also McCoy 1999).

Georgian clusters have also been analyzed articulatorily (Zhgent’i 1956, Chitoran et al. 2002). Cross-linguistically, clusters have been found to feature significantly less overlap word-initially compared to other positions and significantly less overlap in front-to-back clusters compared to...
back-to-front clusters. It has been argued that the causes of these patterns stem from perceptual recoverability: word-initially, clusters lack formant transitions into closure. It is reasonable to assume that less overlap results from the need for more perceptual cues in a position where cues are reduced. A similar explanation has been proposed for the smaller degree of overlap in back-to-front clusters: if back-to-front clusters overlap to a high degree, the release of the first segment in the cluster occurs when the second stop is still unreleased, which again reduces perceptual cues (as summarized in Chitoran et al. 2002).

Georgian clusters that feature a division into two groups (harmonic or recessive and non-harmonic or non-recessive) are perfect for testing these hypotheses. Chitoran et al. (2002) conducted an articulatory experiment involving two speakers of Georgian and provide EMA measurements of cluster articulations. Their study confirms the hypotheses outlined above: there is significantly less overlap in word initial clusters and significantly less overlap in back-to-front (non-harmonic) clusters. Chitoran et al. (2002) also suggest that it is precisely the high degree of overlap in front-to-back clusters that is responsible for the requirement of harmonic clusters to agree in laryngeal features. In other words, a high degree of overlap means that laryngeal features of the first stop are realized at the onset of the second stop, i.e. the members of the cluster share laryngeal features.

The hypothesis outlined above — that perceptual factors influence the degree of gestural overlap — opens an important discussion on the interplay of synchronic and diachronic effects on phonetics and phonology. While it is true that greater overlap in front-to-back clusters could be caused by perceptual mechanisms, this is not the only possible cause. Chitoran et al. (2002) mention that the origin of harmonic clusters are likely velarized obstruents that later segmented into a sequence of two stops (as proposed in Gamkrelidze and Ivanov 2010). However, they do not discuss the implications of this historical development. If Gamkrelidze’s reconstruction holds, then we expect harmonic clusters that go back to single segments to have greater overlap by virtue of their origin, not necessarily because of perceptual recoverability.

That perception cannot be the only cause of differences in degree of overlap is argued in subsequent articulatory work on Georgian clusters. Chitoran and Goldstein (2006) measured overlap in clusters of a stop and a sonorant. There too, back-to-front clusters such as [kl] and [rb] “are less overlapped” than front-to-back clusters (reported in Chitoran and Cohn 2009), although perceptual recoverability plays a much smaller role when sonorants are members of clusters. In fact, as Chitoran and Cohn (2009) point out, “liquids do not rely on their releases in order to be correctly perceived.” The authors then suggest that “perceptual recoverability is not directly encoded in the phonology”. The phonetic differences in the degree of overlap may have emerged for perceptual reasons, but were then phonologized over time, generalized, and encoded as part of phonological grammar. The problem of different degrees of overlap in consonant clusters across different types and positions as well as the question of their causes and origins thus requires further investigation.

Harmonic and non-harmonic clusters of two obstruents represent just a subset of licit clusters. Georgian and other South Caucasian languages permit consonant clusters that are larger than two members: the number can reach up to six members in root-internal word-initial position. Such large clusters can consist of several consonants and two-member clusters. Perhaps the most famous example illustrating Georgian clusters is the word ‘to peel’, /phbtskvna/.

These large clusters are governed by different phonotactic rules in the four SC languages. Several detailed descriptions exist on combinatorial principles in consonant clusters (Pöchertrager 2011, Harris 1991, i.a.). Common to all restrictions is that longer clusters obligatorily include sonorants. For example, Mingrelian clusters of two obstruents can only be preceded by an /r/ or an /n/ and followed by a /v/ (Harris 1991, based on Gudava and Gamkrelidze 1981).
South Caucasian clusters thus pose a problem to the Sonority Sequencing Principle (SSP): if sonorants precede segments with lower sonority such as voiceless stops, SSP is violated quite severely (e.g. in /pʰrtʰskʰvna/). Some analyses of clusters in South Caucasian argue that these sonorants are syllabic and that harmonic clusters are doubly articulated single segments (Butskhrikidze 2002). Under this analysis then, SSP would not be violated. For example, /p'tʃ'ali/ would be analyzed as [p'tʃ'ali] or an initial sequence /bdyvn-/ would be analyzed as [bdyvna-] or possibly [bdyvna-] without SSP violations. There is, however, little external support for such hypotheses: harmonic clusters do not behave as complex segments based on phonetic analysis (as seen above) and syllabification arguments are based solely on speakers’ judgments which are inconsistent (Chitoran 1998).

Phonetic analyses might reveal that sonorants undergo substantial reductions in sonority in the cluster environment, which would mean that SSP would not be violated. This is suggested by the realization of /r/, which is reported as devoiced before voiceless obstruents (Zhgent’i 1956, Butskhrikidze 2013, cf. Sturm 2016). Further acoustic and articulatory studies on the question of phonetic realization of sonorants in clusters are desiderata.

3.2 Svan clusters

Some aspects of cluster phonotactics differ quite radically across South Caucasian languages. Long clusters in Georgian, for example, are restricted to word-initial position: word-finally, clusters are much more limited. This distribution is exactly reversed in Svan: word-initially, Svan allows only one two-member harmonic cluster (or a cluster that historically goes back to a harmonic cluster) with an optional following [w], while no such restriction exists for word-final position (Tuite 1997). Tuite (1997) (via Zhgent’i 1949) illustrates long word-final clusters in Svan: /axeqwsq/ ’you stole up on something’ or /xosqwe/ ’I ordered somebody’.

This restriction against word-initial clusters to the exclusion of word-final clusters only is likely part of active phonology in Svan and is reflected in loanword phonology and morphophonological alternations. For example, initial clusters in loan words undergo epenthesis (/k’aravæt’ < Russ. krovat’ ‘bed’). The restriction against initial clusters is also revealed morphophonologically by processes such as metathesis and epenthesis: /x/ of the prefix /xw-/ is deleted and /w/ is metathesized before a consonant-initial root, e.g. /xw-t’ix-e/ → [t’wixe] ’I return it’ (Tuite 1997). If metathesis is not available, initial clusters are repaired by epenthesis, e.g. /mt’ix-e/ → [mat’xe].

Complex onsets are typologically much more common and articulatorily as well as perceptually more motivated than complex codas. While longer clusters in Svan are not restricted to codas specifically, they are restricted to word-final position, where they necessarily appear in the coda position. In other words, Svan syllables allow many more complex codas than complex onsets, which is a generalization that goes against phonetic motivation. The concentration of clusters word-initially vs. word-finally in Georgian vs. Svan thus opens a discussion on phonetic motivation in phonotactics. We also know that initial clusters are articulated with greater degree of overlap compared to final clusters, which was confirmed by articulatory studies on Georgian. To my knowledge, no articulatory studies of Svan clusters exist; experimental studies on the topic would provide information on whether the universal distribution of the degree of overlap holds true even for languages in which complex codas are more frequent than complex onsets.

3.3 Mingrelian clusters

Mingrelian clusters are much more limited compared to Georgian and Svan, but nevertheless feature intriguing coocurrence restrictions that bear broader implications for the typology of the
Obligatory Contour Principle (OCP) effects (Leben 1973, McCarthy 1986). As Harris (1991) summarizes, Mingrelian non-labial clusters are maximally 4 consonants long and their distributional patterns are described using a subgrouping of consonants, presented in Table 10. Clusters have 2-4 members and strictly follow the ABCD order. They cannot contain more than one member from each of the groups in Table 10. Mingrelian clusters are also governed by “Homogeneity Principle”: they have to agree in laryngeal features (Harris 1991).

| A | sonorant | r, n |
| B | coronal   | d, t’, t’, d, t, ts, ts, da, da, s, z, š, š |
| C | dorsal    | g, k, k’, y, x, q’ |
| D | sonorant  | v |

Table 10: Categories of Mingrelian consonants (Harris 1991, based on Gudava and Gamkrelidze 1981)

In addition to phonotactic restriction within clusters, Mingrelian features cooccurrence restriction of clusters within a word. These cooccurrences are governed by OCP principles (Leben 1973, McCarthy 1986). That languages disprefer sequences of same or similar segments (the so-called OCP constraints) and that strong similarity between sounds is often worse than complete identity has long been established. Mingrelian cluster cooccurrence restrictions follow several OCP principles. Harris (1991) reports that Mingrelian bans sequences of clusters of the structure B_i...B_jC_j: only complete repetition (B_i...B_jC_j) is allowed. Moreover, Mingrelian prohibits two non-identical BC clusters in a word: sequences of clusters such as /tkh...sx/ are prohibited (Harris 1991). This last restriction is especially peculiar, primarily because it bans cooccurrence of perceptually and articulatorily quite different clusters: the only similarity in the two clusters of a banned sequence *dγ...fkh is that both clusters are sequences of a coronal and a dorsal. Several aspects of this restriction require further investigation: synchronic status of these OCP effects would need to be experimentally tested.

3.4 Other families

Consonant clusters are much more restricted in Northeast and Northwest Caucasian languages compared to South Caucasian. The maximal syllable structure in many languages is CVC. Among the languages that have most restricted syllable structure is Ghodoberi (Kodzasov 1996, Saidova 2004): syllables can only be closed with a sonorant or a labial [b] (Kodzasov 1996); tendency towards open syllables is also strong in Tsez (Alekseev and Radžabov 2004). While clusters are more restricted in NEC and NWC (see also Ibragimov 1968), Catford (1977) notes that most “morpheme-initial” clusters in Northwest Caucasian and “intra-morpheme” clusters in Nakh languages observe the same restriction as in SC whereby only harmonic clusters (agreeing in laryngeal features, recessive, and consisting of a stop and a stop or fricative) are licit. The division of clusters to harmonic and non-harmonic thus appears to be a feature of all three Caucasian language families.

4 Processes

Caucasian phonological systems are not characterized by a great amount of active synchronic phonological alternations. Some of the more common processes are vowel deletion in a hiatus, root-final vowel deletion before a grammatical morpheme (e.g. in NEC), vowel deletion in open
syllables, metathesis of \( /v/ \) and \( /r/ \) driven by SSP (e.g. in SC), voice, place, or manner assimilations in clusters, dissimilation of two rhotics (e.g. in Georgian, Abkhaz, Abaza), labialization and delabialization of obstruents before rounded vowels, pharyngealization spreading, and vowel harmony (e.g. in Svan, Bezhta, Tsakhur), to name just a few.

Several processes that have the potential to bear implications for phonological theory have gone unnoticed in Caucasian. This section focuses on these alternations: several descriptions of Caucasian languages were surveyed with the goal to identify processes that are not discussed in phonological literature, but are of high theoretical interest. I present them here with an appeal for further investigation — many of the processes are poorly described. The chapter concludes with one of the most discussed alternations in recent phonological literature: final voicing (FV) in Lezgian.

4.1 Laz identical consonant deletion

In Laz, more precisely in Khopa subdialect in Sarpi, a rule of identical consonant deletion is reported in Holisky (1991) (based on Kartozia 1968). In a sequence VC\(VC_i\), the first consonant is deleted in order to satisfy the OCP constraint: two identical segments are not permissible within adjacent syllables. Deletion of the consonant results in hiatus that remains unresolved, e.g. /mkjapu-pe-k/ \( \rightarrow \) [mkjaupek], /op’ts’op-up-t/ \( \rightarrow \) [op’tsoapult] or /będjaraminonja/ \( \rightarrow \) [będjiamonja] (Holisky 1991, based on Kartozia 1968). In the Optimality Theoretic framework, this alternation is easy to account for: the OCP constraint that penalizes two identical consonants in onsets of adjacent syllables is ranked above the Max constraint. Despite theoretical predictability, such deletion, where a consonant is dropped before an identical consonant in the following syllable, is not typologically frequent.

4.2 Mingrelian nasalization

Harris (1991) describes a rule in Mingrelian whereby /p\(^\prime\)/, /p/, /b/, and /m/ turn into a [n] before a consonant across a morpheme boundary, e.g. [k’ots-ep-i] vs. [k’ots-en-k] and [k’ots-en-s].

\[
\begin{align*}
+\text{lab} & \rightarrow [+\text{cor} +\text{nas}] / - [+\text{cons}] \\
\end{align*}
\]

This alternation is limited to morpheme boundaries: it constitutes a case of Non-Derived Environment Blocking (NDEB, Kiparsky 1993) where, as name suggests, an alternation only operates in derived environments and is blocked elsewhere. Moreover, the alternation is not completely regular: it does not apply in all morphological environments (Harris 1991). The peculiar aspect of the alternation is that it triggers both change in manner from obstruents to nasals as well as change in place from labial to coronal. At first sight, the rule appears phonetically highly unmotivated — nasalization in a non-nasal environment (before a voiceless stop) is not phonetically easy to motivate. However, Mingrelian syllable structure and cluster phonotactics might motivate the rule internally. It is reasonable to assume that the change in manner is motivated by SSP: plateaus such as /p\(^h\)k\(^h\)#/ or /p\(^h\)s#/ are repaired by increasing the sonority of the first element. Note, however, that /p\(^h\)s/ is a licit cluster in Mingrelian. The change in place could be motivated by the restriction on Mingrelian clusters, whereby the first element of a cluster in which second element is an obstruent from the groups B or C (see Table [10]) has to be either /r/ or /n/. Further descriptive and experimental work should provide more conclusive results.
4.3 Focus gemination

One of the more intriguing processes in NEC is focus gemination, reported by Nichols (1994) and Nichols (2011) for Chechen and Ingush. The process targets the “last intervocalic” consonant in Chechen and “first post-vocalic” consonant of a word in Ingush when in focus or emphasized and geminates it, e.g. Ingush /ləqə/ ‘high’ vs. /ləqqə/ ‘high.foc’ (Nichols 2011). This process could also be analyzed as a C-reduplication of focused words, but because another synchronic process, word-final-gemination, produces identical results in Ingush (see 4.5 below), the analysis with gemination seems appropriate. In addition to gemination, focused words receive a special intonational pattern with both vowels around the geminated consonant receiving high pitch and emphasis. This intonation gives an acoustic impression as if the word under focus features two stressed syllables (Nichols 2011).

As Nichols points out, syntactic conditions of gemination are not always clear. For Chechen, this process seems to be “frozen in the lexicon” (Nichols 1994:20), but no such remarks are made for Ingush.

Despite gemination being one of the more peculiar strategies to mark focus, no detailed phonetic or syntactic studies exist for this phenomenon. A similar process is reported in some verbal stems in Abkhaz (Hewitt 1979), where it is sporadic and analyzed as reduplication, e.g. /a-ʰə-ra/ ‘to say’ and /a-ʰə-ra/ ‘to cry’. Functions of the gemination/reduplication in Abkhaz are several: from adding intensive semantic component to having onomatopoetic function (see Hewitt 1979). Further studies of this typologically rare phenomenon are a desideratum.

4.4 Reduplication

Several reduplicative patterns emerge in Caucasian languages, ranging from C- or CV-reduplication to complete reduplication (e.g. Schulze 1997, Butskhrikidze 2002). One pattern that stands out as typologically rare is reported in Hinuq in Forker (2013): reduplication morpheme is of the structure CVC(V), but the initial consonant of the reduplicated stem is replaced by /m/ or /k’/, e.g. [kolə] vs. [kolə-molə], [hali] vs. [hali-malica], [kottu] vs. [k’ot-mottu], [roq’e] for [roq’e-t’oq’e]. The semantics of this reduplication is ‘more emphatic or more extreme’ and can target different parts of speech. This reduplication pattern is highly reminiscent of echo formations or the so-called shm-reduplication, but the semantics of Hinuq reduplication differs radically from the semantics of echo formations. Echo formations are “used to downplay or deride a particular phrase” (Nevins and Vaux 2003); the semantics of Hinuq reduplication are emphatic. The case of Hinuq suggests that echo formation types of reduplication that manipulate root-root-initial segments are possible in the function other than that of downplaying.

4.5 Processes targeting word-final voiceless stops

The rest of this section focuses on two processes in NEC that target consonants in word-final position: final gemination and final voicing. Both processes are typologically rare and arguably unnatural: they target the unmarked segment and turn it into a marked segment in a given environment. While final gemination might be phonetically motivated, final devoicing is indeed “truly” unnatural: it operates against a universal phonetic tendency that devoices word-final voiced obstruents (Beguš 2016). The section discusses the status of FV in Lezgian and implications that it bears for phonological theory.
4.5.1 Final gemination

Nichols (2011) reports that consonants are geminated in word-final position in Ingush. Voiceless obstruents alternate with existing voiceless geminates, whereas consonants without phonemic geminate counterparts get geminated phonetically, e.g. [b₂t⁰] for /b₂t/ or [magʷ] for /magʷ/. Final gemination also targets stops and affricates (but not fricatives) in some, but not all word-final clusters: [för:d⁰] for /för:d/. The application of the rule is morphologically conditioned: it applies in only a subset of morphological forms (Nichols 2011).

Final gemination is typologically an unusual process. A survey of sound changes in Kümmel (2007) found only one case of final gemination, which, moreover, targets only a coronal nasal. Geminates are also articulatorily and perceptually dispreferred in word-final position. Long consonants are in general articulatorily more difficult to produce, but even more so in word-final position where segments have “reduced pulmonary pressure” (Iverson and Simon 2011: 1633). In addition, perceptual cues for closure duration are severely impoverished in word-final position.

While final gemination seems to operate against a universal phonetic tendency, it can also be motivated by a process that lengthens final segments. We know that segments are cross-linguistically phonetically longer in word-final position (Lindblom 1968, Oller 1973). Final gemination can thus be analyzed as a phonetically motivated result of word-final phonetic lengthening. Morphologically limited scope of the rule, however, suggest that the process did not arise from a single sound change. Without further data, Ingush final gemination remains an open issue and, as a typologically highly unusual process, merits further research.

4.5.2 Final voicing

It is safe to say that word-final or coda voicing (/T/ → [D] / _#) is one of the most thoroughly discussed phonological process in the literature. It is assumed to be a highly unnatural process that is either impossible or unattested as a synchronic phonological process. The process is reported in Lezgian (Haspelmath 1993, Fallon 1995, Yu 2004, Gajdarov et al. 2009), where word-final voicing targets final unaspirated stops and voices them, e.g. /rap/ → [rab] ‘needle’.

Its opposite process, final devoicing has clear articulatory and perceptual motivations: phonation is difficult to maintain during closure, and this difficulty is even greater word-finally, where stops are produced “with reduced pulmonary pressure” (Iverson and Simon 2011: 1633, Blevins 2004). Moreover, cues for presence or absence of voicing are perceptually impoverished in final position (Steriade 1997, Iverson and Simon 2011). Passive phonetic devoicing in word-final position is attested even in languages without phonological final devoicing. Word-final devoicing thus fits the bill for a universal phonetic tendency (Beguš 2016): (a) it has a well-motivated phonetic explanation; (b) there exists a phonetic tendency to devoice final stops even in languages without phonological devoicing; (c) it is very common and well-attested cross-linguistically.

Thus, any putative process of word-final voicing would necessarily operate against a universal phonetic tendency. Kiparsky (2006) claims that FV is never attested as a productive synchronic process despite several diachronic scenarios that could lead to it (the scenarios are identified in Kiparsky 2006). In fact, he goes a step further and claims that FV is not only unattested, but also impossible and that cognitive restrictions of synchronic grammar are responsible for this typological gap.

Because of these claims, the (non-)existence of FV has become a battleground for the discussion on factors that influence phonological typology. Non-existence of FV is used as evidence in favor of Analytic Bias approach that claims cognitive restrictions shape the typology: if diachronic explanation (Channel Bias; Moreton 2008) is unable to explain the systematic gap, it has to be
Universal Grammar that rules out FV. Blevins (2004) presents several cases of FV, but Kiparsky (2006) argues that none of the apparent cases of final voicing reported in these qualifies as a case of synchronic final voicing — or, at least, the described processes have competing alternative explanations.

The most robust example of word-final voicing is found in Lezgian (Haspelmath 1993, Yu 2004). Haspelmath (1993) and Yu (2004:77) report that Lezgian distinguishes four stop series prevocally (plain voiced, voiceless ejective, voiceless aspirated, and plain voiceless), which in coda position get reduced to a three-way distinction: the plain voiceless series and voiced series merge into a single voiced series. In other words, Lezgian features a synchronic phonological alternation that targets an unmarked segment, word-final unaspirated voiceless stop, and turn it into a marked segment, voiced stops — final voicing (T → D / __#). This process is limited to monosyllabic words. The examples in Table 11 illustrate FV.

<table>
<thead>
<tr>
<th>Place</th>
<th>V#</th>
<th>V V</th>
</tr>
</thead>
<tbody>
<tr>
<td>bilabial</td>
<td>rab</td>
<td>rapar</td>
</tr>
<tr>
<td>dental</td>
<td>pad</td>
<td>patar</td>
</tr>
<tr>
<td>velar</td>
<td>mug</td>
<td>mukar</td>
</tr>
<tr>
<td>uvular</td>
<td>raq</td>
<td>racar</td>
</tr>
<tr>
<td>dental</td>
<td>warz</td>
<td>wartsar</td>
</tr>
<tr>
<td>post-alveolar</td>
<td>ra</td>
<td>ratfar</td>
</tr>
</tbody>
</table>

Table 11: Final voicing in Lezgian (from Haspelmath 1993)

The phonetic study in Yu (2004), however, shows that underlying voiced and plain voiceless series do not neutralize completely: there exists a statistically significant difference between the outcome of the two series in word-final position. Voiced consonants that derive from underlying voiceless stops have a significantly longer closure duration as well as a longer duration of voicing into closure. If we wish to maintain that Lezgian voices final stops, we must, at the same time, assume that these consonants receive (at least phonetic) lengthening as well. It is unclear from a synchronic perspective why this should happen.

The fact that the two series do not neutralize completely allows Kiparsky (2006) to propose an alternative analysis. He assumes that the Lezgian synchronic phonological system has four series of stops — but unlike Yu, he proposes that the fourth series consists of voiced geminates. Thus, instead of coda voicing, he assumes that the process in 11 is in fact onset degemination and devoicing (/Dσ/ → [T]/ σ[ ]). The two analyses are summarized in Table 12.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>Output</td>
</tr>
<tr>
<td>V</td>
<td>#</td>
</tr>
<tr>
<td>T'</td>
<td>T'</td>
</tr>
<tr>
<td>T'h</td>
<td>T'h</td>
</tr>
<tr>
<td>T</td>
<td>D</td>
</tr>
</tbody>
</table>

Table 12: Different input analyses of Lezgian stops in Yu (2004) and Kiparsky (2006)

Kiparsky’s (2006) analysis, too, has its shortcomings: like Yu, Kiparsky has to devise a two-step process — devoicing and degemination of voiced geminates in onset position — and onset
devoicing is not a particularly common process in its own right. However, this derivation is by no means impossible, and Kiparsky (2008) provides evidence from other languages such as Mordva, Ewondo, and Lac Simon Algonquian (Iverson 1983) demonstrating that initial devoicing is a possible synchronic phonological process. As a sound change, such development may be attested in Anatolian and in Selkup of the Samoyedic group (Kümmel 2007).

In sum, although Lezgian provides an apparently compelling example of final voicing, this analysis encounters two major problems. First, the voicing process is limited to monosyllabic words. Second, the plain voiceless and voiced series do not neutralize completely in coda position; a phonetic difference remains detectable between the two series. These problems pave the way for alternative proposals that analyze the alternating series as underlyingly voiced and assume that the synchronic phonological process in Lezgian is in fact onset devoicing rather than final voicing. Additionally, a lack of speaker data from nonce-word tests makes it difficult to determine how productive this process actually is. Further investigations of this typologically rare case of a “truly” unnatural process are required for more conclusive results: wug-tests would provide information on productivity; dialectal research might reveal dialects that neutralize voiceless and voiced stops completely in word-final position.

5 Conclusion and future directions

This chapter aims to survey the major topics of segmental phonetics, phonotactics, and phonological alternations of Caucasian languages. Discussions of details from individual languages are avoided: descriptions of individual language families and languages are given in subsequent chapters. Instead, the chapter focuses on major issues: aspects that yield broadly relevant results and have received a substantial amount of not only descriptive, but also theoretical or general phonetic literature. It also aims to serve as a collection of those aspects of Caucasian phonology that have not yet received sufficient treatment, but bear the potential for yielding results with broader implications.

Each section reveals that even the major topics in Caucasian phonetics and phonology are understudied and point to those aspects that merit further research. It is in fact surprising that the languages of the Caucasus have not received more attention in phonetic and phonological literature, especially given their rich inventories of typologically highly unusual segments and processes.

The discussion in this chapter suggests several potential research projects that should yield directly relevant results for phonetic and phonological theory. Pharyngeals and pharyngealization, velar lateral fricatives, ejective fricatives, and doubly articulated bilabio-alveolar stops are some of the highly unusual segments of Caucasian languages, yet detailed and systematic phonetic descriptions of these phenomena are still lacking. Standards and technological availability of both acoustic and articulatory research tools and methods have improved dramatically since last major studies of Caucasian phonetics were undertaken: the first next step in phonetic research of the Caucasus should involve instrumental acoustic and articulatory descriptions of at least those dialects that feature typologically unusual segments. Articulatory real-time MRI or ultrasound studies of pharyngeals, epiglottals, pharyngealization, and epiglottization should yield further insights into the phonetics of the radical part of the vocal tract and answer questions such as: is phonemic contrast between pharyngeal and epiglottal place of articulation possible, what are acoustic correlates of pharyngeal vs. epiglottal place of articulation, exactly how many different possible articulations are there in the radical part of the vocal tract and what are their mechanisms, and what (if any) are phonetic differences between consonantal and vocalic pharyngealization.
Articulatory studies of velar laterals, NWC series of sibilants, and ejective fricatives would reveal where precisely in the vocal tract the point of constriction is made for these segments and what are their acoustic correlates. The outstanding question of how we define a doubly articulated stop vs. a secondary articulation and relatedly, are doubly articulated bilabio-alveolar stops even possible should be explored on the case of NWC labialized alveolars: further articulatory and acoustic studies of these segments, especially with respect to timing difference between two constrictions, are a desideratum. To my knowledge, there exist no perceptual studies for any of these typologically unusual segments. Other topics worthy of further investigation include correlation in different acoustic parameters between stops with different laryngeal features (both intra- and inter-dialectally), phonetic typology of ejectives, phonetic effects of ejectives on neighboring sounds, and causes of aspiration dissimilation.

In the area of phonology, several experimental studies have the potential to yield results interesting to the broader theoretical discussion. An experimental study in the form of well-formedness judgments would reveal whether the restriction of large clusters to word-initial vs. word-final position in Georgian vs. Svan is part of active synchronic phonology in these two languages. Likewise, one of the most widely discussed phonological alternation, Lezgian final voicing, has been analyzed phonetically and phonologically in detail, but no experimental studies exist that would test synchronic productivity of the rule: a wug-test of Lezgian final voicing would provide invaluable further insights into the discussion on naturalness in phonology. Similarly, thorough descriptive and experimental studies of other less well-described unnatural or typologically unusual processes discussed in this chapter, such as final gemination, focus gemination, Mingrelian nasalization, Laz identical consonant deletion, or the OCP effects in Mingrelian would also yield further insights for theoretical questions in phonology.

Acknowledgements

I would like to thank Ioana Chitoran, Marika Butskhrikidze, Maria Polinsky, and Lena Borise for their helpful discussion on aspects of Cacuosian phonetics and phonology. Some experimental results are reported for the first time in this chapter. The experiments were funded by the Mind Brain Behavior initiative at Harvard University and the Abby and George O’Neill Research Grant at Davis Center at Harvard University to the author, as well as the NSF grants BCS-1144223 and BCS-1619857 to Maria Polinsky.

Notes

1. The linear mixed effects model was fit using the lme4 (Bates et al. 2016) and lmerTest (Kuznetsova et al. 2016) packages in R statistical software (R Core Team 2016).
2. René Lacroix has also created and shared one of the largest databases of any Caucasian languages: Pan-dialectal documentation of Laz [http://elar.soas.ac.uk/deposit/0070] is an online database of 230 hours of high-quality recording of four dialects of Laz with over 360 speakers recorded.
3. The LaCiTO database is available online at [http://lacito.vjf.cnrs.fr/pangloss/corpus/list_rsc_en.php?lg=Ubykh&aff=Ubykh]; Ubykh recordings from the UCLA Phonetics Lab Archive are available at [http://archive.phonetics.ucla.edu/Language/UBY/uby.html].
4. Recordings of the Burkikhan dialect of Agul with examples of obstruents in the pharyngeal/epiglottal region are available at the UCLA Phonetics Lab Archive at [http://archive.phonetics.ucla.edu/Language/AGX/agx.html].
5. It is possible that experiments with higher power would yield different results with respect to effect of pharyngealization on F0.
References


Grawunder, Sven, Adrian Simpson and Madzhid Khalilov. 2010. Phonetic characteristics of ejectives — samples from Caucasian languages. In Susanne Fuchs, Martine Toda, and


Šaumyan, R. 1941. *Grammatičeskij očerk Agul’skogo jazyka*. Moscow: Izdatel’stvo akademii nauk SSSR.


