SOME REMARKS ON THE FOUR-WAY PHONATION CONTRAST IN BZHEDUGH ADYGHE

Ludger Paschen

Leibniz-Zentrum Allgemeine Sprachwissenschaft (Berlin)
paschen@leibniz-zas.de

ABSTRACT

This paper explores the acoustic properties of the four-way phonation contrast in Bzhedugh, an under-described dialect of Adyghe (Northwest Caucasian). As opposed to most other Northwest Caucasian varieties, Bzhedugh Adyghe exhibits a quaternary system of stops and affricates. We scrutinize durational and spectral measures for each of the four phonation types /D, T, Tʰ, T’/. Results indicate that VOT is crucial for distinguishing /D, T, Tʰ/, while voice quality, spectral energy and vowel duration are the principal acoustic cues of ejectives. Voicing, post-burst frication and glottal pulses are retained in word-final position, making the quaternary system robust across different environments. Bzhedugh unaspirated consonants differ considerably from unaspirated fortis and geminate sounds found in Northeast Caucasian languages.

Keywords: Caucasian languages, Adyghe, VOT, phonation, ejectives

1. INTRODUCTION

The Bzhedugh (alternative spelling: Bzhedug) dialect of Adyghe (Northwest Caucasian, Circassian) is spoken in around 20 to 30 villages in the Teuchezhsky district and neighboring districts of the Republic of Adygea in Russia [13, 22]. With a total of 66 consonant phonemes, Bzhedugh attests the largest sound inventory of all varieties of Adyghe [23]. The reasons for this lie in the distinction of several places of articulation, secondary labialization, and a four-way phonation contrast for a considerable percentage of the stops and affricates. This quaternary system represents an archaic trait of Proto-Circassian that has been simplified to a ternary one in all other contemporary Circassian varieties but Shapsugh [8, 7, 12]. The four types of phonation distinguished in Bzhedugh are voiced, voiceless unaspirated, voiceless aspirated, and ejective. While some consonants show only a ternary (/ẑ, ŝ, ŝ’/), a binary (/q, qʰ/), or no phonation contrast at all (/ħ/), the following series have the full four-way contrast: /b, p, pʰ, p’/, /d, t, tʰ, t’/, /dz, ts, tsʰ, ts’/, /dʒ, tʃ, tʃʰ, tʃ’/, and /dʑ, tɕ, tɕʰ, tɕ’/.

Previous research on the segmental phonetics of Circassian languages and dialects includes data discussed in [18, 2, 5, 8, 6, 14, 22, 9, 1, 21]; a systematic description of the acoustic properties of the quaternary phonation system in Bzhedugh, however, is still missing. Against this background, this study seeks to provide an overview of the most relevant acoustic features associated with the four phonation types for a representative subset of the consonant system.

2. METHODOLOGY

The data analyzed in this pilot study were recorded in 2014 during a field trip to the village of Vochepshiy (a. Вочепший) located in the Teuchezhsky district of Adygeya in Russia. One female adult native speaker of Bzhedugh Adyghe was recorded in a classroom at a local school. The speaker was asked to read out loud the carrier phrase in (1) followed by three repetitions of the target word (“XXX”).

(1) nafsetʰ gʷəɕʰaʔew XXX ɕe qəʔʷaʁ
N. word XXX three:times said
‘Nafset said XXX three times.’

Two positional contexts were tested: words containing the target sounds in word-initial position followed by /ə/ and words containing the target sounds in word-final position. In the case of /p/ and /dʒ/, the speaker did not accept the intended stimuli with /ə/ as V1, in which case words with /e/ or /a/ were employed instead. In total, 20 types (80 tokens) were recorded for the word-initial condition and 18 types (72 tokens) for the word-final condition. Selected lists of target words will be provided in the appropriate paragraphs in Section 3.

The speaker was recorded using a hama EL-80 headset with an omnidirectional microphone plugged into an Olympus LS-10 Linear PCM Recorder. The recording settings were at .wav, 44.1 kHz, 16-bit, stereo. The target words were recorded over three recording sessions in total. Data were then analyzed using Praat [3]. The following acoustic values were measured:

1. Three durational measures: (i) burst and frication duration (l(BF)), (ii) VOT, i.e. Δ_t between voice onset and burst onset, and (iii) duration of the vowel following the initial obstruent (l(V))
2. Spectral energy (Pa^2/sec) of post-burst lag
3. H1-H2CI of the following V, using a modified version of the script in [24] applying the Iseli correction for formant effects [15]
4. Noise-to-harmonics ratio (NHR)
5. F0 (Hz) at the vocalic onset

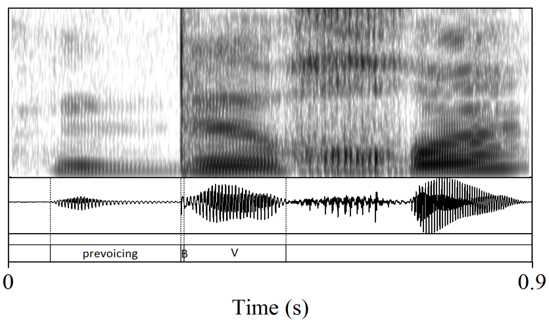
Measurements were taken from the carrier phrase unless the signal was disturbed, in which case one of the isolated words was used instead. Except for frequency range (5.5 kHz) , Praat's default settings were used. In the case of H1-H2CI and NHR, two values were taken: one for the first quartile and the other for the whole duration of the V.

3. RESULTS

3.1. Word-initial plosives

The most striking feature of all voiced obstruents in Bzhedugh is a long section of prevoicing preceding the consonantal burst. For /b/, prevoicing lasts for around 220 ms, which by far exceeds the duration of the following V (see Figure 1). The oral burst is fairly weak and short and is immediately followed by the vocalic onset, with a lag of 0 ms. F0 is lower compared to the non-voiced plosives. The exact acoustic values for the stimuli used are given in Table 1.

Figure 1: Oscillogram and spectrogram (frequency range: 0-6 kHz, dynamic range: 60 dB) of *bəχʷə* 'broad', displaying an extensive phase of prevoicing before the oral burst.



Voiceless unaspirated /p/ is characterized by a short VOT (21 ms) with low-energy noise and absence of prevoicing. Voice quality in the following vowel is modal (H1-H2CI < 0, NHR < 0.10).

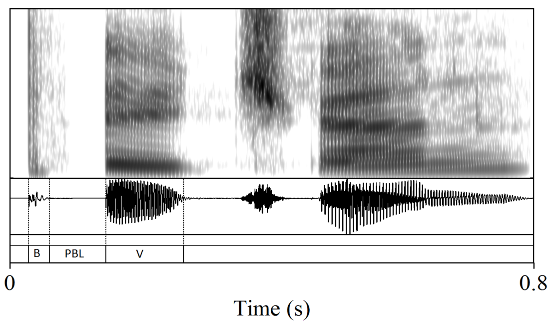
Voiceless aspirated /pʰ/ is characterized by a VOT of around 100 ms with noisy glottal frication. Voice quality in the first fourth of the vowel is remarkably breathy (H1-H2CI > 15, NHR > 0.20), while the overall voice quality is rather modal.

Table 1: Measurements for bilabial plosives.

	/b/ *bəχʷə* 'broad'	/p/ *pane* 'thorn'	/pʰ/ *pʰədzən* 'separate'	/p'/ *p'ətɕ'en* 'freeze'
VOT	-221 ms	21 ms	100 ms	120 ms
l(V)	150 ms	316 ms	96 ms	112 ms
Energy	—	1.4e-8	5.1e-7	4.7e-10
H1-H2CI	13.0, 8.0	1.8, -0.8	16.1, 3.4	2.4, 5.9
NHR	0.16, 0.03	0.17, 0.09	0.21, 0.15	0.11, 0.08
F0	165 Hz	214 Hz	180 Hz	224 Hz

Ejective /p'/ shows a quasi-affricated oral release, which contributes to a longer VOT than its non-ejective counterparts. After the oral burst, the glottal pressure is not released immediately, resulting in an extended post-burst lag (PBL) that comes close to that of /pʰ/. However, unlike /pʰ/, this phase is quasi-silent (lower in energy by a factor of 10^3), and the release of glottal pressure is delayed until the vocalic onset (Figure 2). The glottal release is then distributed over a considerable portion of the vowel. This amounts to a tense or pressed voice quality, a special kind of creak that is typically accompanied by low H1-H2, a low NHR, but not necessarily a low F0 [17]. Note that this particular token of /p'/ had the highest F0 among the bilabial stops. Note further that NHR after coronal ejectives in Bzhedugh was found to be elevated rather than lowered.

Figure 2: One token of the word *p'ətɕ'en* 'freeze'. Note the extended oral release phase, the absence of an isolated glottal release event, and the low inter-harmonic noise in the first vowel.



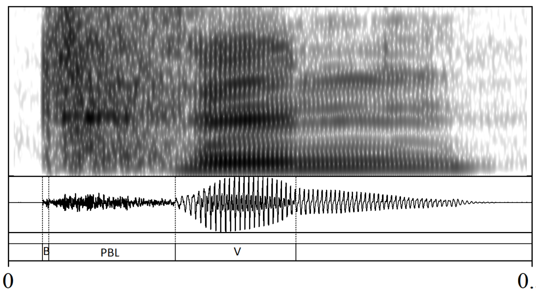
Voiced dental /d/ is similar to bilabial /b/ in that the burst is preceded by a considerable amount of prevoicing. The low F0 and the slightly elevated H1-H2CI compared to its voiceless counterparts also match /b/. /t/ is close to /p/, exhibiting a short VOT

Table 2: Measurements for dental plosives.

	/d/ *dəwe* ‘leech’	/t/ *təʁə* ‘sun’	/t^h/ *t^{h}ən* ‘give’	/t’/ *t’ən* ‘dig’
VOT	-201 ms	16 ms	126 ms	89 ms
l(V)	180 ms	158 ms	112 ms	103 ms
Energy	—	1.2e-7	5.7e-6	7.8e-10
H1-H2CI	10.5, 12.1	7.3, 3.2	14.0, 4.8	-1.6, 0.1
NHR	0.14, 0.05	0.17, 0.07	0.19, 0.07	0.44, 0.15
F0	168 Hz	211 Hz	191 Hz	215 Hz

and modal voice, and /t^h/ parallels /p^h/ with respect to its long and noisy post-burst lag and the breathy initial portion of the following V (see Figure 3). In a similar vein, /t’/ is akin to its bilabial counterpart, showing a quasi-affricated burst release with a high F0 as well as pressed voice at the vocalic onset while lacking an isolated glottal release event. Vowels following ejectives in Bzhedugh tend to be shorter than after non-ejectives, which can be interpreted as a direct consequence of their special phonation.

Figure 3: A token of *t^{h}ən* ‘give’ with high-energy post-burst glottal frication.



3.2. Word-initial affricates

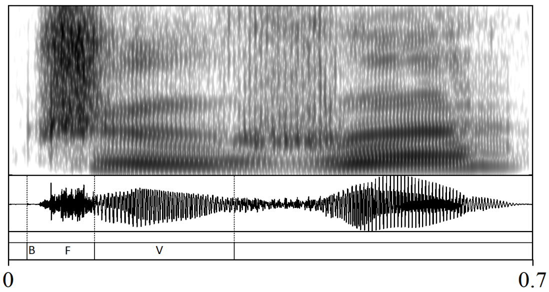
Table 3 shows the acoustic measurements for the postalveolar affricates /dʒ, tʃ, tʃ^h, tʃ’/. Occasionally, a higher NHR in the portion immediately following the vocalic onset can be observed due to partial overlap with consonantal frication noise.

Parallel to the stops, voiced /dʒ/ exhibits a negative VOT and a low F0 at the vocalic onset. Voiceless unaspirated /tʃ/ shows no PBL before the V (see Figure 4). Voiceless aspirated /tʃ^h/ has a typical VOT of around 100 ms with high-energy noise as well a high degree of breathiness at the beginning of the V, while the average voice quality of the whole vowel is fairly modal. Note that the stimulus *tʃ^{h}ə* ‘my brother’ is derived from /s-ʃə/ (1SG-brother) with compensatory aspiration, a process not found in other varieties of Adyghe. Ejective /tʃ’/ has a short and low-energy PBL (by a factor of 10^4 compared to /tʃ^h/) and the following V attests an overall high NHR. As was observed for /t’/, the release of the glottal closure is delayed until the vowel sets in, resulting in pressed voice and a shorter V. However, F0 is not increased after ejective affricates as it is after ejective stops.

Table 3: Measurements for postalv. affricates.

	/dʒ/ *dʒexaħe* ‘sullen’	/tʃ/ *tʃəɣə* ‘tree’	/tʃ^h/ *tʃ^{h}ə* ‘my brother’	/tʃ’/ *tʃ’əg^{w}ə* ‘soil’
l(BF)	42 ms	90 ms	153 ms	105 ms
VOT	-152 ms	90 ms	241 ms	131 ms
l(V)	124 ms	186 ms	178 ms	86 ms
Energy	—	—	4.5e-6	4.4e-10
H1-H2CI	-0.2, -0.1	4.0, 1.4	17.5, 3.2	-5.5, -4.5
NHR	0.27, 0.22	0.13, 0.06	0.14, 0.07	0.68, 0.27
F0	169 Hz	196 Hz	170 Hz	167 Hz

Figure 4: A plain unaspirated postalveolar affricate with zero post-burst lag in *tʃəɣə* ‘tree’.



The dental affricates /dz, ts, ts^h, ts’/ and the alveo-palatal affricates /dʑ, tɕ, tɕ^h, tɕ’/ reveal the same basic patterns as the postalveolar affricates described above, including a negative VOT for the voiced sounds, presence of high-energy glottalic frication and an overall breathier voice quality for the aspirated sounds, and a quasi-quiet post-burst lag as well as tense voice for the ejectives (Table 4).

3.3. Word-final stops and afficates

For word-initial stops and affricates, the relevant acoustic cues are not only located in the signal of the consonant itself but also in the following vowel and the transitional phase. Word-finally, most of these cues are not available, for obvious reasons. One could therefore expect either neutralization of some or all contrasts or phonetic reinforcement to maintain the contrasts. The latter is the case in Bzhedugh.

The contrast between voiceless plain and aspirated sounds is maintained word-finally by the presence

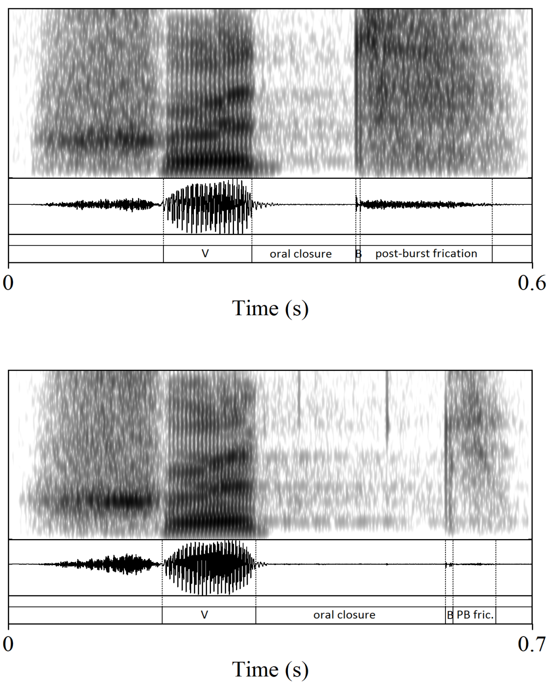
Table 4: Measurements for dental affricates.

	/dz/ *dzəwe* 'bag'	/ts/ *tseȝə* 'gum'	/ts^h/ *ts^{h}ə* 'wool'	/ts'/ *ts'əfə* 'human'
l(BF)	164 ms	166 ms	121 ms	161 ms
VOT	150 ms	166 ms	250 ms	189 ms
l(V)	83 ms	240 ms	125 ms	99 ms
Energy	—	—	3.9e-5	1.4e-9
H1-H2CI	-1.4, 2.7	-1.7, 0.3	13.3, 3.0	-5.1, -2.8
NHR	0.21, 0.08	0.12, 0.08	0.22, 0.13	0.32, 0.16
F0	163 Hz	208 Hz	182 Hz	174 Hz

of a phase of post-burst glottal frication. Figure 5 shows a word-final /t^h/ with noisy PB aspiration of 151 ms and 1.2e-6 Pa^2/sec juxtaposed to a word-final /t/ with PB noise of only 57 ms and 3.3e-8 Pa^2/sec.

Bzhedugh has no process of final devoicing, i.e. voiced Cs retain their voicing in all positions. Word-final ejectives are similar to word-initial ones in that there is a significant interval between the release of the oral and the glottal closure, only that without a following vowel, the release manifests itself in a series of isolated glottal pulses (Figure 6).

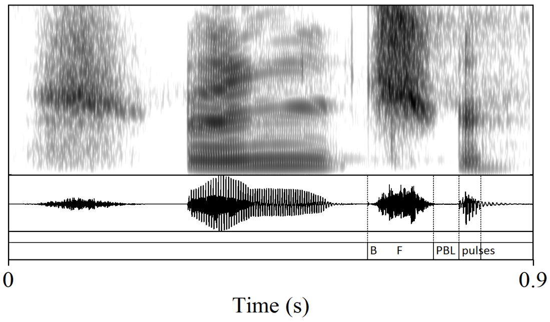
Figure 5: Tokens of the minimal pair *xet^h* 'stands in' (above) and *xet* 'who' (below).



4. DISCUSSION

Non-ejective unaspirated consonants in Bzhedugh have sometimes been transcribed as /C:/ [12]. The findings presented in Section 3, however, support

Figure 6: A token of *ʃk^w'əntɕ'* 'darkness'.



the observation in [8, 6] that these sounds are in fact plain singletons. In general, non-ejective non-continuants in Bzhedugh are well in line with Keating's [16] typology of prevoiced, short-lag, and long-lag stops. This sets Bzhedugh, as well as closely related Shapsugh, apart from its Northeast Caucasian neighbors, where the unaspirated series is fortis/tense or geminated [4, 20, 19, 10]. VOT was not reliably longer for /C^h/ than for /C'/, inviting further inquiry as to how Bzhedugh relates to the east-west divide across this parameter reported in [10]. At the same time, Bzhedugh /C'/ fits the complex picture of ejectives in the Caucasus in general, which is characterized by a plethora of phonetic realizations [11]. The quaternary phonation system proved robust across all sound classes and environments, calling into questions previous claims about an ongoing attrition of the /C∼C^h/ contrast in Bzhedugh [22].

Future research may shed light on the question whether the phonetic cues discussed in this paper are the same for obstruent series with only three or less members, including fricatives. In addition, it would be worth investigating whether there is an empirical basis for the stipulation of an even finer gradation of aspiration. Thus, [18, 22] assert aspiration to be consistently stronger in some (e.g. *tʃ^{h}ə* 'my broter'; cf. Section 3.2) than in other lexical items (such as *xet^h* 'stands in'; cf. Section 3.3).

5. CONCLUSION

In this paper, we have presented a number of acoustic correlates of the four-way phonation contrast in Bzhedugh. Results indicate that VOT and PBL energy are crucial for distinguishing /D, T, T^h/, while pressed voice coupled with a short duration of the following V are the principal acoustic cues for ejectives. Voicing is generally accompanied by a lower F0 in the following V. The quaternary phonation system is fully functional across different sound classes and phonetic environments.

6. REFERENCES

[1] Applebaum, A., Gordon, M. 2013. A comparative phonetic study of the Circassian languages. *Proceedings of the 37th Annual Meeting of the Berkeley Linguistics Society: Special Session on Languages of the Caucasus* 3–17.

[2] Balkarov, B. C. 1970. *Fonetika adygskich jazykov*. Nal'čik: Knižnoe Izdatel'stvo Ėl'brus.

[3] Boersma, P., Weenink, D. 2018. Praat: doing phonetics by computer [computer program]. Version 6.0.43, available from http://www.praat.org/.

[4] Bokarev, A. A. 1949. *Očerk grammatiki čamalinskigo jazyka*. Moskva/Leningrad: Izdatel'stvo Akademii Nauk.

[5] Catford, J. C. 1977. Mountain of tongues: the languages of the Caucasus. *Annual Review of Anthropology* 6, 283–314.

[6] Catford, J. C. 1997. Some questions of N.W. Caucasian phonetics and phonology. *Proceedings of the Conference on Northwest Caucasian Linguistics, Oslo* volume 3 of *Studia Caucasiologica*. Novus forlag 99–113.

[7] Chirikba, V. A. 1996. *Common West Caucasian*. Leiden: Research School CNWS.

[8] Colarusso, J. 1988. *The Northwest Caucasian languages*. New York/London: Garland Publishing.

[9] Gordon, M., Applebaum, A. 2006. Phonetic structures of Turkish Kabardian. *Journal of the International Phonetic Association* 36, 159–186.

[10] Grawunder, S. 2017. The Caucasus. In: Hickey, R., (ed), *The Cambridge Handbook of Areal Linguistics*. Cambridge: University Press 356–395.

[11] Grawunder, S., Simpson, A., Khalilov, M. 2010. Phonetic characteristics of ejectives - samples from Caucasian languages. In: Fuchs, S., Toda, M., Żygis, M., (eds), *Turbulent sounds: an interdisciplinary guide*. Berlin: De Gruyter Mouton 209–244.

[12] Hewitt, G. 2005. North West Caucasian. *Lingua* 115, 91–145.

[13] Höhlig, M. 1997. *Kontaktbedingter Sprachwandel in der adygeischen Umgangssprache im Kaukasus und in der Türkei*. München/Newcastle: Lincom Europa.

[14] Höhlig, M. 2003. *Fonetičeskoe opisanie zvukov adygejskogo jazyka*. Maykop: Kačestvo.

[15] Iseli, M., Alwan, A. 2004. An improved correction formula for the estimation of harmonic magnitudes and its application to open quotient estimation. *International Conference on Acoustics, Speech, and Signal Processing 2004* 669–672.

[16] Keating, P. 1984. Phonetic and phonological representation of consonant voicing. *Language* 60, 286–319.

[17] Keating, P., Garellek, M., Kreiman, J. 2015. Acoustic properties of different kinds of creaky voice. *Proc. 18th ICPhS, Glasgow* paper 821.

[18] Kerasheva, Z. I. 1965. Aspirirovannye i poluabruptivnye smyčnye soglasnye v dialektach adygejskogo jazyka. *Ežegodnik iberijsko-kavkazskogo jazykoznanija* 13, 100–114.

[19] Kibrik, A. E., Kodzasov, S. V. 1990. *Sopostavitel'noe izučenie dagestanskich jazykov: Imja. Fonetika*. Moscow: Izdatel'stvo Moskovskogo Universiteta.

[20] Kodzasov, S. V. 1977. Fonetika arčinskogo jazyka. In: Kibrik, A. E., Kodzasov, S. V., Olovjannikova, I. P., Samedov, D. S., (eds), *Opyt strukturnogo opisanija arčinskogo jazyka* volume 1. Moscow: Izdatel'stvo Moskovskogo Universiteta 185–352.

[21] Paschen, L. 2015. An acoustic study of fricatives in Temirgoy Adyghe. *Proc. 18th ICPhS, Glasgow,* paper 605.

[22] Sitimova, S. S. 2004. *Osobennosti bzhedugskogo dialekta adygejskogo jazyka*. PhD thesis Maykop: Adyghe State University.

[23] Smeets, H. J. 1984. *Studies in West Circassian phonology and morphology*. Leiden: The Hakuchi Press.

[24] Vicenik, C. 2017. Praat Voice Sauce Imitator [Praat script]. Downloaded from http://phonetics.linguistics.ucla.edu/facilities/acoustic/PraatVoiceSauceImitator.txt.