Sound systems in diachrony: Sibilants and affricates in Udi

Jost Gippert
Goethe-Universität, Frankfurt am Main

1 Preliminaries

It is a well accepted fact that the Khoisan languages of Southern Africa possess the largest consonant systems worldwide, with inventories partly exceeding 100 phonemes. It is as well accepted that the next richest inventories are met with in the Caucasus region, in the languages of the (North-)West Caucasian and (North-) East Caucasian families. Different from the Khoisan languages, the richness is not determined by clicks in these languages; instead, we usually find at least three different articulation types of stops and affricates here (voiced, voiceless-aspirated, voiceless-glottalised), extending from labials down to uvulars, pharyngeals, and laryngeals, and partly comprising palatal or labial coarticulations. In this way, a total of 85 phonemes was determined for Ubykh, a West Caucasian language now extinct (the last known speaker died in 1994), and a not much smaller inventory of 82 has been assigned to the Archi language of the East Caucasian stock, which is in general especially characterised by sets of lateral affricates.[1]

A peculiar position within the latter language family is held by Udi, an East Caucasian language spoken by about 5,000 speakers in North-West Azerbaijan, South-East Georgia, North-East Armenia, and elsewhere in states of the former Soviet Union. Udi is not only comparatively poor with respect to laterals as well as uvular and pharyngeal consonants, but it is also the only East Caucasian language whose history can be traced back for about 1,500 years, given that it has been ascertained to be the modern offshoot of the medieval language of the Caucasian "Albanians". With the detection of a large amount of written text in the underwriting of palimpsests from St. Catherine's Monastery on Mt. Sinai, the sound system of this latter language (as spoken by the middle of the first millennium C.E.) has been established with high certainty, and among the 52 letters of the peculiar alphabet used for writing "Caucasian Albanian" (hereafter: CA), 45 have been determined to represent consonant phonemes. In contrast to this, the inventory of modern Udi is slightly smaller, being confined to something between 32 and 38 consonants depending on the source used. A peculiar problem in this

---

connection is the existence of several sets of dental, alveolar and postalveolar sibilants and affricates in modern Udi, which are not always distinguished consistently in the sources and whose nature has been a matter of debate in the linguistic literature. This debate can now be reopened on the basis of a comparison with the CA texts and phonological considerations based thereupon, and on the basis of special fieldwork undertaken in the Udi-speaking village of Oḳṭomberi (now Zinobiani) in East Georgia in March, 2008.[2]

2   The CA evidence

2.1   The CA consonant system

On the basis of the Sinai palimpsests, the consonant system as covered by the CA alphabet has been established as comprising the following elements:

Consonants[3]	Stops			Affricates				Frica-tives		Nasals	Thrill	Laterals	Glides
	vcd.	vcl.	gl.	vcd.	vcl.	gl.		vcd.	vcl.				
Labials	Ⴁ *b*	Ⴎ *p*	Ⴒ *ṗ*					Ⴠ *v*	Ⴈ *f*	Ⴘ *m*			(Ⴑ *w*)
Dental-alveolars	Ⴑ *d*	Ⴇ *t*	Ⴒ *ṭ*	Ⴊ *ʒ*	Ⴀ *c*	Ⴜ *c̣*		Ⴘ *z*	Ⴌ *s*	Ⴄ *n*	Ⴐ *r*	Ⴄ *l*	
Palatalised	Ⴇ *d'*		Ⴏ *ṭ'*	*Ⴔ *ʒ'*	*Ⴆ *c'*	Ⴔ *c̣'*				Ⴊ *n'*		Ⴕ *l'*	
Alveolar-palatals				*Ⴞ *ž̌*	Ⴒ *č*	Ⴏ *č̣*	Ⴞ *ž*	Ⴏ *š*					Ⴢ *y*
Postalveolars				Ⴁ *ž̃*	Ⴓ *ć*	Ⴊ *ç̇*	Ⴣ *ź*	Ⴘ *ś*					
Velars	Ⴚ *g*	Ⴉ *k*	Ⴢ *ḳ*										
Uvulars		Ⴈ *x̣/q*	Ⴍ *q̇*					Ⴝ *ġ*	Ⴈ *x*				
Pharyngeal								Ⴆ *ˤ*					
Laryngeal									Ⴈ *h*				

One type of consonants that is no longer met with in modern Udi is the set of palatalised stops consisting of *d'* and *ṭ'*, with the voiceless aspirate *t'* missing (the set is thus defective even in CA); it is probable that these stops merged with the

---

2   The fieldwork was carried out by the present author in cooperation with M. Tandashvili; the main consultant was Mrs. G. Tizlarishvili, a native speaker of Udi born in the village.

3   Gippert et al. 2008: II-17. In the Table and hereafter, glottalisation is marked by a dot below (or above) a given letter. Vcd. = voiced, vcl. = voiceless-aspirated, gl. = voiceless-glottalised; graphemes that are only attested in the CA alphabet list of the ms. Mat. 7117 (cf. Gippert et al. 2008: II-1-17) are marked with an asterisk.

corresponding palatal affricates (*ǯ* and *č*) as shown by cognates such as CA *ḳod'* and Udi *ḳoǯ* 'house', CA *d'ed'er* and Udi *žežer* 'lip', or CA *aq̇aṭ'i* and Udi *-a(n)q̇ači-* 'naked'.[4] A similar merger can be assumed for *c'*, the only palatalised affricate attested in the palimpsests, if CA *muc'ur* 'pure, holy' can be equated with Udi *mučur* 'cloudless'.[5] A straightforward relation can be set up for the dental-alveolar sibilants and affricates, which are represented by the corresponding Udi sounds; cf., e.g., CA *sa* ~ Udi *sa* 'one', CA *owsen* ~ Udi *usen* 'year', CA *zow*[6] ~ Udi *zu* 'I, me', CA *ayz* 'world' ~ Udi *ayz* 'village', CA *çi* ~ Udi *çi* 'name', CA *viç* ~ Udi *viç* 'ten', CA *cam-pesown* ~ Udi *cam-pesun* 'write', or CA *aci-pesown* ~ Udi *ci-psun* 'throw down'. The only exception in this set is *ʒ*, which is attested in only one cognate, with CA *ʒ* being substituted by Udi *ǯ*; this is Udi *pilinǯ* 'dagger'[7] corresponding to CA *pilinʒ* 'bronze, copper, brass'. As this is an areal *wanderwort* (cf. Arm. *płinj* and Georg. (*s*)*pilenʒ-i* 'id.'), it is not certain whether we have a case of uninterrupted preservation or, rather, repeated borrowing here (cf. also Azeri *bürünc* 'bronze').

In contrast to this, the representation of the sets of alveolar-palatals and post-alveolars in modern Udi is not so straightforward. The problem is that while the "orthography" of CA is very consistent in the assignment of the respective letters, the written sources of Udi are quite inconsistent in this respect and phonetic or phonological descriptions are unequivocal, if not contradictory, as we will see below. Nevertheless there are enough examples of cognates to be established between the medieval language and its modern successor in its two dialects, that of Vartashen (now Oğuz; hereafter: V) and that of Nij (hereafter: N), to provide a basis for the comparison.

2.2 CA (near-to) minimal pairs

Even though the amount of written material we have for CA is rather limited (the palimpsests comprise all in all 121 folios, i.e. 242 pages with CA underwritings, of which about 60% have been deciphered so far) and consists only of Biblical texts (half of the Gospel of John and lectionary texts from other NT books, plus

---

[4] The latter word is attested independently as *aqačin* 'nackt' in the first grammatical description of Udi by Schiefner (1862: 74). In the modern language, only the compound *tur-anq̇ačin* 'barefoot' seems to exist (cf. Gukasjan 1974: 206: "ТУРАНКЪАЧИИН"); an example is found in the recording Education1 (at 2:34:00) of the project "Endangered Caucasian Languages in Georgia" (ECLinG) stored in the DoBeS archive at the MPI Nijmegen, cf. https://hdl.handle.net/1839/00-0000-0000-0014-C211-2@view.

[5] The word seems not to be attested in written sources but was confirmed during the fieldwork; Udi *ačar* 'pure' is barely related. – Whether CA possessed a noun **ḳowḳowc'* 'hen' to be equated with Udi *ḳoḳoç* 'id.' remains unclear (cf. Gippert et al. 2008: IV-24).

[6] The digraph *ow* in CA denotes the vowel *u*.

[7] Attested in the dative form *pilinǯa* in Žeiranišvili (1971: 172).

a short passage from Isaiah and some Psalm verses), it contains enough examples to show the distinction of the "second" and "third" sets of sibilants and affricates, in some cases even in the form of (near-to) minimal pairs. Thus we can contrast the following items:[8]

2.2.1 *š* vs. *ś*

ša 'daughter' or *šan* 'cave, den' vs. *aśal* 'earth' vs. *eśa* 'after' etc.
šel 'good, apt' vs. *eśem* 'tempest' or *eśelown* 'last, final'
šow 'night' or *ha-šow* 'who' or *išow* 'man' vs. *śow* 'member' or *śowm* 'bread' or *śov, śoown* '(sound of a) trumpet'
aš 'work' vs. *laśḳo-* 'marriage'
beši 'our' vs. *eśin* 'then, thus, really, indeed'
mowš 'wind' or *ġowšowy* 'man(kind)' vs. *lowśow* 'all, whole'

The sound value of *š* is clearly determinable by its occurrence in loans such as *-šad-* 'happy, free' (~ Parth. M(iddle) Pers. *šād*), *šambat* 'Sabbath' (~ Parth. *šambat*, Arm. *šabatʿ*, Georg. *šab/pat-*, etc.) or *ašarḳet* 'pupil, disciple' (~ Arm. *ašakert*, < M(iddle) Iran(ian), cf. M.Pers. *hašāgird*), as well as Biblical names such as *Ešaya* 'Isaiah' (~ Syr. *ʾEšaʿyā*) or *Šiloham/Šilohan* (name of the lake 'Siloam', ~ Syr. *Šīlōḥāʾ*). In contrast to this, *ś* is not attested in any foreign word in CA so far.

Only in one case, there is variation between the two letters. This is *xaš* 'light', which is likely to be hidden in the verb *xaś-heǵesown* 'be baptised' if this is a calque of Georgian *natlis-ġeba* 'id.', lit. 'take light'.[9]

2.2.2 *ž* vs. *ź*

For this pair, the evidence is much smaller, both letters being rather rare. What we can contrast is:

žin- 'someone' vs. *źiź-iǵesown* 'be shaken'
žan 'we' or *žaḳ* 'chain' or *žowkowm* 'end, pike' vs. *źe* 'stone'

There is but one possible loanword that can be adduced for *ž*, viz. *ž̃d* 'crowd, synagogue' if this represents, as an abbreviation, Arm. *žołovowrd* 'id.'.[10] There

---

8   For the attestation of all words and forms mentioned below cf. the index to the edition of the palimpsests, Gippert et al. 2008, IV-1-42.

9   In contrast to this, 'baptise' is *xač-besun* in modern Udi, which is an obvious calque on Armenian *xačʿ aṙnel* 'make the sign of the cross', with *xač* being a direct loan from Arm. *xačʿ* 'cross'. – Cf. Gippert et al. 2008: II-8 and II-10 for more details on *š* and *ś*.

10   Cf. Gippert et al. 2008: II-13-14 for more details on *ž* and *ź*.

is but a slight chance that *ǯowmarʒ́* 'neighbour' might represent a M.Iran. compound **ǰud-marz* with a meaning like 'having separate march(es)';[11] if so, we might assume a partial assimilation of *-*z* to -*ʒ́* under influence of the initial *ǯ*.

2.2.3 *ǯ́* and *ǯ*

This pair cannot be contrasted because the presumed letter for *ǯ́* is not yet attested at all in the palimpsests.[12] The letter that is taken to represent *ǯ* appears, apart from *ǯowmarʒ́* 'neighbour' (see above), in *ǯowmo(w)x* 'mouth', *ǯow 'news' (in the compound verbs *ǯow-baq̇al-biyesown* 'ask' and *ǯow-daġesown* 'proclaim (good) news, preach the Gospels'), and *ǯowġ 'Lord', always abbreviated as *ǯ͞ġ*, together with its compounds *bixa-ǯowġ 'God', lit. 'creating Lord', and **k̇od'in-ǯowġ* 'landlord, householder', lit. 'lord of the house'; besides we have *ǯ* in the verb *ća-ǯatesown*, which appears in the 1st person sg. impf. form *ća-ǯatay-zow-hē* in Gal. 1.14 rendering Arm. *yaṙaǰadēm ēi* 'I advanced' and which is likely to contain *ća* 'face'; in addition, we have the formation *axay-ǯalown* 'the last', which seems to be built upon a noun *axay* 'end'.

2.2.4 *č* vs. *ć*

q̇ač 'narrow' vs. *bać* 'hundred'
ačam 'unleavened bread' vs. *k̇aći* 'blind'
če- prev. 'out' vs. *haćex* (dat.sg.) 'right (hand)' (> 'useful, apt')
čin 'nation, tribe' vs. *boći-biyesown* 'stick' (< 'make tight'?)
čohoc 'outside' and *čome* 'untruth, illegality' vs. *ćo-* 'patience' (only contained in the verbs *ćo-biq̇esown* and *ćo-ihesown* 'endure')
ičown 'nappy' vs. *vakahaćown* 'girdle'
bečeown 'cave'[13] vs. *heć-ihesown* 'be helpful'

Clear loanwords are *ṗačar* 'reason' (~ Arm. *patčaṙ* 'id.', < M.Iran.) and *čaṭar* 'temple' (~ Arm. *tačar*, vs. Georg. *ṭaʒar-i*, < M.Iran. **tačar-*); whether *ć* was represented in a loan word **p̣aṭmoćan* 'vesture' (~ Arm. *patmowčan* 'id.', < M.Iran. *patmōčan, cf. Parth. *padmōžan*) remains uncertain.[14]

---

11 Cf. Durkin-Meisterernst (2008: 199) for M.Pers. *jwdy*, *jwd* /*judy*/ and 232 for M.Pers., Parth. *mrz* /*marz*/.

12 Cf. Gippert et al. 2008: II-14 as to the letter.

13 Instead of *beči-q̇a-na-va-bowri* tentatively read as the equivalent of Arm. *i bacʿ kacʿcʿē* 'he should keep out (of the synagogue)' in Jo. 9.22, we most probably have to read *ʿaxi-q̇a-na-va-bowri* 'he should stand far apart'.

14 Cf. Gippert et al. 2008: II-14-15 as to further details concerning CA *č* and *ć*.

2.2.5 *č* vs. *ć*

čalxesown 'know' vs. *ća* 'face'
ičownčow / *ičinčow* '(one)self' (from *ič* 'self') vs. *baćow-biyesown* 'select, choose'
čibowx 'woman, wife' and *čobal* 'sparrow' vs. *ćowdow* 'heaven, sky' and *ćow-pesown* 'spit'
bačxesown 'hide' vs. *baćow-biyesown* 'select, choose'

While *ć* is confined to the words mentioned (with derivates)[15] and thus rather rare, CA *č* is represented in many more words, among them *viči* 'brother' or *ačpē* 'false'. CA *čar* 'time, -fold' might be a loanword from M.Iran. (M.Pers. *ǰār* 'time').

2.3   The CA evidence summarised

It seems clear from the evidence accumulated above that the alveolar-palatal series (*š* etc.) is much better represented in the palimpsests than the 'third' series of sibilants and affricates (*ś* etc.). However, there can be no doubt that the two series were sharply distinguished 1,500 years ago, and so the material can well be contrasted with modern Udi usage.

3   The Udi picture

Many of the CA words mentioned above have their counterparts in the dialects of modern Udi, and most of these have been attested in the written sources available for this language. To account for the different representations, it seems appropriate to proceed from the oldest sources available onwards.

3.1   The written sources of Udi

3.1.1   Klaproth 1814

The first material of modern Udi available is a list of 12 words plus one short sentence published in J. von Klaproth's "Description of the Russian provinces between the Caspian and the Black Sea" (1814: 177-178). This material comprises three of the words contrasted above, viz. "Weib – Schuwuk" (cf. CA *čibux*), "Bruder – Witschi" (also in "Mein Bruder – Bis witschi"; cf. CA *viči/bezi viči*), and "Brod – Schum" (also in the sentence "Iß Brod mit uns, mein Bruder – Mieeke arza schum uka bis witschi";[16] cf. CA *śum*); one more item that is relevant

---

[15] *ʕaxnaćow* 'fight, battle' (vs. *ʕaxnaown* 'id.') remains unclear.

[16] Literally, the sentence means 'Come here, sit down, eat bread, my brother' (Udi *mia eke, arca, śum uka, bez viči*). Note that two items of the word list remain undetermined, viz. "Mädchen – Ssengi" and "Knabe – Galli".

in the given context is "Apfel – Oesch" (see below). It is clear that the clumsy German transcription is in no way apt to render the sounds correctly.

3.1.2 Schiefner 1854, Šopen 1866, Starčevskij 1891

A second word list, which was not much more reliable, was published 40 years later by A. Schiefner (1854: 649-650).[17] It contained a total of 70 entries, again transcribed awkwardly in a German way, with some remarkable errors. Apart from "Weib, tschiwuch" and "Bruder, wili" (sic, obviously representing Cyrillic вічі *vič̌i* misread as вілі),[18] we here find "Gott, bichadsug" (cf. CA *bixaӡowġ), "Mann, ischu" (CA *išow*), "Mond, chasch" (CA *xaš̌*), "Wind, musch" (CA *muš̌*), "Stein, *je*" (CA *ӡe*), "Zehn, wiz" (CA *viċ*), "Hundert, *s*abatsch" (CA *sa* "one" + *baċ̌*), "gut, schel" (CA *šel*), "herauskommen, tsche*s*un" (CA *č̌e-* "out" + *heġesown* "come"), as well as "Apfel, esch" and "Kuh, tschur". The italicisation of s in "*s*abatsch" clearly indicates the voiceless pronunciation of Cyrillic *c* (*sa-* "one-"), whereas that in "*j*e" is likely to represent Cyrillic ж.

The same source is likely to lie behind the list of 48 words published, in Cyrillic script, in I. Šopen's "New Remarks" (1866: 483). Here, we read, among others, Чивухъ (*č̌ivux*) for 'woman', Бихадзухъ (*bixadzux*) for 'God', Ишу (*išu*) for 'man', and Чуръ (*č̌ur*) for 'cow', but also Шумъ (*šum*) for 'bread', Цо́ (*có*) for 'face', and Ласко (*lasko*) for 'marriage'. The word for 'moon' and 'light' is twice misspelt Хамъ (*xam*) for *Хашъ (*xaš̌*).

A third, much more extensive, word list that is likely to have used the same source is the one provided by A. Starčevskij in his "Caucasian interpreter" (1891: 494-508), where it appears together with about 150 sentences (612-615) and a short grammatical sketch of Udi (667-668). The word list, consisting of about 1,900 Russian lemmas with their Udi counterparts, comprises, among others, the words for 'woman' and 'wife' (Чибух and Чубух, чибух), 'God' (Бихадзуг, быхадзуг), 'man' (Ишу), 'moon' and 'light' (Хаш), 'wind' (Муш), 'stone' (Же), 'ten' (Виц), 'hundred' (Бац, сабац), 'good' (Шэл), 'come out' (Чесун), 'apple' (Эш), 'cow' (Чур), as well as 'bread' (Шум), 'face' (Цо), and 'marriage' (Ласко). All these

---

[17] The word list had been published before, together with general remarks on the Udis, in an article "On the Udis" ("Ob″ udinax″") in the journals *Moskovskija vědomosti* (no. 94, 1853; non vidi) and *Kavkaz* (no. 61, 1853, 266); the compiler (noted as "A.S." in Mežov (1894: 251), no. 6128) was probably A.J. Sjögren.

[18] The same error may be responsible for "Schwester, chinli" if representing Cyrillic хінчі (Udi xunč̌i). The item "Tag, tschenachun" must represent Udi ġenaxun 'by day', with Cyr. ҕ or the like misread as ч; "starke Hitze, tscheleitscharych" may contain Udi iġari/ux (V)/eġari/ex (N) 'hot, heat' but the initial part remains obscure (Udi gölö/gele 'very, much'? cf. *iġaruġo gölö/ eġareġo gele* 'in great heat' in Çejrani (1934: 56)).

three word lists have in common that they do not differentiate in any consistent way between the 'second' and 'third' series of sibilants and affricates.

3.1.3 Schiefner 1863 and Erckert 1895

The first treatise that was linguistically founded is A. Schiefner's extensive "Attempt" of 1863. However, this treatise, which comprises a grammar, a set of text specimens ("Sprachproben"), and a vocabulary, was not based upon the author's own fieldwork either but on materials provided by the Udi school teacher Georg Bežanov ("Beshanow"), who had died by 1860, as well as a scholar residing in the Caucasus, Adolph Bergé, who worked with Udi speakers from Vartašen and Nij (Schiefner 1863: 2-3). Schiefner presents these materials in a Latin transcription which uses four types of diacritical marks, viz. a *spiritus asper* ("rough breathing") denoting aspirated stops (e.g., *p̓*), a dot below denoting pharyngealised vowels and uvularised consonants (e.g., *ạ* and *ḥ*), a diaresis above denoting fronted ("umlauted") vowels (e.g., *ä*), and a dot above denoting an alveolar-palatal articulation of sibilants and affricates (*ṡ, ż, ċ, ʒ̇* in contrast to "plain" *s, z, c, ʒ*). In addition, his list (Schiefner 1863: 9) comprises three letters with dots above and below (*ṩ, ẓ̇, ċ̣*), plus a *c* with only a dot below. For *ṡ* and *ż*, he states explicitly that they are "sharp sibilants combining the elements *s + ṡ* and *z + ż* and are pronounced like *sṡ* and *zż*".[19] For the pronunciation of *ċ* and *c*, the reader has to refer to a former work of Schiefner's, his treatise of the Tsova-Tush (or Batsbi) language (Schiefner 1856), where he had introduced his transcription system first; here we read that *ċ* is "eine Verstärkung von *ċ*" equalling Georgian ჰ (i.e., the glottalised affricate *č̣*), whereas *c* is "eine Verstärkung von *c*" equalling Georgian წ (i.e., glottalised *c̣*).[20] In this way, Schiefner's inventory comprises three items less than the CA alphabet for the two sets of sibilants and affricates under concern (*ṡ, ż, ċ, ċ̣, ʒ̇* plus *ṩ, ẓ̇* vs. CA *š, ž, č, č̣, [ǯ]* plus *ś, ź, ć, ć̣, ʒ́*).

From the words appearing in Schiefner's materials, it is clear then that his *ṡ* corresponds regularly to CA *š* as in *ṡel* 'good' (CA *šel*), *ṡu* 'night' and 'who' (CA *šow/ha-šow*), *aṡ* 'thing' (CA *aš* 'work'), *beṡ(i)* 'our' (CA *beši*), *muṡ* 'wind' (CA *mowš*), and *xaṡ* 'moon, light' (CA *xaš*), also in loanwords like *ṡad* 'joyful' (CA *šad-*) or *ṡamat* 'week, Saturday' (CA *šambat̰*). In contrast to this, Schiefner's *ṩ* corresponds to CA *ś* in *ṩum* 'bread' (CA *śum*) and *oṩa* 'behind, after' (CA *eśa*), correctly regarded as a (dative) case form of *oṩ* 'end' (CA **eś*, also represented in *eśin* 'then' and *eśelown* 'last'). The counterpart of CA *aśal* 'earth' appears partly as *oc̣ạl*, partly as *oċ̣ạl*, i.e., with pharyngealised vowels and an affricate instead of the *ṩ*. Of the words with a voiced sibilant, we may identify *że* 'stone'

---

[19]   Schiefner (1863: 9-10): "scharfe Sibilanten, deren ersterer die Elemente s und ṡ, letzterer z und ż vereinigt und die wie sṡ und zż ausgesprochen werden sollen".

[20]   Schiefner (1856: 8), where *c* is misprinted for *ċ*.

with CA *źe*; on the other hand, *ź* in CA *źiź-iġesown* 'be shaken' seems to correspond to *ż* in *żikṗesun* 'shake', while the equivalent of CA *žan* 'we' is *jan* (with *j* indicating the glide; *yan* is the modern Udi form in both dialects).

In the range of affricates, the picture is somewhat blurred due to inconsistencies[21] and to the missing balance between the character inventories. CA *č* is regularly represented by *ċ* in *ċubux* 'woman, wife' (CA *čibux*), *wiċi* 'brother' (CA *viči*), *apċi* 'false' (CA *ačpē*), or *iċ-* 'self' (CA *ič-*), the uninflected form of the latter also appearing as *iċ*. In a similar way, *ç* stands quite regularly for CA *č̣* as in *qaç* 'narrow' (CA *q̇ač̣*) or *çe-* 'out-' (CA *č̣e-*). On the other hand, we find *c* written for CA *ć* in *baç* 'hundred' (CA *báć*) and *boçu* 'tight, dense' (CA *bóći-*), but also *ç* in *ḳaçi* 'blind' (CA *ḳaći*) and *aça* 'right (hand)' (CA *haće-x*). CA *ć̣* is reflected by *c* in *cu-ḳun* 'saliva' (CA *ćow-pesun* 'spit')[22] and *co* 'face' (CA *ć̣a*), the latter once being spelt *ço* (p. 69) and noted as *ċo* for the Nij dialect (p. 89). For CA *ǯ* we may take the two compounds **bixaǯowġ* 'God' and **ḳod'inǯowġ* 'landlord', which appear as *byxaʒuġ-*[23] and *ḳonʒux*, but the latter also as *ḳonǯux* (p. 83); the genitive corresponding to CA *ḳod'in-* is *ḳoʒin* everywhere. CA *ǯowmo(w)x* 'mouth' is represented by *zomox*, but also as *zumox* (p. 93). Thus we see that it is basically the 'third' set of affricates which has no clear counterparts in Schiefner's 'system'. However, we may note that in all the examples listed, Schiefner marks the vowels adjoining these affricates as being pharyngealised; a feature that is in no way visible in the respective CA cognates (different from, e.g., the 2nd person pl. pronoun dat. *waₓ* matching CA *vˁ̣ax* with clear indication of a pharyngeal and thus contrasting with the corresponding 2nd person sg. form, *wax* ~ CA *vax*); we shall return to this observation later on.[24]

Compared to Schiefner's account of Udi, the one in R. v. Erckert's survey of the Caucasian languages (1895) is much less comprehensive. The material provided there consists of about 500 Udi words in a comparative word-list (I: 23-204) and about 200 short sentences plus an extremely rough description of the grammar (II. 60-68), all in a Latin transcription using certain diacritics. Erckert's work has been criticised sharply for its lack of reliability,[25] and there is no information whatsoever on the sources he used; for the present purpose, it may nevertheless be appropriate to quote the following words: *šu* 'night' and 'who' (cf. CA *šow, ha-šow*), *išu* 'man' (CA *išow*), *šel* 'good' (CA *šel*), *muš* 'wind' (CA *mowš*), but

---

[21] Cf. Dirr (1904: V) for a severe critique of the accurateness of Schiefner's material.

[22] In Schiefner's word-list, the verb *cu-ṗsun* is only noted with the meaning 'peel, pluck' ('abschälen, rupfen': p. 89).

[23] *Bixaʒuġon* appearing once on p. 44 is corrected to *Byxaʒuġon* in the corrigenda p. 109.

[24] Schiefner's material was used in Müller (1887: 139-157), where Schiefner's *ṡ, ż, ċ, ç, ǯ, ṣ*, and *ẓ* are represented by *š, ž, tš, tšh, dž, ṣ̌*, and *ž̠*.

[25] Cf. Dirr (1904: III-IV).

also *šum* (besides *ššum*) 'bread' (CA *śum*); *čubux* 'wife' (CA *čibowx*), *viči* 'brother' (CA *viči*), *čoval* 'sparrow' (CA *čobal*); *qaḍč* 'narrow' (CA *q̇ač̣*), but also *baḍč, baċ* 'hundred' (CA *baċ̣*) and *aḍča* 'right (hand)' (CA *haċ̣e-x*); *žže* 'stone' (CA *źe*), but also *žomox* 'mouth' (CA *ǯowmo(w)x*); and *bixadžugh* 'God' (CA **bixaǯuġ*). We see that there is at least a tendency towards denoting sibilants of the 'third' series by gemination (*šš, žž*). To what extent Erckert made use of Schiefner's work remains unclear.

3.1.4 Bežanov 1888 and 1902, Dirr 1904 and 1928

The first Udi materials provided by native speakers were the fairy tale on a shepherd named "Rustam" published by Mixail Bežanov in 1888, and the translation of the four Gospels accomplished by the same author in supporting his brother Semjon, which appeared in 1902. To both these works, printed in Cyrillic letters within the series "Collection of materials for the description of the localities and tribes of the Caucasus",[26] the editor prepended a list of the characters used; the two lists are basically the same, except for the shape of some diacritics. For the sounds under concern here, we find exactly the number of seven relevant letters in the lists as in Schiefner's treatise, viz. ш, ж, ч, ҷ, џ corresponding to Schiefner's *ś, ż, ċ, ̣ċ, ǯ* (and CA *š, ž, č, č̣,* [*ǯ*]), plus ш̆ and ӂ corresponding to Schiefner's *ś* and *ż* (and CA *ś* and *ź*). These correspondences manifest in words like шу 'night' and 'who' (CA *šow, ha-šow*), ишу 'man' (CA *išow*), шел 'good' (CA *šel*), аш 'thing, work' (CA *aš*), беши 'our' (CA *beši*), муш 'wind' (CA *muš*), or шад 'happy' (CA *šad-*), all contrasting with ш̆ум 'bread' (CA *śum*) or лаш̆ко- 'marriage' (CA *laśk̇o-*). For CA *eśa* 'after', the Gospels have ош̆а as expected while the Rustam story provides о̄ш̄а́, with plain ш but the vowels marked as pharyngealised. While ж does not appear at all, we have ӂ in ӂе 'stone' (CA *źe*), ӂикдесун 'stir up' (CA *źiź-*), and ӂомох 'mouth' (but жомо- in Bežanov 1888; CA *ǯowmo(w)x*); CA *žan* 'we' is иан equalling Schiefner's *jan*.

With affricates, things are a bit more complex. First of all, there is a clear correspondence of CA *č* with ч in чалхесун 'know' ~ CA *čalxesown*, чубух 'woman, wife' ~ CA *čibux*, вичи 'brother' ~ CA *viči*, or ич- 'self' ~ CA *ič* (with иш appearing as a variant before certain consonants); in a similar way, ҷ regularly matches CA *č̣* as in каҷ 'narrow' ~ CA *q̇ač̣* or ҷе- 'out-' ~ CA *č̣e-*. For CA *ač̣am* 'unleavened bread', however, we find both the "regular" аҷам and a spelling ӑч̆ам- (Mt. 26.17/Mk. 14.12), with the same diacritic as in ш̆ and ӂ. Indeed, ч̆ appears quite often in the Gospels, although it is missing in the character list; it mostly matches CA *ć* as in ч̆о 'face' (CA *ća*) or ч̆у-ҏ̆сун 'spit' (CA *ćow-pesown*),

---

26   In another article in vol. 14 (1892) of the same journal, M. Bežanov adduces a handful of Udi words denoting meals, measures, etc.; these do not match the material under discussion here.

but also CA *ć* as in *бач̆* 'hundred' (CA *baċ*)[27] or *ача* 'right (hand)' (CA *haće-x*).[28]
ч̆ further appears in *очал* 'earth' (CA *aśal*, cf. Schiefner's *oċal/oċal*) and *кач̆i*
besides *кач̆i*[29] (CA *ḳaċi*). Beyond *ч̆*, the same diacritic is also used with *ц* in the
Gospels, the combination appearing, though not regularly, in the two words
бихаӱуъ[30] (besides more usual *бихацуъ*) 'God' and *конӱух*[31] (besides usual
концух/ъ) 'Lord' (CA *-ʒowġ*). It thus becomes conceivable that the breve-
shaped diacritic was meant to indicate the special articulation of the "third" row
of CA sibilants and affricates, with only the glottalisation of *ć* remaining un-
marked. The description of the pronunciation of *ӂ* and *ш̆* added to the character
lists in the editor's prefaces remains vague though: *ӂ* is "thinner than *ж*, as if *зж*"
and "*ш̆*, a thin *ш*" (1988: 5);[32] both are "sibilants, with the tip of the tongue at the
superior alveolar ridge" (1902: [I]).[33]

The same transcription system as used by M. Bežanov in his article of 1888 is
also applied in A. Dirr's Udi grammar (1904), which appeared in the same jour-
nal. Here, too, we do find *ш̆* and *ӂ* with diacritic marks, but no diacritics on the
corresponding affricates, and again, *ш̆* is declared to be "a thin *ш* (*сшь*)" (p. 2).
Thus, the transcription is the same for many words such as, e.g., *ш̆ум* 'bread',
чубух 'woman, wife', *віч̆i* 'brother', *чалхесун* 'know', *іч̆* 'self', or *ч̆е-* 'out-'; but
there are also some major differences, e.g. in Dirr's *ш̆у* 'night/who' (vs.
Bežanov's *шу* ~ CA *šow*), *іш̆у* 'man' (vs. *ішу* ~ CA *išow*), *аш̆* 'work' (vs. *аш* ~
CA *aš*), *ш̆ел* 'good' (vs. *шел* ~ CA *šel*), and also in *лаш̆ко* 'marriage' (vs. *лаӱко*
~ CA *laśḳo*) and *ӂомо̇-* 'mouth' (vs. *ӂомо-* in the Gospels; CA *ʒowmo(w)x*); in
чо̇ 'face' (vs. *ч̆о* ~ CA *ća*), *очал/очал* 'earth' (vs. *очал*, CA *aśal*), *бач̇* 'hundred'
(vs. *бач̆*, CA *baċ*), but also in *кач̆i* 'blind' (vs. *кач̆i*, CA *ḳaċi*). In *концуъ* 'house-
holder', Dirr has only plain *ц* (vs. the variant *конӱух* in the Gospels, ~ CA
ḳod'inʒ(ow)ġ), and in *'ош̆а́* 'after', he combines the 'pharyngealised' *ош̆а́* of the
Rustam story with the *ош̆а* of the Gospels (CA *eśa*).

Twenty four years later (1928), the same author published a few Udi texts in the
journal *Caucasica*, which was edited by himself. Different from his grammar,
Dirr here uses a Latin transcription, with two different diacritics, a *haček* and an
acute accent, distinguishing the two sets of sibilants and affricates under concern.
The picture thus gained is much more consistent with the data of the Bežanov

---

[27] In Lk. 8.8, we have simple *бач* instead.

[28] In Mt. 20.23 we once have *ача̇* instead.

[29] Only in Mt. 20.30; Lk. 4.18; Jo. 9.1.

[30] Only in Lk. 1.28 and 68 (two out of 143 occurrences in the Gospels).

[31] Only in Mk. 13.35 (one out of 15 occurrences in the Gospels).

[32] "... болѣе тонко, чѣмъ ж, какъ бы зж"; "тонкое ш"; the editor in question was M. Zavadskij.

[33] "... шипящіе; кончикъ языка у верхней альвеолы"; the editor in question ("Z.") was probably
M. Zavadskij, too.

Gospels (and with CA) than that of Dirr's grammar; cf., e.g., *šu* 'night/who', *išu* 'man', or *aš* 'work' contrasting with *śum* 'bread'; *ič* 'self', *viči* 'brother', and *čubux* 'woman, wife'; or *źȯmȯ-* 'mouth' (Gospels жомо-; CA *ӡȯwmo(w)x*). There still remain some discrepancies, however; e.g., we here find *śel* 'good' vs. шел ~ CA *šel*, *laško* 'marriage' vs. лайко ~ CA *laśko-*, *čo* (besides *č'o*)[34] vs. *ŭo* ~ CA *ća*, and *bač* 'hundred' vs. бай (CA *bać*),. The word for 'after' appears as *ośa* (~ CA *eśa*) and *oša* side by side. It should be noted that the four texts in question were not collected by Dirr himself; three of them stemmed from the inheritance of M. Bežanov, written by himself, only the fourth one showing "another hand" (cf. Dirr 1928: 67, 68, 70-72); it is this latter text that provides the divergent *oša*.

3.1.5 Çejrani 1934

A few years after Dirr had edited these texts, the first Udi primer appeared in print. Under the title of "First Lesson" (*samӡi däs*), Th. and M. Ӡeirani published, in Latin script with several extra letters and diacritics, an introduction into writing plus 30 pages of short texts, synoptically arranged in both dialects (Çejrani 1934). For the sounds under concern here, we may quote *şu* 'night' (~ CA *šow*), *aş* 'work' (CA *aš*), *şel* 'good' (CA *šel*), and *şad-* 'happy' (CA *šad-*), vs. *ſum* 'bread' (CA *śowm*); 'after' is *oşa* in both dialects (CA *eśa*). Furtheron we find *ӡħe* 'stone' (CA *źe*, with ħ denoting pharyngealisation of the following vowel); *viꞩi* 'brother' (CA *viči*), *iꞩ* 'self' (CA *ič*), *ꞩoval* 'sparrow' (CA *čobal*), and *ꞩur* 'cow', vs. *ꞇo* 'face' (CA *ꞇ̄a*) and *oꞇal* 'earth' (CA *aśal*); *ce-* 'out' (CA *če-*) ̄vs. *kaꞇħь* 'blind' (CA *kaći*, ь stands for a high central vowel); and *kondʒux* 'householder' (CA **kod'inʒowǵ*) vs. *koç* (N *koz*) 'house' (CA *kod'*). It thus seems that both sets of sibilants and affricates are well represented here, in remarkable consistency with the CA data.

3.1.6 Ӡeiranišvili 1971 and Pančviӡe 1974

In the early 1970ies, two comprehensive grammatical treatises of Udi appeared in Georgian, one (Ӡeiranišvili 1971) with texts and an extensive glossary, the other one (Pančviӡe 1974) amply illustrated with example sentences, including the Nij dialect. In both these works, the Udi materials are presented in a Georgian transcription,[35] and both authors supply a system to differentiate between the two sets of sibilants and affricates under concern here, by denoting the items of the

---

[34]   Three times in Dirr (1928: 69).

[35]   In the Russian summary of Ӡeiranišvili (1971: 271-310), a Latin transcription is applied instead.

"third" series, which they regard as "intensive" or "strong" variants of the "second" one,[36] with an extra mark, thus constrasting ჳ (= *š*) with ჳ*ᶿ*/ჳ̃ (= *ś*) or ხ (= *ž̌*) with ხ*ᶿ*/ხ̃ (=*ž̌*).[37] However, the application of these symbols agrees by far less consistently with the CA data than, e.g., those in Žeirani 1934. Thus we find, in Ž(eiranišvili's) treatise, ჳუ *šu* 'who' (~ CA *ha-šow*), იჳუ *išu* 'man' (CA *išow*), აჳ *aš* 'work' (CA *aš*), and ხაჳ *xaš* 'light' (CA *xaš*) vs. ჳ*ᶿ*უმ *śum* 'bread' (CA *śum*), but also ჳ*ᶿ*უ *śu* 'night' (vs. CA *šow*), and *oša* 'after' (vs. CA *eśa*); P(ančviʒe) in his Nij materials has ჳუმ *šum* 'bread' (vs. CA *śum*) throughout while 'after' (CA *eśa*) is partly ოჳა *oša*, partly ო*ᶿ*ჳ̃ა *oșa* (with pharyngealised *o*). Similarly, we find ჩუბუხ *čubux* 'woman, wife' (~ CA *čibowx*), ვიჩი *viči* 'brother' (CA *viči*), ჩოვალ *čoval* 'sparrow' (CA *čobal*), and იჩ *ič* 'self' (CA *ič*) vs. ჩ*ᶿ*ო/ჩ̃ო *ćo* 'face' (CA *ća*) in both treatises, but also ჩო *čo* for the latter in Ž., ოჩალ *očal* 'earth' in Ž. vs. ოჩ̃ალ *oćal* in P. (CA *aśal*), or ბაჩ*ᶿ* *bać* 'hundred' in Ž. vs. ბაჩ *bač* and -ბაჩ̃ *-baś* in P. (CA *bać*). Ž. further offers ჴ*ᶿ*ე 'stone' (CA * źe*) and ჴ*ᶿ*ომო- 'mouth' (CA *ǯowmo(w)x*), as well as ჴე- *če-* 'out' (CA *če-*) alongside კაჴი *kači* 'blind' (vs. CA *kaći*), and both authors provide კონხუბ *konžux* 'householder' (vs. CA **kod'inǯowǯ*) alongside კოხ *kož̌* 'house' (with the N. variant კოჯ *kož* in P.; CA *kod'*).

3.1.7 Gukasjan 1974

In 1974, V. Gukasjan, a native speaker of Udi, published the first full-fledged dictionary of the Udi language, with both dialects covered and with explanations in both Azeri and Russian. The Udi material is rendered in a Cyrillic-based alphabet, with ъ, I, and an accent-like sign being used as diacritics.[38] For the sounds under concern here, we find ш and шI contrasted in шу 'night; who' (CA *šow*), ишу 'man' (CA *išow*), аш 'work' (CA *aš*), хаш 'moon, light' (CA *xaš*), шел 'good' (CA *šel*), or шад 'open' (CA *šad-*), vs. шIум 'bread' (CA *śum*) or лашIко 'marriage' (CA *laśko*; vs. N. ласкIо); 'after' is оша (vs. CA *eśa*) while the underlying noun 'end' appears as ошI. жI is present in жIеъ 'stone' (CA *źe*), жIиьк-десун 'stir up' (CA *źiź-*), and жIомох 'mouth' (CA *ǯowmo(w)x*), while plain ж appears in the Nij variant кIож of V. кIодж 'house' (CA *kod'*). For both dialects, Gukasjan notes джI in кIонджIух 'landlord' (CA *kod'inǯuǯ*), while both дж and джI occur in V. бихаджугъ and N. бухаджIух 'God' (besides N. бухаджухнутI and -суз 'godless'). ч is found in чибух, чубух (V.) / чугъух (N.) 'woman, wife' (CA *čibowx*), вичи 'brother' (CA *viči*), ич 'self' (CA *ič*), апчи 'false' (CA *aćṗē*), чалхесун 'know' (CA *čalxesown*), or чобал, човал 'sparrow' (CA *čobal*), while

---

[36] The actual terms are *iṭensiuri*/*intensivnyj* (Žeiranišvili 1971: 13-275) and *magari* (Pančviʒe 1974: 29).

[37] The actual diacritics, which are a bit different, cannot be reproduced here exactly.

[38] Apart from the dictionary, V. Gukasjan adopted a similar system in several articles of his, which appeared between 1961 and 1981.

the equivalent of CA *ća* 'face' appears as *чьо*. *чI* is met with in *чIe* 'out' (CA *če-*), *ачIам* 'unleavened' (CA *ačam*), or *къачI* 'narrow'[39] (CA *q̇ač*); in contrast, we have *ҷ̇* in *аҷ̇а* 'right (hand)' (CA *haće-x*), *кIаҷ̇и* 'blind' (CA *ḳaći*), *баҷ̇* 'hundred' (CA *baĉ*), *боҷ̇у* 'dense' (CA *boći-*), but also in *баҷ̇укI* 'crumb(s)' (vs. CA *baćow-biyesown* 'select') and *оҷ̇ал* 'earth' (vs. CA *aśal*). If we leave the inconsistency concerning *дж* and *джI* aside, the correspondence between the CA data and those of Gukasjan's dictionary is thus remarkable.

3.1.8 Ağacani et al. 2011

For the steadily increasing production of Udi materials since the 1990ies, we may take the new translation (in the Nij dialect) of the Gospel of Luke (Ağacani et al. 2011) as an example. In the Latin-based writing system applied here,[40] we see *şu* 'who' (CA *ha-šow*), *üşe* 'night' (CA *šow*), *iş-q'ar* 'mankind' (CA *išow* 'man'), *əṣ̌-* 'work, thing' (CA *aṣ̌*), *xaş* 'light' (CA *xaṣ̌*), and *şad-* 'happy' (CA *šad-*), vs. *šum* 'bread' (CA *śum*) and *laškʼo-* 'marriage' (CA *laśḳo-*); *çuux* 'woman, wife' (CA *čibux*), *viçi* 'brother' (CA *viči*), and *əfçi* 'false' (CA *ačp̄ē*), vs. *ço* 'face' (CA *ća*), but also *baç* 'hundred' (CA *baĉ*) and *oçal* 'earth' (CA *aśal*); *cʼe-* 'out' (CA *če-*) and *qʼacʼ* 'narrow' (CA *q̇ač*) vs. *kačʼi* 'blind' (CA *ḳaći*); and *je* 'stone' (CA *źe*), *jikʼ-desun* 'stir up' (CA *źiź-*) and *jomo-* 'mouth' (CA *ʒowmo(w)x*) vs. *kʼoj* 'house' (i.e. *ḳož*, CA *ḳodʼ*). For 'God', we find three instances of *buxacuǧ-*[41] besides usual *buxačux/ǧ-* (CA **bixaʒowġ*), and for 'Lord', we have three times *qʼoncuǧ-*[42] besides usual *qʼončux/ǧ-* (cf. CA **ḳodʼinʒowġ*). Neglecting the latter variation, we again note a remarkable consistency with the CA notations.

3.2   The nature of the "third" series

Taking the evidence outlined above together, we may state that the differentiation of two series of sibilants and affricates beyond that consisting of "plain" *s, z,* etc. has developed steadily over time in the sources of modern Udi, with the Gospel translation of the Bežanov brothers and the Udi primer of 1934 standing for the most decisive progress; the individual steps in the development are put together in Table I, which includes the *s*-series for easy reference.[43] If we consider that the three series were sharply distinguished in the CA alphabet and that the modern distinction matches that of CA in most aspects (even in cases like Udi *oćal/oĉal*

---

[39]   The derivative *къачIарун* 'narrow(ness)' even appears with the doublet *къачIарун*, which provides the only occurrence of *чI* in the dictionary (Gukasjan 1974: 155-156).

[40]   The writing system is actually based upon the modern orthography of Azeri, in its turn heavily depending on that of Turkish.

[41]   Only in Lk. 7.28; 11.42; 49; 51.

[42]   Only in Lk. 10.2; 11.39; 12.38.

[43]   The Table in Majsak (2008: 456-458) is defective in several points.

'earth' vs. CA *aśal* or Udi *źomox* 'mouth' vs. CA *ʒowmo(w)x*, the series is the same), the question remains why it took so long time until the sound system was rendered adequately. This may be due to the complex phonetic nature of just the "third" series, which is anything but easy to determine.

3.2.1 Attempts at a classification

We have seen that the earlier descriptions of the consonants pertaining to the "third" series were rather vague, with their manner of articulation being determined as being "sharp" (Schiefner 1863), "thin" (Zavadskij ad Bežanov 1888), "intensive" (Ǯeiranišvili 1971), or "strong" (Pančviʒe 1974). The most detailed description[44] is that accompanying the 1902 Gospels, according to which *й* and *жс* are produced "with the tip of the tongue at the superior alveolar ridge". To this we may add the remarks in Gukasjan's dictionary, according to which *жI* and *шI* are "alveolar spirants", while *джI* is a "voiced alveolar-dorsal affricate" in the pronunciation of which the "tongue takes a saddle-like form" (1974: 27-28).[45] In the grammar accompanying the dictionary (ibid.: 256), Gukasjan more explicitly styles *жI* and *шI* "palato-glossal spirants" ("небно-язычные спиранты"), produced with the "tip of the tongue moving backwards, being a bit curved and expanded";[46] in contrast to this, he declares *джI* and *ч́* to be "lamino-palatal, alveolar-dorsal" ("переднеязычный", "альвеолярно-дорсальный") affricates, the first one being produced "with the tongue moving backwards to the middle palate" and the latter, "with the tip of the tongue being curved and ascending towards the alveolar ridge",[47] while *чъ* is a "lamino-palatal alveolar" ("переднеязычный", "альвеолярный") affricate.[48]

In three more recent treatises, W. Schulze (1982: 81-83), T. Majsak (2008: 451-453) and R. Lolua (2010: 35-42) devoted a more extensive discussion to the nature of the sounds under concern. The first-named author concludes that the "middle row of sibilants" has its place of articulation "perhaps between the dental-alveolar and the alveolar region" so that it can be "defined provisionally as pre-

---

[44] The article by D.P. Karbelašvili, "K fonetike udinskogo jazyka", *Jazyk i myšlenie* 3-4, 1935, 259-276 was not accessible to me.

[45] "При произношении альвеолярно-дорсональной (sic!) звонкой аффрикаты **джI** язык принимает седлообразную форму."

[46] "... при произношении шI кончик языка отодвигаетця назад, чуть сгибается у расширяется..."

[47] "... посредством передвижения языка ьазад к среднему небу ... /... кончик языка сгибается и поднимается к альвеоле ..."

[48] In an earlier work devoted to the phonetics of the Nij dialect, the same author had declared джI, ч, чI, and чъ to be "lamino-palatal" ("переднеязычный") affricates, and ч́, an "alveolar" ("альвеоляарный") affricate (Gukasjan 1963: 85).

alveolar".[49] Different from this, Majsak speaks of "velarised" consonants which, however, he marks with the same diacritical *I* as used by him for pharyngealised vowels. Lolua, in challenging both Žeiranišvili's and Majsak's concepts of the sounds being "intensive" or "velarised",[50] comes to the conclusion that pharyngealisation is the main feature of these sounds, an assumption he claims to be supported by his own fieldwork.[51]

3.2.1.1 Pharyngealisation in Udi and CA

Admittedly, Lolua's interpretation seems to match well with the fact that some of the older sources, esp. those published by Schiefner and Dirr, did mark pharyngealisation in many words containing consonants of the "third" series (see above). However, pharyngealisation in Udi has hitherto been regarded as a feature of vowels, not consonants, and it is vowels that are marked as pharyngealised even in those cases, by dots below or above; Gukasjan in his dictionary uses the Russian "hard sign" following the vowel. Different from this, the primer of 1934 introduced a special character to mark pharyngealisation, viz. *ħ*, which was preposed to the vowel in question; cf., from the materials quoted above, *ʒħe* 'stone' (G(ukasjan): *жIэ*ъ, i.e. *ǯẹ*) and *kaeħь* 'blind' (G.: *кIачи*, i.e. *ḳaćí*, without pharyngealisation), but also *ħeŕ* 'apple' (G.: *еъшI*, i.e. *ẹš́*)[52] and *ħek* 'horse' (G.: *еъ*к, i.e. *ẹk*). The latter word proves that the occurrence of *ħ* is not restricted to the environment of consonants of the 'third' series, as do many other words such as, e.g., *ħuq* 'six' (G.: *уъхъ*, i.e. *uq*) or, after another consonant, *vħuɔ* 'seven' (G.: *вуъгъ*, i.e. *vuġ̣*), and there are enough words containing a consonant of the 'third series' that do not show pharyngealised vowels in the primer (or in G.'s dictionary); cf., e.g., *ſum* 'bread' (G.: *шIум*, i.e. *š́um*), *ço* 'face' (G.: *чъо*, i.e. *ćo*), or (N.) *oçal* 'earth' (i.e. *oćalʹ*; G.: *очал*, i.e. *oćalʹ*).

Interestingly enough, the denotation of pharyngealisation in the primer matches that of CA. Here, we have a special character, too, which is preposed to vowels;[53] cf., e.g., *ṗˁa* 'two' ~ *ρħa* (G. *пIаъ*, i.e. *ṗạ*), *vˁa-* 'you (pl.)' ~ *vħa-* (G. *ваъ-*, i.e.

---

[49] "Die 'mittlere sibilantische Reihe'"; "eine Reihe von Phonemen im Udischen, deren Artikulationsstelle offensichtlich zwischen den Dento-alveolaren und Alveolaren angesiedelt ist. Sie sei hier provisorisch als 'präalveolar' definiert" (Schulze 1984: 82-83).

[50] "The question of intensive consonants in the Udi language" ("Inṭensiur tanxmovanta saḳitxi udiur enaši"; Lolua 2010: 35).

[51] "... in most cases where the specialists recognised intensive consonants, we noted pharyngealisation ..." ("... umravles šemtxvevaši, sadac sṗecialisṭṭa mier navaraudevia inṭensiuri, čven davimoçmet paringalizacia"; Lolua 2010: 39-40).

[52] The spelling "Oesch" in Klaproth (1814: 177) might be another attempt to render pharyngealisation.

[53] Except for the u-vowel; for this, we have a special character obviously denoting pharyngealised *u* or *ü* in a digraph with *w* (cf. Gippert et al. 2008: II-11-12).

va̱-), or *kˤaban-* 'desert' ~ (V.) *k̆avh̆an-* 'field' (G. кІаьваьн, i.e. *ḳavan*). There is no peculiar relationship discernible for this character with the consonants of the "third" series; as a matter of fact, it never occurs at all in the neighbourhood of a sibilant or affricate, in none of the three series. The question thus remains what peculiar sort of "pharyngealised" consonants Lolua imagines.

3.2.1.2   The case of Udi *iśa* 'near'

Admittedly again, there is a special case that might be taken to indicate a special relationship between the "third" series and pharyngealisation. This is the case of Udi *iśa* 'near, close' (G.: иьшІа / (N.) ышІа, i.e. *i̱śa/əśa*), the CA counterpart of which is *iˤa* with a plain pharyngeal and no sibilant at all. In a similar way, CA possesses three loanwords that show ˤ instead of an expected sibilant, viz. *mowˤaḵ* 'worker' (~ Georg. *mušaḵ-*, Arm. *mšak*), *vaˤamaḵ* 'cerecloth' (~ Arm. *varšamak*, Georg. *varšamag-*), and *xoˤaḵ* 'heat' (~ Arm. *xoršak*, Georg. *xoršaḵ-*). The latter two examples suggest that we have a special treatment of *-rš- here, which might manifest in the dialect represented by CA in the form of a pharyngeal.[54] On the other hand, Udi possesses quite a lot of ancient loanwords that show *ś* and *ź* instead of *š* and *ž*, cf., e.g., *abreśum* 'silk' (G.: абришІум, армешІум etc.; MPers. *abrēšom*, Georg. *abrešum-*), *niśan* 'sign, betrothal gift' (Lk. 11.29: нишан, N. *nišan*; MPers. Parth. *nīšān*, Georg. *nišan-*, Arm. *nšan*), *xoiś-besun* 'beg' (Mt. 26.53 хоиӣ-б., G. хоишІ-б.; MPers. *xwāhišn* 'desire'), or *źang* 'rust' (G. жІанг, Mt. 6.1 жанг, Arm. *žang*, Georg. *žang-*). Even if in a few such cases, there are sources that mark pharyngealisation as in *peśman* 'repentent' (G. пеьшІман vs. Mt. 21.29 йейман, V. *peśman* vs. N. *p̆eśman* in Çejrani 1934: 51; MPers. Parth. *pašēmān*), it is improbable that these words were taken over with a pharyngeal coarticulation or the like; instead, it is more likely that this coarticulation emerged secondarily within the history of spoken Udi.[55] Instead, the fact that (Middle Iranian, Armenian or Georgian) *š*-sounds were replaced by consonants of the "third" series in early Udi presupposes that this series must have been more similar to the respective series of the yielding language than the "second" one; as none of the languages in question has pharyngealisation, this can be ruled out as the decisive factor.

---

[54]   Cf. Gippert (2009) for a preliminary treatise of these loanwords. – Note that -*rš*- only appears in the Russian loanword *goršečniḳ* 'potter' in Udi sources and, secondarily, for -*rǯ*- in the Nij dialect (e.g. in *boršlu* 'debitor' instead of *borclu*); -*rś*-, -*rž*- and -*rź*- are not attested at all in Udi or CA.

[55]   Note that loanwords like *šad-* 'free, happy' or *šambat* 'Sabbath' show *š*, not *ś*; this may be due to different source languages, different periods of borrowing, or other factors that still have to be investigated. Loanwords from Azeri usually have *š*, not *ś*.

3.2.1.3   Evidence from fieldwork

The recordings undertaken in Oḳtomberi in 2008 (see above) as well as other recordings[56] suggest a different solution for the problem. By contrasting minimal pairs such as the ones discussed above, we arrived at the conviction that the basic difference between the two series of sibilants and affricates consists in the tongue position, quite as described by V. Gukasjan: while the *š*-series is somewhat more palatal, the "third" series is pronounced slightly more towards the velum, with a slight bending of the tongue. This may well be called "velarised"; however, I should prefer to speak of a "retroflex" pronunciation here.[57] Astonishingly enough, the system of Udi thus comes close to that of Sanskrit which distinguishes a palatal and a retroflex *š* (usually denoted by *ś* and *ṣ* in Latin transcription); in Sanskrit, however, this is restricted to the voiceless sibilant.

3.3   Typological outlook

The assumption that the "third" series of sibilants and affricates in Udi is basically retroflex is supported by some more typological evidence from the Indian subcontinent. We have seen that in some cases, a vowel adjoining one of the consonants in question changed from CA to modern Udi (examples are Udi *ćo* 'face' vs. CA *ća*, Udi *ośa* 'after' etc. vs. CA *eśa* etc., and Udi *oćal/oçal* 'earth' vs. CA *aśal*); in all these cases, the vowel was shifted backwards. A similar effect has been witnessed in the history of the Maldivian language (Dhivehi), where the vowel *e* preceding a retroflex was regularly changed to *o*;[58] cf., e.g., the word 'atoll', in itself the only borrowing from Maldivian into European languages, which is *atoḷu* (with retroflex *ḷ*) today but is well attested in the form *ateḷu* in the 12th–15th cc. It is true that Old Maldivian had no retroflex sibilants or affricates; however, modern *ś*, which has developed from retroflex *ṭ*, behaves similarly.[59]

References

Ağacani, R.A., R.A. Danakari, R.B. Mobili & N.R. Rzaeva 2011. *Luk'an exlətbi Mŭq Xavar*. Bakı.

Bežanov, M. 1888. Rustam (udinskaja skazka), Sbornik" materialov" dlja opisanïja městnostej i plemen" Kavkaza 6, pril. 5-28.

Bežanov, M. 1892. Kratkïja svěděnïja o s. Vartašeně i ego žiteljax", *Sbornik" materialov" dlja opisanïja městnostej i plemen" Kavkaza* 14 (1): 213-262.

---

[56]   ECLinG (see above), cf. https://hdl.handle.net/1839/00-0000-0000-0008-26CE-8@view .

[57]   The case of *xaś-heǵesown* 'be baptised' vs. *xaš* 'light' (see 2.2.1 above) may well be explained by assuming a shift backwards introduced by the following laryngeal, *h*.

[58]   Cf. Fritz (2002: 27-28) as to examples.

[59]   Cf. Gippert (2013: 88-89).

Bežanov, S. & M. 1902. Gospoda našego Ïisusa Xrista svjatoe evangelïe = *Sbornik" materialov" dlja opisanïja městnostej i plemen" Kavkaza* 30.

Çejrani, Ţ. & Mixak 1934. *Samçi dəs*. Suxum: Abgiz.

Dirr, A. 1904. Grammatika udinskago jazyka = Sbornik" materialov" dlja opisanïja městnostej i plemen" Kavkaza 33 (6).

Dirr, A. 1928. Udische Texte. *Caucasica* 5: 60-72.

Durkin-Meisterernst, D. 2004. *A Dictionary of Manichaean Middle Persian and Parthian*. Turnhout: Brepols.

Erckert, R. von 1895. *Die Sprachen des Kaukasischen Stammes*. Vienna: Hölder.

Fritz, S. 2002. *The Dhivehi Language. A descriptive and historical grammar of Maldivian and its dialects*. Heidelberg: Ergon.

Klaproth, J. von 1814. *Beschreibung der russischen Provinzen zwischen dem Kaspischen und schwarzen Meere*. Berlin: Maurer.

Gippert, J. 2009. An etymological trifle. In: Sundermann, W. et al. (eds.), *Exegisti monumenta. Festschrift in honour of Nicholas Sims-Williams*. Wiesbaden: Harrassowitz, pp. 127-140.

Gippert, J. 2013. An outline of the history of Maldivian writing. In: Chen, S.-F. & B. Slade (eds.), *Grammatica et verba. Glamor and verve. Studies in South Asian, historical, and Indo-European linguistics*. Ann Arbor, MI & New York: Beech Stave Press, pp. 81-98.

Gippert, J., W. Schulze, Z. Aleksidze & J.-P. Mahé 2008. *The Caucasian Albanian Palimpsests of Mount Sinai*. 2 vols. Turnhout: Brepols.

Gukasjan, V. 1963. Nidžskij dialekt udinskogo jazyka. *Izvestija AN Azerbajdžanskoj SSR, serija obščestvennyx nauk*, 1963 (3): 79-90.

Gukasjan, V. 1974. *Udinsko-azerbajdžansko-russkij slovar'*. Baku: Elm.

Lolua, R. 2010. Udiuri ena (gramaţiḳuli analizi, ţeksţebi, leksiḳoni). Tbilisi.

Majsak, T.A. 2008. Varianty udinskoj orfografii i transkripcii (kratkij obzor). In: *Udinskij sbornik. Grammatika, leksika, istorija jazyka*. Moscow: Academia, 443-460.

Mežov, V.I. 1894. *Bibliografija Azïï* / Mějov, V., *Bibliographica Asiatica*, III. S.-Saint Petersburg: Bezobrazov.

Müller, F. 1887. *Grundriss der Sprachwissenschaft. III. Bd.: Die Sprachen der lockenhaarigen Rassen. II. Abtheilung: Die Sprachen der mittelländischen Rasse*. Vienna: Hölder.

Pančviʒe, V. 1974. *Uduri enis gramaţiḳuli analizi*. Tbilisi: Mecniereba.

Schiefner, A. 1854. Etwas über die Udiner, ein Volk des Caucasus, (Erman's) *Archiv für wissenschaftliche Kunde von Russland* 13: 649-652.

Schiefner, A. 1856. *Versuch über die Thusch-Sprache oder die kistische Mundart in Thuschetien*. Saint Petersburg: Akademie der Wissenschaften.

Schiefner, A. 1863. *Versuch über die Sprache der Uden.* Saint Petersburg: Akademie der Wissenschaften.

Schulze, W. 1982. Die Sprache der Uden in Nord-Azerbaidžan. Wiesbaden: Harrowitz.

Šopen, I. 1866. *Novyja zamětki na drevnija istorïi Kavkaza i ego obitatelej*. Saint Petersburg: Tiblen.

Ʒeiranišvili, Egnaṭe 1971. *Udiuri ena*. Tbilisi: Tbilisis Universiṭeṭi.

Table 1: Sibilants and affricates in Caucasian Albanian and in sources of modern Udi

CA	s	z	ʒ	c	ç	š	ž	ǯ	č	č̣	ś	ź	ʒ́	ć	ç̣
Klpr.1814	ss	s	—	z	—	—	—	—	(sch), tsch	—	sch	—	—	—	—
Sch. 1854	s	s	—	—	z	sch	—	—	tsch	(tsch)	(sch)	(j)	(ds)	—	(tsch)
Šop. 1866	c	з	—	ц	ц	ш	ж	дж	ч	—	(ш, с)	—	(дз)	(ц)	—
Starč. 1891	c	з	—	ц	(ц)	ш	—	дж	ч	(ч)	(ш, с)	(ж)	(дз)	(ц)	(ц)
Sch. 1863	s	z	ʒ	c	c̣	ṡ	ż	ʒ̇	ċ	ċ	ṡ	ż	(ʒ, ʒ̇)	(c, ċ)	(c, c̣)
Erck. 1895	s	z	—	c	tc	š	ž	dž	č	dč	(š), šš	(ž), žž	(dž)	(č)	(dč, c)
Bež. 1888	c	з	дз	ц	ц’	ш	ж	џ	ч	ч’	Ĭ	Ж	(џ)	(ч)	(ч’)
Bež. 1902	c	з	ʒ	ц	ц’	ш	ж	џ	ч	ч’	Ĭ	Ж	ŭ	ў	(ч’)
Dirr 1904	c	з	ʒ	ц	ц’	ш	ж	џ	ч	ч’	Ĭ	(ж)	(џ)	(ч)	(ч’)
Dirr 1928	s	z	—	c	c’	š	ž	ǰ	č	č’	ś	ź	(ǰ)	(č)	(č’)
Çeir. 1934	s	z	dz	Ꜣ	⅂	ş	ẓ	ҫ	ц	c	ʃ	ʒ	dʒ	c̣	c̣
Ʒeir. 1971	ь	ҍ	ә	в	ҕ	ә	ꙃ	χ	β	ʄ	ð ᵊ	ꙃ ᵊ	χ ᵊ	β ᵊ	ʄ ᵊ
Pan. 1974	ь	ҍ	ә	в	ҕ	ö	ꙃ	χ	β	ʄ	ð˜	ꙃ˜	χ˜	β˜	ʄ˜
Guk. 1974	c	з	дз	ц	цI/ц’	ш	ж	дж	ч	чI	шI	жI	джI	чъ	ч’
Ağac. 2011	s	z		ś	s'	ş	j	c	ç	c’	š	ǰ	č	ç	č'