Caucasian Languages

[B.] George Hewitt FBA, Professor of Caucasian Languages, NME Dept., SOAS,

Thornhaugh Street, Russell Square, London WC1H 0XG, UK.

Around 38 languages are deemed to be indigenous to the Caucasus; often difficult demarcation between language and dialect explains the uncertainty. The ancestral homelands are currently divided between:

Russia's north Caucasian provinces (Circassian, Abaza, Ingush, Chechen, Avaro-Ando-Tsezic, Lako-Dargic, northern Lezgic)

de facto independent Abkhazia (Abkhaz, Mingrelian, Svan, Georgian, Laz)

Georgia (Georgian, Mingrelian, Svan, Laz, Bats, Chechen, Avar, Udi)

Azerbaijan (Lezgi, Budukh, Kryts', Khinalugh, Rutul, Ts'akhur, Avar, Udi)

Turkey (Laz, Georgian)

Diaspora-communities of north (especially north-west) Caucasians can be found across former Ottoman territories, particularly Turkey, where the majority Circassian and Abkhazian populations reside and where the term 'Cherkess' often indiscriminately applies to any north Caucasian. Circassians are found in Syria, Israel and Jordan, home also to a significant Chechen population. Speaker-numbers range from 500 (Hinukh) to 3-4 million (Georgian). Many of the languages are endangered.

Three families are usually recognised:

A. South Caucasian (Kartvelian)

Georgian

Svan

Mingrelian (Megrelian)

Laz (Ch'an)

[Scholars in Georgia regard Mingrelian and Laz as co-dialects of Zan]

B. North West Caucasian

Abkhaz

Abaza

Ubykh (extinct from 1992)

West Circassian (Adyghe)

East Circassian (Kabardian)

<u>C. Nakh-Daghestanian</u>

(a) Nakh (North Central Caucasian)

Chechen

Ingush

Bats (Ts'ova Tush)

(b) Daghestanian (North East Caucasian)

1. Avaro-Ando-Tsezic(/Didoic):

Avaric: **Avar**

Andic: Andi, Botlikh, Godoberi, K'arat'a, Akhvakh, Bagvalal, T'indi, Ch'amalal

Tsezic: Tsez (Dido), Khvarshi, Hinukh, Bezht'a (K'ap'uch'a), Hunzib (these last two are sometimes regarded as co-dialects)

2. Lako-Dargic:

Lakic: **Lak**

Dargic: **Dargwa** (Dargi(n)) — some treat K'ubachi, Chiragh, and Megeb as full languages

3. Lezgic:

Lezgi(an), **Tabasaran**, Rutul (Mukhad), Ts'akhur, Aghul, Udi, Archi, Budukh, Khinalugh, Kryts'

Some challenge the Lezgic status of Archi, Khinalugh, Budukh and Kryts'. Mutual intelligibility basically exists between Laz and Mingrelian, Abkhaz and Abaza, West and East Circassian. Only Georgian has an ancient tradition of writing, but during the Soviet period the languages in bold all enjoyed literary status. Publishing in Mingrelian, Laz, Ts'akhur, Aghul, Rutul and Udi was tried in the 1930s but discontinued, though there have been some post-Soviet attempts to publish more widely (including Dido).

Phonetics and Phonology

All Caucasian languages have voiced vs voiceless aspirate vs voiceless ejective plosives, affricates and occasionally fricatives, to which some add a *fortis* series (voiceless unaspirated or geminate). N.W. Caucasian is characterised by large consonantal inventories coupled with minimal vowel-systems, consisting of at least the vertical opposition open /ɑ/ vs close /ə/. Ubykh possessed 80 phonemes (83 if the plain velar plosives attested only in loans are admitted), with every point of articulation between lips and larynx utilised and displaying the secondary features of palatalisation, labialisation, and pharyngalisation — Daghestanian pharyngalisation is normally assigned to vowels:

Consonantal Phonemes for Ubykh

p	b	p'			m	w
pˤ	bˤ	pˤ'			mˤ	wˤ
		f				
		vˤ				
t	d	t'			n	r
tʷ	dʷ	tʷ'				
t͡s	d͡z	t͡s'	s	z		
t͡ɕ	d͡ʑ	t͡ɕ'	ɕ	ʑ		
t͡ɕʷ	d͡ʑʷ	t͡ɕʷ'	ɕʷ	ʑʷ		
t͡ʃ	d͡ʒ	t͡ʃ'	ʃ	ʒ		
			ʃʷ	ʒʷ		
t͡ʂ	d͡ʐ	t͡ʂ'	ʂ	ʐ		
		ɬ			ɬ'	l
					j	
(k)	(g)	(k')	x	ɣ		
kʲ	gʲ	kʲ'				
kʷ	gʷ	kʷ'				
q		q'	χ	ʁ		
qˤ		qˤ'	χˤ	ʁˤ		

q^j	$q^{j\text{'}}$	χ^j	$ʁ^j$
q^w	$q^{w\text{'}}$	χ^w	$ʁ^w$
$q^{\Upsilon w}$	$q^{\Upsilon w\text{'}}$	$\chi^{\Upsilon w}$	$ʁ^{\Upsilon w}$
	h		

Some recent analyses of Daghestanian languages have produced inventories rivalling those of the N.W. Caucasian, though no parallel minimality amongst the vowels is posited. One analysis of Archi assigns it 70 consonants:

Consonantal System of Archi

p	b	p'	pː					m		w
t	d	t'	tː					n	r	
t^w	d^w									
ts		ts'		tsː'	s	sː	z			
ts^w		$ts^{w\text{'}}$			s^w	$sː^w$	z^w			
tʃ		tʃ'		tʃː'	ʃ	ʃː	ʒ			
$tʃ^w$		$tʃ^{w\text{'}}$			$ʃ^w$	$ʃː^w$	$ʒ^w$			
tɬ		tɬ'			ɬ	ɬː	ɮ		l	
$tɬ^w$		$tɬ^{w\text{'}}$			$ɬ^w$	$ɬː^w$				
										j
k	g	k'	kː							
k^w	g^w	$k^{w\text{'}}$	$kː^w$							
q		q'		qː'	χ	χː	ʁ			
q^w		$q^{w\text{'}}$			$χ^w$	$χː^w$	$ʁ^w$			
					ħ		ʕ			
		ʔ			h					

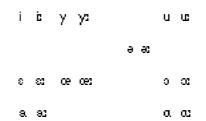Noticeable here is the presence of 10 laterals, though some specialists recognise no more than 3/4.

Kartvelian occupies a mid-position with between 28 and 30 consonants (see *Georgian*). Georgian shares with Avar and Andi the simple 5-vowel triangle:

Georgian-Avar-Andi Vowel-system
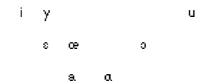
i　　　　　　　　　u

　　ɛ　　　　　ɔ

　　　ɑ

Schwa is added to this in the other Kartvelian languages, whilst the various Svan dialects have length and/or umlaut, Upper Bal having the richest system:
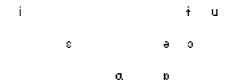
Svan's Upper Bal Vowel-system

i iː y yː　　　　u uː

　　　　　ə əː

ɛ ɛː œ œː　　　ɔ ɔː

æ æː　　　　　ɑ ɑː

Triangular or quadrilateral vowel-systems are attested in Nakh-Daghestanian:

Bezht'a Basic Vowel-system

i　y　　　　　　　u

　ɛ　œ　　　ɔ

　　ə　　ɑ

[All but /y, ɛ, œ/ possess long counterparts, and the nasalised vowels: /ĩ, ɛ̃, ã, ɔ̃, ũ, ũ, ə̃, ỹ/ have also been recognised]

Hunzib Basic Vowel-system

i　　　　　　　ɨ　u

　ɛ　　　　ə　ɔ

　　ɑ　　ɒ

[All these Hunzib vowels have long counterparts, and fluctuating nasalisation on short vowels has been observed]

The simplest (near-)quadrilateral system is attested in Chiragh Dargwa, with four pairs distinguished by length:

Chiragh Dargwa Vowel-system

i(ː)　　　u(ː)

ɛ(ː)

　　ɑ(ː)

Udi has been analysed with:

Udi Vowel-system

i iˤ (y)		u ʊˤ	
ɛ ɛˤ (œ)	ə	ɔ ɔˤ	
(ɐ)		ɑ ɑˤ	

whilst Chechen presents the complicated system:

Chechen Vowel-system

i iː y yː		u uː	
je ie ɥœ yœ		wo uo	
e eː		o oː	
æ æː		ɑ ɑː	

[Most, if not all, of these can be nasalised as a result of the weakening of a following /n/]

Stress is sometimes distinctive (Abkhaz-Abaza) but usually not. Tonal distinctions have been proposed for some of the Daghestanian languages (Andi, Akhvakh, Ch'amalal, Khvarshi, Hinukh, Bezht'a, Tabasaran, Ts'akhur, Ingush and Budukh).

Morphology

N.W. Caucasian sememes are typically C(C)(V), and minimal case-systems combine with highly polysynthetic verbs, which may contain: up to four agreement-prefixes, locational preverbs, orientational preverbs and/or suffixes, interrogative and conjunctional elements, and markers of tense-modality, (non-)finiteness, causation, potentiality, involuntariness, polarity, reflexivity, and reciprocality (see *Abkhaz*). Kartvelian balances a moderate total of cases with reasonably complex verbs, which may contain: agreement with two or three (rarely four) arguments via two sets of agreement-affixes, directional/perfectivising preverbs (the large total in Mingrelian-Laz suggests N.W. Caucasian influence), and markers of tense-aspect-modality, causation, potentiality, version (vocalic prefixes indicating certain relations between arguments), and voice — Kartvelian is the only family to have a full active-passive diathetic opposition. Nakh-Daghestanian has complex nominal systems with both grammatical and sometimes large numbers of locative cases; Lezgi(an), Aghul and Udi apart, nouns fall into one of between two and (depending on the analysis) five or eight (largely covert) classes. Verbs are correspondingly simple: agreement is totally absent from Lezgi(an) and Aghul; elsewhere, verbs with an agreement-slot typically

allow only class-agreement (Andic), though some languages (Bats, Lak-Dargwa, Tabasaran, Akhvakh, Archi, Hunzib, Avar dialects) have added perhaps rudimentary person-agreement, whilst Udi has person-agreement only. Some languages have a small selection of preverbs. Some distinguish perfective from imperfective roots. Some North Caucasian verbs can be construed transitively or intransitively (?passively), depending on the clausal structure; anti-passives are also attested.

Avar illustrates a typical system of locative-cases:

Avar Locative Case-endings

Series	Essive	Allative	Ablative
1. 'on'	-d(.)a	-d.e	-d(.)a.sa
2. 'near'	-q	-q.e	-q.a
3. 'under'	-ƛ̌:	-ƛ̌:'e	-ƛ̌:'a
4. 'in (mass)'	-ƛ̌:	-ƛ̌:e	-ƛ̌:a
5. 'in (space)'	-D (= class-marker)	-D-e	-sa

Ergativity and some other oblique case-function are often merged in a single morph.

Deictic systems range from two-term (Mingrelian, Ubykh, Kryts'), through three-term (Georgian, Abkhaz, Circassian), to five-term in a swathe of Daghestanian, and even six-term (Lezgi(an), Godoberi).

Counting-systems are predominantly vigesimal, at least up to '99' (though Bats is vigesimal throughout), but some systems are decimal.

Syntax

Word-orders are: Kartvelian and Nakh-Daghestanian AN, GN, N-Postposition, SOV, though Old Georgian was rather NA and NG; N.W. Caucasian GN, predominantly NA, N-Postposition, SOV. Some degree of ergativity characterises all the languages but in Mingrelian, where the system was orginally as illustrated for Georgian (*q.v.*), the Ergative case-marker was extended vertically to replace the original Nominative for intransitive (including Indirect) verbs in Series II (Aorist Indicative and Subjunctive), where it functions as a Series II Nominative allomorph, the original

Nominative effectively becoming an Accusative just for Series II — Laz has extended the case-marker horizontally across its three Series for all transitive subjects. Active-Inactive alignment plays a role in some languages (Bats).

A Nominative/Absolutive argument is the obligatory minimum in a clause, and where verbs have class-agreeement, this is the determiner for the class-marker (which in some languages also appears on adverbs and as part of a locative case-exponent); the determiner for person-agreement in languages with class-agreement might be this same or a different argument (e.g. the logical subject), depending on a variety of factors.

Verbs such as 'want', 'have', 'hear' are construed indirectly with the logical subject in an oblique case, but, if Kartvelian and N.W. Caucasian employ just the Dative/general Oblique case for this argument, greater distinctions can apply in Nakh-Daghestanian: Avar employs its Dative case with verbs of emotion ('love'), a locative (Series I Essive) with verbs of perception ('see'), and the Genitive for the possessor in conjunction with the copula.

Only Kartvelian has the category of subordinating conjunctions, naturally associated with full clauses containing indicative or subjunctive finite verbs. Such structures are rare in North Caucasian, where one finds a variety of non-finite (nominalised) verb-forms fulfilling the subordinate role.

Examples:

ilu–di   rik̂ʒ'i   b–ɵʒ–o   vs   rik̂ʒ'i b–ɵʒ–o

mother-Erg   meat.Absol3   3-fry-Past

'Mother fried the meat' vs 'The meat (was) fried'   [Andi]

is–t̂'   ɕi   ʁɑrbɔl–t͡ʃʰɑ

brother-Erg   water.Absol   boil-Pres

'Brother is boiling the water'   [Bezht'a]

vs

is   ɕi–d   ʁɑrbɔl–dɑ–t͡ʃʰ

brother.Absol   water-Instr   boil-AntiPass-Pres

'Brother is regularly engaged in boiling water'   [Bezht'a]

k'ɔtʃ-k   dʒɐʁab-i   kɔ-ø-dʒir-ʊ   vs   dʒɐʁab-k   dɔ-ʁur-u

man-Nom$_A$ girl-Acc$_B$ Prev-her$_B$-see-he.Aor$_A$ girl-Nom$_A$ Prev-die-she.Aor$_A$

'The man saw the girl' vs 'The girl died'   [Mingrelian]

k'ɔtʃ-s   dʒɐʁab-i   ø-a-dʒir-ɐ

man-Dat$_B$ girl-Nom$_A$ he$_B$-Pot-see-her.Pres$_A$

'The man can see the girl'   [Mingrelian]

vs

k'ɔtʃ-s   dʒɐʁab-k   k-ø-a-dʒir-u

man-Dat$_B$ girl-Nom$_A$ Prev-he$_B$-Pot-see-her.Aor$_A$

'The man could see the girl'   [Mingrelian]

insa-u-jɐ   j.as   j-ɔkʼˤ'-u-la

father-Obl-Dat   daughter$_2$.Absol   2-love-TV-Pres

'Father loves (his) daughter'   [Avar]

insa-u-d.a   w.as-ul   r-iʁa-u-la

father-Obl-LocI   son-Pl.Absol   Pl-see-TV-Pres

'Father sees (his) sons'   [Avar]

insa-u-l   tʃι   b-ugɔ

father-Obl-Gen   horse$_3$.Absol   3-be.Pres

'Father has a horse'   [Avar]

lamʃgɐd-χɐn-iʃ   bikw-d   ɐgɐ   la-ø-j-k'wiʃ-ø.   ɐrɐ

shade-from-Gen   wind-Erg$_A$ Prev   Prev-it$_B$-SV-admit-it.Aor$_A$ that

mirɐ   uʃχwar   rɐnɐgɔ   χ.ɔ-l.qmaʃ-a   miʒ

their   each.other.Dat   between   CompPref-strong-CompSuff   sun.Nom$_A$

lə.m.ɐr-ø

apparently.be-it$_A$

'The north wind admitted that the sun was apparently the stronger of them'

                              [Lower Bal Svan]

tɐʁa-ʒɐ-m   tɐʁa-r   jaʒ   naɦ.rɐ.j   naɸ

sun-wind-the.Erg/Obl$_{III}$ sun-the.Absol$_I$ self   much   more

ø-zɐ.rɐ-ɬ aʂɐ-r   ø-qɐ-gˠɐ.rɐ-ø-mɐ-ʔˠa-mɐ

it$_I$-how-strong-Absol.N/F.Stat.Pres$_I$ it$_I$-Prev-Prev-it$_{III}$-not-admit.N/F-if$_I$

ø-mə-χ^wə-n-aɥ   ø-χ^wə-ʁɒ

itʃ-not-happen-Fut-Absʃ   itʃ-happen-Aor.Fin

'It became impossible for the north wind not to admit how/that the sun is stronger

than it'   [Temirgoi West Circassian]

Kinship

Kartvelian is unrelated to any known language(-family), but the debate continues concerning the relationship between the northern families. Linkage to Hattic is postulated for N.W. Caucasian and to Hurrian for Nakh-Daghestanian. Udi has recently been conclusively demonstrated to descend from Caucasian Albanian.

Bibliography

Berg, H. van den. (1995). *A Grammar of Hunzib (with Texts and Lexicon).* Lincom Studies in Caucasian Linguistics 01. München/Newcastle: Lincom Europa.

Berg, H. van den. (2001). *Dargi folktales. Oral stories from the Caucasus and an introduction to Dargi grammar.* Leiden: Research School CNWS.

Berg, H. van den. (2004). The East Caucasian language family, in a special edition of *Lingua*, 147-190.

Boeder, W. (1979). Ergative syntax and morphology in language change: the South Caucasian languages, in F. Plank (ed.) *Ergativity*, 435-480. New York: Academic Press.

Boeder, W. (2004). South Caucasian, in a special edition of *Lingua*, 5-89.

Catford, J. C. (1976). Ergativity in Caucasian languages, in *Papers from the 6th Meeting of the Northeast Linguistics Society*, Montreal, 37-48.

Catford, J. C. (1977). Mountain of Tongues: the Languages of the Caucasus, in *Annual Review of Anthropology 6*, 283-314.

Charachidzé, G. (1981). *Grammaire de la Langue Avar*. Paris: Editions Jean-Favard.

Chirikba, V. A. (1996). *Common West Caucasian. The Reconstruction of its Phonological System and Parts of its Lexicon and Morphology*. Leiden: CNWS.

Deeters, G. (1930). *Das Kharthwelische Verbum*. Leipzig: Kommissionsverlag von Markert und Petters.

Dumézil, G. (1975). *Le Verbe Oubykh. Etudes descriptives et comparatives (avec la collaboration de Tevfik Esenç)*. Paris: Klincksieck.

Dumézil, G. (1976). Compléments et corrections au *Verbe Oubykh* (1975) *I*, in *Bedi Kartlisa (revue de kartvélologie) XXXIV*, 9-15.

Greppin, J. (ed). (1989-2004). *The Indigenous Languages of the Caucasus*. *1: Kartvelian Languages* (edited by A. C. Harris), 1991; *2: North West Caucasus* (edited by B. G. Hewitt), 1989; *3: North East Caucasian Languages, Part 1* (edited by D. M. Job), 2004; 4: *North East Caucasian Languages, Part 2* (edited by R. Smeets), 1994. Delmar, New York: Caravan Books.

Harris, A. C. (2002). *Endoclitics and the Origins of Udi Morphosyntax*. Oxford: OUP.

Haspelmath, M. (1993). *A Grammar of Lezgian*. Berlin/New York: Mouton de Gruyter.

Hewitt, B. G. (1987). *The Typology of Subordination in Georgian and Abkhaz*. Berlin: Mouton De Gruyter.

Hewitt, B. G. (ed.) (1992). *Caucasian Perspectives*. München-Newcastle: Lincom Europa

Hewitt, B. G. (1998). Caucasian Languages, in G. Price (ed.) *Encyclopaedia of the Languages of Europe*, 57-81. [Paperback version 2000]

Hewitt, B. G. 2004. North West Caucasian, in a special Caucasian edition of *Lingua*, 91-145.

Hewitt, B. G. (2004). *Introduction to the Study of the Caucasian Languages.* München-Newcastle: Lincom Europa.

Kibrik, A.E., Kodzasov, S.V. 1988. *Sopostavitel'noe izuchenie dagestanskix jazykov. Glagol* [Contrastive Study of the Daghestanian Languages. The Verb]. Moscow: University Press.

Kibrik, A.E., Kodzasov, S.V. 1990. *Sopostavitel'noe izuchenie dagestanskix jazykov. Imja. Fonetika* [Contrastive Study of the Daghestanian Languages. The Noun. Phonetics]. Moscow: University Press.

Klimov, G. A. (1994). *Einführung in die kaukasische Sprachwissenschaft*, translated and expanded by Jost Gippert. Hamburg: Buske.

Klimov, G. A. (1998). *Trends in Linguistics. Documentation 16. Etymological Dictionary of the Kartvelian Languages*. Berlin-New York: Mouton De Gruyter.

Kuipers, A. H. (1960). *Phoneme and Morpheme in Kabardian*. 'S-Gravenhage: Mouton.

Nichols, J. (1997). Chechen phonology, in A.S. Kaye (ed.) *Phonologies of Asia and Africa (including the Caucasus), 2*, 941-971. Winona Lake: Eisenbrauns.

Nikolayev, S. L. and Starostin, S. A. (1994). *A North Caucasian Etymological Dictionary*. Moscow: Asterisk.

Paris, C. (1974). *Système Phonologique et Phénomènes Phonétiques dans le parler besney de Zennun Köyü (Tcherkesse oriental)*. Paris: Klincksieck.

Smeets, R. (1984). *Studies in West Circassian Phonology and Morphology.* Leiden: Hakuchi Press.

Tuite, K. (1998). *Kartvelian Morphosyntax: Number Agreement and Morphosyntactic Orientation in South Caucasian Languages.* Studies in Caucasian Linguistics 12. München/Newcastle: Lincom Europa.

Abbreviations

Absol	Absolutive
Acc	Accusative
AntiPass	Anti-passive
Aor	Aorist
Caus	Causative
Comp	Comparative
Dat	Dative
Dyn	Dynamic
Erg	Ergative
Fin	Finite
Fut	Future
Gen	Genitive
Instr	Instrumental
Loc	Locative
N/F	Non-finite
Nom	Nominative
Obl	Oblique
P/I	Past Indefinite
Perf	Perfect
Pl	Plural
Pot	Potential
Pref	Prefix
Pres	Present
Prev	Preverb
Res	Resultative
Stat	Stative
Suff	Suffix
SV	Subjective Version
TS	Thematic Suffix
TV	Thematic Vowel