INTRODUCTION

The Great Caucasus mountain range strides for 500 miles across the isthmus from the Black Sea to the Caspian, forming a massive barrier between Europe and Asia. To the ancient Greeks, these mountains and the lands that lay below and around them were a place of mystery and legend. They dominated the farthest recess of the Euxine Sea; to Greek mariners the voyage to the Colchidian city of Phasis (the modern Poti) and up the like-named river (now known by the Svan name of "Rioni") was the "uttermost run." Thither the Argonauts went in quest of the Golden Fleece, and in the mountains behind Phasis, and behind the Greek city of Dioscurias (the modern Sukhumi) Prometheus was chained, and there, too, in the words of Herodotus, there dwelt "many and all manner of nations."

Again and again in the two and a half millenia since Herodotus's day, writers have commented on the ethnic and linguistic diversity of the Caucasus. Strabo, writing about four and a half centuries later, having discounted more exaggerated estimates, affirms that 70 tribes, all speaking different languages, would come down to trade in Dioscurias, and a few decades after Strabo, Pliny claimed that the Romans carried on business in the same city by means of 130 interpreters. Arab travelers in the middle ages bore continuing witness to Caucasian polyglossia, and it was one of them, the tenth century geographer al-Mas'udi, who named the Caucasus jabal al-alsun, "mountain of tongues."

At the present day we count more than 50 languages in the Caucasian mountain valleys, in the foothills, and in the closely adjoining plains. Some of these languages belong to well-known language families. Thus, Indo-European is represented by Armenian and by the Iranian languages Ossetic, Kurdish, Tat, and Talysh, not to mention Russian, the principal lingua franca of the area, and small enclaves of Greeks and speakers of other IE languages. The Turkic languages are Azeri (Azerbaijani), Turkoman, Karachay-Balkar, Nogai, and Kumyk. Mongolic is represented by Kalmyk, and Semitic by "Neo-Assyrian" or Aisor.
All these languages belong to families which have their most characteristic habitat elsewhere and the bearers of which are known to have migrated into the area at various historically attested times. There is, however, a residue of 37 languages which were not imported into the area in historical times and which are believed to have been spoken in the Caucasus area for at least 4000 years. These “indigenous” or “autochthonous” languages comprise the group generally called Caucasian or Caucasian in Western European publications, and most commonly “Ibero-Caucasian” (iberijsko-kavkazskie) in modern Russian sources. It is important to note that the prefix *iber* - in this Russian term has absolutely no reference to the Iberian peninsula at the opposite end of Europe: it refers strictly to Caucasian Iberia, an ancient transcaucasian state in part of what is now Georgia. Some western scholars, it is true, have used the term Ibero-Caucasian to refer quite explicitly to a presumed relationship between Caucasian and Basque, as did Holmer (32), but there is no such presumption behind the Russian term.

There are two major groups of Caucasian languages, Northern and Southern. The North Caucasian languages fall into three groups: Abkhazo-Adyghyan or North West Caucasian (NWC), Nakh or North Central Caucasian (NCC) and Dagestanian or North East Caucasian (NEC). The South Caucasian languages (SC), also known by the Georgian-derived name Kartvelian, form a single major group. Within each of these groups there are subgroups of closely related languages, as well as a few isolated languages.

In the following list the languages are placed within their appropriate groups and subgroups, but are numbered consecutively from 1 to 37. After the name of each language, or in some cases after the name of a whole subgroup, I give a brief indication of where the language is spoken and the number of its speakers in the Soviet Caucasus. These figures are chiefly taken from the *Encyclopaedia Britannica* (25), which quotes the 1970 Soviet Census, supplemented by reference to (40). The Ubykh language is now spoken by only a handful of people at Manyas, Turkey, south of the Sea of Marmara. According to Vogt (87), in 1962 there were no more than 20 speakers, all rather old. I, perhaps optimistically, have guessed the current number to be “about 10.”

A few Caucasian languages have fairly well recognized English names, and in most cases I use these. For the rest I use slightly anglicized Russian forms—which, incidentally, often differ very widely from the native name. Two of the Tsez languages, Hinukh and Hunzib, are often literally transliterated from Russian as Ginukh and Gunzib: I spell them with “H” since this is one of those cases where “G” is used in Russian to represent an [h] sound.

*North West Caucasian (Abkhazo-Adyghyan)*


3. *Ubykh* (near Manyas, Turkey, about 10).

Adyghyan or Circassian: 4. *Adygye* (formerly in most of the region between the Kuban River and the Black Sea; now chiefly in Adygej AO and a few villages near Tuapse and in the Karachay-Cherkess AO—100,000), 5. *Kabardian* (chiefly in Kabardino-Balkar ASSR and in Karachay-Cherkess AO—320,000).
North Central Caucasian (Nakh)


North East Caucasian (Dagestanian)


Lak-Dargi: 23. Lak (Dagestan highlands—86,000), 24. Dargi (E. of Lak—231,000).

Lezgian group


South Caucasian (Kartvelian)

35. Georgian (Georgian SSR—2,792,000), 36. Svan (in the mountains of N.W. Georgia—43,000), 37. Zan with two major dialects: Megrelian (W. Georgia and E. Abkhazia—360,000) and Laz or Chan (on the Black Sea coast of East Anatolia, abutting on Georgia—50,000).

From the above, we see that there are about 5,660,000 speakers of Caucasian languages in the Soviet Caucasus. Outside of the USSR there are perhaps 80,000 Georgian speakers (about 70,000 of them in Turkey and Iran), 50,000 Laz speakers in Turkey, and about 170,000 speakers of NWC languages throughout the Middle East. These last are the people generally known throughout the area as “Circassians,” but include about 10,000 Abazans as well as speakers of Circassian (Adygahan) dialects. They are the descendants of the great number (estimated at nearly half a million) of NWC Moslems who migrated into Turkey around 1865 when the Russians invaded their NWC homeland. Great numbers were settled in Anatolia, where their descendants still live. The Turks settled others around the eastern borders of the Ottoman Empire as frontier guards, hence the present-day Circassian colonies in the Middle East. The 1964 Atlas of Peoples of the World (3) enumerates 105,000 Circassians in Turkey, 25,000 in Syria, 20,000 in Jordan, and 8000 in Iraq. In addition, there are about 2000 in Israel. In the United States there is a small Circassian community in New York City and in New Jersey.

Adding the 300,000 or so Caucasian speakers outside the USSR to the number living in the Soviet Caucasus, we reach a grand total of about 6 million speakers of Caucasian languages.

For an overall view of Caucasian languages very little is available in English. The Dutch linguist Aert Kuipers contributed a survey to the first volume of Current Trends in Linguistics (54). Also valuable, though brief, is the article on Caucasian languages by the Georgian linguists Gamkrelidze & Gudava in the latest edition of the Encyclopaedia Britannica (25). Peoples and Languages of the Caucasus by Geiger et al (27) is a useful reference work, though in some respects (e.g. population statistics) out of date, and, of course, it gives no descriptions of the languages.
For general descriptions of Caucasian languages one must turn to works in languages other than English. Dirr's 1928 Einführung (21) is outdated, though the section on the Caucasian verb is still a useful summary. Deeters' 79-page survey "Die Kaukasischen Sprachen" in the 1963 Handbuch der Orientalistik (15) is quite useful, but the best short introduction to Caucasian languages as a whole is Klimov's Kavkazskie jazyki (Caucasian Languages), published in Russian in 1965 (50) and in German translation in 1969. The German version is more useful in some respects than the original. It contains a much fuller bibliography, better maps, and a useful list of Caucasian alphabets with transliterations. For the reader of Spanish there is Bouda's 85-page "Introducción" (8), which includes texts.

Undoubtedly the most useful general introduction is Volume 4 of the Russian series Jazyki narodov SSSR (Languages of the Peoples of the USSR) (40). This 700-page work contains general introductory material and a short descriptive sketch of from 15 to 22 pages in length of every Caucasian language. These descriptions, covering chiefly phonology and morphology (with a few paragraphs on syntax, lexis, and dialects), are written according to a uniform plan by leading Soviet Caucaso-

Since the major works just mentioned (25, 40, 50) between them provide a fairly complete bibliography up to 1969, I make no attempt to do the same here, mainly citing publications only when they are relevant to the discussion in hand. What I attempt to do in this survey is to provide a general account of Caucasian languages topic by topic rather than language by language.

PHONOLOGY

There is a widespread belief among linguists that Caucasian languages are characterized by enormous numbers of consonants but minimal vowel systems. This is more or less true only of the NWC languages. Among these, Ubykh apparently holds a world's record with 80 consonant phonemes and probably two vowels. The Bzyb or northwest dialect of Abkhaz has 67 consonants (and two vowels), though literary Abkhaz has only 58 consonants. The Bzhedukh dialect of Adyghe has 64 conso-

With three exceptions, the 26 Dagestanián languages have from 30 to 46 consonants and from 5 to 10 basic vowel phonemes. These exceptions are the Andi language Akhwak and the Lezgian languages Artchi and Khinalug. Akhwak has 50 consonants, as against the 38 to 45 of other Andi languages, chiefly because of its full set of 7 laterals, and 4 uvular stop and affricate phonemes. In general, Lezgian consonant systems run from 30 to 46. According to Khajdakov (42), Artchi has 51 consonants, although Mikailov (65) lists only 46. Khinalug appeared already to have some unusual consonantal features for a Lezgian language among its 44 or 45 consonants in Dešeric's grammar (16). The most recent description of Khinalug by a Moscow State University team (47) indicates that at least one consonant /4/ has disappeared from Khinalug. On the other hand, Kibrik et al (47) elaborate
the system, chiefly by recognizing a series of palatalized consonant phonemes, bringing the total inventory to 76. The Kartvelian languages have 28 (Georgian) and 30 (Zan and Svan) consonants plus 5 (Georgian) and basically 6 (Zan and Svan) vowels, though the number of vowel phonemes is higher in some dialects of Svan.

There is as yet no general work on the phonetics and phonology of Caucasian languages, though this is a project toward which the present author is working, but there are, of course, phonological sketches in descriptions of particular languages, notably those in Jazyki narodov (40). We mention here, however, some particularly informative special works on the phonetics/phonology of single languages or language groups.

For the NWC languages there is a useful phonological sketch in French, with tables of the consonant phonemes of Ubykh, Abkhaz, Adyghe, literary Kabardian, and the Bes(le)ney dialect of Kabardian in Catherine Paris’ book on the latter dialect (69). The phonetics of all dialects of the Adyghan languages have been treated by Balkarov (4), the special features of Adyghe dialects by Rogava & Keraševa (72), and of Kabardian dialects in the 1969 symposium edited by Kumakhov (62).

For the NCC (Nakh) languages we have Sommerfelt’s studies published (in French) in 1934, 1938, and 1947 (78). These, like Dešeriev’s 1963 book in Russian (18), have a comparative-historical orientation but also give descriptive information. The most recent work on general NCC phonology is Črelašvili’s System of Consonants of the Nakh Languages in Georgian (14), but available also in Russian in the same author’s 81-page ‘avtoreferat’ or dissertation résumé. There are several descriptive works on Chechen phonetics, notably by Dešeriev (17) and in Dešerieve’s (19) “contrastive analysis” of the phonetics of Chechen and Russian.

Apart from some of the older work of Trubetskoy, such as his “Die Konsonanten- systeme der ostkaukasischen Sprachen,” published in 1931 (84), there are no purely descriptive works on Dagestanian (NEC) languages as a whole. Bokarëv’s pioneering work (7) was, of course, comparative-historically oriented and is the basis of the phonological aspects of later works such as Murkelinskij’s (67) and, to some extent, Khajdakov’s (43). There are phonological sketches in the grammars of individual languages, and in works dealing with language groups such as Gudava’s Consonantism of the Andi Languages (29) and Bokarëv’s Tsez (Dido) Languages of Dagestan (6), both in Russian. One of the most valuable works on Dagestan phonetics is Gaprindašvili’s Fonetika darginskogo jazyka of 1966 (26). Although dealing specifically with the phonetics of Dargi dialects, it contains a great deal of information concerning Dagestanian phonetics in general, and provides some hard data on Dargi in the form of X-ray tracings, kymograms, oscillograms, acoustic spectra, and palatograms. There is a specialized study of the phonetics of the Lezgian language Tsakhur by Ibragimov (33) which is interesting though without instrumental data.

General information on Kartvelian (SC) phonetics and phonology is, of course, contained in comparative-historical works such as those by Schmidt (73), Klimov (49) and Gamkrelidze & Machavariani (24). The structure of the Kartvelian syllable is discussed in Zhghenti’s Comparative Phonetics of the Kartvelian Languages (89), Part I published in 1960 in Georgian, but with Russian and English summaries. As for Georgian, by far the best described of the SC languages, Zhghenti’s Phonetics of
the Georgian Language (88) (in Georgian) contains X-ray tracings, kymograms, and palatograms of Georgian articulations. More accessible are the shorter descriptions of Georgian phonetics and phonology by Robins & Waterson (71) and Vogt (86).

I now propose to characterize briefly the main features of the consonantal and vocalic systems of Caucasian languages.

Consonants

The consonant systems of all Caucasian languages have certain characteristics in common, namely (i) stops articulated at labial, dentalveolar, velar, and uvular locations (types p t k q); (ii) affricates at two locations (types ts tf), fricatives at alveolar, postalveolar, and uvular locations (types s f x). Incidentally, it is noteworthy that in Caucasian languages if there is only one type of dorsal fricative it is always uvular, not velar. The basic consonant inventory is completed (iii) by two nasals (m n), a labial semivowel or fricative (w v), a palatal semivowel (i), an apical trill (r), and all except the Adyghan languages have a lateral approximant (l). In the two Adyghan literary languages all the laterals are fricative: Ԁ ԁ Ԃ .

In all Caucasian languages the stop consonants participate in a voiced, voiceless aspirated, and glottalic triad of the type (d t t'). All three types are always represented at the dentalveolar and velar locations, but one or more members of the triad may be absent at the labial or uvular location: thus Avar has d t t' and g k k', but only b and p and q and q'. In just over half the Caucasian languages the triadic series extend to affricates, i.e. we find dz ts ts'; dʒ tf tf'. However, in most of the Avar-Andi-Tsez languages and Lak and Dargi the voiced member is absent. Finally, the triadic series extends also to some of the fricatives in the Adyghan languages and in dialects of two Andi languages (Bagwali and Chamali).

What we have outlined here is the universal or common-core Caucasian consonant system. Almost exactly this system is, in fact, realized in the Kartvelian (SC) languages, which have the simplest of all Caucasian consonant systems. The Kartvelian consonant system may be represented thus:

<table>
<thead>
<tr>
<th></th>
<th>m</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>d</td>
<td>g</td>
</tr>
<tr>
<td>p</td>
<td>t</td>
<td>k</td>
</tr>
<tr>
<td>p'</td>
<td>t'</td>
<td>k'</td>
</tr>
<tr>
<td>dz</td>
<td>dʒ</td>
<td></td>
</tr>
<tr>
<td>ts</td>
<td>tf</td>
<td></td>
</tr>
<tr>
<td>ts'</td>
<td>tf'</td>
<td></td>
</tr>
<tr>
<td>v</td>
<td>z</td>
<td>ʒ</td>
</tr>
<tr>
<td>s</td>
<td>f</td>
<td>ɣ</td>
</tr>
<tr>
<td>w</td>
<td>j</td>
<td></td>
</tr>
<tr>
<td>l</td>
<td>r</td>
<td></td>
</tr>
</tbody>
</table>

This is the consonant system of Georgian, Svan, and Zan—the aspirated (q) occurs only in Svan, the glottal stop (?) only in Megrelian, and (h) is rare in Georgian.

In all the other Caucasian languages this simple common core system is elaborated in various ways right up to the 80-consonant system of Ubykh. These elabora-
tions for the most part take the following forms. First, there may be an additional term added to triads like *d tt t* or to diadic sets of affricates or fricatives. Additions of this type are usually called **intensive, strong, or geminate** with respect to the 18 Dagestanian languages in which they occur, and unaspirated or **preruptive** with respect to the two dialects of Adyghe (Shapsugh and Bzhedukh) in which they occur. The term "preruptive" (preruptivnyj) for "strong and unaspirated" was apparently coined in contrast to the commonly used Russian term "abruptive" (abruptivnyj) meaning "glottalic." The result of this type of addition is to produce sets like *b p pp p', g k kk k' etc in which the fourth, "intensive" term is realized in various ways. Generally in Lak, Dargi, and the Lezgian languages intensive stops *pp tt kk etc are tense unaspirated, and, when intervocalic, geminate. The corresponding affricates *tt s and *ttf are likewise tense and have a lengthened stop portion and tense unaspirated affrication. In Avar and most of the Andi languages the "intensive" stops are strongly affricated, intensive affricates have lengthened affrication and are unaspirated, and the intensive fricatives are lengthened and unaspirated. In all these languages, the nontense correlates of these phonemes are short and aspirated. Thus in Avar we have the pairs *k-kk, k'-kk', tt-tss, tt'-tss', s-ss, f-ff, etc realized as *[kʰ-kx, k'-kx', tsʰ-tss, ts'-tss', sʰ-ss, etc]. The tense fricatives in Avar, in addition to being longer, generally have a louder hiss than the corresponding nontense fricatives, which betokens a higher velocity of air flow through the articulatory channel, partly achieved by narrowing the channel, as Gaprindašvili's palatograms show for Dargi dialects (26).

The second way in which the consonant system may be augmented is by the addition of more articulatory locations. Thus the stop series is often extended by the addition of a glottal stop. Glottal stop phonemes occur in Abaza and Adyghan, in all three Nakh languages, and in virtually all Dagestanian languages. Indeed a number of languages have more than one glottal stop phoneme. Both of the Adyghan literary languages have plain ? and labialized ?*, and the Abdzhakh dialect of Adyghe has a slightly palatalized ? j as well. This is the reflex of the Proto-Circassian postalveolar glottalic affricates *tj* and *tj' which have lost their oral articulation in Abdzhakh. The Nakh languages and a few Dagestanian languages, notably the Tsez languages and Dargi, contrast a weak glottal stop ?, often realized simply as creak, with a strong glottal stop ??, which involves tight closure of the ventricular bands (as well as the vocal cords) and some constriction of the pharynx. This is sometimes called a "pharyngal stop" in the literature, but is perhaps better described as a (pharyngalized) ventricular + glottal stop. Incidentally, in reading the literature on Caucasian languages one must note that Georgian linguists generally use the Georgian or Russian equivalent of "pharyngal" to mean what is more commonly (e.g. by Moscow linguists and by the International Phonetic Association) called "uvular."

Additions to the fricative series include a voiceless labial fricative in a few languages—labiodental j in the NWC languages and a few Lezgian languages, bilabial p in Ingush and Tabasaran. In addition to the uvular fricatives x and s, most North Caucasian languages except Nakh, Hunzib, and Udi, have at least one velar fricative, and many have at least one pharyngal fricative or approximant.
Other elaborations involve secondary articulations—labialization, palatalization, pharyngalization. The feature of labialization is found in all the NWC languages, sporadically elsewhere. In NWC it applies to stops and to apico-laminal, dorsal, and pharyngal fricatives. It takes various forms, which I have described elsewhere (10) including simple lip-rounding, labiodentalization (of the affricats dz ts ts' in Abkhaz, and also in a few Lezgian languages) and, most interesting, complete labial closure. Thus labialized d t t' in Abkhaz and Ubykh are in fact db tp tp' with simultaneous complete (inner) labial closure. Such "strongly" labialized sounds also occur in some dialects of Lak.

Palatalization is applied to stops and some affricates and fricatives in the NWC languages. Most interestingly, Abkhaz, Abaza, and Ubykh have plain, labialized, and palatalized uvulars (q q' q', xx' xx' etc) in which the primary stricture is indeed at the back-velar or uvular location, but the more anterior part of the tongue dorsum is raised up toward the hard palate, forming a longitudinally extended articulation.

Pharyngalization occurs with labials and uvulars in Ubykh, and with uvular fricatives only in the Bzyb dialect of Abkhaz. In both these languages we have the unusual phenomenon of the co-occurrence of the features labialized and pharyngalized on uvulars.

The features of palatalization and labialization occur with sibilants in NWC (and in Tabasaran, with its labiodentalized s' ts' etc). The elaboration of sibilant articulation can, however, best be described in terms of the distinct articulatory positions involved. In NWC languages there are four basic sibilant types: - apico-or lamino-alveolar, f-apico-postalveolar (slightly velarized), c lamino-postalveolar, palatalized, and the peculiar NWC sibilant s, which occurs in slightly varying forms in Bzyb, Ubykh, and the Adyghan languages. The sound s, described in the Russian literature as "hissing-hushing," is indeed acoustically and physiologically between a typical s and a typical f. In its production the tip of the tongue rests against the alveoles of the lower teeth (as for a laminal s), but the main articulatory channel is at the back of the alveolar ridge (as for a lamino-postalveolar f). It is also often slightly pharyngalized. The greatest proliferation of sibilants is found in the Bzhe-dukh dialect of Adygh, which, by combining these four basic types with the various initiation-phonation types and the "intensity" opposition, achieves no fewer than 14 sibilants, all phonemically distinct, namely z z' z' zw, s' sw, s, f f f, c c c, s s' sw.

One other characteristic Caucasian extension of the consonant system is a set of lateral fricatives and/or affricates. Nineteen Caucasian languages augment their consonant system in this way. The various types of nonapproximant laterals that occur include glottalic, voiceless, and voiced fricatives (the voiced lateral fricative h occurs in the Adyghan languages, to the exclusion of approximant l, which does not occur in these languages), voiceless and glottalic lateral affricates. The intensity opposition is applied to laterals in the Avar and Andi languages, which show the greatest proliferation of laterals, the record being held by Akhwakh which has seven: approximant l plus t h t h, r r r r r r r. All are phonemically distinct. Artchi is exceptional among Lezgian languages in having a set of four lateral fricatives and
affricates, and even more exceptional in that, according to Mikailov (65) and Kibrik (46), these laterals are all velar or slightly palatalized velar with respect to the location of their central obstruction.

In Akhwhakh, the duration of the noise-burst and voicing lag of the short lateral affricates is no more than that of an aspirated stop, while the long or “tense” or “geminate” lateral affricates have long and noisy affrication. One might, therefore, speak of an opposition between “lateral stops” (t and t') and lateral affricates (t' and t'). For actual durations of the voicing lag of these sounds see Catford (13, p. 214.)

To conclude this section on consonants, here are the consonant systems of Avar and Abkhaz illustrating some of the “elaborations” of the core system.

**Avar:**

```
<table>
<thead>
<tr>
<th>m</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>p'</td>
<td>t'</td>
</tr>
<tr>
<td>ts</td>
<td>tss</td>
</tr>
<tr>
<td>ts'</td>
<td>tss'</td>
</tr>
<tr>
<td>tf</td>
<td>tff</td>
</tr>
<tr>
<td>tf'</td>
<td>tff'</td>
</tr>
<tr>
<td>t44</td>
<td>t44'</td>
</tr>
<tr>
<td>v</td>
<td>z</td>
</tr>
<tr>
<td>s</td>
<td>ss</td>
</tr>
<tr>
<td>l</td>
<td>r</td>
</tr>
</tbody>
</table>
```

**Abkhaz:**

```
<table>
<thead>
<tr>
<th>m</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td>b</td>
<td>d</td>
</tr>
<tr>
<td>p</td>
<td>t</td>
</tr>
<tr>
<td>p'</td>
<td>t'</td>
</tr>
<tr>
<td>dz</td>
<td>dz'</td>
</tr>
<tr>
<td>ts</td>
<td>ts'</td>
</tr>
<tr>
<td>tf</td>
<td>tff</td>
</tr>
<tr>
<td>tf'</td>
<td>tff'</td>
</tr>
<tr>
<td>t44</td>
<td>t44'</td>
</tr>
<tr>
<td>v</td>
<td>z</td>
</tr>
<tr>
<td>s</td>
<td>ss</td>
</tr>
<tr>
<td>l</td>
<td>r</td>
</tr>
</tbody>
</table>
```

The symbol ̄u represents a labial plus palatal semivowel, exactly like the initial French sound of *huit*. It is the Abkhaz reflex of the voiced labialized pharyngal approximant of Abaza, ̄u. The Bzyb dialect of Abkhaz has all the above consonants, plus dz dz' ts ts' ̄u ̄u ̄u and the pharyngalized uvular fricatives ̄x ̄x ̄x, giving a total of 67.

Ubykh lacks six of the Bzyb consonants ( g k k' h and h ̄u) but more than makes up for that by adding a set of pharyngalized labials and uvulars, a velar x and ̄u (as well as the uvular fricatives), a glottal ̄u, and the laterals t4' and t4, to achieve its record total of 80 phonemically distinct consonants.
Consonant Clusters

Caucasian languages have a reputation for permitting enormous intrasyllabic clusters of consonants. This is strictly true only of Kartvelian—above all Georgian, and to some extent NWC. The Nakh languages have a moderate number of consonant clusters, and the Dagestani languages, with the exception of Lezgi, admit none but the very simplest clusters of consonant + w, in initial position at least. The considerable number of initial clusters in Lezgi is no doubt of relatively recent origin; at least the highly varied Lezgi clusters do not conform to what seems to be a Caucasian typological norm.

To return to Kartvelian: in Georgian there are syllable initial clusters of from 2 to 6 terms, and final clusters of from 2 to 5. Such forms as prskvna “to peel,” ts’vrtna “to train,” brts’q’inva “to shine,” are not at all rare. Vogt (86) lists all the initial clusters that occur in Georgian. They total 740 (cf English with 40 to 50 initial clusters, according to dialect), of which 4 are 6-term clusters, 21 are 5-term, and 148 are 4-term.

The structure of Georgian consonant-clusters though varied is not random, and there is one frequent and highly typical variety of initial cluster which Georgian linguists call “harmonic complexes” of consonants [see Zhghenti (89)]. These clusters are characterized by the following three formation-rules:

1. They consist of occlusive (stop or affricate) followed by occlusive or fricative, i.e. type pk dts’q’ etc, not *fk *kp etc, except for a limited set of clusters with initial s or f.
2. They are homogeneous with respect to initiation and phonation, i.e. glottalic, or voiced, or voiceless throughout, i.e. types p’k’ ts’q’ bz dts’kp ts’k’ etc, not *p’b *bs etc.
3. They are “recessive,” i.e. in terms of articulation they “recede” back into the mouth, the second term always being posterior to the first term.

There is a relationship of solidarity between the rules such that if any two apply then the third also applies.

Exactly the same rules apply to morpheme-initial clusters in the Adyghan languages, where the only exception is the Adyghe tf (which contravenes rule 3). This, however, is clearly of relatively recent origin and results from the change of Proto-Circassian *xw to f, as evidenced by the corresponding cluster in Kabardian, which is txʷ. Basically the same initial cluster rules apply to Ubykh, Abkhaz, and Abaza, only in the last two languages more complex clusters have arisen, no doubt through widespread loss of interconsonantal vowel.

In the Nakh languages the numbers of initial consonant clusters are small; excluding clusters of foreign (chiefly Russian) origin, Chechen and Ingush have around 20, Batsbiy a few more. What is interesting, however, is that for the most part the Nakh consonant clusters follow precisely the same formation rules as do the Georgian “harmonic complexes” and the NWC intramorphemic clusters. Of these three rules, the third—that is, the rule that the second term in the cluster must have a more posterior articulation than the first—is particularly interesting. The fact that
this rare, if not unique, role of consonant clustering occurs in three distinct sub-branches of Caucasian—NWC, NCC, SC—is undoubtedly of some significance for general Caucasian typology.

Vowels

We mentioned in the general introduction to Phonology that the five NWC languages have minimal vowel systems. These vowel systems have been the subject of a good deal of discussion and controversy, which I can do little more than suggest here. In a NWC language such as Kabardian we can hear many nuances of vowel quality such as itiʊ u e e ə ɔ a ʌ a etc. It soon becomes apparent to any observer, however, that many of these vowel qualities are environmentally conditioned. Thus i or i occur in the neighborhood of j, o u ɔ etc in the neighborhood of w and labialized consonants, and so on. Apparent minimally contrasting sets such as bzi “female,” bzi “ray,” bzu “sparrow,” bze “language,” bi “enemy,” be(i) “rich,” bɛ “much,” ba “kiss,” etc can be shown by simple morphological tests to be analyzable as containing a vowel + j or w. Thus, though bzi “female” + f ’i “good” yields bziʃʃ’ , bzi + f ’i yields bziʃʃ’ (not *bzi:f’), bzu + f ’i yields bzuwɔʃ’ (not *bzu:f’), and so on. In short, one quickly discovers that there seem to be only three distinctive vocalic units: a close vowel which one may write, following Jakovlev, ə (with allophones i t iOU etc), a mid vowel e (with allophones e ʌ a u etc), and a third unit, which we may write a (with allophones a a). In 1923 Jakovlev published a remarkably detailed study of Kabardian phonetics-phonology (36) in which he analyzed the vowel system as ə e a in which there were two vowel qualities, ə and e, a being in fact the long of e. Trubetsky, in his 1925 review of Jakovlev (83), distinguished the vowels as close mid and open. For a long time the Kabardian (and Adyghe) vertical vowel system, ə e a, was regarded as unique, though we now know that vertical vowel systems of this type occur in some languages of New Guinea, as reported by E. Pike (70).

Having reached a vertical 3-vowel system for Kabardian, it is tempting to reduce it still further. Jakovlev, as we have seen, regarded a as the long of e (though measurements show that there is little empirical basis for this assessment—in most environments the difference is primarily one of quality), and he further suggested that since ə alternated with zero in largely predictable ways, it was possible that at some period in its history Kabardian had been a monovocalic language, the single vowel phoneme e contrasting on the one hand with ə~zero and on the other e (=a). Jakovlev was prepared to posit monovocalism only as appertaining to an earlier stage of Kabardian; at no time in 1923 or his later works did he ever suggest that present-day Kabardian is monovocalic. Others, however, have been less cautious. The present writer about 1948 (before having seen Jakovlev’s work) began to write a paper entitled “Kabardian: a Monovocalic language?” but abandoned it for lack of data. In 1960 Kuipers (53) published a book in which, on the basis of complex arguments, he arrives at a monovocalic analysis of Kabardian. Meanwhile Genko (28) and Allen (1) had proposed a monovocalic analysis of another NWC language, Abaza. The hitherto unheard-of possibility of languages with only one vowel phoneme aroused great interest among linguists since it added plausibility to the postu-
lation of a single vowel phoneme for Proto-Indo-European. The monovocalic hypothesis was attacked by Szemerényi in 1964 (81), and strongly defended with reference to the NWC languages by Allen in 1965 (2).

The monovocalic analysis of NWC languages is certainly very attractive. In his 1960 book, however, Kuipers went even further. Pointing out that a syllable-peak with the unique vowel e is distinguished from one with o~zero only by its degree of openness, we could, he suggests, call e simply a "feature of openness." And since this feature of openness occurs only along with consonants, it can be regarded, along with labialization and palatalization, as simply another feature of "shape of the mouth resonator." Kabardian would thus be not merely a monovocalic language, but a phonologically vowelless language—a language with no opposition consonant-vowel (nonsyllabic-syllabic) in its phonology. This extreme position was criticized by Halle (31) and is not widely accepted.

An important advance in the discussion of NWC monovocalism was the publication of an article by the Soviet Adyghian specialist Kumakhov (61). Reviewing all the previous discussion, Kumakhov adduces evidence, for Adyghian in particular, casting doubt not only on the "vowelless" but also on the monovocalic interpretation. Among other things he draws attention to certain clear cases of linguistically relevant oppositions between the close vowel o and zero, which seem to reinstate this vowel as a phoneme. Kumaxov finally posits a three-vowel system o~e~a for Adyghe and Kabardian and a two-vowel system o~a for Abkhaz, Abaza, and Ubykh.

The latest contribution to the discussion is an article by Kuipers in 1976 (56) in which he responds to the criticisms of Halle and Kumakhov and reaffirms his position, and extends the "vowelless" hypothesis to Abkhaz as well as Kabardian. So the problem of the NWC vowel-systems is not yet resolved. At the very least they are certainly all unidimensional "vertical" systems; possibly one or more of them is monovocalic.

The vowel systems of all the other Caucasian languages are more orthodox, though not without features of interest. The Nakh languages have fairly extensive vowel systems. Thus Chechen is said to have a total of 30 vocalic nuclei—15 vowels and 15 diphthongs. The 15 simple vowels, however, break down into 9 basic vowel qualities, roughly i e æ a o u ɪ y of which 6 can be long i: e: a: o: u: y:. Batsbiy, like most of the Andi and Tsez languages, augments its vowel systems by a series of nasalized vowels.

The Dagestani languages have from 5 to 10 basic vowel phonemes, in many cases augmented by a set of long and/or nasalized vowels. In addition, about 10 Dagestani languages have some pharyngalized vowels. The fullest sets of pharyngalized vowels are found in Tsez and Khwarshi and in the Lezgian languages Artchi, Tsakhur, Rutul, and Udi. In Tsakhur, for example, we have a basic set of six vowels i e æ o u ɪ (of which some can occur long), all six also occurring pharyngalized. This pharyngalization takes the form of retraction of the tongue-root, and it appears to induce, as a side-effect, a certain degree of fronting of back vowels, particularly the closer vowels u and o. The exact mechanism of this is not clear, but the fact is that in Tsakhur and Udi the pharyngalized u and o have a distinctly central
quality. Formant frequencies for Tsakhur u show a marked rise (implying a forward shift) in passing from u to pharyngalized u. The averages for a small sample (3 or 4 examples) are u F1 380 F2 760, u F1 400 F2 1050. This "fronting" associated with pharyngalization, or pharyngal articulation, is also found in NWC. As we saw above, the Abkhaz equivalent of Abaza Ϛw is the labial + palatal semivowel ñ (not w). In fact, the rounding applied to Ϛw or hw in Abaza and to hw Abkhaz is precisely of the ñ-type. This association of pharyngalization and fronting constitutes what Trubetskoy (84) termed "emphatische Mouillure" ("emphatic softening"). Moreover, the fronting effect of pharyngalization accounts for a number of curiously "skewed"-looking vowel systems. Thus the vowel systems of Lak, Lezgi, and Tabasaran may be represented as:

<table>
<thead>
<tr>
<th></th>
<th>i</th>
<th>u</th>
<th>i</th>
<th>y</th>
<th>u</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>æ</td>
<td>æ</td>
<td>e</td>
<td>æ</td>
<td>e</td>
<td>æ</td>
<td>æ</td>
</tr>
<tr>
<td>a</td>
<td>a</td>
<td>æ</td>
<td>a</td>
<td>æ</td>
<td>a</td>
<td>a</td>
</tr>
</tbody>
</table>

Lak       Lezgi       Tabasaran

When we realize that the front or central vowels represented by e æ and æ y or æ are slightly pharyngalized, we see that we have basically a triangular vowel system in which the additional vowels result from the probable influence of pharyngal or pharyngalized consonants.

Soviet scholars often refer to vowels of the æ y æ ñ type as "umlauted," but they do not usually result from the particular diachronic process known as "umlaut" in Western Indo-European languages. There are, however, a few Caucasian languages in which a genuine umlaut process—stressed vowel modified by the quality of the unstressed vowel in the following syllable—has occurred. This is the case in some dialects of Chechen, but it is best known in the SC language Svan, on which see Kaldani (41), in Georgian with Russian summary. Apparently in some dialects of Svan the umlaut process has resulted in rounded front vowels, at least the letters û and ù are used in the literature. However, the only two Svan speakers I have worked with have had wi or iw for û and we or ew for ù. Examples of Svan umlaut are: pwir or pirw (cf Geo. puri) "cow," twek' or tekw' (Geo. tok'i) "rope," wep (Geo. opli) "sweat," semi (sami) "three," didæb or didab (Geo. dideba) "glory," etc.

**ORTHOGRAPHIES**

Only the Georgian language has its own native writing system and a rich literature going back to the fifth century. The Georgian alphabet is traditionally supposed to have been invented around 400 A.D. by that same Armenian cleric, Mesrop Masots, who is credited with the invention of the Armenian alphabet and probably also the Caucasian Albanian alphabet. Caucasian Albania was an ancient transcaucasian kingdom located to the southeast of ancient Iberia, largely in what is now Azerbai-
jan. Known about since the end of the nineteenth century, the Albanian alphabet did not become an object of study until after the discovery in 1937 and 1953 of copies of the alphabet in fifteenth and sixteenth century Armenian manuscripts and, more importantly, the discovery in 1948 of a short inscription on a stone slab, and other fragments, at Mingecaur in Azerbaijan. A number of scholars suppose that the language of these fragments in the Albanian alphabet is an older form of Udi, a Lezgian language which is today unwritten and spoken chiefly in two villages in northern Azerbaijan, but more research is required before this can be established with certainty. Klimov (50) outlines the discovery and research on the Albanian alphabet, and other useful sources are by Shanidze (77) and Gukasjan (30).

There are then only two indigenous alphabets for Caucasian languages, and one of them, the Caucasian Albanian alphabet, apparently went out of use very early. During the five or six centuries preceding the Revolution no Caucasian language other than Georgian had an official or widely used orthography. However, that does not mean that these languages were never written. In Dagestan particularly, where there was a long tradition of Islamic culture and of literature in Arabic, various Caucasian languages were written in the Arabic alphabet. The oldest example of this is a group of marginal and interlinear notes in Dargi on an Arabic manuscript of 1243 A.D., believed to be contemporary with the manuscript. There are other early examples from the end of the fifteenth century of Dargi, Lak, and Avar in Arabic script.

By the nineteenth century there was a fairly well established tradition of writing Avar and Lak in the Arabic script, but it was only after the Revolution that the Soviet government created no fewer than eleven new literary languages, providing alphabets for Abkhaz, Abaza, Adyghe, Kabardian-Cherkess, Chechen, Ingush, Avar, Lak, Dargi, Tabasaran, and Lezgi. Most of these languages achieved their literary status in 1918, but Ingush in 1920, Abaza and Tabasaran in 1932.

The first official orthography for Adyghe, Kabardian, Chechen, Avar, Lak, Dargi, and Lezgi used the Arabic alphabet, augmented by additional points and other diacritic marks and by slight modifications of some letter shapes; for instance, the glottal and voiceless lateral fricatives of Adyghe and Kabardian were represented by the Arabic / with, respectively, one and two little ticks added to the letter. In 1928 the Arabic orthographies were replaced by augmented Roman (Kabardian, in fact, went through two different augmented Roman alphabets between 1924 and 1936). Three other languages were written from the start in augmented Roman: Ingush in 1920, Abaza and Tabasaran in 1932. In 1938 (1936 in the case of Kabardian) the augmented Roman alphabets were replaced by Cyrillic. This was a very reasonable move because even though the Caucasians learned to read and write first in their own language, most would eventually learn Russian, and they might as well start out knowing the alphabet. The Cyrillic alphabet introduced at that time and still in use is, in fact, simply the Russian alphabet with no new letters with the exception of capital И, which was, in any case, a letter of the old prerevolutionary Russian alphabet. This capital И, which projects above the line of type in a noticeable way, is called the “paločka” or “little stick” by Caucasians. The paločka serves three purposes: first, by itself, it represents glottal stop; second, following letters represent-
ing a voiceless stop, an affricate, and some fricatives it indicates glottalic initiation; and third, it is used after one or two letters to represent pharyngal or glottal or uvular articulation according to language. Since I is the only additional letter in the Caucasian Cyrillic alphabets, many Caucasian phonemes have to be represented by digraphs, trigraphs, and in one case (the Kabardian voiceless aspirated and/or affricated labialized q") a quadrigraph, _kelas_y. This convention is not as cumbersome as it sounds; one quickly becomes used to reading digraphs and trigraphs as unit phonemes, and the Kabardian quadrigraph is very rare, occurring, according to Balkarov (4), only once in 8000 words of text.

The history of Abkhaz orthography is different from that of all the other newly written Caucasian languages. Abkhaz had been reduced to writing in the nineteenth century by Uslar (85), and a modified form of Uslar’s Cyrillic-based orthography was used up to 1928. Thereafter, till 1938, Abkhaz was written in augmented Roman. At that time, when Cyrillic was introduced for the other languages, Abkhaz, which is spoken within the Georgian SSR, was provided with a Georgian-based orthography—that is, Georgian augmented by a few special letters used either independently or as modifiers. The augmented Georgian alphabet was used until 1954, when an entirely new Abkhaz orthography was introduced, this time on a Cyrillic basis, but differing considerably from all other Caucasian alphabets. The augmented Cyrillic used for Abkhaz not only makes use of digraphs but also includes 13 modified or totally new letters, several of which go back to Uslar’s (85) alphabet of the 1860s. The Abkhaz alphabet makes no use of the palovka; instead it represents glottalic stops and affricates either by specially invented or modified letters or by the ordinary Cyrillic voiceless stop characters. In Abkhaz modified letters are used to represent the voiceless aspirated stops—letters used by Uslar though borrowed by him from Sjögren’s Ossetic grammar of 1844.

One of the decisions which had to be made in developing languages in the Caucasus concerned the selection of a particular dialect upon which to base the literary language. In some cases the problem had already been solved by existing sociolinguistic conditions. Thus Avar was already widely used as a lingua franca in part of Dagestan, and the variety which performed this function was the northern dialect of the Khunzaq area, so this became the new literary language. Among the NWC languages the choice depended partly on phonology. For Adyghe, the Temir-goi (or Chemgui) dialect was selected partly because it had the simplest sound system—only 50 consonants as against the 60 of Shapsugh or the 64 of Bzhedukh. For Abkhaz, the first choice was the Byb dialect, but later the phonologically simpler Abzhu dialect became the basis of the literary language.

Although all Caucasian literary languages, with the exception of Abkhaz, use basically the same alphabet, there are a number of differences in the utilization of particular letters or digraphs. Thus the Russian “soft s” (Sce) character indicates lamino-alveolar palatalized c in Adyghe, the special NWC s in Kabardian, and intensive or geminate /\ in Avar and Lak. The Russian “k + soft sign” stands for glottalic palatalized k in Abkhaz, for aspirated palatalized k in Abaza, the strong glottalic lateral affricate t4z in Avar, and glottalic uvular stop q` in other Dagestan languages. The student of Caucasian languages must thus always remember which
set of conventions he is working with in reading examples from different languages. Moreover, in addition to the official orthographies, publications on Caucasian languages use several different types of phonetic or phonological transcription based on Roman, Cyrillic, and Georgian.

General information on the orthographies of the modern Caucasian literary languages can be found in Musaev (68) and Isaev (35). For Abkhaz, there is the short study by Bgazba (5).

GRAMMAR

Probably every linguist knows that Caucasian languages are characterized by their possession of the ergative type of transitive sentence construction. This is indeed a feature of all Caucasian languages, except the Megrelian dialect of Zan, and it is one of the few general Caucasian features. Though there are grammatical characteristics which stretch over several branches of Caucasian, there are also considerable differences. Thus the NWC languages can rightly be described as “polysynthetic” because of the extreme complexity of their verb forms, which incorporate pronominal and relational indices virtually recapitulating the entire syntactic content of the sentence, while at the other extreme we have the Lezgi verb, which is virtually devoid of indications of anything other than the purely verbal categories of tense, mood, etc.

In spite of some major differences from the well-known Indo-European type of language, Caucasian languages are not so exotic that they cannot be described in terms of the traditional grammatical categories. I propose, therefore, to illustrate some features of Caucasian grammar under such headings as “Nouns,” “Pronouns,” etc.

Nouns

One interesting characteristic of 28 of the 37 Caucasian languages is that nouns are distributed into a number of gender-like classes. In NWC noun classification is minimal; only in Abkhaz-Abaza do we find the distinction human-nonhuman, marked in the form of numerals and in the pronoun system (which also distinguishes masculine and feminine in the second person). Class systems play a larger role in the three Nakh languages and in 23 languages of Dagestan.

Noun classification is for the most part a covert category. That is, it is not indicated on the noun itself but only by the presence of class-markers (prefixes, suffices, or infixes) on words in a concord relationship with the noun, which include adjectives, verbs, pronouns, and adverbs. Avar has three classes: human male, human female, and neither (m., f., n.) marked by w, j, b. These distinctions are not made in the plural, which thus in effect forms a fourth class. Examples: “The good father/mother/table is in the house.”

\[4 \text{ik'}a-w \text{emen ro}_q'o-w \text{ w-ugo}. \quad \text{Good father in-house is.}\]
\[4 \text{ik'}a-j t/f'u \text{ u ro}_q'o-j \text{ j-igo}. \quad \text{Good woman in-house is.}\]
\[4 \text{ik'}a-b \text{ t'ut' } \text{ro}_q'o-b \text{ b-ugo}. \quad \text{Good table in-house is.}\]

In Nakh and Dagestanian languages the number of noun classes ranges from none in Lezgi, Agul, and Udi, through two in Tabasaran to eight in Batsbiy. In terms
of the semantic bases for the classes, there is always a distinction (in the singular, at least) between *human* and *nonhuman*, and this is the only distinction in the two-class system of Tabasaran. The three-class languages (Avar, six of the Andi languages, and Dargi) distinguish *male* and *female* within the *human* class. In the four-class languages (Tsez, Hinukh, Bežti, Lak, and six of the Lezgian languages) the nonhuman class is basically divided into *animals* and *other things*, but the distinction is not carried out with absolute accuracy, e.g. class III may include a few inanimates, and class IV a few animals, particularly insects. There are two five-class languages, Andi and Chamali. In Andi class III is confined to animals, but in Chamali there is no such clear-cut division, animals and all other things being distributed over classes III, IV, and V. In the Nakh languages, Chechen and Ingush each have six classes and Batsbiy eight. All three languages distinguish *human male* and *human female* from everything else, but there is no obvious principle underlying the assignment of nouns to the other classes. The two six-class Dagestani languages, Khwarshi and Hunzib, make a more systematic distribution of nouns in classes III to VI, more or less clearly distinguishing animals from inanimates, and both have a special residual class containing, in Hunzib, the single word “child” and in Khwarshi “child” and “family.” The following table indicates how the classes are constituted for languages with from 2 to 5 classes:

<table>
<thead>
<tr>
<th>Human</th>
<th>Nonhuman</th>
</tr>
</thead>
<tbody>
<tr>
<td>men</td>
<td>women</td>
</tr>
<tr>
<td>2 classes</td>
<td>I</td>
</tr>
<tr>
<td>3 classes</td>
<td>I</td>
</tr>
<tr>
<td>4 classes</td>
<td>I</td>
</tr>
<tr>
<td>5 classes, Andi</td>
<td>I</td>
</tr>
<tr>
<td>Chamali</td>
<td>I</td>
</tr>
</tbody>
</table>

In all languages fewer class distinctions are made in the plural, and it is interesting to see how different languages syncretize or amalgamate class distinctions made in the singular. The following table shows the syncretized plural classes of the four-class languages.

<table>
<thead>
<tr>
<th>Animate</th>
<th>Inanimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>human</td>
<td>nonhuman</td>
</tr>
<tr>
<td>male</td>
<td>female</td>
</tr>
<tr>
<td>Tsez, Hinukh</td>
<td>men</td>
</tr>
<tr>
<td>Bež, Tsakh, Kryz</td>
<td>human</td>
</tr>
<tr>
<td>Lak</td>
<td>animate</td>
</tr>
<tr>
<td>Budukh</td>
<td>men</td>
</tr>
<tr>
<td>Rutul, Arči, Khin</td>
<td>no distinctions in plural</td>
</tr>
</tbody>
</table>
Another major nominal category is case. In only two Caucasian languages, Abkhaz and Abaza, are cases not marked on the surface forms of nouns. In all other Caucasian languages there are systems of case-forms, ranging from only two or three in Ubykh, through four in the Adyghan languages, six or so in the Kartvelian languages, to a maximum of 47 cases in literary Tabasaran (possibly 53 in a Tabasaran dialect). Cases are often divided into two categories: basic or grammatical cases on the one hand, and local cases on the other, and the great exuberance of case-systems in the Dagestanian languages mainly reflects the large number of their local cases. The nearly case-less NWC languages handle location and direction quite differently: chiefly by means of locational preverbs. I discuss both methods of expressing spatial relations together, under “locative expressions.”

To return to the “grammatical” cases: all Caucasian languages distinguish an absolutive or nominative case and an ergative case. The nominative, as in IE, is the case of citation, of the complement of a copula, usually of the subject of an intransitive verb, sometimes the subject of a transitive verb, and finally, unlike IE, the usual case of the object of a transitive verb. The ergative is most commonly the case of the subject of a transitive verb. We have to hedge these remarks with “usually” and “commonly” because the function of these case-forms in Caucasian languages are more varied than is often supposed.

In about 10 Caucasian languages the ergative case-form has the unique function of marking the subject of a transitive verb. In the other languages it combines the function of another oblique case; thus the ergative case-form also functions as an instrumental in Avar and the Andi languages, as genitive in Lak, as dative in the Adyghan languages, as a locative in others.

The Adyghan languages have a simple case-system, embracing nominative, ergative-oblique, instrumental and a fourth case called translative or adverbial which has several functions including the important one of marking the pivotal noun in relativizations. In the NWC languages there is no genitive case, the attributive possessive relationship being expressed by pronominal prefixes on the noun referring to the possessed, thus “the boy's hat” is expressed as k'alem ji-parer, i.e. “the boy his-hat.” In Adyghe, incidentally, but in no other Caucasian language, a distinction is made between “organic” or “inalienable” possession and “material” or “alienable” possession, the latter being indicated by an extended form of the pronominal prefix, thus s-pe “my nose,” but si-par e “my hat”.

Outside of NWC all Caucasian languages distinguish nominative, ergative, genitive, and dative cases, and often instrumental. A few languages have additional “grammatical” cases, one of these being the affective found in six of the Andi languages and in the Lezgian language Tsakhur. The affective case marks the subject of certain transitive verbs of perception—in other languages such verbs usually have their subject in the dative or a locative case. A recent logico-semantic study of case in a Nakh language is Deserieva's Structure of the Semantic Fields of Chechen and Russian Cases (in Russian) (20).

In general, pronouns in Caucasian languages present the same system of case distinctions as nouns, with the exception that in NWC, Kartvelian, and 10 Dagestian languages no formal distinction is made between the nominative and ergative of the
first and second person pronouns. Incidentally, the pronoun systems of Caucasian languages include the distinction between inclusive and exclusive first person plural in Nakh, Avar, the Andi languages, about half the Lezgian languages, and the Kartvelian language Svan. There are no specific personal pronouns of the third person, demonstratives being used in this function. These usually distinguish three degrees of deixis (type "this-that-yon"). In a number of Dagestanian languages, however, the demonstratives also systematically incorporate elements signifying "higher" or "lower," distinguishing, for instance, "that-on-this-level," "that-higher," "that-lower." This is interesting, though not surprising, in languages spoken in auls or mountain villages, where the roof of one house is frequently the verandah of the house above it.

**Locative expressions**

In Caucasian languages there are three different ways of expressing spatial relations: postpositions, preverbs, and local cases. All Caucasian languages make some use of postpositions, which are often transparently derived from (or merely special uses of) nouns, and this is the chief type of locative expression in Kartvelian. The NWC languages, particularly Adyghan, use preverbs, and the Nakh and Dagestan languages local cases.

The locational preverbs of the Adyghan languages are prefixed either to what are called "static" verbs such as be, lie, sit, stand, or to "dynamic" verbs such as go, take, write, etc. The preverbs themselves carry exclusively orientational meanings; that is to say, they specify whether the relation holds toward the orientating object as a whole, to its exterior, to its interior, to its underside, etc. They do not specify whether the relationship is a static one (essive), or involves approach to the orientator (ative), or departure from it (ablative), or both (translative). It is the verb which carries these latter meanings. Thus the Kabardian preverb de, for instance, means simply "relationship to the interior of a horizontally bounded (empty) space," and can thus be translated in, into, or out of; according to what verb it is prefixed to, as in these examples: 1. *He is in the courtyard*, 2. *He went into the courtyard*, 3. *He went out of the courtyard.*

1. ar pš'ant'em de-t-$\ddot{\text{s}}$. "He courtyard in-stand-s."
2. ar pš'ant'em de-ha-$\ddot{\text{s}}$. "He courtyard in-went-to."
3. ar pš'ant'em de-k'as. "He courtyard in-went-from."

Other Kabardian preverbs include xe, "in/into/out of a filled space—e.g. water, sand, a forest, the Party"; tej "on/onto/off," s'e "under/to under/from under," etc.

What NWC performs by preverbs the Dagestanian languages achieve by means of local, or space-relational, case-forms, augmented where necessary (as are the NWC preverbs) by postpositions. The number of local cases ranges from 5 to 6 in one or two of the southernmost Dagestanian languages spoken in northern Azerbaijan to 32 in Lak (which with 8 grammatical cases has a total of 40) and 43 in literary Tabasaran, which with 4 grammatical cases gives a total of 47—though one Tabasa-
ran dialect may still retain an old systematic distinction, raising the total of local cases to 49 and the grand total to 53.

In most Dagestan languages the local cases form an extraordinarily regular system formed by the combination of two sets of case-suffixes. One set, which is always attached directly to the noun stem, carries the orientational meanings (*on, in, under, etc*), and in Dagestanian grammars there are said to be as many “series” of local cases as there are different orientational suffixes. The second set of case-suffixes is attached to the preceding ones and specifies the directional meanings (static, approach, departure, etc). We can illustrate this from a language with a moderate set of local cases, Tsez, which has 4 directional cases—essive, lative, ablative, and translative—in 7 series, which [following Bokarîv’s (6) numbering] are I *on* a horizontal surface, II *in* an empty space, III *in* a filled space, IV *at around*, V *under*, VI *on* a non-horizontal surface, VII *by, with* (e.g. possession).

This gives us a total of 28 cases. Probably no noun takes all 28 of them since there are obvious semantic incompatibilities, but in principle all 28 combined case-forms occur. Here are the forms of only 12 cases, followed by some examples:

<table>
<thead>
<tr>
<th></th>
<th>Essive</th>
<th>Lative</th>
<th>Ablative</th>
</tr>
</thead>
<tbody>
<tr>
<td>I. <em>on</em> (horizontal)</td>
<td>-(o)</td>
<td>-(o)r</td>
<td>l(a)j</td>
</tr>
<tr>
<td>VI. <em>on</em> (non-horizontal)</td>
<td>-t4(o)</td>
<td>-t4’-or</td>
<td>-t4’-aj</td>
</tr>
<tr>
<td>II. <em>in</em> (empty)</td>
<td>-q(o)</td>
<td>-q-or</td>
<td>-q-aj</td>
</tr>
<tr>
<td>III. <em>in</em> (filled)</td>
<td>-a/æ</td>
<td>-a-r</td>
<td>-a-j</td>
</tr>
</tbody>
</table>

Examples: 4. There’s bread on the table. 5. The bread fell off the table. 6. There’s a fly on the wall. 7. He picked a leaf off the tree.

4. istoli-t 4’ betʃɔ magalu. “table-on lies bread.”
5. istoli-t 4-aj bokys magalu. “table-on-from fell bread.”
6. qido-q jo 4’ut’. “wall-on is fly.”
7. qunno-q-aj jet’urno 4ʼeb. “tree-on-from he-picked leaf”

8. There’s water in the jar. 9. Pour the water out of the jar. 10. There’s a worm in the apple. 11. Pour salt into the gruel. 12. A bear is going out of the woods.

8. 4’i tum-æ jo 4’i. “water jar-in is.”
10. henefjo-4 jo-4 bikori “apple-in is worm.”
11. qiqo-4-er tsijo 4’at 4’o. “gruel-in-to salt pour.”
12. tsiqe-4-aj biX zej. “woods-in-from goes bear.”

It will be observed that although the locative expressions of NWC and Dagestanian appear on the surface to be very different, using preverbs on the one hand and local cases on the other, they nevertheless have certain typologically interesting similarities. In both systems the two components of locational meanings, the *orientator* and the *direction*, are separately expressed; in NWC by preverb and verb, in
Dagestan by first case-suffix and second case-suffix. Second, both incorporate certain conceptual distinctions in the system which are not systematically expressed in more familiar languages. These include the distinction in an empty space vs in a filled space, found in both Adyghe and Kabardian in NWC, and in 13 Dagestanian languages, and also the distinction between on a horizontal surface and on a nonhorizontal surface, found in Kabardian, at least, and also in 10 Dagestanian languages. For those who wish to study the locational expressions of Dagestanian languages more closely, there is a valuable article on precisely this topic by Kibrik (45).

**Verbs**

As we indicated earlier, the Caucasian verb exhibits greater complexity in NWC than anywhere else. To some extent there is a trade-off between verbal complexity and case-systems: the more relational information you pack into the verb-form the less you need to express by means of nominal case-forms. This is far from being absolutely true, however, and, as Khadjakov (44) points out, relational preverbs are found even in some of the multicasal Lezgian languages, including Tabasaran.

The complexity of the NWC verb is well illustrated by the Adyghe and Kabardian languages, in which the verb may carry exponents of its static or dynamic status, of transitivity, finiteness/nonfiniteness, polarity (affirmative-negative), person and number of related NPs, up to eight tenses, mood, causative, version (action performed for or against the will of a person etc) association, reciprocity, common action, potential, involuntary action, and a number of other modal nuances. In addition to all of these the verb may carry locational and directional preverbs and up to four pronominal prefixes referring to NPs which may or may not actually be present in the sentence. Kumakhov describes and exemplifies many of these complexities in his books of 1964 (58) and 1971 (60) and also in a useful short article in a 1965 symposium on typology (59). In this latter article he gives a number of Kabardian examples, of which the following is a moderately simple one. 13. If I made you(sing) go in together with them = wadixezhehame which can be analyzed as:

13. w - a - dí - xe - z - ḥ - a - me.
   thee them with in I cause enter past if.

In all the NWC languages there is a single, fairly rigid, surface sequence of NPs in the sentence, namely S O V. There are, however, two distinct sequences for the pronominal prefixes on a verb: O - S - V or S - O - V, the distinction being of great syntactic importance since it distinguishes the ergative construction from the nominative construction.

Though the Nakh and Dagestanian verbs are less complex than those of NWC, they still exhibit a rich variety of tenses, moods, and other forms. In most of these languages some, but by no means all, finite verbs carry class-indices relating to an associated NP. The NP with which the verb agrees in class is always one in the nominative case, namely, the subject of an intransitive or the object of a transitive verb. In a few Dagestanian languages transitive verbs also agree in person with the subject. In two languages, Lezgi and Agul, the verb carries no concord markers whatsoever.
**Syntax**

I can merely sketch the most striking syntactical feature of Caucasian languages—the ergative construction. Typically, in the ergative construction, the *subject* of a transitive verb is in an oblique case (*ergative*), while the *object* is in the same unmarked or *nominative* case as is the subject of an intransitive verb. The ergative construction may thus be represented as $S^E V O^N$ while the intransitive, nominal construction is $S^N V$—an example from Avar: 14. The hunter killed the wolf. 15. The wolf died.

14. $t'\text{fanaqan-as bats' } t'\text{f'wana.}$  
   *hunter$^E$ wolf$^N$ killed.*

15. bats' $\chi$wana.  
   *wolf$^N$ died.*

In Caucasian languages there are many variations on the theme of ergativity. I have described a number of these elsewhere (11, 12) and so will simply outline the most important points here.

Though the configuration $S^E V O^N$, (by the way, the sequence SVO is purely arbitrary—many Caucasian languages require, or else favor, the sequence SOV, but the sequence SVO is very common) represents the basic or unmarked ergative construction, in most Caucasian languages transitive verbs of perception and feeling have their $S$ in a different oblique case. Nine Dagestan languages have a special affective case for this function. In others the subject of such verbs is in the *dative* or *locative* or *ablative* or some other case.

The second widespread characteristic of the ergative construction is that in those Nakh and Dagestan languages which have a class-marker on some verbs, the class-marker always reflects the noun which is in the *nominative* case, i.e. the object of a transitive verb, but the *subject* of an intransitive verb. Thus in Avar: 16. Father loves his son. 17. Father loves his daughter. 18. Father loves his horse.

16. insuje was $w$-ot$^4$ula.  
   *father$^D$ son$^N$ loves.*

17. insuje jas $j$-ot$^4$ula.  
   *father$^D$ daughter$^N$ loves.*

18. insuje $t'\text{u b}$-ot$^4$ula.  
   *father$^D$ horse$^N$ loves.*

but, on the other hand: 19. Father came, mother came, the horse came.

19. emen $w$-at$'\text{ana, ebel j-at}$ $'\text{ana, } t'\text{u b-at}$ $'\text{ana.}$

In three Dagestan languages—literary Tabasaran, Dargi, and to some extent in Lak—in transitive sentences the verb agrees with the object in class, but with the ergative *subject* in person. In southern Tabasaran dialect and in Udi there are no classes, and the verb agrees only with the ergative subject, in person. Udi, incidentally, is perhaps unique among the world's “ergative languages” in that it has an ergative construction $S^E V O^{D/A}$, that is, with the subject in the ergative case and the object in another oblique case—the “dative-accusative.” The most recent descriptions of Udi are in (21a) and (30a). In NWC and Kartvelian the verb agrees
in person with both subject and object, both of which are, in principle, marked in the verb form.

In NWC the sequence of pronominal prefixes is crucially related to ergativity. In the *ergative* construction the prefixes are in the sequence O-S-V. In the contrasting *nominative* construction they are in the sequence S-O-V. In Abkhaz and Abaza, where nouns carry no surface case markings, the distinction between the nominative and ergative construction is made solely by the difference in sequence (and a few differences in form) of the pronominal affixes.

The Kartvelian languages Georgian and Svan are anomalous in that in them the ergative transitive construction $\text{SE} \ V \ \text{ON}$ occurs only with verbs in the aorist set of tenses. In the present series, the construction of transitive sentences is $\text{SN} \ V \ O^{D/A}$, i.e. subject in nominative, object in dative-accusative. In the two Zan dialects, however, the ergative construction has generalized from exclusive use with aorist transitives in two distinct ways. In Laz, *all transitives* (present as well as aorist) have the construction $\text{SE} \ V \ \text{ON}$. In Megrelian, on the other hand, the generalization has gone the other way, so that *all aorists* (transitive or intransitive) have their subject in the ergative, $\text{SE} \ V \ \text{ON}$, $\text{SE} \ V$. Megrelian can thus hardly be said to possess an ergative construction any more, since the ergative case-form is completely dissociated from transitivity.

So far, we have referred only to the construction $\text{SN} \ V$ and $\text{SE} \ V \ \text{ON}$. But most, perhaps all, North Caucasian languages have a second type of transitive, or on both morphological and semantic grounds, a "semitransitive" construction. In the Adyghan languages this nominative semitransitive construction has the configuration $\text{SN} \ V \ \text{O}^{E}$, that is, subject in the nominative case, object in the ergative-oblique case, and pronominal prefixes on the verb in the sequence S-O-V. This opposition of the two transitive constructions, $\text{SE} \ V \ \text{ON}$ and $\text{SN} \ V \ \text{O}^{E}$, is not merely formal; it is functional, i.e. meaningful.

A Kabardian example of the opposition between the ergative (transitive) construction $\text{SE} \ V \ \text{ON}$ and the nominative (semitransitive) construction $\text{SN} \ V \ \text{O}^{E}$ is given in sentences 20 and 21, both of which can be roughly translated as "The boy is reading the book."

20. $\text{s'alem txi'ir je-dʒe}$.  
   boy$^{E}$ book$^{N}$ reads.

21. $\text{s'aler txi'im jew-dʒe}$.  
   boy$^{N}$ book$^{E}$ reads.

Jakovlev has clearly described the semantic distinction between these constructions in both his Adyghe and Kabardian grammars (38, 39) and in an important article published in 1940 (37). In this article he describes the ergative construction as "aim-full," implying the actor's intention to carry the action to completion and "to full penetration into the object," whereas the nominative construction is "aimless"; the action is not necessarily completed, and "the action in this case makes only superficial contact with the object." The construction of example 21 suggests comparison with an English intransitive verb + prepositional phrase: "He's reading in the book," or even better, the Scots expression "He's haein a bit read o' the buik." It would, however, be an error to conclude that example 21 is just an intransitive
verb with a locative complement. We know that this is not so since in the SN V OE construction the verb is in neither the normal transitive form jedze nor the intransitive form madze, but in a special form which I call semitransitive. In a normal sentence with locative complement the verb would be in the intransitive form.

Analogous semitransitive, or "nominative-transitive" constructions occur in all NWC languages, all Nakh languages, and several (perhaps all?) Dagestanian languages. In some they are closely similar to the NWC SN V OE construction. Dargi, for instance, contrasts the regular SE V ON with SN V OE, whereas the oblique case used in the semitransitive construction in Hunzib is instrumental, SN V OI. In the Nakh languages and in Avar, Tsakhrur, and Khinalug (and probably others) the semitransitive construction is SN V ON, i.e. the object, like the subject, is in the nominative case. In all of these examples the opposition between the ergative transitive and the nominative (semi-) transitive is of the same general kind. The ergative construction implies a tight, penetrative association between the verb and its object. The nominative construction, on the other hand, implies a tight relationship between the verb and its subject—it stresses the activity of the subject rather than the effect upon the object.

In the anomalous Kartvelian languages there is no direct opposition between the ergative transitive and the nominative transitive construction. Either, as in Laz, the ergative construction is obligatory with all transitive verbs, or, as in Georgian and Svan, it is excluded from the present tenses. In either case there is no freedom to select the ergative or the nominative construction in order to express different nuances in the relationship between the transitive verb and its associated NPs. The Georgian-Svan restriction of the ergative construction to aorists is comprehensible in light of what was said above. The ergative construction, being effect-oriented, has a certain affinity with pasts and perfectives, whereas the nominative construction, being more action-oriented, has more affinity with presents and imperfectives. In Georgian-Svan the opposition between the ergative and nominative constructions has been absorbed, so to speak, in the tense-aspect opposition and thus has no independent semantic function.

In the Kartvelian languages, the ergative construction is obligatory with transitives, and further restricted in Georgian and Svan to the aorist, and is thus purely formal and meaningless. In North Caucasian languages the ergative construction is in meaningful contrast with the nominative construction and is thus functional. This distinction is not confined to the Caucasus and, as I have suggested elsewhere (12), one can classify all the world's "ergative" languages as possessing either a formal or a functional ergative construction.

It is not surprising that there is a very considerable literature on ergativity in Russian. This includes two collections of articles. One of these (22) brings together many of the older key works on the topic (all in Russian translation), while the other (23) consists of specially written articles of which 10 are on Caucasian languages. There are also two important books on the subject by leading authorities on ergativity: Mēščaninov's work (64) and Klimov's (52) excellent "General Theory of Ergativity."
I would like to mention one other syntactic characteristic of north Caucasian languages, namely the fact that in them virtually all forms of sentential conjoining and subordination, including relativization, are carried out by verbal means—i.e. by the use of special conjunctive verb forms, participles, verbal nouns ("masdars")—and not by the use of conjunctions, which are few, or relative pronouns and adverbs, which are nonexistent. The Kartvelian languages again are anomalous since they possess, and use, conjunctions and relative pronouns.

The following examples give some indication of how relativization works in Adyghe: 22. The girl who brought the old man said (something). 23. The guest whom the boy brought has gone away again. 24. The village where the guest went is pretty.

22. psashew q'ezir zi-ca'cem i?wa'b.
girlA(N) old manN (E)-having-broughtE said-something.

23. hak'ew k'alem qica'ber k*ez{i}be.
guestA(N) boyE hither-having-broughtN gone-again.

24. tqi'lew hak'er zidek*al'er tcile dax.
villageA(N) guestN (E)-having-gone-toN village pretty.

In each of these the pivotal noun is in the so-called "adverbial" case, ending in \( w \) and labeled "A" for "adverbial" with a parenthetic indication of the case—ergative-oblique (E) or nominative (N)—it would have had in the underlying sentence. The participal expression carries the prefix zi- when the related pivotal noun is underlyingly ergative (whether subject, as in 22, or indirect object, or relational object as in 24, etc). The whole phrase is marked by the case-suffix on its last component (the participle) as ergative, -m in 22, or nominative -r in 23 and 24, according to the transitivity of the sentence in which it is embedded.

Finally, here are two sentences illustrating relativization in Avar: 25. The boy to whom Musa gave the money went away, and 26. The day when Ahmad came we had gone to the Institute.

25. musatsa qarat s q'uraw was ana.
MusaE moneyN given boyN went.

26. ahmad w-at /ara-b q'ojasi ni3 institutalde un r-uk'ana.
AhmadN having-come dayE we(excl.)N inst.-to gone were.

Note that in 26 the participle agrees in class, like an intransitive verb, with its subject, \( w-\) (m) but also, like an adjective, with the noun it is attributive to, \(-b\) (n). The ergative/instrumental case on the word for \( day \) is used in its time-marking function.

INTERNAL AND EXTERNAL RELATIONS OF CAUCASIAN

Before discussing the question of the relationships among the Caucasian languages themselves, and between them and other language families, it will be well to summarize some of their characteristic features. This can perhaps best be done by tabulating various phonological and grammatical characteristics which have been discussed
above, indicating by +, (+) (= an ‘attenuated’ +), or −, their presence or absence in the various subgroups: Kart(velian), NWC, Nakh, Dag(estanian), and with an additional column for Indo-European (IE) for purposes of comparison.

<table>
<thead>
<tr>
<th></th>
<th>Kart</th>
<th>NWC</th>
<th>Nakh</th>
<th>Dag</th>
<th>IE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Glottalic consonants</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>very rare</td>
</tr>
<tr>
<td>2. Uvulars</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>rare</td>
</tr>
<tr>
<td>3. Pharyngals</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>4. Labialized consonants</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>(+)</td>
<td>very rare</td>
</tr>
<tr>
<td>5. Non-approximant laterals</td>
<td>−</td>
<td>+</td>
<td>(+)</td>
<td>+</td>
<td>very rare</td>
</tr>
<tr>
<td>6. “Harmonic complexes” of CC</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>7. Minimal vowel system</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
</tr>
<tr>
<td>8. Noun classes</td>
<td>−</td>
<td>(+)</td>
<td>+</td>
<td>+</td>
<td>(+)</td>
</tr>
<tr>
<td>9. More than 2 or 3 “grammatical cases”</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>10. Numerous local cases</td>
<td>−</td>
<td>−</td>
<td>(+)</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>11. Directional preverbs</td>
<td>(+)</td>
<td>+</td>
<td>−</td>
<td>(+)</td>
<td>(+)</td>
</tr>
<tr>
<td>12. “Orientator” and “direction” separately expressed</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>13. Unusual semantic distinctions of on and in</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>14. Ergative construction</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
<td>−</td>
</tr>
<tr>
<td>a. functional</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td></td>
</tr>
<tr>
<td>b. formal</td>
<td>+</td>
<td>−</td>
<td>−</td>
<td>−</td>
<td>(+)</td>
</tr>
<tr>
<td>15. Conjoining, Relativization etc by verb</td>
<td>−</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>(+)</td>
</tr>
</tbody>
</table>

It is clear that there are not many traits common to all Caucasian languages—only 3 out of the 15 listed above. And yet anyone working with these languages receives a strong impression of “family likeness” running through all of them. Whether the 37 Caucasian languages constitute a single “language family” derived from a common Proto-Caucasian, or whether the typological resemblances between them result from the convergence of features in languages of more than one “genetic” origin through millennia of symbiosis, is still an open question. Some Soviet linguists, particularly in Georgia, seem to accept the common-origin hypothesis as self-evident, and merely await definite proof. Others are agnostic and prepared to believe that Caucasian languages constitute a Sprachbund, or language union.

The establishment of relationships between languages requires above all the setting up of regular sound-correspondences between the languages constituting a presumptive “family,” and this task has proved extremely difficult, particularly within the Dagestani and NWC subgroups of Caucasian. We can confidently accept Kartvelian as a genetically related group; much of the evidence is available in (24, 49, 73), to which I have referred above. The relationship of the Nakh languages, too, is well established, on which see (18, 78). The Dagestani languages are more problematic. They obviously fall into a few subgroups: Avaro-Andi, Tsez, Lezgian, and some scholars regard Lak and Dargi as forming a single close-
knit Lak-Dargi subgroup. The publication in 1961 of Bokarëv's *Introduction to the Comparative Study of Dagestanian Languages* (in Russian) (7) was an important advance toward the establishment of the sound-correspondences within Dagestanian as a whole, and Bokarëv's reconstruction of a Proto-Dagestanian sound system is still generally accepted by specialists. The relationships of languages within some Dagestanian subgroups were placed on a solid footing by Bokarëv's work on the Tsez languages (6) and Gudava's on the Andi languages (29), and the study of Dagestanian as a whole has been facilitated by comparative Dagestanian word lists, the one edited by Murkelinskij (67), the other compiled by Khajdakov (43). Between them, these works deal with about 1000 Dagestanian words. However, far from all these items are represented in all Dagestanian languages, and the cognate status of some is dubious. And though the 1971 book lists and exemplifies all the consonantal correspondences of Dagestanian, one is struck by the number of apparent irregularities.

The fact is that in dealing with Dagestanian languages we are faced with a state of affairs quite different from what we find in some of the better known language families, particularly in Indo-European. In Indo-European we know that most languages and language groups have developed for centuries, even millenia, in total separation from each other. We can be virtually certain, for instance, that regular correspondences between Sanskrit and Germanic or Celtic are due to common origin and not to borrowing. In Dagestan (and to some extent in the Caucasian generally) we have a totally different situation, namely a group of languages which have been spoken in very much their present location for probably 4000 years or more. The greatest straight-line distance between any two Dagestanian languages is about 150 miles. While the extremely mountainous terrain probably doubles that maximum distance, nevertheless one must assume some degree of contact throughout the millenia. Strabo explained the great number of different Caucasian languages by the fact that their speakers lived in scattered groups without mixing, because of "their stubbornness and ferocity." But at the very least, there must have been intertribal wars and the taking of prisoners, particularly women. So it is possible that one reason for the difficulty in establishing regular sound-correspondences in Dagestanian is that genetic relationship is heavily overlaid by borrowing throughout millenia of symbiosis.

The NWC languages have so many striking common traits that one is easily convinced of their common origin. They clearly fall into two closely related subgroups, Abkhaz-Abaza and Adyghan, with Ubykh as an isolated language. Thanks very largely to the work of such scholars as Jakovlev and Kuipers, the sound-correspondences of Adyghan are well established. Kuiper's 1975 *Dictionary of Proto-Circassian Roots* (57) is an invaluable aid to the further study of NWC. Nevertheless, we are still far from establishing a thoroughly satisfactory "Common NWC." The Abkhaz linguist Shakryl has published what might be called "comparative-descriptive" works on NWC (75, 76) in which he lists about 500 words which seem to be common to all or most of the NWC languages. However, his lists contain inaccuracies, and no attempt is made to identify borrowings (within NWC or from other languages, e.g. Arabic and Persian). More convincing are the lists of NWC cognates given by Shagirov (74) and Klimov (51). Between them these provide us
with about 120 very plausible cognates. However, as Kuipers (55) has well explained, the problem of establishing regular sound-correspondences is very great when one is dealing with languages in which most, if not all, roots consist of a single consonant or consonant cluster, particularly when the consonant systems are very rich and there are numerous homonyms. In these conditions it is difficult to find series of correspondences which confirm each other, as one can in, for example, Indo-European.

In spite of these difficulties, there can be little doubt that each group within Caucasian—NWC, Nakh, Dagestanian, and Kartvelian—constitutes a genetically related unit. It is the wider relationships within Caucasian as a whole which are more dubious. The number of pan-Caucasian cognates is very small: Klimov (50) lists only 22 of these, embodying 28 sound-correspondences, only one of which occurs in two items.

It is interesting to note that the general impressions of relative closeness of relationships between Caucasian languages are reflected in what little attempts have been made at lexicostatistical comparisons between Caucasian languages. Klimov (50) reports results obtained by Tovar (82) for comparisons between Georgian and Adyghe, Georgian and Avar, and Adyghe and Avar. Klimov himself (48), using Swadesh’s 100-word list, has compared the three Kartvelian languages, and the present writer has carried out some very tentative lexicostatistical comparisons between the NWC languages, and between them and Chechen and Avar, and within Dagestanian, between Avar and Akhwakh, Hunzib, and Lezgi. The last three are very dubious, having been carried out with an incomplete set of 80 to 90 items. For what they are worth, however, I set out the figures for percentage of cognates for these comparisons.

**Within Kartvelian (Klimov)**

- Georgian-Svan 30%, Zan-Svan 30%, Georgian-Zan 44%

**Within NWC (JCC)**

- Abkhaz-Adyghe 27%, Abkhaz-Kabardian 27%, Abaza-Adyghe 31% (average for Abkhaz-Abaza and Adyghan 28%), Ubykh-Abkhaz and Ubykh-Adyghe 36%, Ubykh-Abaza and Ubykh-Kabardian 40%, Abkhaz-Abaza 80%, Adyghe-Kabardian 92%

**Within Dagestanian (JCC)**

- Avar-Akhwakh 48%, Avar-Hunzib 34%, Avar-Lezgi 27%

**NWC-Kartvelian (Tovar)**

- Adyghe-Georgian 7.5–5.14%

**NWC-Dagestanian**

- Adyghe-Avar (Tovar) 12.9–9.75%, (JCC) 14%–10%, Abkhaz-Avar, (JCC) 13–8% (average for NWC-Dag. 13.27–9.25%)

Any attempt to attach absolute time-depths to these lexicostatistical percentages is fraught with great difficulties. The glottochronological standard rate of cognate loss is somewhat dubious in general, and it is likely to be influenced in particular ways in the Caucasus. For instance, if longevity and great respect for elders have been characteristic of Caucasian peoples for many centuries (as they are today), one
that we briefly adumbrated it. Caucasian-Caucasian. The language of the ancient Maikop culture is dated at about 2500-2000 B.C., and one theory is that it was brought about by a northward migration from the Black Sea coast. It is at least possible that the major NWC divergence began at this time, the speakers of what became Abkhaz remaining on the coast. Any glottochronological interpretation of the lexicostatistic data in terms of time-depth of separation of the various groups within Caucasian as a whole is meaningless unless one definitely accepts the view of a common origin for all Caucasian languages.

Finally, I must leave the speculative area of glottochronology to consider very briefly another which is equally speculative—the possible external relations of Caucasian. The ergative construction is the Caucasian trait which has, for a century or more, aroused the interest of scholars and has prompted suggestions of relationship with virtually any language that has an ergative construction, notably Basque, Burushaski, Paleo-Siberian, and among ancient "ergative languages" Sumerian and Urartian.

That the Caucasian languages, isolated in the mountains of the extreme east of Europe, and Basque in the extreme western mountains might be remnants of a language or language-family once spoken all over Europe is not in itself implausible. However, during the century or so that this hypothesis has been current no fully convincing evidence has been forthcoming. The French linguist, Lafon, has published more than anyone on this topic (see 63), but the sound-correspondences which he has adduced are not supported by many examples, and necessarily carry less conviction than they might because there is as yet no well-established Proto-Caucasian with which to compare Basque or any other language.

The relationship with Burushaski, the isolated language of the Karakoram, is based on little more than the fact that Burushaski has an ergative construction and a few rather random word-equations. The Caucasian-Burushaski hypothesis put forward by Bleichsteiner was reviewed and criticized by Morgenstierne (66).

Gamkrelidze & Machavariani (24) have clearly shown similarities between Kartvelian and Indo-European, but they regard them as no more than simply typological resemblances. Incidentally, the student of NWC also cannot but be impressed by the resemblance between certain NWC consonant systems, with their triadic series of stops, labialized velars and uvulars, plain and labialized "laryngeals" and so on, and reconstructed Proto-Indo-European.

The most ambitious attempt at a wider synthesis of Caucasian with other language families is that of Illič-Svityč, who, in a posthumous work published in 1971 (34), develops the "Nostratic Theory" first adumbrated in the latter half of the
nineteenth century. Illič-Svityč purports to show a relationship between Kartvelian, Indo-European, Altaic, Uralic, Dravidian, and Semito-Hamitic. He sets up sound-correspondences not (as so many other have done) between isolated languages of these different families, but between their respective proto-languages. His results are interesting and indeed suggest some form of relationship but whether of a "Sprachbund" or of a "genetic" nature is difficult to say.

It would not be surprising if Caucasian languages were related to one or another of the ancient languages of the Near East, and the most plausible candidate for this relationship is Urartian, the ancient language spoken in the area of Ancient Armenia before the Indo-European speakers of Armenian arrived there. Braun & Klimov (9) have adduced some material correspondences between Urartian and North Caucasian.

The internal and external relations of Caucasian languages are still uncertain, and present an intriguing challenge to linguists—a challenge which is being met by continuous intensive research by Soviet linguists, many of whom are themselves native speakers of Caucasian languages, and by a very small number of scholars in the rest of the world.

Literature Cited

22. Ergativnaja konstrukcija predloženiya 1950. Moscow: Izd. inostrannoj literaturnoj
23. Ergativnaja konstrukcija predloženija v jazykax različnych tipov 1967. Lenin-
grad: Nauka
Britannica, pp. 1011–15
Languages of the Caucasus. Janua Ling-
varum, Series Minor 6. 'S-Graven-
hage: Mouton
29. Gudava, T. E. 1964. Konsonantizm an-
30. Guksjan, V. 1969. Opjet desifrovki al-
banskix nadpisej Azerbaidžana: Izv.
iskusstva No. 2, pp. 52–74
30a. Guksjan, V. 1974. Udinsko-azer-
bažansko-russkij slovar' (with grammati-
cal sketch in Russian). Baku: Elm
31. Halle, M. 1970. Is Kabardian a vowel-
less language? Found. Lang. 6:95–103
32. Holmer, N. M. 1947. Ibero-Caucasian
as a linguistic type. Stud. Ling. 1:11–44
33. Ibragimov, G. X. 1968. Fonetika tsax-
urskogo jazyka. Makhachkala: Dages-
tan. Filial Akad. Nauk SSSR
36. Jakovlev, N. F. 1923. Tablity fonetiki kabardinskogo jazyka. Moscow: Tra-
vaux de la section des langues du Cau-
case septentrional de l'Institut oriental à 
Moscou
37. Jakovlev, N. F. 1940. Drevnye jazy-
kovye sviazii Evropy, Azii i Ameriki Izv.
38. Jakovlev, N. F. 1948. Grammatika lit-
eraturnogo kabardino-čerkesskogo jaz-
41. Kaldani, M. 1969. Svanuri enis ponet'
ik'ta: l umlaut'is sist'ema svanurši (Pho-
netics of Svan I, Umlaut system in Svan —German with Russian summary). Tbilisi: Metsniereba
42. Khajdakov, S. M. 1967. Arčinskij jazyk
Moscow: Nauka
44. Khajdakov, S. M. 1975. Sistema glag-
gola v dagestanskix jazykaz. Moscow: Nauka
45. Kibrik, A. E. 1970. K tipologii pro-
stranstvennyx značenii (на материа-
46. Kibrik, A. E., Kodzazov, S. V. 1970. Printispy fonetičeskoj transkriptsi și transkriptionnaja sistema dla kavkaz-
47. Kibrik, A. E., Kodzazov, S. V., Olov-
jannikov, I. P. 1972. Fragmenty Gram-
matiki Xinaluzskogo Jazyka. MGU
Publ. Otd. Strukt. Frikl. Ling.: Ser. Mo-
ogr., Vyp. 9. Moscow
slovar' kartvel'skih jazykov. Moscow: Akad. Nauk SSSR
Moscow. 1969. Die kaukasischen Spra-
chen. Hamburg: Buske Verlag
51. Klimov, G. A. 1967. Abzaxsko-
adygskie etimologii I. Iskonnyj fond.
Etimologija 1965: 296–306
varum Series Minor 8 'S-Gravenhage:
Mouton
beek, pp. 315–44


89. Zhghenti, S. 1960. *Kartveluri enata shdarebiti ponet’ik’a, I, martvislis age;belebis p;roblemi* (Comparative phonetics of Kartvelian languages I, the problem of syllable structure). Tbilisi Univ Press