Linguistic Areas
Convergence in Historical and Typological Perspective

Edited by
Yaron Matras
University of Manchester

April McMahon
University of Edinburgh

and

Nigel Vincent
University of Manchester
Contents

9 Complex Emotion Predicates in Eastern Indonesia: Evidence for Language Contact? Simon Macgrave
227

10 Another Look at Australia as a Linguistic Area Claire Bowern
244

11 Towards a Typology of the Siberian Linguistic Area Gregory D. S. Anderson
266

Index of Authors

Index of Language Families, Languages and Dialects
Index of Subjects

List of Figures and Maps

Figures

2.1 Degrees of similarity 34
2.2 Languages with overt DET-marking versus languages without overt DET-marking 41
2.3 Competing isoglosses 43
3.1 Output of Network for most conservative sublist, Romance and Germanic 61
3.2 Output of Network for least conservative sublist, Romance and Germanic 62
3.3 Network for Germanic, 'wong' coded as cognate English-North Germanic 63
3.4 Network for Germanic, 'wong' coded as loan into English 64
3.5 NeighbourNet for Indo-Iranian bibi sublist 66
3.6 NeighbourNet (Bryant and Moultion, 2006) for Indo-Iranian lolo data 67
3.7 Neighbour-joining tree of Australian data, drawn using Splittree 69
3.8 Splittree graph for 26 Australian languages 70
3.9 NeighbourNet graph for 26 Australian languages 71
5.1 Genealogical classification of Khoe-Kwadi 102
5.2 Genealogical classification of Tuu 102
8.1 Circum-Baltic languages 183
11.1 Russia and the languages of Siberia 267

Maps

5.1 Distribution of Khoean lineages in the early colonial period 101
8.1 The Circum-Baltic languages 183
9.1 Indonesia, showing the location of the languages discussed 230
On the Roles of Turkic in the Caucasus Area

Lars Johanson

1 Turkic in the Caucasus area

In this chapter I shall discuss briefly the roles of the Turkic varieties in the Caucasus area in order to illustrate some processes in historical contact linguistics. The term ‘Turkic’ will be used for ‘Turkic-speaking’, without any genetic or cultural implications. The general term ‘code’ refers to a language or a variety of a language. The following kinds of Turkic are found in the area today:

Azerbaijani (Azerbaijan: 'Azerbaijan'), is mainly spoken in the Republic of Azerbaijan and in Iran, by at least 20 million people, and there is a sizeable Azeri-speaking population in the south-eastern part of Georgia. Scattered speaker groups are found in Armenia, Daghestan and Nakhichevan. The linguistic borderline between Azeri and Turkish runs through East Anatolia. For Azeri dialects, see Çelebi and Doerfer (1959: 281).

Kumyk (kumyq) is spoken by about 280,000 people north of the Azeri area, in the lowlands on the north-easternmost fringe of Daghestan. The area extends from Derbent in the south to Achi-Su in the north, close to the lower course of the Terek river. In the south, the area is confined to a narrow strip, the middle part is interrupted by a Dargi-speaking zone. Dialects include Reinaq, Khazavysuv and Khaldaq (Benzing, 1959: 392).

Nogay (noy) has fewer than 80,000 speakers. The Black Nogay area extends between the lower reaches of the rivers Terek and Kuma in Northern Daghestan, while the area of Central Nogay is largely situated in the Stavropol territory. Small groups of White Nogay are found in Chechensia and Karachay-Cherkessia, on the Kuban river and its tributaries.

Karachay-Balkar (karaqay-malqara) is spoken by about 150,000 Karachais and 85,000 Balkars in the Alpine ravines on the northern flank of the Greater Caucasus range. The Karachais inhabit the uppermost river valleys in Karachay-Cherkessia. The Balkar live to the south, east and south-east of the Karachay area, in the higher valleys of Kabardino-Balkaria. Dialects include Kirchay, Malqar (Cherek), Balkans, Chegem and Khulum-Bezinga (Pritsk, 1959: 342, Appayev, 1960).

Small groups of Stavropol Turkmen (turgan, Russian türkmen) live near the speakers of Central Nogay of the Stavropol territory. Several brands of Turkic meet in the Caucasus area. Azeri belongs, like Turkish, to the western subbranch of the South-western – Oghuz – branch. Stavropol Turkmen belongs to the eastern subbranch of the same branch. Kumyk and Karachay-Balkar belong to the western, Kipchak-Kumian, subbranch of the North-western – Kipchak – branch. Nogay belongs to the southern – Azalo-Caspian – subbranch of the same branch.

For the purposes of this chapter, the Caucasus area will be divided into a Northern and a Southern sub-area. The Northern sub-area corresponds, from the European point of view, to Ciscaucasia and consists of the northern side of the Greater Caucasus range, the plains north of it, and the northern coastal region. The Southern sub-area consists of the southern coastal region, Transcaucasia, including extensive adjacent zones of Kartvelian, Armenian, Iranian and Neo-Aramaic varieties.

2 Migrations

The Turkic languages are genealogically related to each other in a straightforward way, and exhibit rather well-defined common features. This family's huge geographical area of distribution extends from the south-west, Turkey and its neighbours, over Western Turkistan to the south-east, to Eastern Turkistan and further into China. From there it stretches to the north-east, via southern and northern Siberia, up to the Arctic Ocean; and to the north-west, across western Siberia and Eastern Europe, – today even to Western Europe. In the past, the area also included compact Turkic-speaking regions in the Pontic-Caspian steppes, in the Crimea and so on.

There have been massive displacements of Turkic-speaking groups throughout their known history. Numerous movements away from relatively homogeneous speech communities have caused successive splits into groupings and subgroupings of genealogical descendants. The family has undergone differentiation into a few primary branches, known as Oghuz, Kipchak, Bulgar and so on. Through further differentiation, more specialized kinds of Turkic have emerged as secondary or tertiary branches. From here the family tree branches out further into dialect groups, and regional and local dialects. The continuous migrations have led to the separation of related varieties, so that these do not occur in clear-cut geographic clusters. About twenty Turkic languages, in the political sense, are used today, each consisting of a set of dialects plus a standard variety with a certain area of validity.
3 Contact processes

The continuous movements and displacements of Turkic groups have led to numerous contacts with other codes, both inside and outside the family. When dealing with traditionally nomadic populations, we may reckon with linguistically relatively homogeneous tribes; basic communities, groups of families living together under the same leaders, and sharing certain customs. Contact influence is already operative at this level, because of exogamy and other kinds of encounters with outside groups. There may be influences from mutually intelligible codes and from foreign codes. Similar phenomena typically occur at higher levels, within tribal confederacies. These alliances for common political purposes are not necessarily kin-based, united by ties of descent from Turkic ancestors. In the contact history of Turkic-speaking groups, such confederacies have usually been multidialectal and even multilingual.

The differentiation of Turkic has been caused partly by complicated contact processes, in which Turkic varieties have copied foreign linguistic elements from their neighbours - some more and some less. Intensive contacts with Mongolic, Iranian and Slavic varieties have resulted in convergence of different kinds (Johnson, 2002). Since the spread of Islam among the Turks from the tenth century onwards, most Turkic varieties have been influenced heavily by Arabic, mainly through Persian mediation. Because of contact-induced change, many Turkic codes have become more similar to genetically unrelated codes. Code shift has often taken place. The contacts have also induced numerous changes in non-Turkic varieties, or led to code shift of non-Turkic groups.

4 The introduction of Turkic

When and how did Turkic enter the Caucasian area? The processes of migration have been highly complex, and most answers given so far are partly speculative. There are many open historical questions with respect to the origins of the Turkic populations and their linguistic backgrounds. Turkic does not belong to the indigenous languages of the Caucasian area (Nakh-Daghestanian, Abkhaz-Circassian, Kartvelian), but was introduced relatively late, at least later than the ancestors of Armenian and some Iranian languages.

One early wave of immigration came from the north and included a number of different nomadic complexes. The history of the Eurasian nomads offers many examples of sudden and far-flung political expansion. Numerous tribal complexes with Turkic-speaking elements have passed into the Caucasian area. Turkic-speaking groups probably participated in the campaigns or political organization of the Hunnic tribal groups, which appeared on the northern slopes of the Caucasus in the middle of the fourth century. But it is unknown to what extent the Huns or their ruling groups can be defined linguistically as Turkic. Nothing is known about a Hunnic language.

This case teaches us an important lesson. The old nomadic complexes were linguistically and ethnically heterogeneous, comprising elements of different origin. The known designations refer to the representative groups of the tribal confederacies, but do not tell us which tribes were included. The ethnic or linguistic affiliation of a constituent tribe is not necessarily identical with that of the leading elite group of the complex. Titles are not limited to specific linguistic groups. Given the heterogeneous composition of the nomadic complexes, it is often impossible to determine with which tribes or under which tribes Turkic-speaking groups appeared in the Caucasian area. The same principles apply to conglomerates such as Bulgars, Western Turk, Khazars, Oghuz, Pechenegs, Kuman, Mongols, later Kipchak groups and so on.

- From the end of the fifth century, the Bulgars (Oghuz, Bulgars) appeared in the steppes along the Kuban' river north-west of the Caucasus, and later established the independent state of Great Bulgaria on the eastern Pontic steppes, with its centre on the Kuban' river. We do not know what languages the confederated tribes spoke.
- Under the rule of the large tribal confederacy of the Western Turk, numerous tribes moved westwards and spread out over the steppes as far west as the lower Volga. Again, we know very little about the linguistic and ethnic composition of the tribes dominated by the ruling groups.
- When the West Turk Empire was destroyed in 657-9, tribal groups known as Khazars (Kazani) established a realm that existed from the eighth to the eleventh centuries. It extended over the Volga – Caspian – North Caucasian – Pontic steppes including the Kuban' region and the Daghestanian foothills.
- The Khazars, whose state religion was Judaism, fought against the Arabs in the Caucasian region – for example, at Derbent and Daryal Gorge (Daryal drevis) in the upper Terek valley. One Bulgar conglomerate remained along the Kuban' river, in so-called Black Bulgaria, in the southern part of the Khazar realm. Newcomen belonging to Kipchak Turkic confederacies destroyed the Khazar state and finally absorbed its population, including the groups of Black Bulgaria.
- The Pechenegs roamed on the steppes up to the eleventh century and finally defeated in Bulgaria.
- The Kuman, who appeared at the middle of the eleventh century, destroyed the Khazar state. They entertained good relations with Georgia. Kuman groups served as auxiliaries in the Georgian army and helped to defeat the Seljuks in 1121.
- At the beginning of the thirteenth century, Mongol armies invaded the Pontic steppes, defeating the kumans. Under the so-called Golden Horde,
Derbend and the river Terek in the Caucasus area became the frontier against the Ilkhanids, the Mongol successor-state in Persia.

- Under the overlordship of the Western Turks, Oghuz tribes had migrated westwards rather early, from the end of the ninth century onwards; they were later defeated by Pechenegs and Khazars.

- Oghuz tribes also migrated south-westwards, some of them having advanced greatly by the second half of the tenth century. An Islamic tribe confederacy known as the Seljuk set out for conquests in Iran, the Arab countries and the Byzantine Empire. They seized Persia, overran Transcaucasia (1045), moved on to Mesopotamia and, exerting their dominion over the Caliphate (1055), conquered the greatest parts of its empire, finally taking possession of Anatolia. These federacies led to the formation of the south-western branch of the Turkic in Turkestan, Iran, Transcaucasia, Anatolia and the Balkans. Thus, with the Seljuk conquest in the eleventh century, Oghuz Turkic arrived in Azerbaijan.

The linguistic effects of the old migratory movements are difficult to assess. The early Oghuz in the Pontic-Caspian steppe regions, the Pechenegs and so on, are only known from titles and personal and tribal names. Not a single line in Khazar has been found. The only relatively well-attested old language is Kuman, as documented in the Codex Cumanicus (fourteenth century).

The political expansion of the old nomad complexes was not necessarily tantamount to the expansion of Turkic. Some of the incoming complexes were led by Turkic elites. Others were led by non-Turkic elites, though they contained Turkic subgroups. Interestingly, the armies under nominally Mongol leadership were predominantly Turkic-speaking; the Mongol expansion meant a breakthrough for the expansion of Kipchak Turkic, rather than for Mongolic.

5 The Turkic varieties

Let us summarise some basic facts about the individual Turkic varieties of the Caucasus area:

- Kurnik entered the area in the early Middle Ages. The territory of Kurnik has had vivid contacts with elements from the steppe from the Hunnic era onwards; that is, in the Bulgar, Khazar, Kipchak and Mongol periods. The Kurnik may have Dagestani (that is, Dagestani) origin, including speakers of Nakh-Dagestani who moved to West Kipchak at different periods. The language clearly goes back to the early kind of Kipchak Turkic that is known from the eleventh century onwards and may have spread into the north Caucasian area at the time of the Khazar empire. It became one of the Turkic languages of the lowland areas that functioned as a lingua franca in the region, serving trade and inter-group communication. Kurnik was dominant in the eastern part of the north Caucasian area, often used by Chechens, Avars and other tribes. Today, few non-Kurnikis learn the language, since Russian has taken over the inter-group communication function.

- Noghay arrived relatively late in the area, namely after the break-up of the Golden Horde at the end of the fourteenth century. It is a Kipchak language of a later type, not of the eastern type to which Kurnik belongs. The Noghays once formed a nomadic state on the lower Volga under Russian supremacy; they were later expelled from the area by the Kalmyks and scattered all over the Pontic-Caspian steppes. Most groups were absorbed by other Kipchak elements and gave up their language. Only the parts that lived in the north Caucasian area maintained it. The Noghays, originally Kipchakized Mongols, also comprise Turkicized groups of local origin. Noghay once functioned as a lingua franca in north-western and central Daghestan.

- The speakers of Stavropol Turkmen are descended from Turkomans tribes brought from the Mangyshlak region to the North Caucasus in the eighteenth century.

- The Karachay-Balkar language is an older enterer from the steppes. The ancestors of its speakers have lived in the north Caucasian plain, from where they were pushed into the mountainous areas by the Mongol invasion of the thirteenth century. Their new habitat was the home of Alan tribes, which had remained in North Caucasus since the Khazar era. The Karachay-Balkar developed a close symbiosis with the Alans and were strongly influenced by them. They have even been called Alans themselves by some of their neighbours. In the late Middle Ages, at the beginning of a period of cooling weather that caused highland tribes to move downhill, the Karachay-Balkar were driven by Kabardians and Circassians to poorer locations in the highlands.

Today’s Karachay-Balkar speakers seem to be a mixture of elements from Bulgar, Oghuz and Kipchak tribes, with Iranian Alans and assimilated indigenous Alkhor-Caucasian and even Kartsian (Kvane) elements. In particular, groups of Dguor Ossetic speakers have shifted to Karachay-Balkar. The Kabardians refer to the Karachay-Balkar as Ossetes, whereas the Dguor Ossetes call them As or Assi. The Karachay-Balkar differ from other Turkic groups in terms of specific anthropological and cultural features, sharing, for example, the Nart heroic epics with Ossetes, Kabardians, Alakhor and others. The language of the so-called Black Bulgars living in the Kuban region in the south part of the Khazar empire may be one of the basic elements of Karachay-Balkar. These Bulgars were later assimilated by other Turkic-speaking groups. The ancestors of the Karachay-Balkar were probably Kipchakized at the end of the eleventh century (Pritsker, 1939: 141). However, the attempts to connect bulgar – bulgar – bulgar with the name of the Bulgars seem to rest on a superficial sound resemblance.
Many different ethnic and linguistic elements have been involved in the history of Azerbaijan. Its northwestern part was the ancient Albania, inhabited by the Alans. The line of descent of the present-day Udi speakers, numerous peoples from the east past into the area. The Seljuk conquest in the eleventh century eventually led to its massive Turkicization, with older populations being largely assimilated. The Mongol conquests brought in new waves of Oghuz and other Turkic groups, which additionally advanced Turkicization. Azerbaijan is now an almost exclusively Turkic-speaking country.

6 Contact languages

Turkic has been in contact with most languages of the Caucasian area (Gecadze, 1977). This relatively limited territory displays a remarkable linguistic diversity, with numerous languages spoken under symbiotic conditions for more than 4,000 years, probably with widespread bi- and multilingualism. (For a survey of peoples and languages, see Geiringer et al., 1959.)

-Karachay-Balkar has had close contacts with the Diger dialects of the Iranian language Ossetic, which goes back to varieties of the nomadic complex of the Alans. It has also had contact with the Abkhaz-Circassian languages Abkhaz, Abaza, Adyghe and Circassian, and the Kartvelian language Svani.

-Noglay used to be in close contact with Kalmyk. White Noghay, which formerly interacted with Circassian, Abaza and Karachay, now has little contact with Caucasian languages. Small group near Mineral'nye vody still live in the neighbourhood of Abkhaz and Circassian groups. Black Noghay in Northern Daghestan is in touch with Kumyk.

-Kumyk has been in contact with its immediate Nakh-Daghestanian neighbours; for example, Avar, Lak, Daghi and Khadiq.

-There were many Turkic groups in the Caucasus area and in the adjacent areas.

-Many of these groups have shifted, the existing local populations or the incoming groups.

7 Linguistic effects

When trying to assess the linguistic effects of these encounters, we face, as always in contact linguistics, the problems of distinguishing inherited features from those that are copied, and of determining the source and direction of contact influence. The effects of the old migration waves that reached the Caucasian area from the north are the most difficult ones to interpret. In the following, some ascertainment or potential contact-induced phenomena will be mentioned.

The linguistic effects will be discussed in terms of the model of code-copying: see, for example, Johanson (2002: 8-19). Copying means the insertion of copies of elements of a model code into a basic code. Global copying concerns morphemes or morpheme sequences as a whole with their material, semantic, combinational and frequentational properties. Selective copying concerns individual material, semantic, combinational and frequentational properties. Copying never just means the transfer of elements from one code to another code, but always implies creative adaptation.

In the case of 'take-over' influence (adoption), speaker take over copies from a dominant foreign code in their own primary code. In the case of 'carry-over' influence (imposition), speakers carry over copies from a given primary code into their variety of a foreign code. In the case of code shift, copies carried over from a former primary code may remain as substratum features in the new primary code.

In contact situations, code take-over – with adstratum features – and code shift – with substratum features – yield different results. Many contact-induced phenomena found in varieties of the Caucasian area may result from substratum influence following code shift. In this case, the question is: Which groups have shifted, the existing local populations or the incoming groups?

7.1 Lexical copying

Lexical copying between Turkic and non-Turkic varieties is common in the Caucasian area. The Turkic influence is considerable in various regions. For centuries, until recent times, Turkic varieties have been prestige codes represented by elite groups ruling large parts of the area (Menges, 1968: 176). In North Caucasian, Turkic varieties of groups occupying the lowland areas used to function as lingua francas for trade and inter-group communication. Up to the period of Russian dominance, mountain peoples coming down to the lower regions to trade or to work would learn these Turkic varieties. It was quite common for Caucasians to be at least bilingual. The lowland markets were instrumental in spreading Turkic varieties (Wixman, 1968). As mentioned earlier, nor-western and central Daghestan was dominated by Noghay, while north-eastern Daghestan was dominated by Kumyk. Up to the first part of the twentieth century, there was a tendency to extend Turkicization into the North Caucasian area, for political and economic reasons.

Even more important was the transregional validity of Azeri, also throughout Persia, a situation that continued at least until the eighteenth century. Azeri was used for inter-group communication in south-eastern Daghestan, and even further north-west.

In the Northern Caucasus, Turkic and non-Turkic varieties have exchanged a quantity of loanwords. The foreign lexical impact on Noghay and Kumyk is, however, relatively moderate. For the Turkic lexical influence on non-Turkic varieties, see Sagirov (1989, esp. ch. 2); Dildalalov (1990); Aleskeev and Seyrov (1997: 117-18 and the literature noted there).
In the habitat of the speakers of Karachay-Balkar, a homogenous cultural area with many similarities between its various linguistic communities, the lexical exchange has been relatively balanced. The Osetic vocabulary mirrors long-standing and intensive Turkic contacts. The way of the borrowing may be difficult to determine, since many copied Turkic words are also found in other Ponto-Caspian languages.

Though code shift to Turkic must have been common, there are no indications of heavy substitution effects on Karachay-Balkar. The basic lexicon is of Turkic origin (Semениchenko-Golos, 2000). The loanwords from the adstratum have become the only nouns relating to local conditions of life, flora and fauna, local cultural terminology and so on.

Karachay-Balkar is rich in Osetic and Kartvelian loanwords. The numerous loans from Diger Ossetic are also present in Karachay, although its speakers have not had any direct contact with this vattety during recent centuries. There are certain differences. Thus the Ossetic word ugis' insia, the northern side of mountain, has been copied into Balkar as ь г i t ь. The aspirated sounds occur in meaning, while it means 'forest' in Karachay. The long and strong Alan influence is manifest in numerous ethnoynoms and topographic names. Some Alan terms from the Christian sphere have been preserved (Pritskir, 1959: 341). The Ossetic numerals were previously used by the Balkar (Abaza, 1933: 80–1). An influence from the Caucasian environment (Comrie, 1981: 214) is found in the former Karachay-Balkar and Kumyk vigesimal counting system, with numerals based on the concept of twenty—for example, Karachay-Balkar dchi al'mano bira on ('two twenty with ten') for '60', Kumyk al' yigyme ('three twenty') for '90'.

The situation in the Southern sub-area is quite different. The number of Transcaucasian Turkic is heavily influenced by a substratum: groups speaking Iranian and Caucasian varieties have shifted to Turkic, carrying over local terminology to their new primary code. According to Stilo (2005), the substratum is essentially Tats. Thus words such as ь t e m ь 'seed' and Jeh 'wooden plough; pair' match Tat t a m and Jeh exhibiting the typical vowel changes -em > -m and -dr > -t; compare Persian ь t e m ь and J e h. Persian has for centuries been a culturally dominant adstratum code from which languages of the whole area have copied lexical items relating to a wide range of domains, in particular cultural, erudite vocabulary.

Azeri and Kumyk have exerted strong lexical influence on neighbouring Nakh-Daghestanian languages. For example, Leighton has copied Turkic words extensively, many of which ultimately go back to Persian and Arabic elements; for example, dair 'mountain', sabraba 'patient', rangaz 'colourless'. It has also adopted Turkic suffixes; for example, -d in words such as kolotani 'kohltoz' and the plural suffix -lar in words such as da lar 'mountains' (Haspelmath, 1993: 72, 101, 120).

On the Turkic lexical influence on Georgian, see Golden (1994). There is a very extensive pattern of borrowings from later Ossetic Turkic, coming from the Turkish dialects spoken in the Eastern part of the Ottoman Empire and from the Safavid state in Iran (Golden, 1999: 54).

7.2 Sounds and sound patterns

All Turkic varieties of the area have copied phonetic and phonological features including certain common characteristics of the Caucasian languages. Strongly aspirated voiceless stops, for example, [p'ap'] 'to find', occur in various contact zones. In Azeri dialects, strongly aspirated unvoiced stops occur typically in inutil, but also in final position, for example, [p'ap'] 'cannon', Nakh dialect [p'ap'] 'milk'. Kumyk voiceless stops are aspirated, for example, [p], [t] and [g], except after other consonants and word-finally. In Karachay-Balkar, phonological copying has created a consonant system whose structure is very similar to that of Ossetic. Voiceless obstruents form pairs based on the feature of aspiration: [p] versus [p], [t] versus [k], [g] versus [g], and [p] versus [q]. The aspirated sounds occur word-initially, for example, [q'ati] 'to bring'; [q'at] 'hair'; [q'ur] 'dry'; [q'urul] 'intervocalic position, for example, j z u m 'run' to shout'; [q'urun] 'woman'; and as the second part of consonant clusters, for example, [q'ar] 'passed'; [q'ar] 'to kill'; [q'orq] 'fear' (Pritskir, 1959: 349–50).

Karachay-Balkar has also copied the Caucasian articulation of ejectives, obstruents produced with a glottal egressive airstream mechanism. Karachay- Balkar ejectives are mainly suffix-initial voiceless obstruents, for example, [h'lt] 'uk' (hito 'to sink, past tense), obviously used as particitives that make bound morphemes identifiable in spite of their phonetic variability.

In Kumyk dialects, certain consonants are articulated with a glottal stop, mainly in loanwords from the Caucasian, but also in native words, for example, [q'arq] 'forty'. This feature is typical of the southern Khabad language dialect spoken in the Derbend region but is also to some extent present in the central Kumyk dialect spoken in Burya (Bujnaisky) and neighbouring villages. The northern Azeri dialects of Qazatala and Gaz, spoken in regions bordering on Daghestan and Georgia, exhibit phonological characteristics typical of the Leghidic language Sukhur. In words such as [o'la] 'countryside' and [q'ar] 'three' they obviously function as substitutes for the rounded vowels [a] and [y]; compare Standard Azeri [alae], [yy].

These sound types may be a result of adstratum influence adoption, 'take over' phonology or substratum influence (implosion, 'carry over' processes), or both. In the North Caucasian area, adoption seems more plausible in many cases, for example, the Karachay-Balkar glottalic stops are marginal phonemes without a significant functional load. This is also the case in the neighbouring Ossetic dialect, whose Iranian sound system has not undergone any radical changes.
The situation seems to be different in the Southern sub-area. The innovations found in Azeri dialects of this area seem to be a result of strong substratum influence. Northern dialects have much in common with some southern dialects spoken in Iran—for example, the Khoy dialect. The front stops /k/ and /g/ are palatalized, e.g. [Kim] 'who', [ğyn] 'day'. In certain northern and southern dialects they are pronounced as palatal affricates. Soso suggests that the obstructs /ğ/, /b/, /ğb/, /ék/, /êt/ and /ğ/ are the result of systemic shifts because of a very early substratum. This could be a result of code shifts to Turkic from Iranian varieties of the Tatic type, which in turn, may also use the northern substrata. North-western Persian and Armen-ian exhibit similar consonant systems. According to Ştål, Armenian was probably influenced by Urartian, and the Iranian dialects could have inher-ited the same system. Another Azeri peculiarity that may have emerged in a similar way is the voicing of /g/ to /ğ/, e.g. [g'áru] 'girl'. Some dialects also display a rather un-Turkic fricativization of labial consonants, e.g. [g'áři] 'girl' has come in the Gentje dialect.

Another phonological substratum effect may be the infringements on the systems of sound harmony in various Turkic varieties in the Caucasus area. Deviations of this kind are found in some Karachay-Balkar dialects as a result of contacts with Abkhaz-Circassian languages (Abkaz, 1963: 47); they are also typical of Kuma dialects and of Azeri dialects spoken in zones of close contact with Nasid Daghestanian languages such as Dargi, Lak and Tush-speaking, for example, Khadiq [gi'man] 'I enter' (Dmitriev, 1940, 1962; Kerimov, 1967). In Azeri dialects, the sound harmony is disturbed by fronting tenden-cies. Many suffixes are invariable—that is, exhibit non-harmonic forms, for example, in the Nukha and Baku dialects.

On the other hand, Turkic-like sound harmony tendencies may be found in some non-Turkic varieties. Thus the emergence of a front versus back vowel harmony and a rounded versus unrounded vowel harmony may be observed in the Lezgic languages of Ud and Lezghian as well as in some other Caucasus languages.

7.3 Grammar

Contacts between Turkic and non-Turkic varieties have led to changes in grammatical structures in the sense of adstratum and substratum influence. The Turkic varieties have, however, essentially retained their basic grammatical properties. Grammatical morphemes are mainly of Turkic origin, though many are used in new ways according to new models. Even where new grammatical structures have been copied, native morphological material has normally been employed. The Turkic varieties do not share the common morphosyn-tactic features of the Caucasian Sprachbund, namely the characteristic ergative alignment. On the other hand, it seems that Turkic has exerted considerable grammatical influence on many non-Turkic languages of the area.

7.3.1 The Northern sub-area

Some North Caucasian areal phenomena may have resulted from, or have been stimulated by, contacts with Turkic. Although Ossetic has essentially preserved its Iranian character, it shows certain signs of foreign grammatical influence. It lacks the complex system of verbal suffixes typical of Turkic, but its nominal inflection tends towards the agglutinative type, for example, with case endings following the plural endings.

One North Caucasian areal feature is the syntactic subordination by means of non-finite verbal forms, for example adverbiale clauses with verbs in -g 'instead of finite verbs. For example, for the use of the Ossetic verb in -g may have been supported by contact with Turkic in the sense of frequentive copying that is, shift in relative frequency under the influence of an adverbial element in a model like: A leh, Abkhaz-Circassian and Ossetic verb forms of the verb ducni to say are used as citation particles to convey reported speech or thought, which also matches the typically Turkic use of converb forms such as Karachay-Balkar dẹb 'saying', for example, Ossetic dẹb 'g. Following Turkic models, Ossetic has also developed postpositions such as ke, g. of verbs to 'look', meaning 'according to', for example, meneve ge g. 'according to me' (to me looking); compare Karachay-Balkar kine.

7.3.2 The Southern sub-area

The situation is the Southern sub-area is, again, rather different. In a geographically contiguous area we are confronted with certain common tendencies of the individual languages, independently of their linguistic relatedness. I shall comment on some cases in which Turkic varieties have played more or less important roles.

7.3.2.1 Intra terminals.

Several languages of the Southern sub-area exhibit similarities in their aspect-tense structures (for Kartvelian, see Christópoulos, 2004). Intra terminals—that is, aspects and imperfects—are often formed with intransitive constructions based on locative metaphors meaning to be inside, for example, Armenian երտամ i am going: Talysh kinde i in going: Lezghian iwa i am going: Bgavali i/g you i'm going: Neo-Aramaic էրց i am going. The development of these items may well be interpreted as aspectual convergences in the direction of Turkic, where intraterminality is a stable grammatical category. Items such as Azeri gedgok and Turkish dekko go to the futureative-presentative: 1 i am going are widespread in the Turkic family, for example, Modern Uyghur gilmaq i am doing, gilmaq i am doing. These are numerous parallels between the verbal systems of Azeri and some adjacent non-Turkic varieties. Nakh-Daghestanian languages such as André, Avar, Archi and Lezghian possess phenomena that have undergone delexicalization, such as the Azeri old present in -av, for example, yazar 'will write, writer'; compare the new focal present in -av, for example, yazar is writing, writer (Johnson, 2000: 92–3). The Lezghian da form has,
by virtue of a similar value, general and partly modal functions interpretable as 'habitual', 'future', and so on, for example, *fida 'goes, will go', *fida 'went', would', etc. It has therefore been characterized as a 'future with a habitual/prescriptive polysemicity' (*Hausmann, 1993: 276). The item kafir 'threw, will throw' of Khwarshi (Daghhestanian Tar-Svan group) displays analogous uses. The Iranian languages Tabylh and Tat exhibit similar items with habitual and modal (mostly potential) readings, for example, Talysh *naad 'read, used to read, would read'. The relevant similarities between Azeri and Tat are discussed by Grigorew (1962); see the comparative table pp. 17-18.

7.3.2.2 Evidential: Evidentials, categories that state the existence of a source of evidence for a propositional content (Alkhieváld, 2004: 11), are widely acceptable in areal contacts. Turkic languages, which possess stable evidentials of the indirective type (Johnson, 2000, 2003) are known as sources of their diffusion. Indirectives appear in contact areas such as the Balkans, Anatolia, the Volga region and Central Asia (Johnson, 2000a: 83–4; 2006: 285–9). If they appear in a number of languages of the Caucasian area, one is justified in considering possible Turkic influences.

The expression of indirectivity is closely connected with postterritoriality, a view on events from an orientation point posterior to their inherent relevant imit (Johnson, 1971: 283; 2000: 29, 102–4). Expressions of postterritoriality vary according to the degree of frequency, the relative narrowness of the time interval around the orientation point. With high-focal postterritorials, the interval is narrow, which yields static or resultative meanings, for example, English is gone. With low-focal postterritorials, the interval is wider, which yields 'perfect' meanings, for example, English has gone. The development from postterritoriality to evidentiality can be accounted for as a process of defocalization. Focal postterritorials may lose their focality and finally even their postterritorial character, which often leads to the emergence of new high-focal postterritorials.

The widespread development of postterritorials into evidentials in the Caucasian region is not necessarily a result of Turkic influence, but contact with Turkic may have supported it in the sense of frequentional copying and indirect impressions towards evidentiality (Johnson, 1996: 87).

In most Kartvelian languages, defocalized postterritorials have evidential functions. Modern Standard Georgian does not possess any special morphological category to code evidentiality, but one of the meanings of the well-formed perfect is regarded as being evidential, for example, *sirta 'it has obviously rained'. This is thought to be a rather recent phenomenon, since the Old Georgian perfect was purely resultative. Christophe (2005) assumes that Kartvelian evidentials have evolved from high-focal postterritorials through processes of cyclic grammaticalization. The Georgian static item *sc'rida 'it is written' and the resultative dac'rida (< Old Georgian dac'ridar) originally differed with respect to the degree of focality of the postterritorial view. The former item had the highest degree, whereas the latter became increasingly defocalized. Mingrelian resultatives are products of the same diachronic process of cyclic grammaticalization by which postterritorials turn into evidentials, leaving a gap that is filled by new postterritorials. The new resultatives take over the original function of the older ones, but only to become increasingly evidential themselves.

Laz, Mingrelian, Svan and some adjacent western dialects of Georgian have developed elaborate Turkic-like evidential systems that also include *mari, imperfect and future items (Haris, 1985: 296–300). The Georgian particle t'urbe 'apparently', for example, t'urbe toveli murls 'it has apparently snowed', also expresses evidentiality with non-past events, which is reminiscent of the use of Tushub and Afar imit. Some Georgian dialects, for example, Ingılı, spoken in an Azeri surrounding, too the particle qipil (a), a perfect form ('it has been') based on the past participle of 'to be'. It might be interpreted as a semantic copy of Azeri imit, compare bamsbux qipil 'cullbe is apparently coming' (Boeder, 2000: 314, at 17) with Azeri qe'ter imit. The difference lies in the fact that qipil is an invariable particle, added to inflected verb forms, whereas imit is added to non-finite forms and bears the personal endings. Boeder still sees good reasons for considering qipil an areal phenomenon due to code copying (2000: 284). The inflected copula is suffixed to finite verb forms in several Kartvelian varieties, for example, qipil-m in Khvuz, an East Georgian dialect that has been in close contact with Chechen-Ingush. The Laz dialects of the Findikli-Arhabi region in Turkey use the evidential particle doren, a form of 'to be', after finite verb forms. In the Hopa region, however, the corresponding element is used as a finite copula verb, with the preceding verb form lacking personal endings; this matches the use of imit.

The use of the East Armenian perfect to express evidential meanings corresponding to those of the Georgian perfect may be because of contacts with Turkic and Iranian (Johnson, 1992: 245, 282; 2002: 99, 146–7; Kozitsinswa, 2000: 414). West Armenian, formerly spoken in Eastern Turkey and thus a contact language of Turkic, has a distinct evidential particle vur developed from the inferential participle of the verb 'to be'; see the above-mentioned Ingılı particle qipil. Like qipil, t'urbe ari imit, etc., vur is a more general evidential particle that is not only used with the perfect (Dwomohedan, 1996: 95–7). Tat displays the evidential particle mü, a copy of Azeri imit. The Tat perfect may have evidential readings itself, but the particle is often added to it (Grunberg, 1968: 88).

Turkic-like evidentiality systems exist in numerous Nakh-Daghestanian languages. Defocalized postterritorials display evidential functions in Avar. The Nakh languages Chechen and Ingush use the participle kaxar 'been' of kaxar 'to be', for example, a. t'evaxar kaxar-a: 'he has apparently gone out' (Boeder, 2000: 314). This particle is reminiscent of the Khvuz particle qipil, but, like Azeri imit, it bears the inflection marker.
Abkhaz-Circassian languages possess evidentials that seem to have already existed in Proto Abkhaz and Proto-Circassian in the eighth and ninth centuries. If this is true, it is very likely that the Abkhaz-Circassian evidential system was inherited from Proto-Georgian. The Abkhazian grammatical system stresses that there was ‘no discernible’ Turkish presence in Abkhazia prior to the sixteenth century (Chirkiba, 2003: 260).

On the other hand, some tend to see the basis for a Turkish influence that is limited in time and place. For example, some suggest that the Turkish impact was in general City of Tbilisi. The discussion on possible contact-induced effects has so far focused on the evidence for the contribution of Oghuz Turkish—that is, Turkish and Azeri—and very few exclusively on contacts during the period of Ottoman domination. According to Kobakhidze (2000: 227) the evidentiality system of Mingrelian may partly be a result of contacts with Turkish. Still, there are only a few remote from Turkish domination, may include the relict system that, after Mingrelian (see Sumbatov, 1995; Friedman, 2000: 357). Although it is not yet clear whether the evidential variants of the Georgian perfect aspect, as many as we know (2000: 297-8).

While this may be true, the presence of evidentials is certainly not a convergence phenomenon in which only Oghuz Turkish has been involved. The emergence of Turkic-like evidential categories may result from, or both may be supported by, other Turkic varieties in the area. These are the Geogams that have dealt with virtually every Turkic group that entered the North Caucasus region (1999: 96). We are confronted with the much wider issue of Balkan-Turko-Caucasian-Caspian contacts. The Bulgarian evidential system points clearly to pre-Ottoman contacts with Turks (Johansson, 1996). Certain Kipchak Turkic languages may use the particle eken, once widespread in the area (Johansson, 2001: forth. in press), as a preterite-future particle added to main clauses, without carrying personal information. Karakas (2003: 31) deals with Negy sentences in which the personal markers are carried by eken, and sentences in which they are carried by the preceding verb, for example, A Rulgen ellef? Why do they seem to have complementary? Nege kegelgen eken? I wonder why they have come home, Turkish and Azeri imit only represents the first model. As we have seen, however, the second model is found in many evidential constructions in the Caucasian area.

7.3.2.3 Syntax. The Southern sub-area exhibits many common syntactic features. Azeri dialects exhibit strong Iranian influences in the syntax. Sito (2000) discusses possible substratum structures such as the use of the subjunctive after modal verbs for example, istir basmar ‘I want to go’, copied Iranian function words, loss of the question particle mi (replaced by intonation), cohesion of subjunctive and optative forms, for example, istir gada – istir getim ‘wants to go’ (also typical of Northern Talysh), and certain loan translations. The unmarked constituent order in Turkish is verb-final. This is also the case in Georgian, Armenian, Nakh-Dagestani and Abkhazian languages such as Dargi, Lezghian, Chechen, Ingush and so on. Georgian and Armenian have, however, switched from verb-object to object-verb order since the medieval period. This typological reversal may well be the result of areal contacts with languages of the opposite type. In Old Georgian, adjectives generally followed their nominal head. Turkic languages normally have preposed relative clauses (Johansson, 1992: 274-5; 2002: 135-7); this is also the case in Nakh-Dagestani. Ossetic is the only Iranian language of the area that sometimes exhibits preposed relative clauses. Postposed relative clauses of the Iranian kind are normally found in spoken Georgian, Armenian, Circassian and Azeri (Stilo, 2004). Spoken Azeri also makes heavy use of postverbal subordinate clauses and regularly places constituents expressing destination/goal after the verb, for example, golamem ever ‘I go home’ (Stilo, 2004).

The decline of ergative constructions and the emergence of Turkic-like syntactic structures in some Caucasian varieties are not necessarily contact-induced phenomena, but may be rooted in internal tendencies that have been supported and generalized through contacts with Turkish in the sense of frequentional copying. For the loss of ergativity in one dialect of Udi (Gulatian, 1973: 43); Azeri influence seems to have been the decisive triggering factor.

8 Similar and different relatives

The languages of the Caucasian area as defined in the present chapter certainly do not qualify as a single Sprachbund. The sub-areas, however, rather constitute continuous communication areas exhibiting a good deal of bi- or multilateral convergence. It is often impossible to judge the direction of influence. The spread of areal features cannot always be accounted for by one source of diffusion. If we had access to detailed historical data, however, cases of multilingual convergence would probably turn out to consist of series of unilateral and bilateral code-coping processes (Johansson, 1992: 279-84; 2002: 141-5). Many common features observed in the sub-areas may be because of contact-induced frequentational copying that has supported latent internal tendencies.

The Turkic languages of the North and Southern sub-areas differ considerably from one another with respect to their relative position.

8.1 The Northern sub-area

In spite of a long, symbiotic coexistence with their non-Turkic neighbours, the Turkic languages of the Northern sub-area display relatively little foreign influence. Kumyk and Karachay-Balkar have essentially maintained the features of Early West Kipchak – ‘Kipchak-Kuman’ – as documented in the Coassiamic languages (Magomedov, 1966: Xabibov, 1966: Berta, 1998). Nogay exhibits little influence from the Caucasian area beyond loanwords, although...
The Roles of Turkic in the Caucasian Area

Menges speaks of a strong influence of Circassian (1959: 436). None of the
languages shows any major syntactic peculiarities.

The situation in Caucasian-Balkar, which has long developed in relative
isolation from other Turkic languages, is particularly interesting. In the 1930s,
the Marr school of linguistics focused on this language as an exemplary
crossbreed of Turkic and Caucasian (Jafry's) elements. According to Marr's
theory, languages arise by processes of interweaving and assimilating -
linguistic development is seen as constant 'mixing'. The alleged 'mixed'
status of Caucasian-Balkar, however, by no means supported by the data.
The neighboring languages have left relatively little imprint on it. It has
retained the character of its Turkic subgroup, just as Ossetic has largely
retained its Iranian character. The influence exerted by Ossetic and the
North Caucasian languages on Caucasian-Balkar is much more limited,
owing to the reverse influence. As already mentioned, while language shift must
have been common through the centuries, the overall picture is not dominated
by the above-mentioned influence.

8.2 The Southern sub-area

The situation in the Southern sub-area is quite different. Transcausasia and
northern Iran constitute a multilingual convergence area whose languages
share typological traits at all levels. Azerbaijani displays numerous deviations
from the classical type of Turkic, much more than the Turkic varieties of the
Northern subarea.

Transcausasia was largely Turkicized during the Seljuk era. Oghuz Turks
began to move into Azerbaijan in the eleventh century. Turkicization
began in the north and continued into the south by the twelfth century
(Golden, 1992: 225). At the time of the immigration, numerous mutually
unintelligible varieties were spoken in the area, for example, various brands of
Turkic: Caucasian Albanian, Kurdish, Armenian, Avaric and possibly
variants of Persian. According to Stifo (2004), these varieties had probably already
formed an 'intermediate' dialect, as can be seen today. The linguistic areas
facilitated a shift to a Turkic lingua franca, which soon became the primary
code of large parts of the population. There was probably a rapid shift to Turkic, a
process which subsequently expanded into other Iranian-speaking areas.

Many Caucasians had probably shifted to Iranian, Armenian and Avaric
before they shifted to Turkic; others may have shifted directly to Turkic.
Many features from Iran and its Caucasian substrata were imposed on the
Caucasian Turkish varieties which had already been subject to Iranian
influence in Central Asia and Khorasan.

Stifo points out that Persian was not a spoken language in the area at the
time of Oghuz immigration. It has, however, served ever since as a highly
influential adstratum language because of its cultural dominance.

The number of Turkic newcomers who had moved into Azerbaijan,
replacing the old ruling class and causing the code shift, seems to have been

small. The shift was not based on a massive migration, and thus took place
with little genetic impact. Today's Turkic-speaking Azeris are related most
closely to their nearest non-Turkic neighbors, that is genetically closer to
Armenians and some Caucasian-speaking groups than to other Turkic-speaking
populations (Nadzide and Stoneking, 2001). This closeness is certainly
cased by shared Caucasian substrata.

9 Reasons for the linguistic impact

What was the reason for the strong linguistic impact of the small Turkic-
speaking elite groups? The situation may be compared with that of other areas
in which Turkic has played major role. Many originally non-Turkic groups
in Southern Iran have abandoned their primary codes in favor of Turkic varieties.
In Central Asia, sizeable Iranian-speaking groups shifted to Turkic. In the north of the
Turkic world, Tungusic-speaking groups adopted Yakut (Pakendorf, 2001), and in
southern Siberia offers similar scenario, with Southern Samoyed, Ob-Ugrian
and Yenisei Ostyak (Ket) substrates. Old Bulgar groups went up the Volga
River to the Volga-Kama confluence area and established a state there. Some
of these Bulgars imposed their language on Finno-Ugrian populations, which
led to the emergence of Chuvash (Johnson, 2000b). On the other hand,
Turkic-speaking groups have shifted to non-Turkic languages. One of the
important Bulgar movements, caused by pressure from the Khazars, was the
migration of one group to the Jambule, the Byzantine frontier. These Balkan
Bulgars very soon gave up their primary code in favor of a Southern Slavic
variety, today's Bulgarian. In old Eastern Europe, other Turkic groups of
unknown size, Bulgars, Kumans, Pechenegs and others, adopted Slavic varieties.
In recent times, Turkic groups have massively replaced their primary codes
with Russian.

Contacts leading to shift depend on social dominance relations. The size
of the incoming groups relative to the local groups may be relevant, but
diffusion is not always accompanied by major population move-
ments. The proportions may vary considerably. Elite minorities may impose
their codes on comparatively large local indigenous populations. Shifts of
this kind occur without major demographic changes.

The introduction of Turkic into the Caucasian area has rather been the
result of minor movements, but the influential roles of the Turkic varieties is
obvious. One reason for the linguistic success of Turkic elite groups was
their advanced political organization, which contributed significantly to their
dominance. As mentioned, they were prestige languages and served as
lingua francas in bilingual intercourse.

The situation of the Turkic-speaking newcomers in the Caucasian area
differed essentially from that of the Turkic groups on the Balkans. Both
Caucasian sub-areas were characterized by an extreme diverseness of mutually
unintelligible codes which mainly exhibited rather complex structures.
Prestige languages with structural properties that made them relatively easy to learn (Johannsen, 1992: 199–206; 2002: 41–8) could be used to bridge the gaps between the local speech communities.

10 Balkan–Pontic–Caucasian–Caspian contacts
What must be kept in mind is that Oghuz Turkic is not the only source of Turkic impact on the non-Turkic varieties, and of Turkic contributions to common area tendencies. The highly probable influences from various older Turkic tribes that invaded the area should not be ignored: Bulgar element from the end of the fifth century onwards, Western Turk elements from the eighth century onwards, Khazar elements from the seventh century onwards, and Kipchak Turkic elements of the earlier and the later type from the tenth century onwards. It is not possible to establish, as Gazdileva (1979) has attempted, one single type of early Turkic for the Caucasus area, but it is necessary to deal with the linguistic situations within a more general framework of Balkan–Pontic–Caucasian–Caspian language contacts.3

Acknowledgement
Work on this chapter, which is a contribution to the Special Research Project (SRP) 295 of the Deutsche Forschungsgemeinschaft, was carried out while the author was Guest Scientist at the Max Planck Institute for Evolutionary Anthropology, Leipzig. The chapter is based partly on a talk given at a meeting of the Genetics-Linguistics Group (geneling) on 2 April 2002. I profited very much from the paper ‘Languages and Prehistory in the Caucasus’ given by Johanna Nichols on 18 April 2002. I am particularly grateful for the fruitful discussions I had with Helma van den Berg, the ardent investigator of Daghestanian languages, whose life ended only one year later, on 11 November 2003 in Derbent, Dagestan.

Notes
1 As is well known, counting by twenties is also found in some West European languages. Thus Old French quatre vins discret (modern French quatre-vingt-discret) for ‘ninety’ is modeled on Basque laurtzatur (I.e. lauratx-tatx-tara-tatx).
2 A further question is whether the incoming groups have been predominantly male, mating with indigenous females who have shifted their code. New methods of population genetics will finally allow answers to questions of these kinds.
3 It is not possible to deal with linguistic similarities between the Balkan and Caucasian areas in this chapter. It is, however, interesting to note that, according to Nedjaakov (2002), Kara-Tatar-Balkar is the only Turkic language in which verbs marked with the reciprocal suffix -yv ‘have a competitive meaning ‘to do something competing with each other’ this meaning happens to be similar to that of the Bulgarian reflexive elicit.

References
Akiberalov, Ç. V. (1913) Fenikstil Dialektleri Kavkaz-olyskyn yazikam (Cakermas: Karta-cavro-Cerkesske Knjizna izdateljstvo).
Dmitriev, K. N. (1940) Grammatica kamyokska yazyka (Moscow:Kinesyedija: Akademia nauk SSR).