Sociolinguistics in the Caucasus

Victor A. Friedman

Introduction

As a geo-political region, the Caucasus can be divided into the North Caucasus and the South Caucasus (Transcaucasia). Transcaucasia consists of the Republics of Georgia, Armenia, and Azerbaijan—including polities whose status is still disputed as of this writing: Abkhazia and South Ossetia in Georgia and Nagorno-Karabakh, which seceded from Azerbaijan to join Armenia in 1988, followed by a war lasting until 1994. In strictly geographic terms, Georgia and Azerbaijan also include some north Caucasian slopes in their territory. In the south, the political borders of Turkey to the west and Iran to the east form a convenient demarcation, although speakers of relevant languages extend into and/or migrated to both these countries. The geo-political North Caucasus is entirely within the Russian Federation and consists of a series of Republics (from west to east): Adygea, Karachay-Cherkessia, Kabardino-Balkaria, North Ossetia (Alania), Ingushetia, Chechnya, and Dagestan. Adygea is surrounded by Krasnodar Krai, which, with Stavropol Krai and Kalmukia, forms the northern administrative border of the remaining North Caucasian republics, whose southern borders are defined by Georgia and Azerbaijan. The Black Sea defines the western border, and to the east is the Caspian.1

The Caucasus, long known for its linguistic diversity, is home to three indigenous language families as well as representatives of Indo-European, Altaic, and Afro-Asiatic. The indigenous families are Kartvelian (South Caucasian), Nakh-Daghestanian (North-east Caucasian), and Abkhaz-Adyge (North-west Caucasian). These three were assumed to form a larger Ibero-Caucasian family, but that idea is no longer generally accepted owing to the lack of any plausible reconstruction. Attempts to unite the North Caucasian languages into a single family present serious problems of data and methodology. The time depth for North-east Caucasian alone is estimated as approximately that of Indo-European (Nichols 1992: 14).

The Kartvelian languages are Georgian, Svan, Mingrelian, and Laz. The North-west Caucasian languages are Abaza, Abkhaz, Adyge (West or Lower Circassian), Kabardian (Cherkes, East or Upper Circassian), and Ubykh. Tsezesk, the last fluent Ubykh-speaker, died in Turkey in 1992, and Abkhaz-Abaza and Circassian are sometimes each
treated as single languages. The Nakh (or Vaynakh) languages are Chechen, Ingush, and Tsvoa-Tush (Batsbi). Chechen and Ingush also form a continuum with the transitional Galanchozh dialects being claimed by both. The Dagestani languages can be divided into three groups: Avaro-Ando-Tsezic, Lak-Dargic, and Lezgic. The Andic languages are Andi, Akhvakh, Bagwala, Botkhi, Chamalal, Godoberi, Karata, and Tindi; the Tsezic languages are Tsezi (Dido), Bezha, Hinukhi, Hunzib (Kapsa), and Khwarshi, while Avar constitutes its own sub-group. Lak is unique in the Lak-Dargic group, its dialects being relatively close to one another, while the dialects of Dargwa (Dagi) are differentiated between a core and peripheral dialects, some of which latter—Kubachi, Katag, Megeb, and Chirag—are sometimes considered separate languages (Tsirkha and Itsam are likewise peripheral but counted as Dargwa). The main sub-group in Lezgic is the Samuritan: Lezgian, Rutul, Tsakhur, Agul, Tabasaran, Budukh, and Kryz. Archi, Udi, and Khinalug, are each so peripheral that their relationship within the rest of Dagestanian was, until recently, considered isolated. Of these, Archi is closest to Samurian, Udi is peripheral, and it now appears that Khinalug is a separate branch of Nakh-Dagestani subsequently influenced by Lezgic.

Of the Altaic languages, Azeri and Turkish belong to the Oghuz Turkic group, while Karachay-Balkar, Kumyk, Noga, and Tatar are Kipchak Turkic. Almost all are Muslim. Ossetian, one of two surviving North-Caucasian Iranian languages (the other, Yaghnobi, is spoken in Tajikistan), has two dialects, Digor (Muslim) in North Ossetia and Iron (Christian and Muslim), the basis of standard Ossetian, in both North and South Ossetia as well as Georgia proper. Talysh and Kurdistan are North-west Iranian, while Tat (Northern or Caucasian Tat) is South-west Iranian (like Persian). Armenian is an isolate within Indo-European (with its own church), as is Greek (Orthodox, Pontic dialect). Armenian has two standards, Eastern, which is official in Armenia, and Western, which is the variant used by the Armenian diaspora and those in Turkey whose ancestors survived the massacres of the early twentieth century. We can also mention here Lomavren (Bodha), a para-Romani language whose grammar is Armenian with significant Indic vocabulary. Finally Assyrian (Neo-Aramaic) is North-west Semitic and spoken by Christians using Syriac as their liturgical language.

All the republics except Dagestan have titular linguistic nationalities, with speakers of other languages as minorities. All Kartvelian-speakers in Georgia are counted as (and consider themselves to be) ethnic Georgians and most live in Georgia except the Laz, who have a distinct consciousness and, with the exception of a single village in Georgia, are across the border in Turkey. The estimates are 350,000 for Mingrelian, and between 80,000 and 35,000 for Svan. Both use Georgian as their literary language. Estimates for Laz vary from 90,000–250,000. The Laz are Muslim, and like other Muslim linguistic minorities in Turkey, have no language rights. Other Kartvelian-speakers are mostly Georgian Orthodox, except the Adjarians, who are Muslim Georgian-speakers with an autonomous (but not breakaway) republic in south-western Georgia. They currently identify as Georgian on the basis of language, although under Ottoman rule they identified as Turks on the basis of religion. Tsvoa-Tush survives in the village of Zemo Alvani in Georgia (estimated at 3000). Unlike most other Nakh-speakers, who are Muslim, Tsvoa-Tush-speakers are Christian. The Mskhetian Turks are Muslim Turkish-speakers from Georgia who were deported by Stalin.

The Dagestani languages are all spoken in Dagestan, except Budukh, Kryz (Azerbaijan only), and Udi (two villages in Azerbaijan and one in Georgia); Lezgian, Rutul, and Tsakhur are spoken in both Azerbaijan and Dagestan, while Bezha and Hunzib are spoken in Dagestan and Georgia. Dagestanian-speakers are mostly Muslim, except Udi-speakers, who are Armenian or Georgian Christians. Nogai is spoken in Karachay-Cherkessia and Chechnya in addition to Dagestan (where it is one of 14 official languages together with Kumyk, Azeri, Tat, Chechen, Russian and 8 Dagestani languages: Avar, Dargwa, Lezgian, Tabasaran, Lak, Tsakhur, Rutul, and Agul). Chechen is also spoken in Ingushetia and Ingush in Chechnya and Alania. Azeri extends into Iran, Armenia, and Georgia, Georgian into Turkey, Azerbaijan, and North Ossetia, and Armenian is also spoken in Georgia, Azerbaijan, and parts of the North Caucasus. Abaza is spoken in Karachay-Cherkessia, and Ossetian is spoken in Georgia and Ingushetia in addition to its titular republics. Useki atlases of the Caucasian languages include Korjakov (2006) and Cipper and Schulze (2008).

History of sociolinguistic research in the Caucasus

Although the Caucasus is intimately connected with the beginnings of recorded history, the crucial events for the current sociolinguistic situation were the Russian conquest and annexation of the nineteenth century, the subsequent Soviet renewal of Russian hegemony after the October 1917 Revolution and its aftermath (during which the Caucasian republics were briefly independent), and the break-up of the USSR in 1989 and its reconstitution as the Russian Federation plus the CIS. As a result of Russian conquest in the nineteenth century, an estimated 1.2 million Muslims left the region, and an estimated 800,000 survivors settled in Ottoman Turkey. During the same period, the Russian Empire engaged in an intense campaign of Russification, aimed even at languages such as Georgian, which had a Christian majority and a literary tradition five centuries older than that of the Slavs. As Wissen (1980: 21–30, 121–69) argues, Soviet national policy was also ultimately aimed at Russification. Part of that goal was the creation of modern national identities, with concomitant standard languages, to serve as vehicles of literacy, modernization and, eventually, Russification.

Given the political system, Soviet sociolinguistics was dedicated to serving the ends of the state. Studies focused on questions of language planning and bilingualism, usually comparing non-Russian languages with "the language of inter-national communication," i.e., Russian. Russification was framed as "mutual enrichment" while western sociolinguistics was labeled "bourgeois" and its results " falsifications" (e.g., Treskov 1982: 133–4). As Kreindler presciently observed of late Soviet sociolinguistics:

There is a frantic, almost hysterical quality in the campaign waged on behalf of one of the most powerful world languages, which by all logic would seem in no need of special support. The spirit as well as much of the rhetoric of the campaign bears remarkable resemblance to the tsarist Russification campaigns in the last decades of the regime.

(1985: 356)

Under Soviet rule, studies of social variation were discouraged on ideological grounds: a classless society with perfect gender equality that was moving forward on the road to the elimination of all social (including, ultimately, national or ethnic) distinctions does not support investigations of actually existing class, gender, regional, and other social differences (unless aimed at eliminating them). Madieva (1975) looked at differences in the acquisition of Russian by Avar-speakers according to traditional sociolinguistic parameters such
as education, age, and social group (farmers, functionaries, students)—but not gender—at work, school, and home, complete with graphs and percentages. Unsurprisingly, younger, more urban, more educated Avars are more likely to use Russian. Medieva concludes that special efforts must be made to prepare children in a non-Russian environment to learn Russian, and that the development of bilingualism among Avars (which includes a complete shift to Russian among some) is in keeping with the Program of the Communist Party of the Soviet Union.

The four most important sociolinguistic works for the Caucasus during the Soviet period are Lewis (1972), Wixman (1980), Kreindler (1985), and Kirkwood (1989). For Dagestan, Dědělava (2005) covers a broad range of topics spanning the Soviet and post-Soviet periods. For Circassian, Smeets (1994) gives a thorough and realistic survey through the Soviet period. He observes that while in 1920 there were more Circassian-speakers in Turkey than in the Caucasus, 70 years later language preservation was better in the Russian Federation than in Turkey, where there was no support of Muslim minority languages. Nonetheless, because urban Circassian-speakers are a minority in their republics, Russian being the dominant language in cities, Smeets cautions that Circassian faces the fate of Breton in France or Friulian in Italy. Weitenberg (1990) gives a useful survey of both Western and Eastern Armenian language reform. T.A. Laragyulian (Agay 1981: 1-119) and H.L. Zak’aryan (Agay 1981: 120-271) give concrete data on modern colloquial Eastern Armenian, and, interestingly, Zak’aryan finds that despite the huge impact of Russian on colloquial Armenian due to Soviet policies, the degree of use of Russian loans in colloquial speech diminishes with increase in academic education and degree of association with language work—from most to least likely to use loans: workers, officials, natural scientists, social scientists, language professionals. Puristic tendencies have also been documented recently for Georgian (Apondidze 2003).

**Variationist studies, pragmatics, discourse, and conversation**

For the most part, studies of variation in the Caucasus are in their infancy or pre-natal stage. Many languages of the Caucasus are only now being adequately described, and considerable effort is directed at such basic tools as grammars and dictionaries as well as problems of dialectology. Questions of discourse, pragmatics, and conversation have been minimally explored. The extremely promising project on language contact and variation in the Dargic regions of centrall Dagestan by Helma van den Berg was aborted by Helma’s tragic death in 2003.

Although Vitkovskaja (2005) contains abstracts of articles with promising titles on discourse, communicative structures, politeness, etc., in Circassian, Karachay-Balkar, Abkhaz, and Ossetian—as well as the usual Russian-oriented “bilingualism” studies—in fact, there is very little concrete data. Solovev and Maz’al’n’ko (1996) present 17 abstracts (out of over 130) that pertain to sociolinguistics in the North Caucasian and Azeri. Of these, nine deal with North-west Caucasian (one includes Ossetian), four treat Dagestan, one for Azerbaijan and one for the three breakaway republics (Abkhazia, South Ossetia, and Nagorno-Karabakh), one gives an overview of the shift from Arabic to Latin to Cyrillic orthographies in the North Caucasus, and one provides a comparative survey of Dagestan, Karachay-Cherkessia, and Kabardin-Balkaria. A number of the abstracts promote the use of Russian, either as an inter-nationality language (Dagestan) or as a means of conflict resolution in the breakaway republics. The article on Azeri notes the decline of Russian and the advance of English, criticizes the new Latin alphabet, and notes the use of Armenian in Nagorno-Karabakh and the “greater role” of minority languages. The articles on North-west Caucasian deal with language policy, language planning, and discourse phenomena, mainly for ease. Alekseev and Peresvetsk’aja (2000) and Balan-mamedov (1992) contain useful conversational data from Dagestan. After the fall of the USSR, it was common to hear complaints that the Soviet system had destroyed Russian stylistics by forcing clichéd and wooden language on all writers. Almost nothing has been written on speech registers in the Caucasus, although Rayfield (1992) gives a trenchant account of polemical style in Georgian in the Soviet period and in 1990.

**Language contact, bi-/multilingualism, diglossia, code-switching**

Although the Caucasus is sometimes described as a *Sprachbund*, Tuite (1999) argues that it does not fit the model of areas such as the Balkans, which originally gave rise to the concept. The Caucasus is not characterized by a single lingua franca and does not share the kind of morphosyntactic features that are typical of linguistic areas. That said, however, we can point to the fact that glottalization is shared by all the indigenous and—more significantly—most of the non-indigenous languages (or some of their dialects), and a number of bilateral or areally restricted multilateral contact-induced phenomena have been noted, mostly calques and lexical borrowings, but also, e.g., convergence in the use of personal pronouns (Nichols 1992; Dědělava 2005).

One could argue that as an area where languages accumulate without being replaced, the nature of linguistic areality differs from that of the Balkans, where the time-depth of accumulation is shallower, the expansion of lingue franca has been more frequent, and complete replacement of pre-Indo-European languages has occurred. Of course, the linguistic elephant in the room when discussing language contact in the Caucasus is Russian.

In the Soviet period, language contact studies were dominated by a Russifying agenda. The 2002 Russian census (Russia 2004) gives data on the knowledge of the language and of declared nationality and of Russian, but also provides some data for assessing multilingualism, e.g. Table 4.5 gives figures for the 24 nationalities over 400,000 (including Avar, Armenian, Azeri, Chechen, Dargi, Iagush, Karbardian, Kumyk, Lезgian, Ossetian, Tatar) for knowledge of languages other than Russian including, for example, Avar, Chech, Georgian, Karachay-Balkar, Lak, Ngats, Tagasaran. Table 4.6 gives the breakdown of these figures by administrative division, including the republics.

The 1999 Azerbaijan census (Azerbaijan 2006) gives data for knowledge of ethnic language, Azeri, Russian and English (Table 5.2). Of the ethnicities enumerated (figures in thousands of a total of 7953.4 of which 7205.5 were Azeri), Avar (50.9), Lезgian (178), Udi (41), Talysh (76.8), Tat (10) Muslims and 8.9 Mountain Jews, Tskhur (15.9), and Georgian (Ijigoli) (14.9), are covered in Clifton (2002, 2003), as are Budukh (3000 Jurcheva et al. 1999: 228), Kryz (10-5 000 Jurcheva et al. 2001: 154), and Kinhag (2500 in 1976 and the same in Jurcheva et al. 2005: 319), which were not included in the census separately.7 Mentioned in the census but not covered in Clifton are Armenians (1207; down from 390.5 in 1989), Tatars (30), Turks (43.4), and Kurds (13.1).8 For most nationalities, reported fluency in the ethnic languages was 95 percent or greater.9

The 2001 Armenian census (Armenia 2003) gives figures for Armenian, Yejidian,10 Russian, Ukrainian and “other” mother tongues all correlated with nationality and broken down by rural and urban. Other tables give correlations of ethno with gender, education,
age, urban/rural, using a different set of ethnicities: Armenians (3,145,354 = 97.89 percent), Assyrian (3409), Yezydin (40620), Greek (1716), Russian (14660), Ukrainian (16333), and Kurdsh (1519), Other (4640).
The 2002 Georgian census (Georgia 2002) gave only ethnicities broken down by munici-
palities. The total figures were Georgian (3,661,173), Abkhaz (3527), Ossetian (38,028), Armenian (248,929), Russian (67,671), Azeri (284,761), Greek (15,166), Ukrainian (7039), Kist (7110), and Yezydin (18,329).
Nicholls (1998) gives an excellent presentation of verticality in Caucasian multilingualism prior to the Russian conquest, which I quote here:

[In highland villages many people knew the language(s) of lower villages, but not vice versa. This was because markets and winter pasture were to be found in the lowlands, while the highlands afforded few economic advantages. The male population of highland villages was largely transhumant and spent perhaps half of its working life in the lowlands. Naturally, under these conditions, lowland languages tended to gradually spread uphill, reducing highland languages to islands and eventually replacing them entirely. At present and for all known history and known prehistory, languages with large numbers of speakers have both lowland and high-
land ranges and a generally elongate vertical distribution; these are economically advantageous and/or culturally prestigious languages that have spread uphill. Lang-
guages with small numbers of speakers, including several one-village languages, are mo-
stly found in the highlands. This pattern apparently predominated during the Little Ice Age (late middle ages to mid-19th century), a period of global cooling in which highland farms and pastures were economically precarious and the lowlands more prosperous. Prior to that, there is evidence that highland communities were larger and more prosperous and their languages spread downhill, and that highland communities formed and maintained lowland colonies. Chechen-Ingush isoglosses, and the discontinuous distribution of language families like Chechen-Ingush, Avar, and Lakh all point in this direction. Overall, then, geography and size of speech community are correlated, and this is explained by verticality, economy, and climate change.

After conversions to Islam, Arabic served as a lingua franca for educated classes. In the northern lowlands of Daghestan Kumyk was used, in the southern lowlands Azeri, while the Av婴-Komratov “army language” was used in the north-east highlands. Lak in the central highlands, Tabasaran and Dargwa to the south-west. Peoples living along the Caucasus ridge are likely to know languages from both sides of the ridge. Since political organization was by clan, lineage, and commune—although there were at various times an Avar khанate, a Lak shahkhahate, Kajtag and Tabasaran utsemates, etc.—allegiances were fluid and determined by desire and necessity. These coupled with patterns of endogamy contributed to a linguistically complex situation. We can also note that in the Muslim-majority republics of the Caucasus, as elsewhere in the Muslim ex-USSR, there is a resurgence and revival of Arabo-Perso-Turkic loans that had been eliminated by Soviet Russifying policies.

Language status, language planning, language and identity

Although a few languages of the Caucasus have long literary traditions and others have varying degrees of earlier written documentation, modern identity formation is connected

with the nation-state-building processes of the nineteenth and twentieth centuries that affected other parts of Europe and the former USSR. A primary source of identity for peoples in the North Caucasus was the fact of being a mountaineer as opposed to a lowlander. Attempts at polities in the North Caucasus approached language issues in terms of potential linque franche (Arabic, Kumyk, Azeri; under Shamyl, Avar) rather than a mono-ethnic nation-state language. Even today, multiplicity of languages can lead to a multiplicity of identities that are not necessarily language-determined. Friedman (1998) discusses how identity in Daghestan has been determined by ruling polity, economic access, or school language rather than home language. The 1926 Soviet census attempted to measure objectively the linguistic and ethnic composition of the USSR’s population, but every subsequent Soviet census reduced the number of language and identity categories from the decade before in connection with the ideology of gradually creating a single Soviet (and de facto Russophone) nationality. For the 2002 census, social scientists at the Russian Academy of Sciences worked with the Bureau of Statistics to create the most complete and accurate possible list of language and ethnic categories. Unfortunately, in Daghestan, where access to resources and power continue to be allotted by ethnicity (e.g. the larger the ethnicity, the more seats in par-
liament), ethnic elites risked erosion of their power bases if speakers of unwritten lan-
guages declared their ethnicity on the basis of mother tongue and/or declared their home language rather than school language as mother tongue. In response, a rumor was spread that Moscow politicians and scholars were seeking to destabilize Daghestan and dismember it. As Tilkov and Kistiev (2005: 286) state: “Many interpreted this as requiring that they affirm their Avaro-Dargi and general Daghestanian loyalty.” As a result, the 2002 census fails to reflect the true ethno-linguistic picture of Daghestan.
For example, in 2002, Botlikh, which numbered 3379 in 1926 and 4100 in 1938, had zero declaring Botlikh nationality and 54 in Daghestan declaring Botlikh mother tongue. Tilkov and Kistiev (2005: 279–80) give the following linguistic estimates: Ahd (c. 40000), Archi (1200), Akhwhak (6000+), Bagwal (5000+), Bezhta (6500), Botlikh (4500), Himukh (under 1000), Godoberi (c. 4000), Hunzib (c. 800), Tse (c. 8000), Karata (c. 7500), Tindi (c. 10000), Khwarshi (2500), Chamalal (c. 10000), Kubachi (23000) (see also M.E. Alekseev’s articles on these languages in Mixal’en’ko et al. 2003, where he cites estimates of 20,000 for Kaitag; the estimate for Megeb was 1500 in 1982). The official 2002 figures for the other Daghestanian languages of Daghestan are: Avar (784840), Lak (153373), Lezgin (397310), Rutul (29383), Tskhvir (9771), Agul (29399), Tabasaran (128391), and Dargwa (503523).
For the North Caucasus, Mixal’en’ko et al. (2002, 2003) is an extraordinary resource. Included are all the indigenous languages of the North Caucasus as well as the main Turkic (Karachay-Balkar, Kumyk, Nogai, Tatar) and Iranian (Ossetian and Tat) lan-
guages. Kabardino-Cherkessian (Circassian) is treated as one language, Kubachi and Kaitag each have separate entries, but Azeri is not included despite being official in Daghestan. Census statistics are from 1989 and the partial census of 1994. Each entry has 20 sections: (1) language names; (2) detailed statistical and geographic data, e.g. numbers of monolinguals, numbers and ages of bilinguals by gender, numbers considering the language native broken down by declared ethnicity; (3) general linguistic data including classification, distinctive structural and typological features, dialects, contact languages, and regional variants; (4) literacy and orthography (both current and historical); (5) status including attitudes of speakers, regional literature, legal status (including texts); (6) lit-
erature, including languages into which and from which works are translated, publication
endangered. At issue is not necessarily the absolute number of speakers but rather the fragile ecology of these languages owing to processes of modernization, urbanization, and the on-going effects of Russification or official language dominance. Caucasian endangered languages generally display vitality in the villages and morbidity in the towns, where Russian (or Georgian, Azeri, or Armenian in Transcaucasia) is the lingua franca, the language of higher education, the key to upward mobility and prosperity, and the majority urban language and the unmarked language in public contexts. In 2007, out of approximately 70 major languages in the Dagestanian Languages Department of Dagestan State University (where Russian is the language of instruction)—a group representing the future of Dagestani languages and language pedagogy in the republic—only four were from urban backgrounds (fieldwork observation). Another factor in language endangerment is the Caucasus violence and its consequences. There are three crucial periods that have altered the linguistic landscape of the Caucasus by disrupting patterns of language transmission: the Russian wars of conquest in the mid-nineteenth century, the famines and Stalin’s deportations of the mid-twentieth century, and the wars, rebellions, pogroms, and ethnic cleanings of the end of the twentieth century. Even in rural areas some small languages are losing ground to slightly larger ones owing to the breakdown of endogamy; Conrie (2008), citing D. Fisker, observes that Hinukh is losing ground to Bezha and Tsez owing to intermarriage.

Conclusion

Any field of sociolinguistic research in the region will prove fruitful for the future. Issues such as language standardization, language policy, linguistic identity, and to some extent language vitality and language contact are better served at this point than variationist studies, conversational analysis, code-switching, etc., and so one could point to precisely those areas with the most gaping lacunae in the Soviet period as areas still in need of the most work. At the same time, however, those areas that have traditionally been the focus of Caucasian sociolinguistics will also benefit from increased attention. In this respect, the shortcomings and/or absence of recent official census data both point to the most problematic areas and leave open the possibility of future advances.

Acknowledgments

I am grateful to Winfried Boeder, Jasmine Dum-Tragut, June Farris, Jost Gippert, Johanna Nicholls and Wolfgang Schulze, as well as John Colarusso, Anaïd Donabedian, Alice Harris, and Fatima Tislova who responded to my queries and helped me track down sources. All errors are my responsibility.

Notes

1 Kranodar, Stavropol and parts of Kalmykia are sometimes included in historical accounts of the region, 2 Schulze (2004: 9) includes Kalmyk (Mongol) transhumants (20,000) in northern Dagestan and Trukhman (Karapakhi; East Oghur, 5000 in 14 villages) in southern Dagestan. 3 The term Tats was used by (originally nomadic) Turkic-speaking peoples to refer to (Iranian-speaking) settled farmers from Crimea to Central Asia. As a language name, it currently applies to
References

Sociolinguistics in the Iranian world

William O. Beeman

Introduction

The Iranian language sphere consists of three currently recognized “core” varieties, and a number of peripheral varieties that diverge significantly enough to be thought of as separate “languages.” The core varieties are Standard Modern Persian,1 Dari and Tajik. Other Iranian languages are Kurdish, of which there are several varieties; Baluchi; Modern Sogdian, known today as Yaghnobi; and a variety of so-called Eastern Iranian languages, including Nuristani, Pashto, Pasha’i. Included as well are the languages of the Pamir mountains: Shugni (Shugnani), Yazgulami, Ishkashimi and Wakhi.

There have been a number of informal debates among Persian language specialists concerning the status of Tajik and Dari vs. standard Modern Persian. All linguists know that speech communities utilize a continuum of varieties of speech, and that the term “language” is more a political appellation than a scientifically accurate descriptor. The range of variation in Persian, Dari and Tajik communities is quite extensive, embodying regionalisms and borrowings from other language families. The term “register” has a special status in describing languages in that it represents a speech variety that is marked for particular specific occasions. Whereas Modern Persian and Dari are very close in form, Tajik has more divergent discourse structures. Based on fieldwork carried out in Tajikistan, I theorize that standard Persian as spoken in Iran has become a special register of Tajik marked for formal occasions such as political speech making, wedding orations, news broadcasts, and elevated scientific discourse. In this way the opposition between all the varieties of colloquial Tajik and standard Persian in Tajikistan resembles the diglossic opposition between dimitiki and katharevousa in modern Greek. In the discussion below I provide several examples, and speculate on the concretion and meaning of such shifting diglossia in the use of vocal speech registers.

A controlled comparison of “Persian,” “Dari” and “Tajik” is not very productive, since there is considerable regional variation within the three varieties. It is far more productive to explore the social and cultural relationships between these language varieties, and to provide a sketch of the development of these varieties in recent years.