seriously endangered by the *UNESCO red book on endangered languages*: Gagauz (Moldovan), Crimean Tatar, Noghay (Nogai), and West-Siberian Tatar

- *Caucasian*: Laz (a few hundred thousand speakers), Georgian (30 000 speakers), Abkhaz (10 000 speakers), Chechen-Ingush, Avar, Lak, Lezghian (it is unclear whether this is still spoken)
- *Indo-European*: Bulgarian, Domari, Albanian, French (a few thousand speakers each), Ossetian (a few hundred speakers), German (a few dozen speakers), Polish (a few dozen speakers), Ukranian (it is unclear whether this is still spoken), and these languages designated as seriously endangered by the *UNESCO red book on endangered languages*: Romani (20 000–30 000 speakers) and Yiddish (a few dozen speakers)
- *Neo-Aramaic* (Afroasiatic): Tūrōyo and Sūrit (a few thousand speakers each)
- *Languages spoken by recent immigrants, refugees, and asylum seekers*: Afroasiatic languages: Amharic, Somali, and Tigrigna; Niger-Congo: Lingala, Swahili, and various languages spoken in Nigeria; Indo-European: Russian and Farsi; and Altaic: Kazakh and Kirghiz (hundreds of thousands of speakers)
- *Foreign languages taught in secondary schools*: English, French, German, and Italian

See also: Arabic; Armenian; Azerbaijanian; Caucasian Languages; Endangered Languages; Greek, Modern; Kurdish; Sign Language: Interpreting; Turkic Languages; Turkish.

Bibliography

Andrews P A & Benninghaus R (1989). *Ethnic groups in the Republic of Turkey*. Wiesbaden: Dr. Ludwig Reichert Verlag.

Aydın Z (2002). 'Lozan Antlaşmasında azınlık statüsü; Farklı kökenlilere tanınan haklar.' In Kaboğlu İ Ö (ed.) *Azınlık hakları (Minority rights)*. (Minority status in the Treaty of Lausanne; Rights granted to people of different origin). İstanbul: Publication of the Human Rights Commission of the İstanbul Bar. 209–217.

Çağaptay S (2002). 'Otuzlarda Türk milliyetçiliğinde ırk, dil ve etnisite' (Race, language and ethnicity in the Turkish nationalism of the thirties). In Bora T (ed.) *Milliyetçilik (Nationalism)*. İstanbul: İletişim Yayınları. 245–262.

Grimes B F (ed.) (2000). *Ethnologue* (14th edn.). Dallas: SIL International.

Sadoğlu H (2003). *Türkiye'de ulusçuluk ve dil politikaları (Nationism and language policies in Turkey)*. İstanbul: İstanbul Bilgi University Publications.

Şener C (2004). *Türkiye'de yaşayan etnik ve dinsel gruplar (Ethnic and religious groups in Turkey)*. İstanbul: Etik Yayınları.

Turkic Languages

L Johanson, Universität Mainz, Mainz, Germany

Development and Classification

The Turkic language family was first attested in 8th century inscriptions. Turkic-speaking groups first appeared in the Inner Eurasian steppes, from where they moved to Central Asia, Eastern Europe, the Middle East, Siberia, etc. Because of their high mobility, Turkic expanded over a huge area.

The Proto-Turkic network of varieties was dissolved by an early split of Oghur or Bulgar Turkic. Its modern representative, Chuvash, a descendant of Volga Bulgar, differs from Common Turkic by specific phonetic representations, e.g., *r* and *l* instead of *z* and *š* in words such as *śěr* 'hundred' and *śul* 'year' (Turkish *yüz* 'hundred,' *yaş* 'age'). A second split is represented by Khalaj, which retains a reflex of Proto-Turkic **p-* as *h-*, e.g., *hadaq* 'foot.' Dialect splitting has led to further differentiation of Common Turkic.

There is no mutual intelligibility throughout the family today. The following division combines the current areal distribution with genealogical and typological features.

1. The Southwestern or Oghuz branch contains a western subgroup comprising Turkish, Gagauz, and Azerbaijanian (Azerbaijani, Northern and Azerbaijani, Southern), a southern subgroup comprising dialects of southern Iran and Afghanistan, and an eastern subgroup comprising Turkmen and Khorasan Turkic.
2. The Northwestern or Kipchak branch has a western subgroup comprising Kumyk, Karachay-Balkar, Crimean Tatar, and Karaim, a northern subgroup comprising Tatar and Bashkir, and a southern subgroup comprising Kazakh, Karakalpak, Kipchak Uzbek, Nogai, and Kirghiz (of different origin, but strongly influenced by Kazakh).
3. The Southeastern or Uyghur-Karluk branch has a western Uzbek subgroup and and eastern Uyghur subgroup.

4. The Northeastern or Siberian branch has a southern heterogeneous subgroup comprising Sayan Turkic (Tuvan, Tofan), Abakan (Yenisei) Turkic (Khakas, Shor), Chulym Turkic, Altai Turkic (Altai, Northern and Southern), and a northern subgroup comprising Yakut (Sakha) and Dolgan.
5. Chuvash is geographically situated in the north-western area (Volga region).
6. Khalaj is geographically situated in the southwest-ern area (central Iran).

Deviant languages in China are Salar, of Oghuz origin, Yellow Uyghur (Yugur, West) and Fu-yü (Manchuria), both of south Siberian origin.

One traditional classificatory criterion is the final consonant of the word for 'nine.' Its representation as *r* in Chuvash *tăxxăr* separates Oghur from Common Turkic (Turkish *dokuz*). The intervocalic consonant in the word for 'foot' divides most Northeastern languages, Chuvash, Khalaj, etc. from the rest, which exihibits -*y*- (Turkish *ayak*), e.g., Tuvan *adaq*, Khakas *azax*, Chuvash *ura*. Oghuz Turkic differs from the rest by loss of suffix-initial velars, e.g., *qal-an* [remain-PART] instead of *qal-ɣan* [remain-PART] 'remaining.' Final -*G* is devoiced in the Southeast (Uyghur *taɣ-liq* [mountain-DER] 'mountainous'), preserved in southern Siberia (Tuvan *daɣ-lïɣ* [mountain-DER]), and lost elsewhere (Turkish *dağ-lı* [mountain-DER]).

Most older linguistic stages are insufficiently known. Written sources, where available, provide no direct information on spoken varieties. Early Oghuz and Bulgar (East Europe, 6th–7th centuries) are unknown. There are no texts in the language of the Khazars (7th–10th centuries). Pecheneg and Kuman, predecessors of West Kipchak, are only known from loanwords, titles, and names.

Written Varieties

Turkic literary varieties have emerged in various cultural centers. Many older Turkic empires, how-ever, used foreign languages for administration (Sogdian, Persian). Muslim Turks often used Persian for poetry, and Arabic for religious and scientific writing. Russian has played an important role for many groups. The following main stages of written Turkic may be distinguished.

1. An older pre-Islamic East Old Turkic period (8th century–), is represented in inscriptions, manu-scripts, and block prints. East Old Turkic proper is documented in stone inscriptions (Orkhon Valley), which celebrate the rulers of the Second Eastern Türk Empire, in other inscriptions found in Mongolia and the Yenisei and Talas valleys, and also in a few manuscripts. The Old Kirghiz inscriptions are of this type. Old Uyghur is first recorded in the period of Uyghur rule over the Eastern Empire. Early Old Uyghur is attested in runiform inscriptions and manuscripts. From the 10th century on, Old Uyghur became the medium of a flourishing literary culture in the Tienshan-Tarim area, attested in texts of Buddhist, Manichaean, and Nestorian content.

2. A middle Turkic period comprises various early Islamic varieties.
   The first East Turkic written language, Karakha-nid (11th century–), developed in Kashgar, is close to Old Uyghur but lexically influenced by Arabic and Persian. Maḥmūd of Kashgar provides informa-tion (1073) on Karakhanid and other contemporary Turkic varieties.
   Khorezmian Turkic, used in the 13th–14th cen-turies in the Golden Horde and Mamluk Egypt, is based on the older languages but contains Oghuz and Kipchak elements.
   This tradition is continued in Chaghatay (15th century–). Early Chaghatay contains regional ele-ments of the Timurid area. Later, Chaghatay became the dominant written language of Central Asia, eventually conquering an immense area of validity and developing regional varieties.
   The first West Turkic written language is Volga Bulgar, insufficiently known from epitaphs of the 13th and 14th centuries. Information on early Kipchak Turkic is given in the Codex Cumanicus, compiled by Christians, and in dictionaries and grammars written in Mamluk Egypt and Syria.
   Oghuz Turkic is first represented by Old Anato-lian Turkish (13th century–), which was a subordi-nate written medium until the end of Seljuk rule. Old Ottoman is the initial stage of Ottoman, which begins with the foundation of the Ottoman Empire in 1307. In Azerbaijan a literary language developed from the 15th century on.

3. A premodern period (16th century–) begins with the development of regionally influenced written languages. Middle and Late Ottoman became the leading written language with an abundantly rich literature. Chaghatay continued to play a major role and remained the literary language of all non-Oghuz Muslim Turks until a century ago.

4. A modern period begins in the second half of the 19th century with the formation of regional written languages. The political division of the Turkic-speaking world in the 20th century and the language policies pursued in the Soviet Union, Turkey, China, and Iran had dramatic effects that increasingly obstructed transregional linguistic contacts. A dozen 'national' languages with a narrow radius of validity emerged. In Turkey,

Ottoman was replaced by modern Turkish. The social importance of many Turkic languages was very limited. After the recent political developments, their significance is rapidly increasing, but the varieties spoken in Iran, Afghanistan, Iraq, etc., still have poor possibilities to develop.

Various scripts and script systems have been applied to Turkic. A specific runiform script was created for Old East Turkic. Most Old Uyghur texts are written in Uyghur script, originating in the Near East and later taken over by Mongols and Manchus. It is similar to the Sogdian script, which is also used in Buddhist texts. A few Buddhist manuscripts are written in Brahmi script, Manichaean texts in Manichaean script, and Nestorian texts in Syriac script. Arabic script was used for the languages of the Islamic era (still used in China for Uyghur and Kazakh). A unified Roman-based script was introduced for several languages in the early Soviet period, but later replaced by different Cyrillic-based scripts. A Roman-based alphabet was introduced in Turkey in 1923. Most of the newly established Turkic republics have introduced or are introducing Roman-based scripts.

Contacts

The massive displacements of Turkic-speaking groups throughout their history have led to various phenomena induced by contacts with Iranian, Slavic, Mongolic, Uralic, etc. Speakers of Turkic have copied lexical, phonetic, morphological, and syntactic elements, whereas non-Turkic (e.g., Iranian, Greek, Finno-Ugric, Samoyedic, Yeniseian, Tungusic) groups shifting to Turkic have exerted substrate influence by copying native elements into their new varieties. Languages such as Chuvash, Yakut, Salar, Yellow Uyghur, Khalaj, Karaim, and Fu-yü have long developed in isolation from their relatives, preserving old features and acquiring new ones in their environments. Long and intense interaction with Iranian in Central Asia, Iran, Afghanistan, etc., has led to profound convergence phenomena. Massive foreign influence has sometimes caused considerable typological deviations, e.g., drastic structural changes in Karaim and Gagauz under Slavic impact.

Most written languages have been strongly influenced by Persian and Arabic. In Chaghatay (Chagatai) and Ottoman, lexical borrowing contributed to a remarkable richness of the vocabularies, whereas grammar was much less affected. The overload of Persian and Arabic in Ottoman led to strong puristic efforts in the 20th century to create a so-called Pure Turkish.

Internal convergence processes have resulted in leveling of languages of the central area. Several Turkic koinés have been used as transregional codes for trade and intergroup communication, e.g., Azerbaijanian in Iran and the Caucasus region.

Linguistic Features

Despite their huge area of distribution, Turkic languages share essential phonological, morphological, and syntactic features.

They have a synthetic word structure with numerous highly applicable derivational and grammatical suffixes, and a juxtaposing technique with clear-cut morpheme boundaries and predictable allomorphs. These agglutinative principles yield considerable morphological regularity and transparency. Exceptions include traces of vowel gradation in the pronominal declination, e.g., Turkish *ben* 'I,' *ban-a* [I-DAT] 'to me.' The agglutinative structure is partly deranged in languages of the northeast and southeast. Some languages, e.g., Uzbek, even display borrowed prefixes.

The syllable contains minimally a vowel with maximally one preceding and one subsequent consonant. Vowel hiatus and consonant clusters are avoided.

Most languages exhibit eight short vowel phonemes, *a, ï, o, u, e, i, ö, ü,* classified according to the features front vs. back, unrounded vs. rounded, and high vs. low. Proto-Turkic long vowel phonemes are preserved in Turkmen, Yakut, and Khalaj. Iranian and Slavic phonetic influence has sometimes affected the front vs. back distinctions. Tatar, Bashkir, Chuvash, and Uyghur exhibit systematic vowel shifts. Chuvash, Gagauz, Karaim, etc., have developed palatalized consonants, e.g., Karaim *mén* 'I'. Tuvan and Tofan exhibit a glottal element signaling strong obstruents, e.g., *aʔt* 'horse' vs. *at* 'name.'

The most general sound harmony phenomenon is an intrasyllabic front vs. back assimilation. An intersyllabic front vs. back harmony causes neutralization of the front vs. back distinction under the influence of the preceding syllable. If applied consistently, it excludes back and front syllables in a word, e.g., Turkish *ev-ler-im-e* [house-PL-POSS.1.SG-DAT] 'to my houses,' *at-lar-ım-a* [horse-PL-POSS.1.SG-DAT] 'to my horses.' Some languages only display this kind of harmony, whereas others also apply a rounded vs. unrounded harmony, neutralization of the distinction rounded vs. unrounded in high suffix vowels, e.g., Turkish *el-im* [hand-POSS.1.SG] 'my hand,' *gül-üm* [rose-POSS.1.SG] 'my rose.' Languages such as Yakut and Kirghiz apply this harmony to low-vowel suffixes as well, e.g., *börö-lör* [wolf-PL] 'wolves.' There are numerous exceptions to harmony rules in loanwords. Further allomorphs are created by various consonant assimilations.

The rules of word accent vary. A high pitch accent, interacting with a dynamic stress accent, mostly falls on the last accentable syllable of native words.

The morphological structure has remained relatively stable through the centuries. The main word classes are nominals (nouns, adjectives, pronouns, numerals) and verbals. The primary stems can be used as free forms, e.g., *at* 'horse,' *at!* 'throw!.' From verbal and nominal stems, which are sharply distinguished, expanded stems are formed. Nominals take plural, possessive, case, and specific derivational suffixes. Grammatical gender is not marked. The verbal morphology comprises markers of actionality, voice, possibility, negation, aspect, mood, evidentiality, tense, person, interrogation, etc. Voice is expressed by passive, reflexive-middle, causative, and cooperative-reciprocal suffixes. The order and combinability of suffixes is basically common to all Turkic languages.

Constructions with postposed auxiliary verbs (postverbs) express actional modifications. A few constructions have developed into aspect-tense categories, e.g., Turkish *gel-iyor* [come-PRES] 'comes' < *gel-e yorï-r* [come-CONV run-AOR] ('runs coming'). Possibility markers are formed with auxiliary verbs such as *bil-* 'to know' and *al-* 'to take,' e.g., Kirghiz *ber-e al-* [give-CONV AUX.POTEN] 'to be able to give.'

Turkic languages share many syntactic characteristics. With respect to relational typology, they adhere to the nominative-accusative pattern. They have a head-final constituent order, with dependents preceding their heads. The unmarked order of clause constituents is subject + object + predicate (SOV). Adjectival, genitival, and participial attributes precede the head of the nominal phrase. Postpositions are used instead of prepositions. There is no agreement in number or case between dependents and heads. The focus position is in front of the predicate core. The unmarked constituent order is often deviated from for discourse-pragmatic reasons. Contact-induced word order changes are common, e.g., in Gagauz, which has become an SVO language.

Preposed subordinate clauses are based on verbal nouns, participles, and converbs. The use of postposed subordinative patterns with conjunctions are typical effects of Iranian and Slavic influence. Most languages possess conjunctions, even coordinative ones meaning 'and,' 'or,' and 'but' of Persian, Arabic, or Russian origin.

Turkic lacks definite articles. The indefinite article is formally identical with the numeral 'one' Genitival attributes, expressing a possessor, stand in the genitive, whereas their head, indicating a possessed entity, carries a possessive suffix, e.g., Turkish *at-ın baş-ı* [horse-GEN head-POSS.3.SG] 'the head of the horse.' The dominant type of nominal compounds follows the pattern noun + noun + possessive suffix, e.g., Turkish *el çanta-sı* [hand bag-POSS.3.SG] 'handbag.'

All Turkic varieties exhibit numerous loanwords. Arabic and Persian loans are frequent in all Islamic-Turkic languages. The Iranian influence is strong in Uyghur, Uzbek, and varieties of Iran and Afghanistan. Many languages have been subject to considerable Mongolic and Slavic influence. Loans and calques from European languages have become increasingly important. The Turkic languages spoken in China exhibit old and recent loans from Chinese.

See also: Afghanistan: Language Situation; Arabic; Azerbaijan: Language Situation; China: Language Situation; Chinese; Iran: Language Situation; Iranian Languages; Russian Federation: Language Situation; Slavic Languages; Turkey: Language Situation; Uralic Languages; Uzbek.

Bibliography

Deny J *et al.* (eds.) (1959). *Philologiae turcicae fundamenta, 1*. Aquis Mattiacis: Steiner.

Johanson L (2002). *Structural factors in Turkic language contacts*. London: Curzon.

Johanson L & Csató É Á (eds.) (1998). *The Turkic languages*. London: Routledge.

Menges K H (1995). *The Turkic languages and peoples. An introduction to Turkic studies*. Wiesbaden: Harrassowitz.

Róna-Tas A (1991). *An introduction to Turkology*. Szeged: University of Szeged.

Turkish

R Underhill, San Diego State University, San Diego, CA, USA

Turkish (natively *Türkçe*), the official language of the Republic of Turkey, is spoken by a large proportion of the Turkish population. There are also Turkish speakers in the Balkans, particularly in Greece, Bulgaria, and the former Yugoslavia, although there has been extensive population inflow from those countries into Turkey, and there is a substantial minority of Turkish speakers in Cyprus. There are Turkish-influenced Turkic dialects in Iraq in the region of Kirkuk, where the speakers are called

Oghuz Turkic
Old Anatolian Turkish
Old Uyghur
Ottoman Turkish
Pecheneg
Sakha
Salar
Samoyedic
Sayan Turkic
Shor
Slavic
Sogdian
Tofan
Tungusic
Turkic languages
Tuvan
Uralic
Uyghur-Karluk
Volga Bulgar
Yakut
Yellow Uyghur
Yenisei Turkic
Yeniseian