1. How do we classify the Turkic languages? Which are the six main branches? Which languages belong to these branches?

Classification

The dynamic history of the Turkic-speaking peoples makes it difficult to set up a classification of modern Turkic languages that combines geographic and genealogical criteria. The following rough scheme represents an attempt to combine the current area distribution with genetic and typological features. It primarily distinguishes six branches, some of heterogeneous origin:

1. A southwestern (SW) branch, Oghuz Turkic
2. A northwestern (NW) branch, Kipchak Turkic;
3. A southeastern (SE) branch, Uyghur Turkic;
4. A northeastern (NE) branch, Siberian Turkic;
5. Chuvash, representing Oghur or Bulghar Turkic;

Subdivision

The six branches may be subdivided as follows:

Southwestern

SWw, a West Oghuz group comprising Turkish, Gagauz and Azerbijanian.

SWe, an East Oghuz group comprising Turkmen and Khorasan Turkic.

SWs, a South Oghuz group comprising dialects of Iran (Kashkay, Sonqori, Aynallu etc.) and Afghanistan (Afshar).

Northwestern

NWw, a West Kipchak group comprising Kumyk, Karachay, Balkar, Crimean Tatar and Karaim.

NWn, a North Kipchak or Volga-Ural group comprising (Kazan Tatar, Mishar, West Siberian) and Bashkir. West Siberian Tatar (Tura, Baraba, Tomsk, Tümen, Ishim, Irtish, Tobol, Tara and other dialects) is partly of different origin.

NWs, a South Kipchak or Aralo-Caspian group comprising Kazakh, Karakalpak, Kipchak Uzbek and Noghay. In spite of strong genetic bonds with Altay Turkic of the NEs group, modern Kirghiz may also be placed in the NWs group, since recent changes have brought it close to Kazakh. Though the so-called Fu-yü language of Manchuria is considered a Kirghiz dialect, it is closely related to South Siberian Turkic, e.g. Yenisey Turkic.

Southeastern

SEw, a western group comprising modern Uzbek and its various dialects, i.e. Oghuz Uzbek.

SEe, an eastern group comprising modern Uyghur, Taranchi and different Eastern Turki dialects of Kashgar, Yarkand, Khotan, Kerya, Cherchen, Aksu, Kucha, Turfan, etc. Deviating languages belonging to the SEe area are Yellow Uyghur, probably of NEs origin, and Salar, historically developed from the SW branch.

Northeastern

NEn, a North Siberian group comprising Yakut (Sakha) and Dolgan.

NEs, a heterogeneous South Siberian group comprising three subgroups:

1. Sayan Turkic consisting of Tuvan (Soyot, Uryankhai) and Tofa (Karagas);
2. Yenisey Turkic comprising Khakas, Shor and related dialects (Saghay Qaca, Qizil);
3. Chulym Turkic comprising dialects such as Kızır.

4. Altay Turkic comprising Altay (Oyrot) and dialects such as Tuba, Qamanda, Qu, Teleut, Telengit. Some Altay Turkic dialects are rather close to Kirghiz (NWs). As noted above, Yellow Uyghur may be of NEs origin, and Fu-yü displays similarities with Yenisey Turkic.

2. What are the traditional classificatory criteria? What features characterize the Oghuz languages?

One traditional classificatory criterion is the treatment of the final consonant in the word for ‘nine’. Its representation as r in Chuvash (tär) separates Oghur from Common Turkic, e.g. Noghay tär. The treatment of the intervocalic consonant in the word for ‘foot’ divides most NE languages from the rest, e.g. Tuvan adaq, Yakut ata/i Khakas aza/i vs. SW, NW, SE ayaq. Note, however, Khalaj hadaq, Yellow Uyghur azaq, Fu-yü azï. This criterion is less important than traditionally assumed. The Oghuz languages are divided from the rest by the loss of suffix-initial G, e.g. qalan instead of qalyan ‘remained’.

3. What are the main criteria for establishing genealogical relatedness between languages? What is a regular phonological correspondence? Give some examples of regular phonological correspondences between Turkic languages.

Languages change over time. As a result, when two groups of speakers that speak the same language are separated geographically, their languages will develop independently, yielding different dialects, and eventually different languages. Such languages are genealogically related. This is opposed to languages that are just areally related, i.e. languages that share similarities just because they are spoken in the same area. Assuming that sound change is regular, when we discover that words with similar meanings from two different languages show regular phonological correspondences, we take this as evidence that the two languages are genealogically related. We call such pairs of words cognates, a. cognates: words from different languages that are evidence for genetic relatedness because they (i) mean similar things and (ii) show regular phonological correspondences. Once it has been concluded that two or more languages are genealogically related, it is possible to attempt reconstruction of their parent language. This is done by means of the comparative method: once a regular phonological correspondence has been established between
3. What do you know about short and long vowels in Turkic? Which languages have maintained the old long vowels? Have how new length distinctions developed in Turkic languages? Which Turkic languages have diphthongs? Give some examples!

**Quantity.** There are quantitative distinctions between short and long vowels. Turkic had, at an early stage, a distinction between short and long vowel phonemes. Some researchers even assume that there were three quantitative types, long, half-long, and short vowel phonemes. But at least we must reckon with two, e.g. "a" and "a:, relatively short vowels as opposed to relatively long ones. A few languages at the periphery, Yakut, Turkmen and Khalaj, have preserved this opposition between long and short vowels. Turkish and the other languages have not preserved these oppositions. The present distinction between the short and long vowels in da: ‘also’, and daq ‘mountain’, pronounced [da] and [da:], is a secondary development.

**Proto-Turkic:**

<table>
<thead>
<tr>
<th>Length</th>
<th>Front Unrounded</th>
<th>Front Rounded</th>
<th>Back Unrounded</th>
<th>Back Rounded</th>
</tr>
</thead>
<tbody>
<tr>
<td>High</td>
<td>i</td>
<td>ü</td>
<td>ü:</td>
<td>u</td>
</tr>
<tr>
<td>Low</td>
<td>e</td>
<td>o</td>
<td>o:</td>
<td>a</td>
</tr>
</tbody>
</table>

We should put an asterisk, a small star, in front of these segments in order to show that they are not documented, but reconstructed.

Words with an a that goes back to a long a:

- Khalaj has a kind of diphthong, that is, there is a movement within the articulation of the vowel. Khalaj ha: ‘ra ‘interval’, cf. Turkmen a:ra, ta: ‘r ‘narrow’, cf. Turkmen da:r.
- Also in Turkmen and Yakut, the long vowels are mostly diphthongs, thus a: ‘, e.g. Turkmen da: ‘r ‘narrow’. The existence of old long a: is also mirrored in early documents. In the East Old Turkic inscriptions. And in Mahmūd al-Kāfî’s compendium from the 11th century (1073).

Concerning the development of the other long vowels, see your lecture notes.

5. What are the correspondences of Turkish "e", "u", "i", "ö", "o", "ü" and "u" in the other Turkic languages? Illustrate the correspondences with examples. (30)

- The short a sound is a low, back and unrounded vowel.
- The word for 'horse', a: (e), has this form in most Turkic languages. Also in East Old Turkic.
- There are two exceptions. Labialization is the first one. In some languages, a tends towards lip rounding. It is mostly a slight rounding. This labialized a is represented by the sign å.
- Uzbek (NE) has, e.g., åt ‘horse’, båš ‘head’. There are many exceptions (e.g. barmåq ‘finger’), which are hard to explain. Note that this å is not o, which might be suggested by the official orthography. For example, åt ‘horse’ is written ot, which is quite misleading. It has been assumed that the labialization in Uzbek is due to Iranian influence. We find a similar vowel å in Tajik. In Khalaj, all -a-sounds are labialized, e.g. ådå ‘horse’. This may be due to Persian influence. Tatar and Bashkir (NW) have å, stronger labialized in the first syllable, e.g. åt ‘horse’, båla ‘child’. In Chuvash, which is spoken in the same area, the labialization goes further, so that we have a development over å to o and in the standard language to u, e.g. ut ‘horse’. This is a rather recent development.

**High vowel.** The other exceptional phenomenon we find here is that the vowel corresponding to Turkish a is the high counterpart i. We encounter this in Tuvan (NE), Yakut (NE), and Chuvash. Tuvan (NE) tür- ‘to draw’ (cf. tär-), sip ‘handle’ (cf. sap), Yakut (NE) sît- (cf. yat-) ‘to lie down, to lie’, tîmîr ‘vein’ (cf. damar), Chuvash (aut) šîr- ‘to write’ (cf. yaz-), pîl ‘honey’ (cf. bal), idlan ‘gold’ (cf. altın), sûpå ‘handle’ (cf. sap), ôyt- ‘to say’ (cf. ayr-). These are phenomena occurring in certain words. The conditions under which i corresponds to a are unknown. See your lecture notes on the other vowels.