In August 1999 the North Eurasia Group of SIL International and the Institute for Bible Translation held a workshop for translators from ten Turkic languages: Azerbaijani, Turkmen, Tatar, Karakalpak, Kyrgyz, Uzbek, Khakas, Tuvan, Altai, and Yakut. None of the translators were trained linguists. Non-native speakers of the languages who worked with the translators also participated.

The purpose was to help the translators consider discourse-related structures relevant to translation. These included anaphora, use of pronouns, participant reference, focus and topic, word order, sentence length, background/foreground distinctions, use of conjunctions, cohesion, and the use of tense/aspect.

After each interactive lecture, participants worked in groups according to language. During these sessions, the participants applied the material that had been presented to actual texts.

The papers that developed from the workshop were published in Comments on Discourse Structures in Ten Turkic Languages, edited by Clifton and Clifton. Two other papers in the volume, those by Gray and by Humnick, represent work by members of the North Eurasia Group who were not able to attend the workshop. All the papers are presented as work papers. Each paper includes at least one interlinear text upon which many of the observations are based.

As much as possible, we have left the analyses as they were developed by the authors. All the interlinear texts are presented in a 3-line format consisting of the text as written, a line with morpheme breaks, and a line with glosses. We have not, however, standardized the representation of morphemes. Some authors use abstract forms, while others use the surface allomorphs.

We have standardized grammatical terminology as much as possible. For example, in a number of the languages we have made a basic distinction in the participles between perfect and imperfect rather than between past and present, and have consistently referred to aorist forms as present/future. We also decided not to break off derivational morphemes.

We have used gerund, the traditional term from Turkic linguistics, for forms which have been referred to elsewhere as converbs or serial verbs, but differentiate between various types. For example, the most neutral form is referred to as a conjunctive gerund, while other forms are simultaneous gerunds, sequential gerunds, gerunds of manner, and so on.

We trust these papers, especially the interlinear texts, will be of use to others involved in the analysis of Turkic languages.
Alphabets of Ten Turkic Languages

John M Clifton

The following tables show the realization of the various phonemes in ten Turkic languages: Azerbaijani, Turkmen, Tatar, Kumyk, Karakalpak, Kyrgyz, Uzbek, Khakas, Tuvan, and Yakut. In the case of Azerbaijani, two alphabets are shown: The current Latin-based alphabet and the previous Cyrillic-based alphabet.

Equivalent phonemes are not always phonetically identical. For example, the most common realization of the phoneme /s/ is [s] in most of the languages, but is [θ] in Turkmen. Such instances are indicated in the footnotes.

In some languages, multiple phonemes are realized by a single grapheme. For example, both /y/ and /u/ are realized as ゥ в Uzbek. Therefore, ゥ is in both columns for Uzbek.

In other cases, variants of a single phoneme are realized with different graphemes. These instances are shown by lightly shaded cells.

Finally, graphemes that are used only in borrowed words are indicated by more heavily shaded cells.

1 Where alternates are given, the first is used word-initially; the second is used word-medially.
2 Phonetically a mid vowel.
3 In literary Tatar, /a/ is labialised in word-initial syllables
4 Phonetic [i].
5 Phonetically between [o] and [u].

Comments on Discourse Structures in Ten Turkic Languages, pp 293-295
The following table illustrates the phonetic inventory of the languages covered in this section. The phonemic notation is that of the literature, supplemented in some cases by additional allophones: the phonemes are given in bold, the allophones in italics. The phonetic notation is that of the IPA.

---

6 'z' is used syllable finally; 'b' is used syllable initially (e.g., in мәсьәлә [mæsʔælæ] ‘matter’).

7 The affricate is used in Russian loans and in some dialects. In literary Tatar this is a palato-dental fricative.

8 The affricate is used in some dialects. In literary Tatar this is a palato-dental fricative.

9 Allophone of /k/ in the environment of back vowels. In loans, when [q] occurs in the environment of front vowels, η is used syllable-initially followed by the equivalent back vowel (e.g., /вәкаләт/ [вәкәләт] ‘authority’ is written вәкаләт, while кә is used syllable-finally (e.g., /тәкдим/ [тақдим] ‘suggestion’ is written тәкдим). In noninitial position this is [j].

10 The most noticeable distinction between /p t d/ and /b d/ is aspiration, not voicing.

11 Phonetically between velar and uvular.

12 Realized as [y] in certain environments.

13 The phonemes /s z/ are realized phonetically as interdental [θ ð].

14 All allophone of /q/ in the environment of back vowels in Tatar words. In loans, when [y] occurs in the environment of front vowels, η is used syllable-initially followed by the equivalent back vowel (e.g., /вәкаләт/ [вәкәләт] ‘miracle’ is written вәкаләт), while кә is used syllable-finally.

15 Uvular point of articulation.

16 In most dialect, this is pronounced as [h].

17 Uvular point of articulation.

18 Uvular point of articulation.
This final chart represents sequences of the semivowel /j/ followed by a vowel. These are not independent phonemes. In an number of cases, however, the sequence is represented by a single grapheme. Therefore, I am including this table. While this table is undoubtedly incomplete, especially in terms of the realization of phonemic sequences realized as sequences of graphemes, it should be useful in interpreting graphemes like Юю, Ёё, and Яя.

<table>
<thead>
<tr>
<th>Language</th>
<th>j</th>
<th>jy</th>
<th>je</th>
<th>ju</th>
<th>jо</th>
<th>jо</th>
<th>jo</th>
<th>ja</th>
<th>jæ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azerbaijani</td>
<td>М</td>
<td>M</td>
<td>N</td>
<td>N</td>
<td>L</td>
<td>L</td>
<td>r</td>
<td>j</td>
<td>w</td>
</tr>
<tr>
<td>Azer Cyr</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>J</td>
<td>J</td>
<td>j</td>
<td>w</td>
<td></td>
</tr>
<tr>
<td>Turkmen</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>Й</td>
<td>Й</td>
<td>Й</td>
<td>Й</td>
<td>Й</td>
</tr>
<tr>
<td>Tatar</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>Kumyk</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>Karakalpak</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>Kyrgyz</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>Uzbek</td>
<td>М</td>
<td>М</td>
<td>N</td>
<td>N</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
<td>Ы</td>
</tr>
<tr>
<td>Khakas</td>
<td>М</td>
<td>М</td>
<td>НЬ</td>
<td>НЬ</td>
<td>ЛЛ</td>
<td>ЛЛ</td>
<td>РР</td>
<td>Йй</td>
<td>Йй</td>
</tr>
<tr>
<td>Tuvan</td>
<td>М</td>
<td>М</td>
<td>НЬ</td>
<td>НЬ</td>
<td>ЛЛ</td>
<td>ЛЛ</td>
<td>РР</td>
<td>Йй</td>
<td>Йй</td>
</tr>
</tbody>
</table>

Where alternates are given, the first is used word-initially; the second is used word-medially.

Used between syllables, or in compound words between words, if the second syllable begins with /j/. For example, Марьям /marjæm/ ‘Mary’.